Received: 18 February 2022 | Revised: 28 July 2022 | Accepted: 2 September 2022

DOI: 10.1111/mila.12446

SUBMITTED ARTICLE

WILEY

How to think about higher-level perceptual contents

Daniel C. Burnston

Philosophy Department, Tulane Brain Institute, Tulane University, New Orleans, Louisiana, USA

Correspondence
Daniel C. Burnston, Philosophy Department, Tulane University, 1229 Broadway St., New Orleans, LA 70118, USA.
Email: dburnsto@tulane.edu

The standard assumption for what perception must do in order to represent a "higher level" content—say, *tiger*—is that it must represent the kind *as such*. I argue that this "as such condition" is not constitutive of what it means for a content to be "higher-level", and that embracing it produces a range of unfortunate dialectical consequences. After offering this critique, I give an alternative construal, the "extended perceptual space" view of higher-level contents. This view captures the phenomena targeted by the "higher-level content" thesis, is empirically supported, and avoids the negative dialectical consequences of the as such condition.

KEYWORDS

categorical perception, higher-level content, mental architecture, perceptual spaces, perceptual learning

1 | INTRODUCTION

The "higher-level content thesis" (HLCT) has been a focus in philosophy of perception for over 15 years. According to the HLCT, perception represents *kinds*—that is, it represents categories of objects "over and above" basic, agreed-upon perceptual properties like *redness*, *squareness*, and *brightness*. According to a canonical formulation (Siegel, 2006), the HLCT is meant to explain a subject's *recognitional dispositions*—their ability to recognize objects immediately on the basis of their perception. According to the HLCT, at least some of these dispositions are best explained by subjects' perceptions representing the recognized kinds.

But what *is* it for perception to represent a higher-level content? The standard answer in the literature is that in order to represent a higher-level content, perception must represent a kind *as such*. If perception represents kinds as such, then the HLCT is true; if not, then it is not. Unfortunately, given its dialectical centrality, there is no explicit definition or extensive discussion of this "as such" condition in the literature. And, I will argue, non-coincidentally, there is nothing approaching consensus as to whether the HLCT is true, what the limits of the phenomenon are, and to what philosophical uses the thesis can reasonably be put. Nonetheless, the

thesis *has* been put to a wide range of philosophical uses, including in debates about ethical perception (Audi, 2013; Reiland, 2018; Werner, 2016), perceptual justification (Siegel, 2013; Silins, 2013), theory of mind (Gallagher, 2008; Lavelle, 2012; Spaulding, 2015), social cognition (Ludwig, 2022; Toribio, 2018), and aesthetics (Majeed, 2018; Stokes, 2014).

It is thus time for a reassessment. In particular, I will argue that the adoption of the as such condition has led to a wide range of negative consequences for the debate. And, moreover, it is not definitive of "higher-level" content. So, it should be abandoned.

I proceed as follows. In the next section, I will attempt an articulation of the as such condition based on the literature. I contend that the standard construal of the as such condition is that it involves representation of a kind *independently of its features*. In Section 3, I outline a range of dialectical shortcomings of the condition. In Section 4, I articulate my proposed alternative construal, which I call the extended perceptual space view (EPS; for a related view, see Gauker, 2017). Whereas on the as such condition, perception must represent a category *independently* of lower-level features; on the EPS view perception represents a category as a *pattern of variation* amongst those features. I argue that the empirical evidence supports the existence of higher-level contents, so construed. In Section 5, I show how the EPS view overcomes the dialectical problems for the as such condition. In Section 6, I conclude with some thoughts about the philosophical upshot of embracing the EPS view of higher-level contents.

2 | THE "AS SUCH" CONDITION

The HLCT is an explanatory thesis. It assumes that, sometimes, subjects develop dispositions to categorize objects immediately on the basis of their perceptual experience. It then seeks to explain those dispositions as the result of perception representing kinds.[1] Often, this comes along with an intuition, which I broadly endorse, that the kinds must be in some sense represented "over and above" the features of the members. It is important to note, however, that this characterization leaves open exactly how perception might represent kinds. There is no stricture here about what type of representation perception must employ (e.g., the representation's structure, or whether it is conceptual or non-conceptual), or what semantic relationship the representation must have to the kind (e.g., whether it represents by description, by ostension, via structural mapping, etc.; for an explicit statement of this neutrality, see Siegel, 2006, p. 485).

Yet, there is a substantive assumption made by the literature, namely, that in order for perception to represent a kind, it must represent the kind *as such*. This "as such condition" has received little explicit formulation, and indeed is often delineated purely negatively. In order to represent a kind as such, the literature suggests, perception must *not* represent the kind solely as a combination of lower-level features. If, for instance, the perceptual representation of cypress trees comprises solely a binding or "congery" (Burge, 2014) of features that are distinctive of cypress trees—for example, the shape of their leaves, the width of their trunks, the color of their bark, and so on—then this cannot be higher-level content.

Further, in order to represent kinds as such, proponents suggest that perception must do more than represent *gestalts*, which are taken to be abstract features that are specific to a category. (Gestalts, in this sense, are distinct from the traditional Gestalt psychology principles taken to govern object and scene perception). While views differ on exactly what defines gestalts, the common

[1]There is a very complicated relationship, dialectically, between the HLCT and the idea that cognition permeates perception (Brogaard & Chomanski, 2015). I will not address this here, as I consider the relationship between categorical perception and cognitive permeation in detail in other work (Burnston, forthcoming).

properties attributed to them are that they are *abstract* and *invariant* (for the most complete formulation of this view, see Landers, 2021). That is, each member of the category instantiates, perhaps in different determinants, the *same* abstract gestalt property. On the gestalt hypothesis, subjects do not perceive higher-level contents as such, but instead learn to perceive category-specific gestalts.[2]

While I do not know of any explicit articulation of the as such condition in the literature, the dialectic suggests the following:

> The as such condition: A subject represents a kind, K as such when they represent K *independently of its features*.

What I mean by "independently" here is that the *representation of the kind* is distinct from and dissociable from the representation of an object's features. This is compatible with, in the perceptual case, the need for a kind representation to be affixed somehow to lower-level representations, or the view that perception can only recognize the kind on the basis of first representing lower-level features. The key is that the kind representation is functionally and semantically distinct from the feature representations also involved in representing an object.

I suggest that this characterization captures the extant dialectic. Proponents of the HLCT attempt to *exhaust* any explanation in terms of lower-level features, their combinations, or gestalts that they instantiate, leaving higher-level contents as the last hypothesis standing. This suggests that the higher-level content representation must be functionally distinct from those feature representations.

The as such condition is not always named explicitly, but is ubiquitous nonetheless. It is explicitly endorsed by Bayne (2009, 2016), who says that opponents of the HLCT "might claim that ... perceptual experience doesn't represent pine trees as such but instead represents only the distinctive spatial gestalt of pine trees" (pp. 116–117). Byrne (Siegel & Byrne, 2017), an opponent of the thesis, uses the "as such" terminology, in the way I have discussed, in his recent exchange with Siegel on the HLCT. Stokes (2014), in his discussion of aesthetic perception, suggests that learned art viewers come to recognize aesthetic properties like *gracefulness* as such.

Others who do not explicitly use the "as such" terminology *do* endorse the dialectic as I have laid it out. Siegel, for instance, does not use the term, but she also does not object to Byrne's construal of the debate in terms of the as such condition. Others employ language that goes very naturally with the condition. Fish (2013), for example, suggests that the question of higher-level contents is the question of whether there is an "interpretational, cognitive element to perception", suggesting that this latter function adds something on to distinct, non-interpretational components, such as feature representations. Burge (2014) and Block (2014), in their exchange, agree that the question is whether vision represents "higher-level attributives", with Burge linking these to conceptual representations. Mandelbaum (2017) also makes this connection.[3]

[2]I do not deny that we sometimes perceive gestalt properties, but I do not think these properties can account for the data surrounding categorical perception, as I will show below.

[3]Mandelbaum's view merits more attention than I can give it here. Mandelbaum argues that perceptual systems are modular, on the basis of being ballistic (i.e., fast and mandatory), and that subjects' ability to rapidly categorize items based on short presentation times suggests that perception outputs concepts. I disagree with both points—on an alternative conception of modularity (Burnston & Cohen, 2015), what makes a module is not its speed but the delimitability of its inputs. On this view, categorical perception of the kind I will outline below is paradigmatically *non*-modular. Moreover, fast presentation times do not entail fast perception, if one rejects the ballistic notion of modularity; hence, the argument from speed to conceptuality does not follow. This is of course an outline of general disagreement, but not an argument. I discuss the disagreements between mine and Mandelbaum's views in slightly more detail in Section 6.

Those who do not explicitly endorse this kind of conceptualism about higher-level contents (Siegel, 2013; Silins, 2013), do however conclude that higher-level contents are *similar enough* to concepts to directly justify the contents of beliefs.

I do not wish to engage with the conceptualism/nonconceptualism debate here, since doing so would require much more discussion of the nature of concepts. For provisional use, however, we can think of higher-level contents that meet the as such condition as a kind of "label" applied to an object by perception. This kind of representation would meet the as such condition, since a label is functionally and semantically distinguishable from lower-level features of the labeled object. It would also for that reason account for the "higher-level attributive" rendering, as well as those characterizations which explicitly build in a closer association with conceptual representation. The question is whether a view like this is the right way to characterize the HLCT, and under which to have the debate. The rest of this article argues that it is not, and provides an alternative.

Before going further, it is vital to point out that, if one wishes, one can simply *define* higher-level content in terms of the as such condition. If that is how one defines the issue, then any view that discusses higher-level contents on a different construal will simply be talking about something different, and any defense of the HLCT in other terms will simply be changing the question. Since this point is largely definitional, it is hard to give a knock-down argument against it, and it has some backing given the universality of the as such condition in the literature. However, counteracting this construal of the debate are the facts that: (i) the *phenomenon* that higher level contents are posited to explain, namely, subjects' developing recognitional dispositions on the basis of changes to their experience, is characterizable independently of the as such condition; and (ii) early discussions of the view by theorists such as Siegel (2006) and Lyons (2005) were non-committal about the structural and semantic properties of higher-level contents, so long as they represent categories/kinds. On my view, if there is some mental representation that is perceptual, represents kinds over and above their features, and explains the kinds of phenomena that the HLCT is posited to explain, then the thesis is established, the as such condition be damned. This is especially true if the alternative construal is interestingly revelatory for thinking about mental architecture.

So, I will assume that it is possible to have alternative construals of higher-level content. In the next section, I explain the negative dialectical consequences of the as such condition, and in Section 4 give my alternative proposal.

3 | THE DIALECTICAL BADNESS OF THE AS SUCH CONDITION

In this section, I will spell out some negative dialectical consequences of the as such condition. While I phrase them here as problems for the dialectic as a whole, I think they specifically accrue to proponents of the HLCT; the adoption of the condition leaves a difficult-to-defend view without clear boundaries and explanatory import. Before diving into specifics, it is worth talking about why the problems arise in general. Basically, they are all due to a problematic mental architecture implied by the condition. On the as such condition, the HLCT, if true, implies that perception *first* represents a set of lower-level features—everyone agrees that recognizing a tree, for instance, must proceed partially in virtue of representing its leaves, trunk, and so on—and *then* applies a separate, feature-independent representation of the category to them. It is this latter representation that is the higher-level content.

The central problem with this view is that it posits a representation that is a lot like a conceptual representation (in being independent of particular instances of feature combinations), but insists that it exists *in addition* to conceptual representations, and further that it must be represented *within* perception. But what constitutes the boundaries of perception and cognition is itself subject to debate (Beck, 2018; Green, 2020; Phillips, 2018), and the view that a label must be applied to already present feature representations makes the idea that it is really *perception* doing this hard to adjudicate. There are (at least) four specific ways in which this negatively affects the dialectic.

3.1 | Dialectical stalemate

There are (at least) two very popular types of objections to the higher-level content thesis. One type we may call *supervenience* objections, versions of which are given by Dretske (2015) and Burge (2014). These objections start from the architecture presented above, specifically the view that any application of a higher-level content must be to an already, independently, represented set of features. The objection then asks how a subject's perceptual system might learn to apply a higher-level content. Presumably, given that they could perceive the features perfectly well before, without recognizing the kind, the subject's new ability must come from some *change* to their feature representations. Either they must recognize new features, learn to combine previous feature representations in a new way, or their previous feature representations must change somehow—perhaps by relevant features being enhanced or more easily attended to. But, the objection continues, if change to recognitional dispositions must depend on a change in feature representations (hence the "supervenience"), higher-level contents are explanatorily inert. We can just directly explain the disposition with lower-level feature representations.

The second type of objection is *judgment-based* (Dretske, 2015; Reiland, 2014). Judgment-based objections suggest that there is a very straightforward answer to the question of how subjects learn a recognitional disposition—subjects learn *that* a certain combination of features is a reliable cue to category membership. But, the objection proceeds, this is exactly how *concept application* and *judgment* work. One applies a concept by recognizing that the membership conditions of a category are met and judging that the instance is a member of the kind. Even if this process becomes particularly *quick*, and even if it is *passive* (Reiland, 2014, 2015), the best explanation for category recognition is not a higher-level content within perception, but regular old concept application.

The two types of objection work compellingly well together—either recognitional dispositions are best explained via (perhaps modified) representations of lower-level properties, or they are best explained as (perhaps passive) judgments. In response to these kinds of arguments, proponents of the HLCT generally fall back on phenomenal considerations (Siegel & Byrne, 2017). But it is *very* doubtful that phenomenal methods are fine-grained enough to establish the presence or absence of multiple similar mental representations, one each on either side of the perception/cognition divide (cf., Butterfill, 2008; Pautz, 2008). The result is an unhelpful dialectical stalemate.

3.2 | Orthogonality to "higher-level"

Just as there is insufficient attention paid to what makes a content higher-level, there is similarly insufficient agreement about what makes a *kind* higher-level (cf., Bayne, 2009;

Brogaard & Chomanski, 2015; Lyons, 2005). Intuitively, the kind *tiger* is somehow more rarefied than the kind *striped object*, but there is no well-established view on what makes this so. Here are a couple of provisional considerations. The first is metaphysical: The category *tiger* is more unified than the category *striped object*. While there may be some subcategories of *tiger*, they are all closely related. The subcategories of *striped object* are both more numerous and more ontologically distinct. The second consideration is about mental architecture, and is based on the notion of representational dependence, the idea being that, as discussed above, the kind *tiger* can only be recognized on the basis of perception in virtue of first representing properties like *striped object*.

Let us grant for argument's sake that *tiger* is a higher-level category than *striped object* in both of these senses. The problem is that the as such condition is simply orthogonal to this notion of "higher-level". This is because any time the mind represents a property or kind on the basis of previously represented properties, we can ask the question of whether it represents the kind as such. Take a traditional "hierarchical" picture of perception, on which perception begins with representation of the simplest possible features—wavelengths, luminance, and displacements (Clark, 2004; but cf., Burnston & Cohen, 2013). Even the property of being striped will have to be represented in virtue of *some* combination of these more basic level features. So, is *striped object* ever represented as such? If we can ask the question of whether perception represents lower-level features as such, then the question of as such representation has nothing specifically to do with higher-level contents.[4]

To further cash this out, consider the possibility that perceptual features are represented in terms of "quality spaces" (Burnston, 2017; Clark, 2000; Rosenthal, 2010). The standard view of color perception, for instance, is that colors are represented within a space, defined by the dimensions of hue, saturation, and brightness (Berger, 2015, 2021; Isaac, 2013, 2014). On this view, the way a color, say *orange*, is represented is as a particular set of values along dimensions that, at other values, would also comprise other colors. It is not clear that representation of *orange* in a color space is representation of *orange* as such, since the representation of *orange* is not independent of the representations of the more basic features/dimensions that define the space. Indeed, this argument is the basis for my positive view of higher-level contents, fleshed out in much more detail below.

3.3 | Lack of limits

Would establishing that perception can represent *some* higher-level kinds entail that perception can represent *any* higher-level kind? While some would embrace this view (Carruthers, 2015; Prinz, 2006), it is dialectically undesirable for the relationship between a limited version of the HLCT and an unlimited version to be one of entailment. Someone for whom it is perfectly intuitive that perception can represent *tiger* or *airplane* might still reasonably balk at the idea that it represents *morally wrong*, or *philosophy major*. A reasonable construal of the HLCT should at least give us the resources to ask when and under what conditions the HLCT might extend to some contents, but not others.

It is very questionable whether the as such condition construal can do so. On the as such condition, a higher-level content is a type of label that perception has learned to apply to a

[4]I thus disagree with Bayne's claim that "it is beyond doubt that one can phenomenally represent squares as such" (2009, p. 401). It is not beyond doubt; it is very much a substantive assumption.

particular set of lower-level feature representations that is indicative of the kind. Since the label is distinct from the feature representations, there is no particular combination of or relation between features that needs to be in place for the categorical percept to occur—any feature combination sufficient to signal the kind could in principle be learned by perception.

This is exactly the starting point for Prinz's argument that perception can represent *philosophy major*. Prinz asks us to envision the following scenario: Suppose you ask all of the philosophy majors in the room to raise their hands. What you have done is established a reliable link between a feature, *hand-raisedness*, and a kind, *philosophy major*. According to Prinz, once you have learned this link, you can just *see* that the folks with hands raised are philosophy majors. Just like in the learning scenario where perception picks up on a reliable link between *orange*, *striped*, and *tiger*, here the learning scenario picks up on a similarly structured, if less intuitive, informational relationship. One might reasonably suspect that there is *something* wrong with this argument, but it is hard to see precisely *what* is wrong with it, if one construes the HLCT according to the as such condition.

3.4 | Lack of demarcation

If the HLCT is true, then representation of higher-level contents should be a distinctive mental phenomenon. That is, we should be able to tell it apart from other, perhaps closely related mental phenomena. Consider what I will call "bare association". This is where, through habits of associating particulars with kinds, one can quickly categorize them together. Suppose Fran and Sasha are both philosophy majors. I know this because they come to department colloquia, have taken multiple philosophy classes, and so on. Any time I see either Fran or Sasha, I will quickly and easily be able to recognize that they are philosophy majors. But, offhand, this is a very different phenomenon from that posited by the HLCT. One is inclined to say that there is nothing *perceptually* distinctive about the set, Fran and Sasha, that picks them out as philosophy majors. Similarly, consider the category *US presidents*.[5] We can certainly easily classify members of this kind together. And while there may be some features that the majority of them share in common (mostly being stuffy old white dudes), one might reasonably suggest that there is nothing distinctively "presidential" about their appearances (cf., Lyons, 2005; Reiland, 2015).

Arguably, the as such condition cannot account for a distinction between higher-level contents and bare association. There is no restriction, on the as such condition, as to how many features might be indicative of a kind, or on how unified they need to be. I might recognize a cypress tree by its distinctive leaves *or* by its distinctive bark. Similarly, I might recognize *philosophy major* by recognizing *Fran's face or* by recognizing *Sasha's face*. The as such condition is not in a position to distinguish between these two kinds of recognition. And yet, recognizing both Fran and Sasha as philosophy majors seems obviously a case of bare association, rather than higher-level content.

4 | THE EXTENDED PERCEPTUAL SPACE VIEW

Any proponent of the HLCT has to show how the explanatory aims of the HLCT are met—that is, explain a change in recognitional dispositions as due to a change in perceptual experience,

[5]Thanks to Max Coltheart for first pushing me to consider this example.

the best account for which is that perception is representing a kind. We can further endorse that it must, in some sense, explain how perception does this "over-and-above" representing feature congeries, or gestalts.

In this section, I will draw on empirical evidence to show how these explananda can be met in a way that is distinct from the as such condition. The view starts from the idea of a perceptual space, introduced above (Section 3.2). A space, most basically, is an ordering of instances. In the color space, particular instances of colors are put in ordered relationships. What puts them into those relationships is the "dimensions" that comprise the space. A dimension, in broad terms, is simply a set of ordered values. In a representational space, a dimension comprises an ordering of instances along some continuously varying quantity—say, the quantity of "hue" in the color space. Other dimensions that vision may represent include motion energy/velocity, orientation, and depth (cf., Green, 2020).

Let us suppose for the sake of discussion that lower-level properties, such as color, are represented via perceptual spaces. I propose that a version of the HLCT can be based on what I call *extended perceptual spaces* (EPS). To flesh this out, consider the following: What a dimension represents is *variation* among some quality. The "hue" dimension in color space is a way that colors can vary—varying in that way is part of what determines the phenomenal relationships between colors. What it means to be an *extended* space, as I define it, is to be a space comprising *higher-order* dimensions—ones that capture patterns of *co-variation* across other dimensions. The notion of higher-order dimensions is extremely common in the motor control literature (Burnston, 2021), and is, I suggest, implicit in the literature on categorical perception. According to the EPS, perception represents categories as locations within a space defined by these higher-order dimensions.

The key difference between the EPS version and the as such version of the HLCT is how they capture the intuition that a kind must be represented "over and above" its features. Where the as such condition suggests that the kind must be represented *independently* of its features, the EPS view suggest that a kind is represented as a distinctive *pattern of variation* amongst those features. In the next subsection, I will outline the relevant empirical methodologies used to study categorical perception, and in Section 4.2 will argue that the EPS meets the explanatory aims of the HLCT.

4.1 | Morphspaces and categorical perception

The studies I will discuss are based on the creation of a "morphspace". To produce a morphspace, experimenters begin with individual stimuli, called "parents". They then use an algorithm to continuously vary the features of one parent until they match the features of the other, generating new examples along the way. This can be done between two distinct pairs of parents at the same time, such that every example is a combination of the two sets of parents. That is, each example takes up a particular place both along the dimension defined by morphing parent 1 into parent 2, and along the dimension defined by morphing parent 3 into parent 4. The morphspace is the set of examples generated in this way. So, for instance, in the face space shown in the upper left in Figure 1, each individual face in the space is a combination of variation along "Dimension A", which consists of continuous feature variation between parents 1 and 2, and of variation along "Dimension B", which captures the same variation between parents 3 and 4.

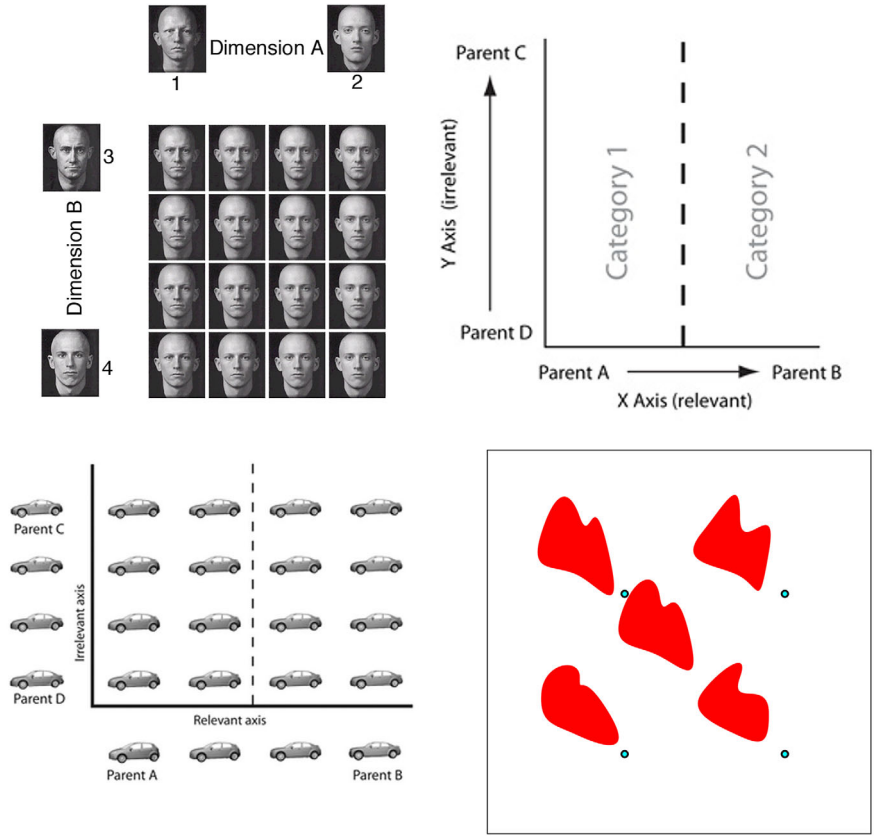

FIGURE 1 Morphspaces of faces, cars, and random blobs. (Top left and bottom right panels courtesy of Rob Goldstone, top left after Goldstone & Steyvers, 2001; top right and bottom left panels from Folstein et al., 2012)

Other morphspace examples are shown in the rest of Figure 1. I will often discuss face spaces (Valentine et al., 2016), and some have suggested that face perception is unique. But faces are not unique in the morphspace methodology. Morphspace-based studies have been used for a variety of other kinds of stimuli, from simple colored squares (Goldstone, 1994) to cars (Folstein et al., 2012), to fish, to random blob-shaped objects (Goldstone, personal communication) and all of them produce similar results. So, conclusions based on morphspace studies generalize across types of stimuli.

In a morphspace-based study, experimenters define arbitrary categories within the space of examples, and subjects learn to recognize members of those categories. In the examples in Figure 1, the experimenters divided the category by drawing a vertical "line" directly through the middle of the space, creating two categories, *Category 1* and *Category 2*, on either side of the line. Thus defined, it is the place along the dimension between parents A and B, referred to as the "X Axis" in Figure 2, upper right, that determines which category a face belongs to. I will call the dimension that determines category membership the "discriminating dimension".

Perceptually representing discriminating dimensions, I suggest, is what constitutes higher-level content. As I noted above, higher-order dimensions track patterns of covariation across lower-level features. For more controlled stimuli, such as the cars and shapes, the lower-order dimensions are easy to identify (shape, curvature, tilt). Faces, of course, are complex stimuli,

and there is not a universal account of what the relevant lower-order dimensions are for faces. Some features that are certainly relevant are the shapes of individual facial features (eyes, nose, ears) and the distances between them. I will not offer a full account here, nor need I do so for this project. What is important is that there *are* lower-level dimensions that vary, and that it can be proven that subjects' perceptual systems learn to track covariation in them.[6]

In the studies, subjects are trained via feedback. A subject sees an example, guesses the category to which it belongs, and is told whether they are correct. Subjects are generally at chance at the beginning of the experiment, but quickly learn to recognize the categories—usually not perfectly, but to a statistically significant degree. Importantly, subjects can learn to parse a range of different categories in a morphspace. That is, the "line" between categories could be drawn in a number of different ways, and thus different discriminating dimensions can be studied within the same morphspace.

The fact that subjects come to represent discriminating dimensions is generally shown by *transfer*. A subject representing a discriminating dimension needs to be distinguished from their simply learning that individual examples belong to a particular category. This is important because, in order to genuinely represent the category, a subject's recognitional ability must *generalize* beyond the examples already given—one must be able to recognize a new example as the same kind of thing as previous category members. Goldstone and Steyvers (2001) trained subjects on a morphspace comprising two dimensions, an X and a Y dimension, such that dimension X was the discriminating dimension. After subjects had learned the classification, the experimenters created a new morphspace, comprising dimension X and a novel dimension, Z, where X was still the discriminating dimension. Subjects could immediately make category distinctions along dimension X in the new morphspace, even though the faces, as a whole, were distinct in each space. To put it simplistically: The subjects could recognize the "X-ness" that separated the categories even in the novel stimuli.

Here is some context for this. Suppose one wants to recognize *angry* faces (Block, 2014). Not every angry face is equally angry, but there is a distinctive pattern of variation that marks off angry faces from non-angry ones. It is *this* pattern that one must recognize to pick out different angry faces, and one must recognize it despite all the other kinds of ways that faces can vary. What the experiment shows is that, even with a relatively short period of training, subjects can learn to represent a discriminating dimension and use it for recognition, even in novel stimuli where other dimensions differ.

Another, less intuitive sense of transfer is also important. Goldstone and Steyvers had some subjects learn an initial categorization in which the horizontal dimension was the discriminating dimension, and then categorize in a transfer condition using the same morphspace, but where the vertical dimension was the discriminating dimension. In the transfer condition, the discriminating dimension has "rotated" 90°, and is now a dimension that was irrelevant in the learning phase (Figure 2, left panel). Somewhat surprisingly, there is transfer in these conditions, compared to a rotation of 45°. Why would this be? The explanation is that learning differentiates not only the discriminating dimension, but the dimension orthogonal to it (cf., Jones &

[6]The lower-level dimensions for faces can be explored in a variety of ways. First, one can create *schematics* of faces, which eliminate much of the variation and attempt to show face perception effects based on idealized, simplified stimuli. Second, one can explore *adaptation* and *caricature* effects. That is, one can create a space where one knows the lower-level dimensions, and look for whether there are adaptation effects for the facial category of interest (individuals, expressions), or whether one can create caricatured faces by taking a particular face and "extending" it outward in the space by increasing the magnitude of certain dimensions. Each of these show the existence of face-space phenomena. (For fuller discussion, see Valentine et al., 2016).

Goldstone, 2013)—learning the category in the first condition involves separating the independent patterns of variation along the X and Y dimension. Since this discrimination is made already in the training, all subjects have to learn in the transfer is that the two dimensions they can already distinguish have switched relevance.

It is important for what follows that this kind of result is not obtained when faces only vary on lower-level features. Goldstone and Steyvers (2001) explicitly controlled for this possibility by comparing the previously mentioned studies with ones where only the mouths and eyes varied—that is, the X dimension was variation in eyes and the Y dimension was variation in mouths. They showed that, in these conditions, the transfer benefit for a 90° rotation, as opposed to 45°, did not obtain. The explanation for this is that what the original recognition requires is distinguishing two independent patterns of variation across the *entire* faces. This supports in general the notion that categorical perception is holistic (Gauthier & Tarr, 2002), and dependent on correlational and configurational relationships across the lower-level features of the stimuli.

So, I suggest that learning to represent discriminating dimensions underlies recognitional dispositions. But these processes also affect perceptual experience, as is shown by effects on *similarity judgments* and *discriminability*.

Several results suggest that objects are perceived as *more similar* when they belong to the same category. Subjects persistently judge objects as more similar when they have been categorized together. Of course, this result needs to be differentiated from strategic judgment—that is, from subjects simply rating objects as more similar because they have been grouped under the same label. Goldstone et al. (2001) tested this with a variation on similarity judgments in which subjects judged similarity between learned objects and a *neutral* object, which was not present during the training set. The idea was that, if within-category objects are perceived as similar to each other, then their perceived similarity would shift in a similar way relative to the neutral object. That is, they would both be perceived as either more similar or less similar to the neutral object after training. These results were indeed found, even though the neutral object belonged to no category, and hence there was no motivation for strategically modifying similarity judgments.

The ability to perceptually discriminate objects is often tested by asking subjects whether two sequentially presented stimuli are the same or different. Traditionally, the idea was that, if categorical perception occurs, discriminability should go *down* for intra-category objects, since they look more similar after training (Goldstone & Hendrickson, 2010). It turns out the situation is more complicated than this, however. Folstein et al. (2012) showed increased discriminability across the *entire* discriminating dimension in a morphspace, albeit with the greatest difference for stimuli immediately on either side of the category boundary. Similarly, Gureckis and Goldstone (2008) showed that, as categories are further divided into subcategories, a kind of balance is struck on which subcategories become discriminable from each other, whereas within-subcategory similarity goes down. While details remain to be cashed out, the point is that representations of discriminating dimensions are *pliable*—they can be "stretched" to accentuate relevant category similarities and differences, and this stretching affects subjects' perceptual experience of category members.

Recently, progress in the field has largely consisted in expanding and clarifying these results, and investigating their neural underpinnings (see, e.g., Folstein et al., 2015). Importantly for what follows, the kind of learning demonstrated by subjects is *flexible*, but not *unlimited*. Flexibility is shown in the fact that subjects not only can learn to differentiate arbitrary dimensions in the space, but that they can do so on the basis of a variety of exemplar sets. So,

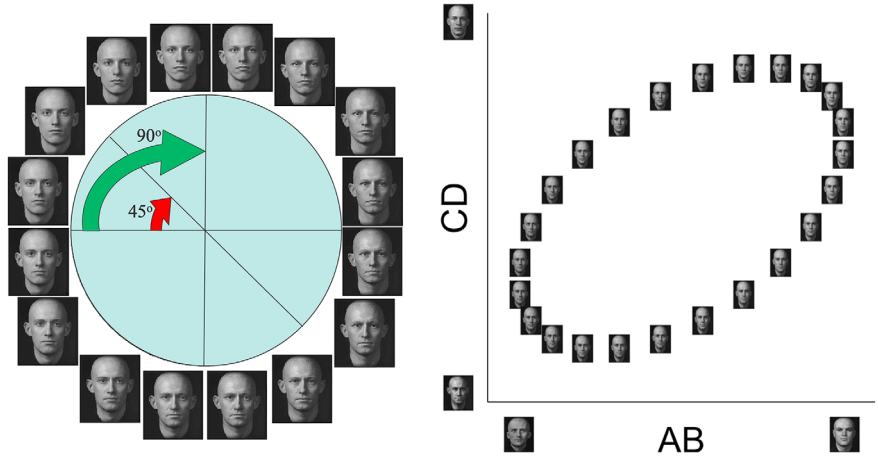

FIGURE 2 Flexibility in category learning (Images courtesy of Rob Goldstone; left panel after Goldstone & Steyvers, 2001; right after Jones & Goldstone, 2013)

the examples used in training can comprise a variety of "shapes" within the exemplar space, so long as the example set varies consistently with the discriminating dimension.

The two panels in Figure 2 show distinct samplings from complete morphspaces. That is, these are the faces that subjects were shown during training, drawn from the entire exemplar set. So long as there *is* a pattern of variation that constitutes a discriminating dimension within the exemplar space, subjects can learn that dimension even if the specific examples they are trained on differ.

While this flexibility is impressive, there are important constraints on it. Specifically, there have to be dissociable patterns of variation in the morphspace. In a morphspace like the ones pictured in Figures 1 and 2, there are distinct dimensions corresponding to the morph between pairs of parents. Given that "factorial" structure, subjects can learn to recognize a range of discriminating dimensions within the space. But in a "blended" space, such as the ones pictured in Figure 3, every parent is morphed into every other. In these circumstances, subjects cannot learn any discriminating dimensions. I will return to this point further below, when I revisit the question of the limits of higher-level content. For now, the question is whether the kind of representations evidenced here constitute evidence for the HLCT.

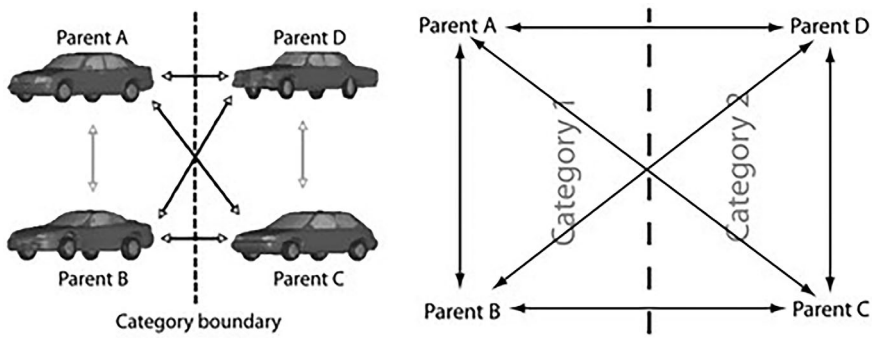

FIGURE 3 A blended morphspace (from Folstein et al., 2012)

4.2 | The EPS view and the HLCT

The subjects in morphspace studies develop new recognitional dispositions—for example, they come to recognize *A faces* or *B faces* (/cars, fish, etc.) that they could not before. If the interpretation of the similarity and discriminability studies above is correct, these abilities involve changes in subjects' perceptual experience. The question is whether these changes are best explained by perception representing the categories, over-and-above their lower-level features. We can break the question down further into three questions: (i) do subjects represent categories? (ii) are those categories represented over-and-above their features? and (iii) are the categories represented by perception? I suggest that the answer to all three questions is "yes".

Begin with (i). Subjects' recognitional abilities correspond to specific categories—that is, they learn a discriminating dimension that picks out A faces, discriminates A faces from B faces, and so on. This ability is distinct from the ability to recognize variation in lower-level features—that is why lower-level feature variation does not generate the same transfer and generalization results as variation across the whole of the exemplars. So, the recognitional disposition corresponds to a representation of the kind. (I say more about how to specify *which* kind is represented below.)

To drive this argument home, it is important to remember that, as discussed in Section 2, the explanatory aims of the HLCT do not specify *how* the kind must be represented; that is, they do not specify the semantic relationship that the representation must have to the kind. One might object to the EPS that a representation of a differentiating dimension does not result in a representation of *A faces*, but only in a representation of *An A-face-relevant pattern of feature variation*. But this objection is out-of-bounds, because it is not incumbent on the defender of the HLCT to specify exactly the semantic properties of the perceptual representation of the kind, only that the kind is represented over-and-above its features. To insist that the kind must be represented independently of its features would, of course, simply re-assume the as such condition, and hence beg the question.

Now, onto (ii). If subjects represent categories via discriminating dimensions, then their perceptions cannot be explained in terms of feature congeries or gestalts. Note that there is no specific feature congery that subjects recognize when they represent a kind via a discriminating dimension. Indeed, there is not even any particular *set* of feature congeries that they recognize. Remember, subjects can learn to represent discriminating dimensions from a wide range of different exemplar sets, so long as those sets vary along the discriminating dimension (Figure 2). So, no particular set of feature congeries is necessary for recognizing the kind. Moreover, representing a discriminating dimension is distinct from representing a gestalt held in common across all instances of a kind. A differentiating dimension is not defined by constancy, but by variation—it is that exemplars vary in a specific way, not that there is something they all share, which defines a kind representation on this view.

The answer to (iii) is the flip side of this point. A higher-order dimension is defined as a pattern of variation amongst lower-level features. Thus, a higher-order dimension is an "extension" of the space, because it comprises a new ordering within that lower-level space, one that captures a higher-order variation. But it is an extension *of that space*—representing a higher-order dimension requires representing the lower-level features that vary. Everyone admits that the lower-level properties are represented by perception, and the space represented here is an extension of that space, defined according to the same basic qualities in that space (albeit higher-order relationships between them). Hence, if the lower-level features are represented by perception, then the higher-order dimensions are too.

This is, admittedly, not a full argument to this effect. A full argument would require assessing the different proposals for how to construe the perception/cognition border, how to construe the percept/concept distinction, and so on, and I pursue this analysis in detail elsewhere.[7] For now, even if provisional, these considerations should provide significant motivation for counting the morphspace results as evidence in favor of the HLCT—as construed in the EPS version. Next, I will show how the EPS version overcomes each of the negative dialectical consequences of the as such condition.

5 | THE DIALECTIC REVISITED

5.1 | Dialectical stalemate

In Section 3.1, I argued that a stalemate has developed between proponents and opponents of the HLCT. Supervenience objections, recall, argue that the development of a recognitional capacity must go through some change to lower-level feature representations, and hence that purported higher-level contents are explanatorily inert. Judgment-based objections state that the kind of process involved in category recognition, on the as such condition, is best explained as a kind of (perhaps passive) judgment or concept application. The question is whether the EPS version of the HLCT enables perspicuous answers to these objections. I argue that it does.

The response to the supervenience objection relies on the answer to point (ii) above. It involves noting that, as revealed by the morphspace framework, categorical perception does not rely on any particular set of feature combinations, or any particular feature, to recognize a kind. Similarly, changes to particular feature representations do not result in representation of discriminating dimensions. It is only by capturing a pattern of co-variation *across* feature changes that these categories come to be represented, and this does not supervene upon specific lower-level representations.[8]

The response to the judgment objection relies on the answer to point (iii) above. If the positive characterization of discriminating dimension representations as occurring within extended perceptual space holds, then there *is* an explanatory role for rich percepts in the account of categorical perception, and it is not best explained as an instance of judgment or concept application. Now, it is of course true that subjects *do* learn a new concept in the course of these studies, for instance the concept of an "A-face". However, possession of this concept is not sufficient to explain the recognitional ability—the only way that subjects can accurately apply that concept is if they come to represent the discriminating dimension, and the discriminating dimension is distinct from the lexical concept of an "A-face". The lexical concept itself does not explain *any* of the particular transfer and generalization results that are evidence of

[7]One might note here that there are views of *concepts* that pick them out as part of a representational space, particularly those of Churchland (1996) and Gardenfors (2004). So why not just say that extended perceptual spaces are where concepts reside? One important consideration is that such views do not meet many of the standard conditions on concepts, such as the generality constraint. There are also kinds of representations that may not clearly fall into either the perceptual space type of representation I have discussed, or into conceptual representation—Beck (2014) gives a compelling argument to this effect in discussing analogue magnitude representations.

[8]Of course, variation in lower-level features may be sufficient to represent *some* categories; for instance, Goldstone (1994) constructed categories of objects defined solely by variation in size or brightness, and showed results similar to the morphspace cases discussed above. But the sufficiency of lower-level feature variation for some categories does not entail that they are sufficient for all categories, as Siegel (Siegel & Byrne, 2017) points out.

representation of the kind. So, the concept application relies asymmetrically on the presence of the higher-level perceptual representation.

5.2 | "Higher-level"

I argued in Section 3.2 that the as such condition is orthogonal to the notion of higher-level kinds. I offered two possible notions of "higher-level", both of which I think are implicitly operative in the literature. The first was the notion of metaphysical unity—the kind *tiger* is more unified than the kind *striped object* because there are fewer distinct subkinds comprising the former than the latter. The second is the idea of representational dependence, on which one kind is higher-level relative to another if one must be able to represent the second before one can represent the first.

The EPS view captures "higher-level" in both of these respects. Take face space again as an example—any face in a face space belongs to a broad category, namely *faces*. This category comprises *many* different subtypes (discussed further below). However, once one comes to recognize *A-faces* that kind comprises many fewer subtypes than the more general *face*, and is hence higher-level on the first construal. The second characterization is also met. Ability to represent the discriminating dimension requires, first, having the ability to represent the lower-order dimensions that comprise the objects in the space. If one could not recognize these individual patterns of variation, then one could not learn to parse the patterns of co-variation that constitute the discriminating dimension. Hence, the EPS version makes sense of the representation of kinds as "higher-level".

5.3 | Limits and boundaries

I will assess the last two dialectical consequences—lack of limits, and inability to distinguish higher-level contents from bare association—together, since the answer to both lies in the kind of representational structure that comprises higher-level contents on the EPS view. Recall that discriminating dimensions cannot be represented *ex nihilo*—they cannot be learned without the right kind of variational structure amongst the exemplars. The key to assessing whether a particular kind can be represented in perception is to consider whether there is a distinctive pattern of variation corresponding to that kind. If there is none, then we are likely to use other mechanisms, such as bare association, to recognize it. Bare association occurs when exemplars are grouped together without any distinctive pattern of variation connecting them and discriminating them from other kinds.

Contrast two examples: the kind *angry faces* versus the kind *US Presidents*. I suggest that the EPS view gives the resources for ruling in the former as a higher-level content, and ruling out the latter. The key to assessing whether there is a discriminating dimension is that, if a kind is represented as part of an EPS space, there will be certain patterns of *generalization* and *interpolation*. Generalization means that new instances will be recognized, without any further information, as members of the kind. Interpolation means that new instances will be represented as in ordered relations to other examples. So, when one represents *angry faces*, a new face will not only be immediately recognized as angry, even if no further information is presented along with it, but the new face can be recognized as more or less angry than some other face. It

is the ordering along the discriminating dimension that produces this kind of outcome (and, arguably, the kind of adaptation effects discussed in this case by Block, 2014).

With *US presidents*, however, there is no distinctive pattern like this. Just grouping some examples under the kind—say, recognizing the US presidents from the last 70 years, plus Lincoln—entails nothing about the ability to recognize new instances. Further, there is no ordering or variational structure in the examples such that a new instance will be seen as more or less "presidential" than another. I suspect that most Americans will not recognize a picture of Van Buren, for instance, as a US president, and even if they are told he was a president they would likely puzzle at the question of whether he looks more or less presidential than Carter. The EPS version gets the right answer in this case—on this view, *angry face* is a good candidate higher-level content, and *US President* is not. What allows us to categorize an object as a US President, when we can, is that it is one example we have lumped together with the other instances. That is, it is a case of bare association.

So, what makes the kind *angry face*, rather than some other kind, the one represented in the angry face case? It will again be helpful to compare the two cases, along with the following stricture: one should identify a perceptual content with the *lowest-level kind* for which there *are* distinctive patterns of generalization and interpolation. Notice that there *are* kinds that (many) US presidents are instances of, that might be represented perceptually—there is a lot of evidence that there are distinctive patterns of variation for *male* faces and *Caucasian* faces, for instance (Valentine et al., 2016). For these kinds, examples of US presidents *will* fall into patterns of generalization and interpolation. (A face can be more or less stereotypically male, Caucasian, etc.) But there is no such pattern for *US president*. In the *angry face* case, conversely, the patterns of generalization and interpolation are arguably specific to that kind. There is no lower-level category that will explain this specific pattern, the ability to differentiate these particular faces from similar-level kinds, such as *sad face*, and so on.

This leaves open a lot of interesting questions about how rich perceptual contents can get in particular instances. I am going to suggest, speculatively, that in most realms of human expertise there will be developed perceptual categories, but that these will not correspond to the most abstract categories that could characterize the domains. In addition to facial perception, we should expect a wide range of other biological and social categories to have correlates in perception (Marchi, 2015; Toribio, 2018). For instance, the widespread phenomenon of biological motion shows that it distinguishes a number of fine-grained categories. In point-light studies, subjects recognize not only a *kind of act* (e.g., walking, kicking, running), but can recognize the age and gender of the person performing it (Pollick et al., 2005).

This suggests a general pattern of variation in bodily movements that characterizes walking, with sub-clusters (as in Gureckis & Goldstone, 2008, discussed above) for when those are performed by differently gendered or aged individuals. We might posit other candidates in other modalities, for instance hearing a remark as *rude* or *cutting*, although a similar argument would have to be made in each case. While social perception may involve very rich contents, then, it is questionable whether this will extend to social/moral properties like *benevolence* or *goodness*.

Again, I am not trying to fix a specific limit on higher-level contents here, but to show how the EPS view allows us to have empirically tractable dialogue about specific proposals. Observe Figure 4:

Consider the far-left and far-right paintings, which are early Renaissance and Mannerist, respectively. While it is not easy to characterize just *one* feature that distinguishes them, there are many patterns of variation that correlate to make the one distinguishable from the other. People who are practiced in viewing these artforms, I suggest, might not only easily distinguish

FIGURE 4 Madonnas and children. (left) Filippo Lippi, Madonna and child; (center) Domenico Ghirlandaio, The virgin and child, (right) Parmigianino, Madonna with the long neck

one from the other, but also "place" the Ghirlandaio in the center as more like the Lippi on the left than the Parmigianino on the right.

So, is there a perceptual category for Mannerist artworks? I think it is likely that there is, but I do not want to push the argument too far in this venue. What is important is that the EPS view gives us ways of *determining the answer* that go beyond intuition pumps or phenomenal reflections. On the EPS view, what is required to show the existence of higher-level contents is to describe a feature space that the paintings inhabit, and to offer evidence that subjects' perceptions—shown, for instance, in their ability to categorize novel exemplars, in transfer effects, and in similarity/discriminability judgments—are best explained by their learning to track higher-order dimensional variation in that space. In the Mannerist artwork case, art historians suggest that the features that are distinct from earlier genres are the richer *contrast* of the colors, the extended *depth* in the scenes, and the *torsion* of the human figures in the depictions.[9] On the HLCT (EPS version), practiced art viewers' perceptions could be shaped by covariation along these distinctive features. If so, then there is a higher-level percept for Mannerist artworks.

6 | PERCEPTUAL ARCHITECTURE, PERCEPTUAL JUSTIFICATION, AND "DIRECTNESS"

One motivation for endorsing the HLCT has been to espouse a kind of mental architecture characterized by what we might call the "directness" thesis, on which enhanced perception feeds directly into cognition by outputting higher-level contents to belief-formation mechanisms. I suspect, although I have not argued here, that this posited direct-take-up-ability by cognition is part of what inspires the as such condition. Further, it motivates the close association between higher-level contents and conceptual contents which is standard, if occasionally non-explicit, in the literature.

[9]Thanks to Holly Flora for discussion of this category.

The "directness" claim has two further specifications—an epistemological and an architectural one. In epistemology, the issue is how we construe perceptual justification. If perception can output higher-level contents that are directly taken up into beliefs, then one might think that the range of non-inferentially justified beliefs is expanded (Silins, 2013), or that perception itself is evaluable in terms of belief formation (Siegel, 2013). In cognitive architecture, one might think that the role of perception is to give us the quickest route possible to concept tokening, so that thought is immediately enabled by perceptual engagement with the world. On both of these motivations, one is inclined to posit as close a relationship as possible between the content of a percept and the content of the concept/belief it occasions.

The EPS view sits poorly with the directness thesis, for a couple of related reasons. First, morphspaces do not come neatly packaged into discrete categories. A *grouping* within a psychological space is not the same as an explicit belief that an exemplar falls under the space. Hence, there is an architectural gap between the percept and the belief. If the EPS view is right, then this space needs to be theorized about, and its upshot for epistemic justification and cognitive architecture debated. Second, psychological spaces are not independent from each other. Consider a given face, say, Fred's face. Depending on what Fred looks like, there is not just *one* distinctive perceptual space he will fall into, but in fact many. Fred's face may look to be *angry*, *middle-aged*, and *Caucasian*, not to mention simply looking like *Fred*. This suggests that there is no *single* concept or belief that is determined by a given percept, but that a given percept is compatible with a range of concept tokenings in a given instance.

So, is the directness thesis true? There is evidence on either side. For instance, Mandelbaum (2017) argues for directness on the basis of the speed of category judgments, and their specificity to "basic level" categories—for example, on the basis of a very fast presentation, one might be able to see "guitar" without perceiving "Les Paul" or "Stratocaster". To contravene this view, the EPS proponent can make (at least) two responses. First, the EPS proponent can question whether these results are evidence of specificity in the relationship in all circumstances, or only in speeded presentations. Fast presentation times may allow for only the extraction of low-frequency information, which only supports distinctions in a lower-level space or via gestalts. Second, higher-level contents are fundamentally the result of perceptual *learning*, and studies of expert perceivers, as well as perceivers who have had training with specific morphspaces, suggest that after learning subjects begin to spontaneously categorize at more specific levels (Gureckis & Goldstone, 2008; Tanaka et al., 2005). If so, then which concepts are tokened on the basis of which percepts is at least malleable.

This is not the place to finally adjudicate these issues. Again, my primary purpose here is to show that having the EPS version of the HLCT on the table opens up useful dialectical space that is missed when one endorses the as such condition.

7 | CONCLUSION

I have argued that extant discussions of higher-level content in the literature erroneously define it in terms of the "as such" condition, and that this has left lacunae in the space of available positions on the issue. The alternative, extended perceptual space view that I propose fills this dialectical space, offering an empirically supported account of higher-level contents that allows for philosophically interesting analysis in both epistemology and philosophy of psychology. It may have further upshot for other philosophical uses to which the HLCT has been put.

ACKNOWLEDGEMENTS

This article benefitted from discussions with many colleagues over the long course of its development. The first draft of the article was produced during a visit with the "From Perception to Belief and Back Again" Research Group at Ruhr University in Bochum. I am exceedingly grateful to Peter Brössel, Insa Lawler, Giulia Martina, Luca Moretti, Stephan Pohl, Jim Pryor, Filippo Vindrola, and Alfredo Vernazzini for helpful comments and discussion. Subsequent drafts were presented at The University of Edinburgh, the University of Mississippi, and the 2022 meeting of the Southern Society for Philosophy and Psychology in Mobile, Alabama. I am grateful to the audiences from each venue, and especially to Lesley Walker for her insightful commentary at the SSPP. The final version benefitted greatly from comments from Jake Beck, Alistair Isaac, Dylan Ludwig, Don Oxtoby, and Indrek Reiland.

DATA AVAILABILITY STATEMENT

There are no data available.

ORCID

Daniel C. Burnston https://orcid.org/0000-0001-8281-3256

REFERENCES

Audi, R. (2013). *Moral perception*. Princeton University Press.

Bayne, T. (2009). Perception and the reach of phenomenal content. *The Philosophical Quarterly*, *59*(236), 385–404.

Bayne, T. (2016). Gist! *Proceedings of the Aristotelian Society*, *116*(2), 107–126. https://doi.org/10.1093/arisoc/aow006

Beck, J. (2014). Analogue magnitudes, the generality constraint, and nonconceptual thought. *Mind*, *123*(492), 1155–1165.

Beck, J. (2018). Marking the perception–cognition boundary: The criterion of stimulus-dependence. *Australasian Journal of Philosophy*, *96*(2), 319–334.

Berger, J. (2015). The sensory content of perceptual experience. *Pacific Philosophical Quarterly*, *96*(4), 446–468.

Berger, J. (2021). Quality-space functionalism about color. *The Journal of Philosophy*, *118*(3), 138–164.

Block, N. (2014). Seeing-as in the light of vision science. *Philosophy and Phenomenological Research*, *89*(3), 560–572.

Brogaard, B., & Chomanski, B. (2015). Cognitive penetrability and high-level properties in perception: Unrelated phenomena? *Pacific Philosophical Quarterly*, *96*(4), 469–486.

Burge, T. (2014). Reply to block: Adaptation and the upper border of perception. *Philosophy and Phenomenological Research*, *89*(3), 573–583.

Burnston, D. C. (2017). Cognitive penetration and the cognition–perception interface. *Synthese*, *194*(9), 3645–3668. https://doi.org/10.1007/s11229-016-1116-y

Burnston, D. C. (2021). Bayes, predictive processing, and the cognitive architecture of motor control. *Consciousness and Cognition*, *96*, 103218. https://doi.org/10.1016/j.concog.2021.103218

Burnston, D. C. (2022). Perceptual learning, categorical perception, and cognitive permeation. *Dialectica*.

Burnston, D. C., & Cohen, J. (2015). Perceptual integration, modularity, and cognitive penetration. In *Cognitive influences on perception: Implications for philosophy of mind, epistemology, and philosophy of action*. Oxford University Press.

Burnston, D., & Cohen, J. (2013). Perception of features and perception of objects. *Croatian Journal of Philosophy*, *12*(3), 283–314.

Butterfill, S. (2008). Seeing causings and hearing gestures. *The Philosophical Quarterly*, *59*(236), 405–428.

Carruthers, P. (2015). *The centered mind: What the science of working memory shows us about the nature of human thought*. Oxford University Press.

Churchland, P. M. (1996). *The engine of reason, the seat of the soul: A philosophical journey into the brain*. MIT Press.

Clark, A. (2000). *A theory of sentience*. Oxford University Press.

Clark, A. (2004). Feature-placing and proto-objects. *Philosophical Psychology*, *17*(4), 443–469. https://doi.org/10.1080/0951508042000304171

Dretske, F. (2015). Perception versus conception: The Goldilocks test. In J. Zeimbekis & A. Raftopoulos (Eds.), *The cognitive penetrability of perception: New philosophical perspectives* (pp. 163–173). Oxford University Press.

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

Folstein, J. R., Gauthier, I., & Palmeri, T. J. (2012). How category learning affects object, representations: Not all morph spaces stretch alike. *Journal of Experimental, Learning, Memory & Cognition*, *38*(4), 807–820.

Folstein, J. R., Palmeri, T. J., Van Gulick, A. E., & Gauthier, I. (2015). Category learning stretches neural representations in visual cortex. *Current Directions in Psychological Science*, *24*(1), 17–23.

Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and Cognition*, *17*(2), 535–543.

Gardenfors, P. (2004). Conceptual spaces as a framework for knowledge representation. *Mind and Matter*, *2*(2), 9–27.

Gauker, C. (2017). Three kinds of nonconceptual seeing-as. *Review of Philosophy and Psychology*, *8*(4), 763–779. https://doi.org/10.1007/s13164-017-0339-2

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 431–446.

Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, *123*(2), 178–200.

Goldstone, R. L., & Hendrickson, A. T. (2010). *Categorical perception*. Wiley Interdisciplinary Reviews: CognitiveScience, 1(1), 69–78. https://doi.org/10.1002/wcs.26

Goldstone, R. L., & Steyvers, M. (2001). The sensitization and differentiation of dimensions during category learning. *Journal of Experimental Psychology: General*, *130*(1), 116–139.

Goldstone, R. L., Lippa, Y., & Shiffrin, R. M. (2001). Altering object representations through category learning. *Cognition*, *78*(1), 27–43.

Green, E. (2020). The perception-cognition border: A case for architectural division. *Philosophical Review*, *129*(3), 323–393.

Gureckis, T. M., & Goldstone, R. L. (2008). The effect of the internal structure of categories on perception. Presented at the *Proceedings of the 30th Annual Conference of the Cognitive Science Society*.

Isaac, A. M. (2013). Quantifying the subjective: Psychophysics and the geometry of color. *Philosophical Psychology*, *26*(2), 207–233.

Isaac, A. M. (2014). Structural realism for secondary qualities. *Erkenntnis*, *79*(3), 481–510.

Jones, M., & Goldstone, R. L. (2013). The structure of integral dimensions: Contrasting topological and Cartesian representations. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 111–132.

Landers, C. (2021). Specialized visual experiences. *The Philosophical Quarterly*, *71*(1), 74–98. https://doi.org/10.1093/pq/pqaa018

Lavelle, J. S. (2012). Theory-theory and the direct perception of mental states. *Review of Philosophy and Psychology*, *3*(2), 213–230.

Ludwig, D. (2020). Social eyes: Rich perceptual contents and systematic oppression. *Review of Philosophy and Psychology*, *11*(4), 939–954.

Lyons, J. (2005). Clades, Capgras, and perceptual kinds. *Philosophical Topics*, *33*(1), 185–206.

Majeed, R. (2018). Do gestalt effects show that we perceive high-level aesthetic properties? *Analysis*, *78*(3), 440–450. https://doi.org/10.1093/analys/any002

Mandelbaum, E. (2017). Seeing and conceptualizing: Modularity and the shallow contents of perception. *Philosophy and Phenomenological Research*, *97*(2), 267–283.

Marchi, F. (2015). Cognitive penetrability of social perception: A case for emotion recognition. *Review of Philosophy and Psychology*, *6*(4), 617–620. https://doi.org/10.1007/s13164-015-0258-z

Pautz, A. (2008). What are the contents of experiences? *The Philosophical Quarterly*, *59*(236), 483–507.

Phillips, B. (2018). The shifting border between perception and cognition. *Noûs*, *53*(2), 316–346.

Pollick, F. E., Kay, J. W., Heim, K., & Stringer, R. (2005). Gender recognition from point-light walkers. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1247–1265.

Prinz, J. (2006). Beyond appearances: The content of sensation and perception. In T. S. Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 434–460). Oxford University Press.

Reiland, I. (2014). On experiencing high-level properties. *American Philosophical Quarterly*, *51*(3), 177–187.

Reiland, I. (2015). Experience, seemings, and evidence. *Pacific Philosophical Quarterly*, *96*(4), 510–534. https://doi.org/10.1111/papq.12113

Reiland, I. (2018). On experiencing moral properties. *Synthese*, *198*, 315–325.

Rosenthal, D. (2010). How to think about mental qualities. *Philosophical Issues*, *20*(1), 368–393. https://doi.org/10.1111/j.1533-6077.2010.00190.x

Siegel, S. (2006). Which properties are represented in perception. In T. S. Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 481–503). Oxford University Press.

Siegel, S. (2013). The epistemic impact of the etiology of experience. *Philosophical Studies*, *162*(3), 697–722.

Siegel, S., & Byrne, A. (2017). Rich or thin? In B. Nanay (Ed.), *Current controversies in philosophy of perception* (pp. 59–80). Routledge.

Silins, N. (2013). The significance of high-level content. *Philosophical Studies*, *162*(1), 13–33.

Spaulding, S. (2015). On direct social perception. *Consciousness and Cognition*, *36*, 472–482.

Stokes, D. (2014). Cognitive penetration and the perception of art. *Dialectica*, *68*(1), 1–34.

Tanaka, J. W., Curran, T., & Sheinberg, D. L. (2005). The training and transfer of real-world perceptual expertise. *Psychological Science*, *16*(2), 145–151.

Toribio, J. (2018). Visual experience: Rich but impenetrable. *Synthese*, *195*(8), 3389–3406.

Valentine, T., Lewis, M. B., & Hills, P. J. (2016). Face-space: A unifying concept in face recognition research. *The Quarterly Journal of Experimental Psychology*, *69*(10), 1996–2019.

Werner, P. J. (2016). Moral perception and the contents of experience. *Journal of Moral Philosophy*, *13*(3), 294–317.

How to cite this article: Burnston, D. C. (2022). How to think about higher-level perceptual contents. *Mind & Language*, 1–21. https://doi.org/10.1111/mila.12446