The Philosophical Quarterly Vol. 0, No. 0 *2021*
ISSN 0031-8094 https://doi.org/10.1093/pq/pqab054

MAPPING THE VISUAL ICON

By Sam Clarke[1,2]

It is often claimed that pre-attentive vision has an 'iconic' format. This is seen to explain pre-attentive vision's characteristically high processing capacity and to make sense of an overlap in the mechanisms of early vision and mental imagery. But what does the iconicity of pre-attentive vision amount to? This paper considers two prominent ways of characterising pre-attentive visual icons and argues that neither is adequate: one approach renders the claim 'pre-attentive vision is iconic' empirically false while the other obscures its ability to do the explanatory work, which motivates positing pre-attentive visual icons in the first place. With this noted, I introduce the (heretofore unarticulated) notion of an 'Analogue Map' and argue that it provides a superior characterisation of pre-attentive vision's iconicity. I then argue that this forces a reassessment of debates which have traditionally presupposed the iconicity of pre-attentive vision, emphasising ramifications for the viability of a format-based perception-thought border.

Keywords: iconic representation, analogue representation, pictures, maps, visual representation, format.

I. INTRODUCTION

Philosophers and cognitive scientists have long debated the format of mental representations. For instance, they have long debated the existence of a language of thought (Fodor 1975), the use of mental models in reasoning tasks (Johnson-Laird 1980), the conjecture that mental images 'depict' (Kosslyn 1980), and the involvement of analogue structures in numerical cognition (Dehaene 1997). In each case, these debates concern the formal properties that mental representations possess—the question of how mental representations are structured to make content 'explicit' and 'accessible' (Marr 1982: 20–2).

Such concerns don't come from nowhere. Mainstream cognitive science conceives of the mind as a computational system, whose inferential transitions proceed (partially, perhaps entirely) in virtue of the formal properties mental representations possess (von Neumann 1958/2013). Thus, apprehending the format of mental representations has been expected to explain otherwise puzzling patterns of performance in human psychology (Fodor & Pylyshyn 1989; cf. Aizawa 1997; Kosslyn *et al.* 2006; cf. Anderson 1978) and to help demarcate psychological kinds that could otherwise be obscured by an advancing cognitive science (Quilty-Dunn 2019a; cf. Gross & Flombaum 2017).

In what follows, I'll consider the format of *pre-attentive visual representations*. To be clear, 'pre-attentive vision' is a provocative label. It's provocative because some deny that any stage of visual processing proceeds independently of attention (Mack & Rock 1998). Regardless, the label is routinely used to reference a stage of early visual processing, which is marked by its characteristically high-processing capacity; processing information in parallel across large portions of the visual field, without requiring *focussed* attention to the items and features it encodes (Wolfe 2018). Thus construed, 'pre-attentive vision' denotes a significant psychological kind irrespective of its status as *fully* pre-attentive (Palmer 1999). And since the format of our pre-attentive visual representations features heavily in debates about non-conceptual content (Fodor 2007), phenomenal consciousness (Block 2011), and the architecture of the mind (Dretske 1981), an obvious objective in the philosophy of psychology is to apprehend these representations' formal structure.

With this in view, it's striking to find that existing discussions reflect broad consensus on this matter. Debates rumble on regarding the format of many mental representations. But when theorists examine the formal properties of pre-attentive visual representations, they typically converge on the conjecture that these have an *iconic* (roughly: picture-like) format. For many, this explains pre-attentive vision's characteristically high-processing capacity (Dretske 1981; Fodor, 2007, 2008; Neisser 1967; Quilty-Dunn 2019a) and is implied by both retinotopically arranged maps in early visual cortex (Zeki 1993) and an overlap in the mechanisms of early vision and imagery (Block 1983a; Quilty-Dunn 2019b). Even Pylyshyn—perhaps *the* leading critic of picture-like representations elsewhere in cognition—concedes that certain pre-attentive visual representations are iconic, and states that this much is 'generally accepted' (2003: 29).

Despite this broad consensus, I believe we lack an adequate characterisation of pre-attentive vision's (alleged) iconicity. The trouble is not that characterisations of 'iconic representation' have not been provided. It's that these prove inadequate in this context—on inspection, they render the claim 'pre-attentive vision is iconic' empirically false or they obscure its ability to do the explanatory work which motivates positing pre-attentive visual icons in the first place (Section 2). As such, I'll seek to provide an improved characterisation. On my account, pre-attentive visual representations are iconic (or picture-like) by having a distinctive kind of cartographic structure (Section 3). While related suggestions have been made (Burge 2018; Treisman 1988), these remain underdeveloped. To advance matters, I'll introduce the (heretofore unarticulated) notion of an 'Analogue Map', note that it reflects a point of departure from existing characterisations of pre-attentive visual representations (Section 3.1), and argue that it provides a superior characterisation of pre-attentive vision, avoiding the shortcomings which afflict rival accounts (Section 3.2). I'll then suggest that this forces a reassessment of debates which routinely presuppose icons in pre-attentive vision. To illustrate, I'll emphasise ramifications for ongoing debates over the viability of a format-based perception-thought border (Section 4).

II. TWO ACCOUNTS OF PRE-ATTENTIVE ICONICITY

To begin, let's consider two prominent ways of characterising 'iconic representation'. For brevity, I'll label these 'Pictorialism' and 'Analogism', respectively. Both have been seen to capture the sense in which pre-attentive visual representations are 'iconic' or 'picture-like', but neither is entirely adequate in this regard. In showing this, I'll demonstrate the need for an improved account, and highlight desiderata on an adequate characterisation.

II.1 Pictorialism

Pictorialism may be the most prominent approach to characterising 'iconic representation' and pre-attentive vision in the philosophical literature. Die-hard proponents include Fodor (2007) and Quilty-Dunn (2016, 2019a, 2019b). But many slip into a kind of Pictorialism when describing pre-attentive vision and related representations (e.g. Burge 2014: 493 [cf. Burge 2018]; Carey 2009: 135; Dretske 1981: 137–8; Kosslyn et al. 2006: 13).

For the Pictorialist, iconic mental representations function like realistic pictures. Of course, the question of how realistic pictures function is, itself, vexed. But for the Pictorialist, this involves iconic mental representations conforming to a strong reading of the Parts Principle (PP) and Holism Principle (HP) at a psychological level of analysis.

Minimally, PP concerns the suggestion that if R is an icon representing X, then:

> parts of R function to represent parts of X with structural relations between parts of R representing structural relations between parts of X.

Thus, when we consider Fig. 1 (a paradigmatic icon), we're invited to note that (spatial) parts of the vehicle represent (spatial) parts of the depicted cat (e.g. the cat's nose, front left leg, and so forth). Furthermore, structural (spatial) relations between vehicle parts mirror structural (spatial) relations between cat parts depicted. For those emphasising PP in their characterisation of an icon, this is true of icons generally—photographs offer but a convenient illustration.

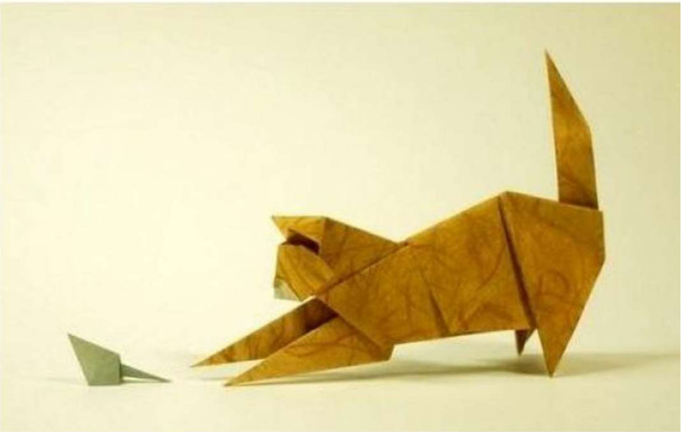


Figure 1. An origami cat.

Notably, Pictorialists go beyond this minimal reading of PP. They hold that interpretable parts of the representation must not only represent parts of that represented by the icon as a whole; they must *picture these*. As Fodor puts it: if P is an icon representing X 'then parts of P are pictures of parts of X' (Fodor 2007: 108). If nothing else, this involves semantically interpretable parts of P representing spatial parts of X by themselves conforming to some version of PP (this holds until some point at which the parts in question constitute representational 'primitives' which no longer comprise semantically relevant parts; Fodor 2008: 173; Quilty-Dunn 2019b: fn.8).

While controversial (Burge 2018), it's at least intuitive to think this *stronger reading* of PP further likens mental icons to realistic pictures. For, returning to the photograph in Fig. 1, and the fact that it depicts an origami cat by having vehicle parts represent cat parts, it seems that these vehicle parts represent the cat parts in question by picturing them. This involves them comprising vehicle parts which, themselves, represent the cat parts' parts (e.g. nose parts, leg parts, and so forth), and through the spatial arrangement of these vehicle part parts mirroring the spatial arrangement of the depicted cat part parts in question. So, if we cut out these vehicle parts, they would depict cat parts by picturing them, in a PP conforming manner, as opposed to (say) describing them with linguistic labels.

To see why this matters, consider the Pictorialist's second mark of iconicity: HP. HP concerns the suggestion that semantically interpretable parts of an icon (implicated in PP) 'encode various properties at once' (Quilty-Dunn 2019b: 4). Thus, when we consider the part of Fig. 1 that depicts the cat's tail *as triangular*, we're invited to note that it (naturally, perhaps inevitably) depicts this as having a certain orientation relative to the rest of the picture. Depiction of shape and orientation are, thus, depicted in an HP conforming manner. And when the vehicle part depicts the tail's colour (as in Fig. 1), we're invited to note that it (naturally, perhaps inevitably) depicts this colour as bound to a spatial region with a relatively determinate shape (Dretske 1981; Quilty-Dunn 2019b). So, while a black and white picture might depict an item's shape without depicting its colour, Pictorialists hold that when an icon depicts an item's colour it (more or less inevitably) takes a stand on the colour's spatial extension. Indeed, it's the assumption that this should be true of our psychological icons (like those found in pre-attentive vision), which leads Pictorialists to regard evidence that post-attentive object representations encode colour and shape separately to indicate non-iconic underpinnings (Green & Quilty-Dunn 2017; Quilty-Dunn 2019b).

None of this is meant to be arbitrary stipulation. Pictorialists consider HP conformity a natural consequence of PP. Thus, PP and HP are seen to naturally cluster together, demarcating a natural kind of mental representation (Quilty-Dunn 2019b: 4). Critically: this relies on the stronger reading of PP. A representation (e.g. a map) might represent parts of that which is represented by the representation as a whole using spatially arranged, linguistic labels. Provided that structural (e.g. spatial) relations between these labels mirror, and thereby specify structural (e.g. spatial) relations between the parts being labelled, the representation would meet the minimal notion of PP conformity outlined above. But critically, this would not imply the HP conformity Pictorialists demand. Part-representing labels might simply signify things like 'red here!' without taking a stand on the shape, orientation, or size of the coloured region. Thus, what implies the HP conformity Pictorialists emphasise is conformity to the *stronger reading* of PP wherein part-representing-parts of the vehicle *picture* the depicted parts. On this view, colour properties are bound to vehicle parts with a (relatively determinate) shape-representing-shape. Furthermore, these shape-representing-shapes depict the shapes they do by having spatially arranged parts mirror relevant parts of the shape (and their spatial arrangement) within a functional/physical space. As such, depicted orientation emerges (naturally, perhaps inevitably) out of relations between parts of the relevant shape-representing vehicle parts and the rest of the representation, and it is (by hypothesis) hard, if not impossible, to depict these properties independently (cf. Block 1983b).

II.2 Problems with pictorialism

Pictorialism reflects a prominent approach to characterising the iconicity of pre-attentive vision. Unfortunately, it is empirically inadequate in this regard.

Pictorialist icons are posited in pre-attentive vision to explain key aspects of pre-attentive vision's performance profile. To this end, Pictorialists routinely emphasise pre-attentive vision's characteristically high-processing capacity; the

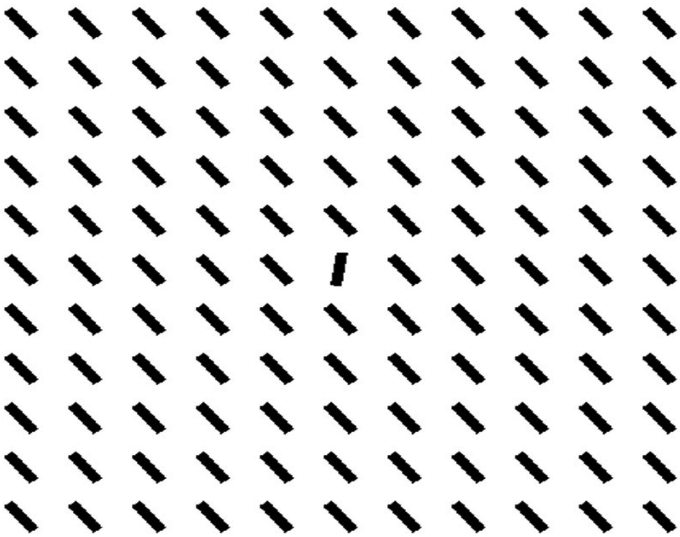


Figure 2. The 'oddball' *pops out* of the array.

fact that pre-attentive vision encodes, and efficiently processes, the distinguishing properties/features of many items, in parallel, across large portions of the visual field. For Neisser (1967), who coined the term 'pre-attentive vision', this was illustrated by Sperling's (1960) pioneering work on iconic memory. However, experimental studies investigating stereoscopic vision (Fodor 2008) and ensemble perception (Quilty-Dunn 2019b) tell a similar story. In each case, Pictorialist icons are invoked to explain item-unlimited computations. Thus, Fodor (2008) posits that the iconicity of pre-attentive vision explains how 'thousands of dots' (191) are processed (and not simply encoded) efficiently and in parallel during random-dot stereogram experiments, while Quilty-Dunn (2019b) posits that the iconicity of pre-attentive vision explains how the visual system extracts 'statistical regularities in scenes by computing over many more than four objects' during ensemble perception (8).

For my purposes, 'disjunctive visual search' offers a convenient illustration of this general suggestion. In disjunctive visual search tasks, subjects are tasked with identifying 'an oddball' in an array of distractor items (Fig. 2). Crudely: if the oddball is suitably distinguished by one or more appropriate visual 'features'—e.g. colour, size, motion, or orientation (Wolfe & Horrowitz 2017)—performance is (largely) unaffected by set size. That is: subjects are as fast and accurate at locating oddballs whether they are located in arrays containing thirty items or just three (Treisman 1986: 117). Since there needn't be any way for subjects to pre-empt the distinguishing feature or location of oddballs, these results indicate that some stage of early visual processing succeeds in representing all of the potentially distinguishing features (and feature locations) in a large array (containing many features and feature locations) as quickly and accurately as it succeeds in representing all (or nearly all) of the potentially distinguishing features (and feature locations) in a comparatively small array (containing less features and, hence, feature locations). In turn, these findings suggest that the formation of these early visual representations cannot depend on the serial deployment of focused attention to the individual features/items encoded, hence why they have traditionally been said to reflect 'pre-attentive' visual processing (Treisman 1985).

Pictorialists invoke Pictorialist icons to explain such results. To see why, suppose our pre-attentive visual representations are not Pictorialist icons and have a descriptive, language-like format instead. On this view, pre-attentive visual processing would (seemingly) need to produce and process extra symbols when representing large arrays, for extra symbols would be needed to independently encode the extra items and potentially distinguishing features contained therein. (To illustrate: contrast the number of words in 'that item is red and triangular' with those in 'that object is red and round and upright, and that object is. . . '.) Hence, we might expect pre-attentive visual processes to decrease in performance when processing large arrays—a prediction that's not borne out.

But suppose the relevant visual representations are Pictorialist icons. Here, it's unclear why accurately depicting extra items, features, and properties would involve extra complexity in the representations themselves. By conforming to a strong reading of PP, parts of *a single symbol* could picture spatial regions of the seen array, irrespective of whether these regions contain entire items, item parts, or empty space. In so doing, structural relations between content-bearing parts of the symbol would encode spatial relations between depicted parts of the array, all without increasing demands on representational resources. And, through HP conformity, perceptible features/properties (like colour, shape, and orientation) could be holistically bound to, or emergent out of, depicted regions and their spatial relations, marking oddballs out from the crowd. Hence, the intrinsic complexity of the vehicle would not correspond to the total number of features/items being represented. So, in the same way that a photograph, like Fig. 1, might depict all the cats (and cat features) in a large clowder, without containing extra pixels or taking longer to develop, Pictorialists claim that the iconicity of pre-attentive vision explains how and why pre-attentive vision successfully *encodes* and (through sensitivity to the formal properties of its representations) *processes* large numbers of seen items, item features and item locations, in parallel, without incurring a cost to performance (Dretske 1981; Fodor 2008; Neisser 1967; Quilty-Dunn 2019a).

Unfortunately, there's a problem with this suggestion. While visual search is (often) unaffected by the number of distractor items in *disjunctive* search tasks of the above sort, this isn't true of other search types. For instance, in

Figure 3. A disjunctively defined oddball pops out of the left-hand array; a conjunctively defined oddball in the right-hand array does not.

'conjunctive' search tasks, oddballs are defined by a *conjunction of features*. Thus, an oddball could differ from other objects in the array by being the only grey and horizontal line, but not by being the only grey line nor the only horizontal line (Fig. 3). In tasks of this sort, search efficiency might be increased in various ways (Humphreys, Hodsoll & Riddoch 2009; Wolfe, Cave & Franzel 1989), but set size invariably matters. Thus, detecting oddballs in conjunctive search tasks is *always* more time consuming and error prone when oddballs are located in large arrays, *ceteris paribus* (Treisman & Gelade 1980; Wolfe 2018).

While conjunctive search depends on the allocation of focussed attention, and thus concerns a later ('attentive') stage of visual processing, it places pressure on the Pictorialist hypothesis. If Pictorialist icons are to explain the lack of set-size limitations in *disjunctive* search tasks, then they must not only represent the potentially distinctive properties of items found in observed arrays—they must also pick out, or flag, oddballs *as* oddballs, by virtue of representing their potentially distinctive properties. This is crucial because iconicity needs to explain how many items, properties, and locations are depicted and processed efficiently such that oddballs are detected. But, unless the representation, itself, flags the oddball, *as such*, we're left with the task of explaining how the Pictorial icon (encoding this information) is scanned, such that the oddball is detected in a timely manner.

But here's the problem: if oddballs are represented in an HP conforming manner, as Pictorialists insist, and pre-attentive visual icons flag oddballs *as* oddballs *by* depicting their potentially distinctive properties, then oddball detection should also be efficient in *conjunctive* search tasks. For if representing (say) an oddball's colour, as such, involves representing it *as* coloured *and* possessing-a-given-shape (such that there is no separation in the depiction of these properties), and (furthermore) representing the item as possessing *either* of these properties suffices to flag the item as an oddball in disjunctive search tasks (such that it *pops-out* of the array), then represented conjunctions of properties should facilitate pop-out in conjunctive search tasks, where oddballs are marked by this conjunction of properties. But, as we have just seen, they do not. Thus: insofar as HP conformity (borne out of a strong reading of PP) serves as a necessary condition on a representation's qualification as an icon (as Pictorialist's insist), and insofar as the characteristically high processing capacity of pre-attentive vision serves as a key source of evidence for the postulation of Pictorialist mental icons (as it does: Dretske 1981; Fodor 2007; Quilty-Dunn 2019a, 2019b), visual search presents a problem for the Pictorialist hypothesis.

In response, Pictorialists might distance themselves from the assertion that iconicity explains pre-attentive vision's characteristically high-processing capacity. This would be a significant concession since it is this which standardly motivates the postulation of Pictorialist icons in pre-attentive vision (Fodor, 2007, 2008; Quilty-Dunn 2019a, 2019b). In any case, further findings compound the problem.

Perhaps most famously: if pre-attentive visual representations of shape size and colour are encoded independently, in non-HP conforming ways, we might expect pre-attentive vision to occasionally mis-combine these. Thus, we might expect the (surprising) existence of *illusory conjunctions*. The existence of these is well-documented (Prinzmetal 2012) but puzzling for the Pictorialist.

In one experiment, subjects were presented with an array containing three coloured letters. Provided that presentation times were brief (e.g. 200 ms), and subjects were prevented from directly attending to the shapes, they would regularly report incorrect letter-colour combinations. For instance, in an array containing a blue 'T', a green 'X', and a red 'O', subjects might report seeing a blue 'O', or a red 'T', and so forth (Treisman & Schmidt 1982: Exp.1). That subjects were significantly less likely to report colours and letters absent from the arrays suggests that subjects' reports reflected genuine mis-combinations of *perceived* features as opposed to *misperceptions* of the features themselves. And since these results were found to obtain in simultaneous matching tasks (Exp.3), it's generally agreed that they don't simply reflect post-perceptual failures of memory. Rather, mis-combination occurs within visual processing itself (Prinzmetal 2012). Similar findings are obtained when the relevant features are all spatially defined and (e.g.) individual edges of a triangle are mis-combined with 'S' shapes to produce illusory dollar signs (Fig. 4; Treisman & Patterson 1984).

These results are puzzling unless the representation of relevant feature types enjoys an 'independent psychological existence' (Treisman 1986: 117) that gives rise to a 'Binding Problem' (Feldman 2013). If pre-attentive vision specifies features (e.g. colours and shapes) accurately, such that these are not misperceived (in themselves), and it processes and represents these in an HP conforming way (such that depictions of colour cannot float free of depicted shape), it's hard to see why mistaken conjunctions would reliably occur. Meanwhile, positing non-HP conforming representations in pre-attentive vision makes easy sense

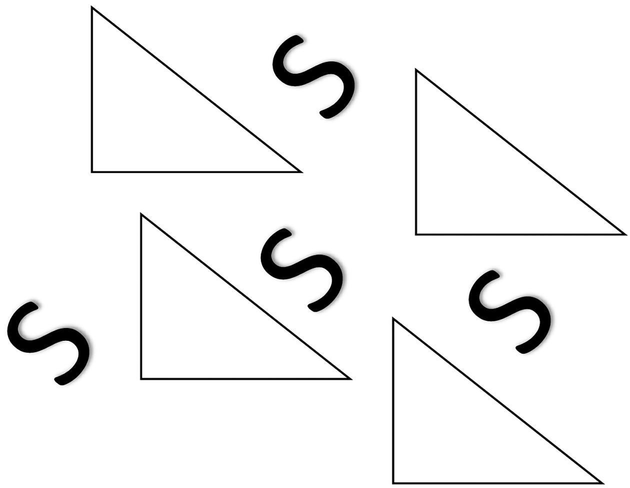


Figure 4. Treisman and Patterson (1984) found that, under certain conditions, 'S' shapes would be perceptually combined with the hypotenuse of triangles, so as to form illusory dollar signs.

of this: if independent features of single items are encoded independently (in non-HP-conforming ways) then they can (of course) be incorrectly combined. Indeed, this is independently predicted by the hypothesis (Treisman & Schmidt 1982).

A Pictorialist might respond that HP conformity is not meant to render an icon's independent encoding of these features impossible, only unlikely. As Green and Quilty-Dunn (2017) acknowledge, to assume otherwise would commit them to 'the photographic fallacy'—to (erroneously) assume that a pictorial representation must be determinate with respect to *every* visual feature (Block 1983b). Thus, Pictorialists might respond that even if the preceding results do demonstrate the existence of distinct icons, specifying the relevant feature types independently of one another, these will still (somehow) qualify as Pictorialist icons.

Such a response would be unsatisfactory on two fronts. First, Pictorialists propose that the independent encoding of colour and shape in (post-attentive) object perception demonstrates that post-attentive object representations are not Pictorialist icons (Green & Quilty-Dunn 2017; Quilty-Dunn 2019b; see also Fodor & Pylyshyn 2016). Thus, it's not clear how they can dismiss these concerns in the case of pre-attentive vision. But notice: even if they could, HP would still fail to *usefully* characterise pre-attentive visual representations. Since pre-attentive visual representations are supposed to exemplify the iconicity Pictorialists describe (Fodor, 2007, 2008; Quilty-Dunn 2019a, 2019b), finding that pre-attentive visual representations encode relevant features separately would still suffice to show that HP is not a helpful way of characterising the kind.

Since HP is supposed to be a natural consequence of the PP-conformity Pictorialists posit, Pictorialism should then be rejected as a *useful* characterisation of iconicity in this context, even if it can somehow accommodate the above.

II.3 Analogism

Pictorialism is empirically inadequate as a (useful) characterisation of pre-attentive visual format. But, *prima facie,* an alternative way of characterising icons—*Analogism*—can avoid its shortcomings.

Like Pictorialists, Analogists articulate a sense in which icons are picture-like (hence, they don't simply change the subject). But, unlike Pictorialists, Analogists deny that icons must be picture-like in their PP or HP conformity. Instead, they hold that the iconicity of a representation simply depends on the representation exploiting an isomorphism between content-bearing properties of the vehicle and the properties these depict (Beck 2019; Burge 2018; Maley 2011). As such, iconicity need only involve content-bearing properties of the vehicle *mirroring* the properties they represent, like how a mercury-in-glass thermometer's mercury-level *mirrors* the temperature.

This mirroring relation could be fleshed out in various ways. It might involve content-bearing vehicle properties varying as a *linear* function of represented properties (Maley 2011), a *logarithmic* function of these (Dehaene et al. 2003), or simply as a *monotonic* function of these (Beck 2019). But note: while the realistic pictures (emphasised by Pictorialists) seem to count as iconic on an unqualified Analogism of this sort—pictures tend to depict properties by having content-bearing properties of the vehicle *mirror* these (hence why moving points of a picture further apart tends to increase the depicted distance between these)—many representations which do not qualify as iconic by Pictorialist standards do qualify as iconic by Analogists'.

For instance, Burge (2018) notes (approvingly) that even simple colour chits qualify. A colour chit represents 'through its colour's being the same as the colour that is represented' (81). Thus, there is a natural 1–1 mapping between the colour of the vehicle and the colour depicted, and this mapping preserves relations between vehicle values at the level of content. Indeed, it is by doing so that colour chits represent what they do. As such, colour chits qualify as iconic by Burge's analogist standards since content-bearing vehicle properties mirror their contents. But colour chits do not have vehicle parts which represent parts of the property they depict (Clarke forthcoming). Nor does their depiction of a given colour imply their depiction of anything else. Indeed, two distinct chits (e.g. a 'colour chit' and 'orientation chit') might independently represent distinct properties/features of the same item/region/feature, even if these would be holistically bound together in a realistic picture. Thus, Analogue icons need not conform to PP or HP in the way Pictorialists demand. This allows Analogism to avoid the aforementioned problems afflicting Pictorialism in this context.

II.4 Problems with analogism

Alas, Analogism faces problems of its own. Most notably, it's not clear how Analogue icons are meant to do the explanatory work that motivates positing pre-attentive visual icons in the first place. This is not to suggest that Analogism is false as a characterization of pre-attentive visual representation. It simply shows that if an appeal to 'iconic representation' is to do the explanatory work which originally (Neisser 1967) and standardly (Fodor, 2007, 2008; Quilty-Dunn 2019a, 2019b) motivates the postulation of icons in pre-attentive vision, Analogism needs to be qualified in some (as yet unspecified) way.

To illustrate, recall that icons are regularly posited in pre-attentive vision to explain its high-processing capacity (Section 2.2). Here, the explanatory purchase Pictorialist icons offer stems from their PP and HP conformity. Crudely: it stems from the idea that pre-attentive visual icons have arbitrarily many (spatially arranged) vehicle parts, depicting arbitrarily many (spatially arranged) regions of the visual array (with relevant properties holistically bound to these), irrespective of what these regions contain. This offers to explain why the number of items/item features fails to dictate the intrinsic complexity of the visual representation and thus, fails to dictate the speed/accuracy of the visual processes involved in producing/manipulating it. Of course, commitment to a strong reading of PP and HP got Pictorialism into trouble as a characterisation of pre-attentive visual representation. But, in abandoning these principles, it's unclear how the Analogist intends to explain these results. For if the Analogist allows us to posit distinct 'chits' that independently mirror each feature in the display, the total number of features being depicted will dictate the complexity of the representation. Thus, short of qualifying Analogism in some (as yet unspecified) way, the explanation for pre-attentive vision's characteristically high-processing capacity no longer goes through.

Similar points apply to key findings in the mental imagery literature. These are standardly explained by appeal to mental icons (Kosslyn et al. 2006). And, given a neuropsychological overlap in the mechanisms involved, many consider this reason to deem pre-attentive visual representations similarly iconic (Block 1983a; Carey 2009: 458; Quilty-Dunn 2019b).

Take Shepard and Metzler's seminal (1971) work on mental rotation. Here, subjects were presented with pairs of line drawings, each depicting a three-dimensional shape. In each pair, depicted shapes were either identical, albeit oriented at different angles (like Pair A from Fig. 5), or mirror images of one another (like Pair C from Fig. 5). Subjects pressed a button indicating which of these possibilities was true of successively presented pairs and did so as quickly and accurately as possible. In so doing, it was found that the speed

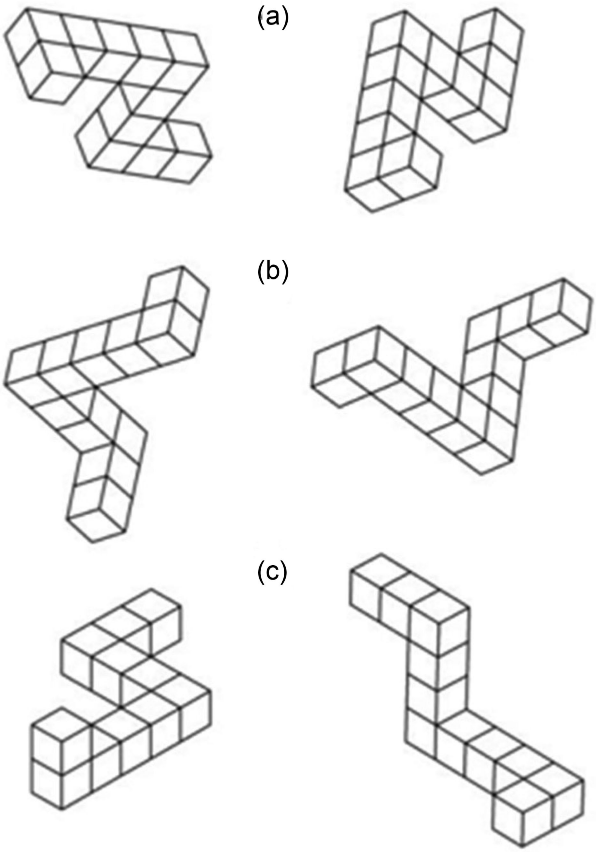


Figure 5. Stimuli of the sort used in Shepard and Metzler's (1971) groundbreaking mental rotation study.

with which subjects correctly identified shapes from a pair *as identical*, albeit rotated at different angles (if and when they were), was proportional to the angular difference in the shapes as depicted on the screen. To make sense of this, Shepard and Metzler proposed that subjects performed the task by, first, forming a mental *image* of one or other of the shapes and, next, 'rotating it' to check for a match (or lack thereof) with its perceived partner. The idea was that if the mental image represented the shape like a literal picture (or a three-dimensional model), rather than a lingua-form description, and 'rotation' of the image proceeded at a steady rate, angular differences would wind up proportional to the time taken to bring paired shapes (imagined and perceived) into correspondence, to confirm a match, thereby explaining the above result.

The point to note is that (once again) an unqualified Analogism leaves this obscure. This is not to suggest that Analogism is wholly irrelevant to these results. The mirroring relation articulated by the Analogist implies that two Analogue icons, representing distinct orientations, will have relevant content-bearing properties differ by an amount that (in some way) mirrors the difference in orientation depicted. Thus, it could explain why bringing two depicted orientations into correspondence requires an amount of vehicular manipulation (and, hence, time) that is proportional to the difference in orientation being circumnavigated (Maley 2011). The trouble is: while this helps explain the temporal profile of the rotation process that is (by hypothesis) performed in these tasks, it's irrelevant to the identification of a match unless the rotation process brings relevant spatial parts of the shapes into correspondence. So, at this point, Pictorialists will maintain that explanations of this sort still require that the relevant orientation representations be holistically bound to the depiction of spatial parts, and hence that the mental images conform to PP and HP. Indeed, this seems particularly pressing, for if the depiction of shape is encoded independently of orientation, it's unclear why rotation would even be helpful (Quilty-Dunn 2019b: 8). Representations of shape might just be directly compared, independently of (distinct) constituents representing orientation, generating the (false) prediction that angular deviation would fail to predict response time.

These are just two examples. But both demonstrate that if an appeal to 'iconic representation' is to do the explanatory work which standardly motivates positing icons in pre-attentive vision, an unqualified Analogism is inadequate. This is not to say that Analogism is false as a characterisation of pre-attentive vision. Many Analogists mean to characterise a broader class of perceptual representation than the class of pre-attentive visual representations (Beck 2019; Clarke forthcoming), and they are often explicit that the total set of Analogue icons subsumes numerous more determinate representational kinds; beyond realistic pictures and colour chits, Burge (2018: 89) lists maps, musical notations, diagrams, bar graphs, abacuses, and hieroglyphs. Nonetheless, what these examples show is that if an appeal to 'iconic representation' is to do the explanatory work that's routinely expected of it in this context, the Analogue apparatus must be qualified in some (as yet unspecified) way(s).

III. A CARTOGRAPHIC CHARACTERISATION

I've now considered two ways of characterising 'iconic representation' and found that neither makes adequate sense of the claim that 'pre-attentive visual representations are iconic'. Pictorialist characterisations look empirically inadequate, while (an unqualified) Analogism fails to provide the explanatory purchase motivating the postulation of pre-attentive visual icons in the first place. For those wedded to the iconicity of pre-attentive vision, this presents a challenge: specify a relevant format type which provides the explanatory purchase needed in this context, without inheriting the empirical problems afflicting Pictorialism.

With this challenge on the table, I'll now argue that pre-attentive visual representations possess a specific type of cartographic format. This cartographic format builds on an unqualified Analogism of the above sort and is related to existing characterisations of pre-attentive vision in the philosophy (Matthen 2005) and psychology (Treisman 1988) literature. But since there are various types of cartographic structure, I'll begin by clarifying the cartographic kind I have in mind: *The Analogue Map*. What's distinctive about Analogue Maps is that they are analogue through and through – they situate analogue constituents within an analogue (spatial) structure. In this way, they differ from other kinds of map, which might be thought of as hybrid representations, containing both analogue and non-analogue (digital/discursive) elements. At the same time, I'll show that Analogue Maps avoid collapsing into, and thereby inheriting problems associated with, Pictorialist icons. In making this conjecture, I'll distinguish my proposal from related proposals (Section 3.1) and highlight reasons to think pre-attentive visual representations cartographic in my recommended sense (Section 3.2).

III.1 Map-like structures

In a penetrating discussion, Camp notes that maps are marked by a spatial principle of composition (Camp 2007: 158). Constituents are located and related in a spatial structure which *mirrors*, and thereby represents, the spatial relations between these. In this respect, maps are iconic by Analogue standards and akin to realistic pictures (Burge 2018). However, they differ from realistic pictures in that their spatially related constituents need not picture/resemble the entities they represent. This allows that a cartographic characterisation of pre-attentive vision might avoid the HP conformity which proved problematic for Pictorialism.

This is not to suggest that maps cannot utilise pictorial constituents. Fig. 6a *maps* a park by *picturing* park structures (e.g. a pond) from a bird's eye perspective and locating these within a spatial array. In so doing, the figure conforms to a strong reading of PP (it represents what it does by having spatially arranged parts of the icon *picture* parts of the region depicted) and HP (in depicting the park's pond, the map takes a stand on its [relative] size, shape, and even colour). As such, there are maps, like 6a, which qualify as Pictorialist icons. The point is just that some maps are not like this. For example, Fig. 6b *maps* the seating arrangement of students in a seminar by arranging representational constituents in a spatial array. These constituents specify parts of the class represented by the map as a whole and so the map conforms to a minimal reading of PP. But the constituents are not pictures; they are linguistic

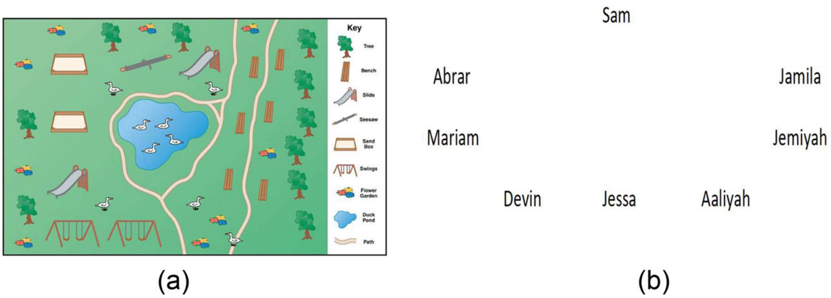


Figure 6. Some maps are Pictorialist Icons (a), and some are not (b).

labels, like 'Mariam', which bear arbitrary relations to the students they represent. As such, spatially arranged constituents represent the students they do without specifying student parts, *pace* the Pictorialist's strong reading of PP, or properties, *pace* HP.

This is relevant to our concerns in this paper. Pictorialism fails to provide an adequate characterisation of pre-attentive vision since it maintains that the pre-attentive visual representation of certain features (e.g. an item's colour) should take a stand on others (e.g. its shape, size, and/or orientation). This is an empirical conjecture we found grounds to reject, and one which stems from the idea that pre-attentive visual icons represent things by picturing their parts. But, since maps, like 6b, represent items and item locations without picturing them or their parts (e.g. by using digital constituents to specify entities), they can specify individual feature types independently of others. For instance, a map like 6b, might use the symbol 'R' to flag the location of red regions in space without specifying the shape or size of these. Thus, a broadly cartographic conception of pre-attentive vision can avoid the commitment to HP which spelled trouble for Pictorialism in this context.

This observation is not new. When Treisman first proposed that individual feature types enjoy an 'independent psychological existence' in pre-attentive vision, she proposed that individual feature types get represented in syntactically distinct 'feature maps' (Treisman 1988). Thus, she proposed that pre-attentive vision involves independent, spatially organised representations, charting distinct feature types. To clarify this suggestion, philosophers like Matthen (2005) and Clark (2001) have suggested that these function as spatially arranged networks of nodes where each node in the network functions as a standalone digital symbol, signalling the presence or absence of a designated feature type at its indexed location (a bit like Fig. 7a). Thus, nodes in the network function like arbitrary symbols specifying things like RED HERE! at their indexed location without taking a stand on further features of the red item/region. On this view, pre-attentive visual representations are Analogue icons in that nodes are located in a functional space which mirrors a depicted space, but

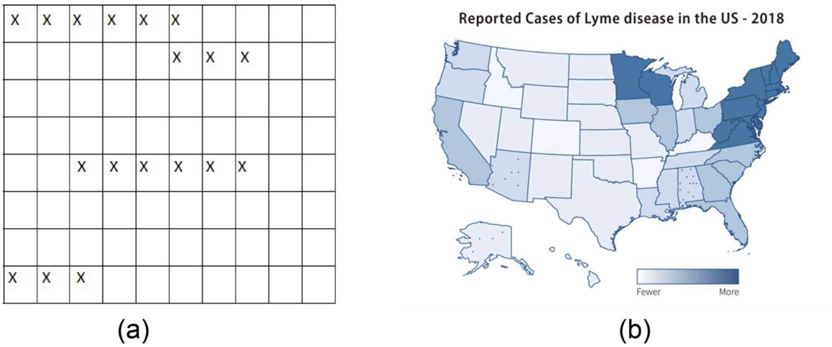


Figure 7. a: A Digital Map that uses the digital symbol 'X' to flag the location of a feature type (e.g. 'red', or 'horizontal edge') within a depicted space. b: An Analogue Map—each US state can be thought of as a node which functions as a standalone analogue icon, varying in blueness/saturation as a function of reported cases of Lyme disease.

they situate non-iconic, *digital* symbols within this space. For this reason, let's call these hybrid representations 'Digital Maps'.

For reasons we will discuss, this proposal is on the right track. But critically, it's not the only way of characterising pre-attentive feature maps which respects the independent encoding of distinct feature types. This is because what enables Digital Maps to avoid the HP conformity emphasised by Pictorialists is not the fact that constituent parts are *digital* symbols, which bear arbitrary relations to that which they represent. It's the fact that spatially arranged constituents are not full-blown *pictures*, which invariably take a stand on multiple features of a represented entity at once.

Sometimes philosophers conflate the non-pictorial with the digital (Fodor 2007; Sober 1976). But we have already seen that this is a mistake. Simple Analogue icons, like colour chits, differ in important ways from digital representations in that they function to mirror the properties they depict and, thus, do not bear entirely arbitrary relations to what they represent. At the same time, they do not amount to full-blown pictures—at least not PP or HP conforming pictures. Consequently, on the proposal I wish to advance, we should reject the suggestion that pre-attentive visual representations are Digital Maps, which use spatially arranged *digital* constituents to denote the presence and location of individual feature types. Instead, pre-attentive visual representations are cartographic networks of nodes, where each node in the network functions as a standalone Analogue icon, mirroring the properties/features it denotes at its indexed location (Fig. 7b). On this view, pre-attentive visual representations are doubly Analogue: depicted spatial relations are mirrored by the spatial relatedness of nodes in a functional space, and the nodes within this functional space mirror the feature values they depict (i.e. function as standalone Analogue icons, like independent mercury-in-glass thermometers). Thus, they are *analogue through and through*, containing *no digital/discursive constituents* whatsoever. For this reason, I'll call maps of the sort I am positing 'Analogue Maps', in contrast with the 'Digital Maps' noted above.

III.2 Pre-attentive visual representations are analogue maps

To motivate my conjecture that pre-attentive visual representations are Analogue Maps, I'll now argue that it does the explanatory work expected of icons in this context, without inheriting the empirical problems afflicting Pictorialism.

Let's begin by considering the illusory conjunctions discussed in Section 2.2. These raised problems for Pictorialist characterisations of pre-attentive vision, since they suggested that individual feature types are independently encoded in pre-attentive vision. This indicates that pre-attentive visual representations do not conform to HP in the way Pictorialists suggest. It is, however, something that a cartographic characterisation of pre-attentive vision straightforwardly accommodates (see above). This is because distinct feature maps (utilising non-pictorial constituents) can independently depict distinct feature types such that each map accurately depicts a given feature type independently of all others (Treisman 1988). In turn, this allows that such maps (whether analogue or digital) can be mis-combined. So, when a subject is presented with a 'red circle' and 'blue triangle', one feature map might (correctly) register the presence of blue and/or red and one feature map might (correctly) register the presence of circularity and/or triangularity (or, more likely, the edge orientations corresponding to these shapes), yet further resources might mis-combine these, such that they misperceive the circle as blue and the triangle as red.

This charts a logical possibility. However, the empirical details support it. For one thing, illusory conjunctions are not random. Rather, the likelihood that two distinct features are mis-combined is predicted by their proximity in objective space (Prinzmetal 2012; cf. Treisman & Schmidt 1982). This suggests that features are indexed to locations within a vehicular space that mirrors this. For if 'locations' in vehicular space mirror objective spatial locations, as a cartographic characterisation implies, then noise in the encoding of features at relevant locations would naturally imply comparable levels of noise at the level of the representation's spatial content. Thus the (noise-ridden) placement of features would be expected to average around veridicality, and features would be more likely to get (mis)combined with nearby items in objective space. This is something you would not get with, say, messy (i.e. noisy) handwriting that merely *describes* feature locations, suggesting that the relevant pre-attentive representations really do have a spatial structure, like a map.

In this way, illusory conjunctions suggest that pre-attentive visual feature representations function like spatially structured maps. But, while illusory conjunctions speak against these representations having *pictorial* constituents, which exhibit the HP conformity we found grounds to reject (see Section 2.2), these results are silent on whether pre-attentive feature maps are *Analogue Maps*, as I propose, or *Digital Maps*, as others assume.

To adjudicate these possibilities, consider the disjunctive visual search tasks discussed in Section 2.2. Recall that *disjunctive* search for an oddball (suitably defined by a single visual feature) often seems to be free of set-size limitations (to an interesting degree). This highlights the characteristically high-processing capacity of pre-attentive vision and is the sort of finding that the iconicity of pre-attentive visual representation has been expected to explain. A broadly cartographic characterisation of pre-attentive vision—which builds in considerably more structure than an unqualified Analogism—can do this. Since a feature map, with a cartographic (spatial) structure, can specify features in parallel across the whole area being mapped, disjunctively defined oddballs can be efficiently marked out by individual feature maps irrespective of total array size. In the same way that parts of a photograph picture items at the regions of the scene they depict (irrespective of how many items are located there or elsewhere in the scene) nodes in a cartographic network can function to flag the presence of relevant features at their indexed locations, in parallel, irrespective of how many features or items are located in the array as a whole. This allows that individual feature maps can mark out disjunctively defined oddballs, with the distinctive property they encode.

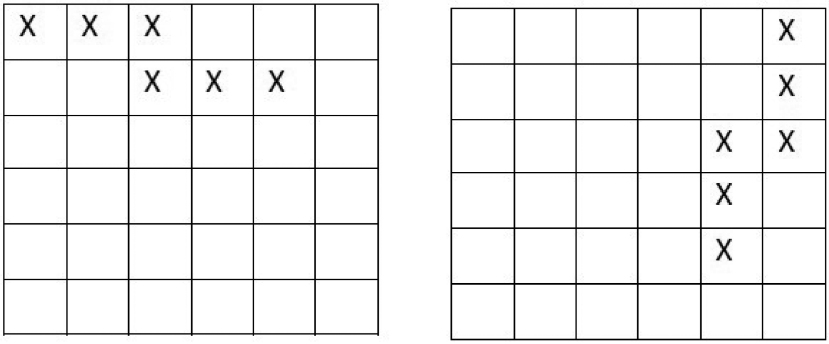
In this way, the characteristically high-processing capacity of pre-attentive vision can be explained by a broadly cartographic characterisation of pre-attentive vision, on which pre-attentive vision involves a series of spatially arranged feature maps. But crucially, it is important to note that disjunctive visual search is not always fast, efficient, and item unlimited. For one thing, oddballs do not pop-out unless they differ from distractor objects (with respect to their defining feature) by a suitably large degree (Duncan & Humphreys 1989). And when differences are sufficiently pronounced pop-out is often asymmetric. For instance, an item may only pop out if it is, itself, *brighter* than other (distractor) items populating the array, with similar points holding for magnitudes like length and even features/properties like orientation, curvature, and closure (Wolfe 2018).

A point which has been missed in existing discussions is that a cartographic characterisation of pre-attentive vision can elucidate these complications if we suppose that the relevant cartographic representations are *Analogue Maps*, with Analogue constituents mirroring the properties/features they depict/encode at indexed locations, rather than *Digital Maps*, containing digital/discursive constituents. To appreciate this, consider that if content-bearing properties of constituents represent the properties they do *by mirroring these*, then similarities in content will be mirrored by similarities in the content-bearing vehicle properties themselves. So, just as the vehicles representing 99° and 100° in two mercury-in-glass thermometers may differ by a barely discriminable amount, explaining why their contents get easily confused, the same is true of the Analogue constituents involved in Analogue feature maps (a point which would be compounded by the noise that is inevitably involved in their production). Since you wouldn't expect this if the representational constituents are digital, and bear nothing but arbitrary relations to the things they represent (e.g. the digital symbols '99°' and '100°', have vehicular properties that seem to differ more than '99°' and '90°' despite the former symbols' relative proximity in content) positing that the constituents in our pre-attentive feature maps are, themselves, Analogue icons can explain why oddballs do not reliably pop-out unless discriminated (at the level of content) by a suitably large margin. This is one reason to favour the suggestion that pre-attentive visual feature maps are Analogue Maps, over the suggestion that they are Digital.

Better still, the postulation of Analogue Maps can elucidate the aforementioned search asymmetries. If the constituents of a feature map depict properties like brightness at their indexed location by having their content bearing properties mirror these, then nodes depicting the brightest points in the array can function as if they are shouting louder than other nodes. Or, put in computational terms, it allows that a consumer system involved in oddball detection could simply be sensitive to this content bearing property of the representation, allocating resources to regions of the feature map oscillating/spiking/firing most/least intensely. This would be analogous to the same way one might find the longest (but not the shortest) piece of spaghetti in a pack by tapping the butt of the pack on the table and seeing which piece pops out furthest. Once again, this is an aspect of visual search that is left unexplained by the suggestion that pre-attentive feature maps have digital constituents. So, again, it is reason to favour the conjecture that pre-attentive feature maps are Analogue Maps.

Finally, the postulation of Analogue Maps can explain why conjunctive search tasks are often slow and inefficient. In this respect, the hypothesis inherits the virtues of a cartographic characterisation more generally. For while individual feature maps suffice to mark out suitably defined disjunctive oddballs, additional resources are required to appropriately bind and consider content from distinct maps in tandem (Treisman 1988). So, in sum: positing *Analogue Map*s in pre-attentive vision inherits the virtues of a broadly cartographic characterisation of pre-attentive vision while avoiding the empirical problems afflicting Pictorialism in this context. Better still, it explains aspects of visual search which are left obscure by an unqualified Analogism or by the postulation of Digital Maps.

What about mental imagery? In Section 2.4, we saw that those positing pre-attentive visual icons often assume that visual imagery has an iconic format, where a neurofunctional overlap in the mechanisms involved evinces the iconicity of pre-attentive vision. Such reasoning is indirect. But a similar line of thought might motivate an inference in the opposite direction: for if



(a) (b)

Figure 8. Rotating map (a) 90 degrees clockwise will bring indexed features into correspondence with indexed features in map (b).

pre-attentive visual representations are Analogue Maps, an alleged overlap in the mechanisms of vision and visual imagery might imply that visual imagery be underpinned by representations of a comparable sort. This raises the question: could the Analogue Maps I'm positing explain performance in (say) mental rotation tasks?

This question is pressing given the shortcomings of an unqualified Analogism in this context. However, the answer seems to be 'yes'. Indeed, there are respects in which Analogue Maps offer a superior explanation of the relevant phenomena.

To illustrate, note that the representations I am positing have a fundamentally spatial structure, indexing features to spatial locations. As such rotating feature maps (as a whole) can still bring indexed features into correspondence. For instance, placing Fig. 8a over Fig. 8b and rotating Fig. 8a 90° clockwise will bring indexed features of the former into correspondence with the latter.

Admittedly, this (alone) does not suffice to establish a match in the spatial structure of depicted shapes. This is because the correspondence of indexed spatial features among rotated feature maps is consistent with these spatial features differing in marked ways (for instance, their correspondence does not establish that the represented features [e.g. edges] are now oriented in the same ways). Nevertheless, positing that the constituents at these indexed locations depict the spatial features they do (e.g. oriented edges) by *mirroring* these (e.g. having some property of the vehicular constituent vary as a function of the orientation being depicted) implies that when an overlap in the constituents occurs, differences in the values of relevant constituents will be consistent at the level of both content and vehicle. This allows that a match in shape be confirmed (or rejected) via a simple computation, which tracks a consistent difference in the content-bearing properties of overlapping constituents. And since feature maps only specify the features of shapes depicted (e.g. edges as opposed to shapes as a whole), the account explains why mental rotation would still be needed to confirm a match.

An attractive aspect of this proposal is that it can explain why subjects seem able to abstract away from certain feature types in mental rotation tasks. For example, Khooshabeh and Hegarty (2008) found that subjects with relatively poor spatial acuity perform better in mental rotation tasks when matching pairs of shapes are consistently coloured. By contrast, subjects who perform well in mental rotation tasks enjoyed no such benefit. Indeed, *inconsistent* colouring failed to interfere with their performance altogether, suggesting that they were able to disregard colour entirely.

Both results can be explained by my characterisation of the underlying representations. Those with poor spatial acuity could have identified matching shapes by simply rotating spatially arranged colour maps and observing whether this brought relevant colour constituents into correspondence. Meanwhile, those with better spatial acuity may have completed the task by rotating relevant spatial feature maps (as above). But since spatial maps are independent of colour maps, subjects were able to disregard colour maps entirely, avoiding Stroop-like interference (something which would seem surprising on a thoroughgoing Pictorialism).

IV. RAMIFICATIONS

This paper has examined the suggestion that pre-attentive visual representations have an iconic (picture-like) format. Section 2 considered two prominent ways of understanding this suggestion and argued that neither is entirely adequate: Pictorialism was suggested to be *empirically* inadequate, while (an unqualified) Analogism failed to do the *explanatory* work which motivates positing iconic representations in the first place. This challenged us to provide a characterisation of pre-attentive visual format that avoids these shortcomings; a challenge which I have argued is met on the conjecture that pre-attentive vision comprises functionally independent *Analogue Maps*. On this view, pre-attentive visual representations are doubly Analogue: they situate (non-Pictorial) Analogue constituents within a functional space which (itself) mirrors objective space and possesses no digital/discursive elements whatsoever. In so doing, the proposal stands in contrast to the postulation of Pictorialist icons in this context, builds on an unqualified Analogism, and improves upon existing cartographic characterisations in the vicinity.

To close, I wish to stress that, if correct, my proposal has ramifications for debates which typically assume the iconicity of pre-attentive vision. To illustrate, consider recent debates over the existence and mark of a perception-cognition border. Here, a venerable tradition (henceforth *The Common Format Hypothesis*) holds that the perceptual is somehow marked or demarcated by the format of its representations (Block 2014). This hypothesis is notoriously controversial (Fodor 2015; Quilty-Dunn 2019b). But both sides in the debate have proceeded on the assumption that since pre-attentive visual representations are iconic, the viability of a format-based perception-cognition border requires that all perceptual representations take this form (cf. Clarke forthcoming). As such, it's notable that opponents of a Common Format Hypothesis have typically assumed a thoroughgoing Pictorialism about pre-attentive vision's iconicity. This has led them to propose that if certain perceptual representations are not Pictorialist icons, then The Common Format Hypothesis fails.

To this end, Quilty-Dunn has recently deemed perceptual object representations 'the most striking' counterexample to the Common Format Hypothesis (2019b: 9). He finds several problems with the idea that these might be Pictorialist icons. However, an initial worry stems from the suggestion that perceptual object representations track objects in abstraction from low-level properties—e.g. colour, shape, and size (see, van Marle & Scholl 2003)—something which is 'inconsistent with [the HP conformity of these representations]' (Green & Quilty-Dunn 2017: 8) and, more dramatically, the Common Format Hypothesis itself (Quilty-Dunn, 2016, 2019b).

The problem with this should now be clear: this alleged lack of HP conformity only speaks against the iconicity of perceptual object representations if one assumes a Pictorialist notion of iconicity. If one, instead, takes a cartographic characterisation as their starting point, as is motivated by the preceding discussion, proponents of a Common Format Hypothesis have no reason to expect, let alone demand, HP conformity in these representations. This is striking, since there seems to be little obscurity to the idea that perceptual systems might track and represent objects in a map-like format. After all, perceptual object tracking appears to involve an indexing of depicted objects to *spatial* locations (Pylyshyn 2007) and this is readily explained in cartographic terms since maps conform to a principle of spatial isomorphism (Camp 2007; Treisman 1988).

Admittedly, if perceptual object representations really do abstract away from the depiction of *all* object features (bar spatial location) then differences between pre-attentive visual representations and post-attentive object representations seem likely to persist. For one thing, it is likely that the constituents of our spatially arranged object maps will not be Analogue icons (as I have suggested is true of pre-attentive visual feature maps). Rather, constituents will need to function in a digital/discursive manner, signalling a sharp (non-graded) distinction between the presence and absence of given objects at their indexed locations. Even still, the relevant perceptual kind may remain unified by its broadly cartographic structure—or, perhaps, by a broader Analogue characterisation of iconicity on which cartographic representations constitute but an important sub-type (Burge 2018; Clarke forthcoming). Indeed, with these points in view, it's natural to suppose that the deepest challenge to a Common Format Hypothesis may stem from the existence of post-perceptual, *cognitive* representations with a cartographic or Analogue structuring (e.g. cognitive maps used in spatial navigation, the depictive representations used in mental imagery/imagistic thought and, perhaps, the analogue magnitude representations employed in primitive numerical computations). In any case, a Cartographic Conception of pre-attentive vision seems to command a fundamental reformulation of these debates: soothing problems which have traditionally presented problems for the Common Format Hypothesis, while highlighting neglected challenges in need of further examination.[1]

REFERENCES

Aizawa, K. (1997) 'Explaining Systematicity', *Mind & Language*, 12: 115–36.

Anderson, J. R. (1978) 'Arguments Concerning Representations for Mental Imagery', *Psychological Review*, 85: 249–77.

Beck, J. (2019) 'Perception is Analog: The Argument from Weber's Law', *Journal of Philosophy*, 116: 319–49.

Block, N. (1983a) 'Mental Pictures and Cognitive Science', *The Philosophical Review*, 92: 499–541.

—— (1983b) 'The Photographic Fallacy in the Debate about Mental Imagery', *Noûs*, 17: 651–61.

—— (2011) 'Perceptual Consciousness Overflows Cognitive Access', *Trends in Cognitive Sciences*, 15: 567–75

—— (2014) 'Seeing-As in the Light of Vision Science', *Philosophy and Phenomenological Research*, 89: 560–72.

Burge, T. (2014) 'Reply to Rescorla and Peacocke: Perceptual Content in Light of Perceptual Constancies and Biological Constraints', *Philosophy & Phenomenological Research*, 88: 485–501.

—— (2018) 'Iconic Representation: Maps, Pictures and Perception', in S. Wuppuluri and F. Doria (eds) *The Map and the Territory*. Cham: Springer.

Camp, E. (2007) 'Thinking with Maps', *Philosophical Perspectives*, 21: 145–82.

Carey, S. (2009) *The Origin of Concepts*. Oxford: OUP.

Clark, A. (2001) *A Theory of Sentience*. Oxford: OUP.

Clarke, S. (forthcoming) 'Beyond the Icon: Core Cognition and the Bounds of Perception', *Mind & Language*, doi: https://doi.org/10.1111/mila.12315.

Dehaene, S. (1997) *The Number Sense: How the Mind Creates Mathematics*. Oxford: OUP.

Dehaene, S., Piazza, M., Pinel, P. and Cohen, L. (2003) 'Three Parietal Circuits for Number Processing', *Cognitive Neuropsychology*, 20: 487–506.

Dretske, F. (1981) *Knowledge and the Flow of Information*. Cambridge: MIT Press.

Duncan J.Humphreys, G. W. (1989) 'Visual Search and Stimulus Similarity', *Psychological Review*, 96: 433–58.

Feldman, J. (2013) 'The neural binding problem(s)', *Cognitive Neurodynamics*, 7: 1–11.

[1] I am extremely grateful to Jake Beck, Denis Buehler, Christina Busse, Steve Butterfill, Laurenz Casser, Martin Davies, Casey Landers, Kevin Lande, Michael Martin, Matt Parrott, Ian Phillips, Jake Quilty-Dunn, Nick Shea, Joulia Smortchkova, and Brandon Tinklenberg for comments on earlier drafts of this paper. The resulting paper also benefitted from discussions with Lauren Emberson, audiences in Miami, Paris, Princeton, NYC, and Oxford, and unsuspecting students who were forced to read a draft for Jake Beck's graduate seminar at York and for Denis Buehler's graduate seminar in Paris. I was supported in writing this paper by York University's Vision: Science to Applications Program; thanks in part to funding from the Canada First Research Excellence Fund.

Fodor, J. (1975) *The Language of Thought*. Cambridge: MIT Press.
—— (2007) 'Revenge of the Given', in B. P. McLaughlin and J. D. Cohen (eds) *Contemporary Debates in Philosophy of Mind*, 105–16. Oxford: Blackwell.
—— (2008) *LOT2*, Cambridge: MIT Press.
—— (2015) 'Burge on Perception', in S. Laurence and E. Margolis (eds) *The Conceptual Mind: New Directions in the Study of Concepts*, 203–22. Cambridge, MA: MIT Press.
Fodor, J. and Pylyshyn, Z. (1988) 'Connectionism and Cognitive Architecture: A Critical Analysis', *Cognition*, 28: 3–71.
—— (2016) *Minds without Meanings*. Cambridge, MA: MIT Press.
Green, E. J. and Quilty-Dunn, J. (2017) 'What is an Object File?', *British Journal for the Philosophy of Science*, 72: 665–99.
Gross, S. and Flombaum, J. (2017) 'Does Perceptual Consciousness Overflow Cognitive Access? The Challenge from Probabilistic, Hierarchical Processes', *Mind & Language*, 32: 358–91.
Humphreys, G. W., Hodsoll, J. and Riddoch, M. J. (2009) 'Fractionating the Binding Process: Neuropsychological Evidence from Reversed Search Efficiencies', *Journal of Experimental Psychology: Human Perception and Performance*, 35: 627–47.
Johnson-Laird, P. (1980) 'Mental Models in Cognitive Science', *Cognitive Science*, 4: 71–115.
Khooshabeh, P. and Hegarty, M. (2008) 'Differential Effects of Colour on Mental Rotation as a Function of Spatial Ability', in *International Spatial Cognition Conference*.
Kosslyn, S. (1980) *Image and Mind*. Cambridge, MA: Harvard University Press.
Kosslyn, S., Thompson, W. and Ganis, G. (2006) *The Case for Mental Imagery*. Oxford: OUP.
Mack, A. and Rock, I. (1998) *Inattentional Blindness*. Cambridge: MIT Press.
Maley, C. (2011) 'Analog and Digital, Continuous and Discrete', *Philosophical Studies*, 155: 117–31.
Marr, D. (1982) *Vision*. Cambridge, MA: MIT Press.
Matthen, M. (2005) *Seeing, Doing and Knowing: A Philosophical Theory of Sense Perception*. Oxford: OUP.
Neisser, U. (1967) *Cognitive Psychology*. Oxford: Psychology Press.
Palmer, S. (1999) *Vision Science: From Photons to Phenomenology*. Bradford: MIT Press.
Prinzmetal, W. (2012) 'At the Core of Feature Integration Theory: on Treisman and Schmidt (1982)', in J. Wolfe and L. Robertson (eds) *Oxford Series in Visual Cognition: From Perception to Consciousness: Searching with Anne Treisman*, 211–6. Oxford: OUP.
Pylyshyn, Z. (2003) *Seeing and Visualizing: It's Not What You Think*. Cambridge, MA: MIT Press.
—— (2007) 'Multiple Object Tracking', *Scholarpedia*, 2: 3326.
Quilty-Dunn, J. (2016) 'Iconicity and the Format of Perception', *Journal of Consciousness Studies*, 23: 255–63.
—— (2019a) 'Is Iconic Memory Iconic?' *Philosophy & Phenomenological Research*, 101: 660–82.
—— (2019b) 'Perceptual Pluralism', *Noûs*, 54: 807–38.
Shepard, R. N. and Metzler, J. (1971) 'Mental Rotation of Three-Dimensional Objects', *Science*, 171: 701–3.
Sober, E. (1976) 'Mental Representations', *Synthese*, 33: 101–48.
Sperling, G. (1960) 'The Information Available in Brief Visual Presentations', *Psychological Monographs: General and Applied*, 74: 1–29.
Treisman, A. (1985) 'Pre-Attentive Processing in Vision', *Computer Vision, Graphics, and Image Processing*, 31: 156–77.
—— (1986) 'Features and Objects in Visual Processing', *Scientific American*, 255: 114–25.
—— (1988) 'Features and Objects: The Fourteenth Bartlett Memorial Lecture', *The Quarterly Journal of Experimental Psychology Section A*, 40A: 201–37.
Treisman, A. and Gelade, G. (1980) 'A Feature-Integration Theory of Attention', *Cognitive Psychology*, 12: 97–136.
Treisman, A. and Patterson, R. (1984) 'Emergent Features, Attention, and Object Perception', *Journal of Experimental Psychology: Human Perception and Performance*, 10: 12–31.
Treisman, A. and Schmidt, H. (1982) 'Illusory Conjunctions in the Perception of Objects', *Cognitive Psychology*, 14: 107–41.
van Marle, K. and Scholl, B. (2003) 'Attentive Tracking of Objects Versus Substances', *Psychological Science*, 14: 498–504.
von Neumann, J. (1958/2013) *The Computer and the Brain*. Newhaven: Yale University Press.

Wolfe, J. M. (2018) 'Visual Search', in J. T. Wixted (ed.) *Steven's Handbook of Experimental Psychology and Cognitive Neuroscience*, Fourth Edition. Hoboken, NJ: John Wiley & Sons, Inc.

Wolfe, J. M., Cave, K. R. and Franzel, S. L. (1989) 'Guided Search: An Alternative to the Feature Integration Model for Visual Search', *Journal of Experimental Psychology: Human Perception and Performance*, 15: 419–33.

Wolfe, J. M. and Horrowitz, T. S. (2017) 'Five Factors that Guide Attention in Visual Search', *Nature Human Behaviour*, 1: 0058.

Zeki, S. (1993) *A vision of the brain*. Cambridge, MA: Blackwell Scientific Publications.

[1] *Department of Philosophy and Centre for Vision Research, York University, Canada*

[2] *Departments of Philosophy, Psychology, and Linguistics, University of Pennsylvania, USA*