Asian Journal of Philosophy (2022) 1:50
https://doi.org/10.1007/s44204-022-00039-x

ORIGINAL ARTICLE

Three grades of iconicity in perception

Jack C. Lyons[1]

Received: 8 March 2022 / Accepted: 1 July 2022
© Crown 2022

Abstract
Perceptual representations are sometimes said to be iconic, or picture-like. But what does this mean, and is it true? I suggest that the most fruitful way to understand iconicity is in terms of similarity, but there are three importantly different grades of similarity that that might hold between perceptual representations and their objects, and these should be distinguished. It is implausible that all perceptual representations achieve even the weakest grade of iconicity, but I speculatively suggest a "Kantian" view, whereby all perceptual representation involves fillers placed in a "strongly iconic" spatial framework. Those fillers will likely be hybrid structures, with varying grades of iconicity combined with non-iconic elements.

Keywords Representation · Iconic · Analog · Perception · Pictorial representation

It is often held that perceptual states are iconic—in fact, some hold that this is part of what distinguishes perceptual states from other, cognitive, states (e.g., Block, 2014, forthcoming; Burge, 2010, 2014; Raftopoulos, 2019). But what does it mean to say that perceptual states are iconic, and what evidence is there for thinking it's true? In what sense(s) might perception be distinctively iconic?

The basic idea, in somewhat crude form, starts with a contrast between *iconic*, or picture-like representations, on the one hand, and *symbolic*, or language-like representations on the other hand. I won't try to say much in advance about what's meant by "picture-like" or "language-like," hoping that the relevant likenesses will emerge as we progress. The general characterization should be quite familiar from the existing literature. Iconic representations include maps, charts, diagrams, models, drawings, sculptures, and, most paradigmatically, photographs and realistic paintings. I'll use "symbolic" to refer to language-like representations, hoping the idea is clear enough for now.[1]

[1] Some writers prefer "propositional" or "discursive" to "symbolic". I'll assume that all non-iconic representations (or non-iconic aspects of representations) are symbolic, though I won't be putting any weight on that assumption.

✉ Jack C. Lyons
jack.lyons@glasgow.ac.uk

1 University of Glasgow, Glasgow, Scotland, UK

Published online: 14 November 2022

I have two main aims here. One is to better understand iconic representation—how it works, what forms it takes, how it relates to other forms of representation. The other is to see what kind of connection it has to perceptual representation.

Some of the literature on iconicity in the philosophy of mind has focused on paradigm cases like photographs and paintings, ignoring important features of the larger class. Some focusses on the very broadest, most abstract construal of iconicity, ignoring important differences between, say, maps and vector representations in connectionist networks. In fact, depending on how strictly we want to use the term "iconic," we can distinguish three nested classes. Genuinely iconic representations are a special case of what I will call "co-similar" representations, and some of the more interesting features of more paradigmatically iconic representations are features they share with co-similar representations more generally. Even this broadest category, however, does not plausibly distinguish perception from cognition. Although I doubt that all perceptual representation is fully iconic—even in the most liberal sense of "iconic," I think an intriguing but underexplored possibility is that all perceptual representation involves the placement of fillers in a spatial framework—where, although those fillers might not be at all iconic, the framework itself exhibits the highest grade of iconicity.

Although the most natural and obvious way to understand iconicity is in terms of similarity, or resemblance, several authors have tried to understand it in other terms. In Sect. 1, I briefly consider some popular attempts to cash out iconicity in non-similarity terms. All of these are limited in important ways. In Sect. 2, I turn to similarity. Similarity is often glossed in terms of isomorphism, but this is unhelpful unless isomorphism is constrained in certain ways. Articulating these constraints allows us to distinguish the first two grades of iconicity. In Sect. 3, I offer a more rigorous and more general view of iconicity, distinguishing the three grades of interest here. Section 4 turns to the question of iconic *mental* representation and the question of whether mental representations can themselves have the kind of spatial properties needed for the strongest grade of iconicity. In Sect. 5 I use all the foregoing to ask whether and to what extent various perceptual representations might plausibly be iconic. I suggest a kind of Kantian view whereby there is a strongly iconic element in many or all perceptual representations, but whereby there will very likely be non-iconic elements in most or all of them as well.

Some clarificatory and terminological points are in order before proceeding. First off, my concern is not with our use of the *term* "iconic." I will sometimes make appeal to the kinds of representations we normally take to be iconic, but our normally taking them thus is not intended as evidence for anything, unless there's independent reason—as I hope to provide later on—for thinking that they constitute a natural or otherwise interesting category, or family of categories, with interesting similarities among the principles that govern them.

Second, many representations are hybrid; there are symbolic representations with iconic elements (like rebuses), and iconic representations with symbolic elements (most maps, for example, have words written on them, and even many of the picture-like elements—so-called icons—seem, like the gender markers outside of restrooms, more symbolic than iconic). Until the possible hybridity of perception takes front and center, I won't worry about hybridity or purity. I'll normally treat maps and the like as if they were purely iconic, with the understanding that the discussion is only aimed at and only applies to the iconic components.

Finally, I want to make very explicit that, until Sect. 4, when I turn to mental representation, all my talk about pictures, maps, and the like (even sentences, for that matter) is aimed at ordinary, literal, physical, *non-mental* representations. We do, of course, want to know whether and how mental entities can be iconic or symbolic, but I'll want these questions to be kept separate initially, for reasons that will emerge below.

1 Non-similarity-based distinction criteria

What distinguishes pictures and other iconic representations from language and other symbolic representations? The most obvious answer is that pictures are in some way *similar to* the things they represent. However, some very different answers have been at least as influential in recent philosophy of mind, and I want to discuss some of these alternatives briefly before returning to similarity in more detail.

Not everyone who is comfortable with representation talk is equally happy with talk about representational *vehicles*. If we could capture iconic representation in terms of contents, we could avoid reference to vehicles and thus avoid attributing to representational states any kinds of properties other than what they must—as representational states—have: semantic properties. Although I'm not sympathetic, I'll first consider distinction criteria that are concerned only with iconic *contents*, before turning to those that involve vehicle structure. For now, I want to just briefly gloss the candidate criteria without worrying about how successful they might me. All have some *prima facie* plausibility.

1.1 Distinctive "skeletal" contents

Haugeland (1991) has argued that the distinguishing feature of iconic representations is that they have special contents that differ from the contents of non-iconic representations. To boost the credibility of this claim, he distinguishes between the "skeletal content"—what is explicitly represented—and the "fleshed out content"—what an informed user could easily extract from the representation. Thus, the fleshed out content of a photograph might represent something as a horse, in that most viewers would be able to see that it was a picture of a horse; but what it represents strictly—its skeletal content—is just the instantiation of a certain highly determinate spatial arrangement of color patches. We might invoke the old cliche that a picture is worth a thousand words and claim that iconic representations have a content so rich and detailed that they are difficult or impossible to replace with expressively equivalent symbolic representations.

1.2 Mandatory commitment

A point that goes back to Berkeley (1710/1975) and Hume (1739/1978) is that iconic representations have mandatory representational commitments, in a way that language-like representations don't. A photograph of a car can't be neutral about the shape of the car, or what direction it's facing, or whether it has wheels (if shown from the correct viewpoint and unoccluded); the word "car" is silent on all of these (Cummins, 1989). A picture is always from a certain angle, and the features visible from that angle have to be either determinately rendered or deliberately obscured; they can't simply be left "unmentioned." Two distinct phenomena are worth pointing out here. One is "mandatory specificity": a picture cannot be general and neutral in the way that symbols can, but must represent its objects as having highly determinate and specific properties. Another is "representational entanglement": *if* a picture represents something as having one kind of property, then that same representation has to represent it as having another kind of property. You can't pictorially represent the orientation of some item in an array without representing its location in that array (Quilty-Dunn, 2020); you can't depict the color of an object without depicting its shape, etc.[2]

These are quite different from the previous point about rich contents. Haugeland's point was about what pictures can do that symbols can't; the current point is about what pictures can't do that symbols can.

If we're willing to talk about representational *vehicles* and their properties, there are a number of other ways we might try to distinguish iconic representations.[3]

1.3 Analogicity/continuity

One obvious feature that distinguishes, say, photographs from linguistic representation is that the former are continuous, while the latter are discrete. I myself am tempted to characterize this difference using the terms "analog" and "digital," respectively, but these terms are also used in very different ways, including one whereby "analog" means at least approximately what I will later mean by "co-similar," rather than "continuous."[4] To avoid confusion and terminology disputes, I'll just (mostly)

[2] Mandatory commitment is connected to Dretske's well-known distinction between analog and digital information-bearing. For Dretske (1981), if a signal *r* carries the information that *s* is *F*, it does so digitally whenever *r* carries no additional information not already nested in the information that *s* is *F*; it carries that information in analog form otherwise. Most of the information a picture carries, it carries in (Dretske-) analog form; not so with symbols. The notions of mandatory specificity, representational entanglement, and Dretske-analogicity are all open-ended with respect to just which properties must be additionally represented. As it should be, for these will vary from one representational system to another. A black-and-white photograph must represent shape and brightness, but not color; a color photograph must represent color (though perhaps only apparent or perspective-relative color) as well. There are likely iconic (/non-symbolic) representations accessible to other sense-modalities besides vision. An auditory representation of, say, a dog barking must represent pitch, timbre, and number of barks, but not color, shape, or orientation, etc.

[3] Before leaving the topic of differences in content, it should be noted that maps are sometimes said to be non-predicative, or to predicate in ways different from symbolic representation (Camp, 2018; Casati & Varzi, 1999; Kulvicki, 2015; Rescorla, 2009). I haven't discussed this here partly because there's too much disagreement among these authors about what the relevant difference is. Also, because this view of maps as fundamentally involving the placement of markers at coordinates, it makes the "predication" primarily a matter of ascribing property instances to spatial locations, in a way that would apparently distinguish maps from pictures. As such it's not a good candidate for distinguishing iconic from symbolic representation.

[4] E.g., Beck (2019), Clarke (2022), Maley (2011). Dretske's (1981) analog/digital distinction is, as he admits, an "unorthodox" (p. 137) use of the terms. Note that 'analog' when contrasted with 'digital' almost always denotes continuity, especially when used outside of philosophical contexts. Counting with one's fingers (/digits) will be "analog" in the sense of "co-similar" but obviously "digital" in the sense of "discrete".

eschew the terms “analog” and “digital,” with the understanding that a representation will be “continuous” in roughly the sense that analog electronics or recordings are continuous (similarly for discrete/digital). Of course, although the concepts themselves are absolute, we normally use a bit of latitude in applying them. A water clock is a paradigmatically continuous device, although water is of course composed of discrete molecules. An hourglass using grains of sand is continuous in the relevant sense, even though grains are discrete, etc.

The continuity of pictures seems to be a feature of the pictorial vehicle itself, not of its content. We might be agnostic about what kinds of contents pictures have while still finding it obvious that they are continuous. And continuities are easily represented in symbolic form, as is common in mathematics. Thus, is the vehicle, not the content, whose discreteness or continuity is at issue.

1.4 Picture principle

Fodor (2007, see also Carey, 2009) holds that iconic representations are just those that satisfy the “Picture Principle”:

(PP) If *R* is a picture of *x*, then every part of *R* is a picture of some part of *x*.

Parts of words, on the other hand, typically don’t represent anything and certainly don’t always represent parts of what the whole word represents. But, says (PP), a picture remains a picture, no matter how you cut it.

1.5 Limitations of non-similarity-based distinction criteria

These four candidate distinction criteria are prima facie plausible accounts of what distinguishes, say, photographs from linguistic representations, but they don’t hold much promise of capturing a more general class of iconic representation.

Consider subway maps, for example. A typical subway map consists of circles representing stations or stops, and lines (sometimes curved) connecting them. Distances between the circles don’t mean anything, nor do the shapes of the lines. But the *order* of the circles connected by a line represents the order of stops on that subway line. A subway map is less similar to the actual subway system than an aerial photo (with ground-penetrating radar) would be—or more abstract in its similarity—but it still resembles the system that it represents, in a way that a complex verbal description does not. In fact, the similarity and the abstractness of the similarity combine to make subway maps so effective: the abstractness prunes away all irrelevant information, other than the order of the stops, location of transfer points, etc.; at the same time, the picture-like-ness of the representation makes that information user-friendly in a way that the verbal description would not be.

Subway maps violate Fodor’s picture principle. The left half of a given circle doesn’t represent half—or any part—of the stop that the whole circle represents. Parts of the circles and the lines don’t represent anything at all, contrary to (PP); they function similarly to submorphemic parts of words. Not because subway maps represent arbitrarily, as words do, however, but because subway maps are discrete

(/digital), as words are. Fodor thinks (PP) shows that iconic representations lack canonical decompositions into constituents; but this is simply mistaken. It's only *continuous* iconic representations that lack canonical decompositions; subway maps have clearly defined, canonical, constituent structure: only *some* cuts result in a smaller map. The same is true for other discrete iconic representations, like the Tinker-toy molecule models from high school chemistry class.

Subway maps also refute the idea that iconic representations have distinctive contents. The contents—the truth- or satisfaction-conditions—of a subway map can be given without loss or addition by a complex verbal description. The verbal description is, of course, much harder to use, but this is just an instance of the very general and familiar point that the form of representation matters.

Nor are subway maps bound by mandatory commitment or representational entanglement. They are noncommittal about shapes and distances in a way that aerial photographs cannot be; one can represent the order of stops without representing *anything* else that isn't entailed by that.[5]

It's not just subway maps that cause trouble for these criteria. The same points could have been made with molecular model kits, seating charts, pie charts, bar graphs, histograms, etc., (although I'll leave those arguments as an exercise for the reader). In short, the non-similarity-based distinction criteria are initially plausible not for iconic representations in general, but only for *continuous* iconic representations. If continuity is not a necessary or even a very deep feature of iconic representation, then these distinction criteria are no good.

2 Similarity/isomorphism

This brings us to the elephant in the room. The four distinction criteria just glossed steer clear, apparently deliberately, of similarity. Yet, this is one of the most obvious features of pictures, and a striking point of contrast between them and symbols, whose representational relations are saliently *arbitrary*, in the sense that it's causation, or convention, or something else extrinsic to the representation that determines content. A salient fact about pictures is that they are *like* their representanda. This is true not only of photographs and realistic paintings but of subway maps, charts, graphs, and everything else that seems to be a clear case of iconic representation.

Resemblance, or similarity, is normally and most naturally understood in terms of sharing concrete properties. But the philosophical literature standardly prefers the technical concept of isomorphism, which is a sharing of abstract rather than concrete properties. I'll come back to this in more detail later, but the rough idea is that A and B are isomorphic just in case the *structure* of A mirrors the *structure* of B, that is, that the relations among the elements that make up A can be put into correspondence with the relations among the elements that make up B. Iconic representations

[5] To return to Dretske's terminology: all the information carried by a subway map is carried digitally; and it's exactly the same information as carried by some verbal (hence symbolic) description.

are then held to be those that represent in virtue of isomorphism (e.g., Burge, 2018; Cummins, 1996a).

Pictures exhibit two importantly different kinds of isomorphism. Pictures are isomorphic both at the level of the *system* of representation—the representational scheme—and also at the level of the individual representation. The idea of a representational scheme should be pretty intuitive (although I'll return to it in more detail below). A Venn diagram scheme is different from a Euler diagram scheme, a color pictorial scheme from a monochromatic scheme, a Roman numeral scheme from an Arabic numeral scheme, etc.

The first kind of isomorphism just mentioned, which I will call "scheme isomorphism," is much more widespread than iconic representation. The internal states of a calculator can represent the numbers only because the causal relations among those states correspond to mathematical relations among the numbers (Block, 1995; Cummins, 1992). There is therefore an isomorphism between the scheme, consisting of all the possible internal representational states of the calculator, and the domain, consisting of numbers in Plato's heaven. The *system* of representation (the set of internal states of the calculator) mirrors the representandum domain, although there is no interesting sense in which the individual representations mirror their individual representanda. Similarly with Arabic or Roman numeral systems, the ordering relations among the representations map onto successor relations among the numbers, which makes the former capable of representing the latter. But the Arabic numeral "4" isn't in any interesting sense similar to, or isomorphic to, the number 4. No more than it is to the number 27, anyway. So the relevant isomorphism here holds between representational scheme and representational domain, not between individual representation and representandum. These schemes are paradigmatically symbolic, despite this scheme isomorphism.

By contrast, pictures, maps, and the like also exhibit "vehicle isomorphism:" the individual representation is isomorphic to the thing it represents. It's important to keep scheme isomorphism and vehicle isomorphism separate, as it's only the latter that is distinctive of iconic representation.

I think the central distinguishing feature of iconicity is similarity, or resemblance, but of a sort that's obscured, rather than illuminated, by simply calling it isomorphism. Paradigmatic iconic representation (photographs, paintings, etc.) is essentially spatial—that is, the vehicles themselves are essentially spatial, and so too are the representanda. The representation and representandum might diverge in some of their spatial properties—one is bigger and three-dimensional, etc.—but they will also share many spatial properties; the internal, relational spatial properties of the photo (the arrangement of the elements) will be not merely isomorphic to but literally shared by the representandum.

Discussion of iconic representation often starts with the intuitive idea of similarity. But, maybe because we're hoping to eventually connect our talk about photographs with *mental* representation, and maybe because the technical notion is more precise, this intuitive notion is quickly displaced by talk about isomorphism. But it's not (mere) isomorphism that makes a map a map; it's that it has the very spatial properties that it represents the geography as having. Iconic representations *are* isomorphic to their representanda, but the bare isomorphism isn't explanatory of their contents, or of much of anything. What's explanatory is the very particular and rather rare species of isomorphism that obtains when two things are similar in virtue of sharing concrete properties.

Note that representations and (at least some) representanda are concrete particulars. The former might include things like brain states, or neural events, or—more relevant to our current focus on non-mental representations—pieces of paper or canvas with markings on them, and the like; the latter are things like cities, babies, scenes, sporting events, and so on. Isomorphism—when it's not just a five syllable replacement for "resemblance"—is a relation that's only defined for relational structures, which are not concrete particulars but ordered structures consisting of a set of objects and a set of relations over those objects. One relational structure $\langle A, R_1,... R_n\rangle$ is isomorphic to another relational structure $\langle B, S_1,... S_n\rangle$ just in case there is a function, *f*, that maps A onto B, and for every m-tuple of elements in A, and every R_i and S_j, $\langle x_1,... x_m\rangle \in R_i$ iff$\langle f(x_1), \ldots f(x_m)\rangle \in S_j$.

To say that two concrete particulars are isomorphic, then, can only mean that they *instantiate* relational structures that are isomorphic to each other.[6] The representation has a number of parts (pieces, components, elements), and those parts bear various relations to each other, which together specify a relational structure; same for the representandum. The particulars are isomorphic just in case these relational structures are. We shouldn't require that every part of *r* has to map to every part of *o* in order for *r* to represent *o* (i.e., that *every* relational structure instantiated by *r* is isomorphic to every relational structure instantiated by *o*).[7] That would make isomorphism extremely rare. The obvious alternative, however, makes isomorphism ubiquitous. My bicycle can be divided into two parts (right/left or inside/outside) and a spatial relation between those parts. The fall of the Roman Empire can be divided into two parts (the earlier part and the later part) and a temporal relation between them. So my bicycle is isomorphic to the fall of the Roman Empire. Clearly, if we only require that *some* relational structure instantiated by *r* be isomorphic to some relational structure instantiated by *o*, then pretty much anything is going to be isomorphic to pretty much anything else.

There are various ways to try to make the general notion of isomorphism less trivial, including appeal to canonical decompositions and/or "natural" relations.[8] But I think it's more helpful to go directly to the paradigm cases and broaden our view outward from there. The discussion so far ignores one salient and highly significant

[6] It follows that two items can be nontrivially isomorphic to each other only if they both have a significant amount of internal structure. Items without internal structure could still be similar to each other, even though not isomorphic. If colors are unstructured, and if one thinks, with Burge (2018), that paint swatches are iconic representations of colors, then this might make it seem that a generic resemblance view is better than one that analyzes similarity in terms of isomorphism. I'm unconvinced that swatches are representational in any robust sense. I might take a fabric sample to the store to find one that matches, just as I might take a glass slipper around the kingdom to find a foot that fits it. Imbuing the paint chip, or the fabric, or the slipper, with genuine semantic content seems unnecessary, unilluminating, and unparsimonious.

[7] For this reason, homomorphism may be a better model for the representation relation than isomorphism. But since concrete particulars will instantiate a large number of relational structures, this distinction doesn't matter for the present purposes.

[8] Burge (2018) pursues this line, arguing that iconic representations have canonical or natural decompositions revealed by our visual processing of them.

fact about paradigm iconic representations: "pictures and maps *have* the very properties they represent their objects as having." Maps don't just represent spatial properties by having properties that *behave like* space; they represent spatial properties by having those very same spatial properties. Similarly with photos and paintings. In general, pictures satisfy a "shared property principle":

(SPP):*r* represents *o* as *F* only if *r* is itself *F*.[9]

This point is hardly new or unnoticed. But it's oddly absent from discussions about what distinguishes iconic representation, and I think its implications have been underappreciated.

The obvious kinds of properties pictures and maps share with their objects are spatial properties (e.g., length, shape, angle, ordering, orientation, etc.; some of these will be viewpoint-dependent spatial properties (Hill, 2016, 2021; Schellenberg, 2008)).[10] We can think of pictorial representation as isomorphism, if we want, but it's the very special case of isomorphism under a mapping relation of spatial similarity. This, of course, requires both vehicles and representanda to have spatial properties and relations.

If we want to allow charts, graphs, diagrams, and the like to count as iconic representations, as I think we do, we need to allow for the representation of non-spatial properties by the spatial properties of the representation. The angle (or area) of a wedge in a pie chart represents the percentage of the budget spent on that category; the height of a bar in a bar chart represents the murder rate for that country, etc. In these cases, we have spatial properties representing non-spatial *quantities*. The isomorphism required here is less constrained than SPP requires, but it's still highly constrained; we have a spatial quantity on the one end, and on the other end, a property that must be in some sense analogous to that spatial quantity. Such representations satisfy a weaker principle than SPP, an "analogous property principle":

(APP):*r* represents *o* as *G* only if *r* has some property *F*, that is analogous to *G*.

Obviously, "analogous" will have to be unpacked in some detail, which I will do shortly, but the examples render it clear enough for the time being.

Thus, we can distinguish our first two grades of iconicity: "strong iconicity" and "weak iconicity." The former requires the representation and the representandum to share properties, while the latter requires them to have analogous properties. Sticking with these familiar examples, which rely on the spatial properties of the representations, strongly iconic representations use their spatial properties to represent their objects as having those very same spatial properties, while weakly iconic representations use their spatial properties to represent their objects as having some distinct but space-analogous quantity. Both are constrained forms of vehicle isomorphism.

[9] Representations that satisfy SPP will be, as Millikan (2004, 2017) calls them, "relative reflexive signs," although this concept is quite a bit broader than iconicity. Indexicals like "yesterday" are also relative reflexives but don't satisfy SPP.

[10] They sometimes share color properties as well, but I will set those aside for now.

3 A broader view

I've been focusing on the spatial properties of the representational vehicles, partly because these are the noteworthy representational elements of the most familiar iconic representations: pictures, maps, charts, diagrams, sculptures, etc. But I mean to be inclusive here, and what's distinctive about, say, strongly iconic representations is not their reliance on spatial properties per se but their satisfying SPP. I don't know if audio recordings and crash test dummies count as representations of musical performances and car passengers, but if they do, then we have temporal, pitch, timbre, and weight, articulation, and flexion properties of one thing representing something else as having those very same properties.

It is easier to think of non-spatial but clearly representational examples of weaker forms of iconicity. Geiger counters use click frequency to represent radiation intensity; metal detectors use tone pitch to represent proximity of a metallic object, etc. We could use light intensity to represent volume, or simple melodies to represent contours. All of these cases are non-symbolic, in that they trade on a kind of resemblance, or isomorphism between representations and representanda.

I complained above that *mere* isomorphism is fairly trivial. To make isomorphism at all explanatory, we need to make explicit some features that have only been implicit so far. In all of these cases, from pictures to pitches, the representational properties—the properties of the vehicle that do the representing, and the properties that the representandum is represented as having—are, as we might call them, "dimensionable": they are determinables the determinates of which can be laid out in nontrivial and nonarbitrary orderings in a one-or-more-dimensional similarity space.[11] In addition, the mappings from representations to representanda are "order preserving": it's not enough that there be a distinct pitch for every distance represented— the ordering of the pitches mustn't criss-cross the ordering of the distances. We could have higher pitches representing longer distances or higher pitches representing shorter distances, but if pitches A, B, and C represent distances A*, B*, and C*, respectively, then B is between A and C iff B* is between A* and C*.

Dimensionability and order-preservingness can now be read back into weak iconicity to give substance to the earlier brief appeal to properties "analogous" to vehicle properties. Spatial and color properties are dimensionable, and identity is order preserving, so the paradigm examples of strongly iconic representations will automatically satisfy dimensionability and order-preservingness.

For the last three paragraphs, I've been deliberately postponing the distinguishing of two importantly different phenomena. Using pitches to represent distances differs from using melodies to represent contours, in that the latter involves vehicle isomorphism, while the former only involves scheme isomorphism. We can use middle-C to represent 7 ft, and C# just above middle-C to represent 8 ft, etc. But these same pitches could just as easily (and no more or less iconically) represent 7 ft and 6 ft, respectively, or 3 in and 6 in, etc. This is because C# doesn't resemble 8 feet in any interesting sense, any more than it resembles 6 ft or 6 in. There is resemblance in the pitch-distance case, but the resemblance—the isomorphism constrained by dimensionability and

[11] A dimensionable property can't always be measured on an "ordinal scale" (Stevens, 1946), although anything ordinally scalable is dimensionable. Hues and pitch classes, for example, form (roughly) circular similarity spaces, where there isn't any *first* element; they are nevertheless dimensionable in the relevant sense.

order-preservingness—resides at the level of the scheme, not the individual representational vehicle. And it couldn't really be otherwise, since vehicle isomorphism requires the representation to have internal structure, which pitches don't have in any significant sense. Melodies obviously do, which is why a melody might represent a particular contour in virtue of being isomorphic to it. I want to count melody representations of contours as more picture-like than pitch representations of distance, because the former rely on vehicle isomorphism, while the latter rely only on scheme isomorphism. I'll call the latter "quasi-iconic," a term intended to highlight the difference between vehicle isomorphism and scheme isomorphism.

Here, then, are three grades of iconicity:

- A *quasi-iconic* representation is one that represents at least partly in virtue of being an element of a representational scheme that is isomorphic, under an order preserving mapping, to its representandum domain, in respect to the dimensionable properties of each.
- A *weakly iconic* representation is one that represents at least partly in virtue of being vehicle-isomorphic, under an order preserving mapping, to its representanda, in respect to the dimensionable properties of each.
- A *strongly iconic* representation is one that satisfies the conditions for weakly iconic representation, where the mapping function is not merely order preserving, but one of identity.

Because it doesn't exhibit vehicle isomorphism, quasi-iconicity is the weakest of the three grades, the most un-picture-like. However, quasi-iconicity counts as (somewhat) iconic on my taxonomy because of a very important feature it shares with the stronger grades. Although scheme isomorphism without restriction is too commonplace to be of much interest, quasi-iconicity involves a rarer form of scheme isomorphism, one that is constrained by dimensionability and order-preservingness. In any such system, similarity relations among the representations that constitute the scheme are mirrored by similarity relations among the contents of those representations. Semantically similar representations (i.e., representations with similar contents) will be structurally similar, and vice versa. Call this feature *co-similarity*.[12]

[12] I think that quasi-iconicity either is or is very close to what Shepard meant by "second-order isomorphism" (Shepard 1968, 1975, 1978; Shepard & Chipman, 1970). In these classic papers, Shepard says things about isomorphism that would be very puzzling if we take the term in its mathematical usage but not at all if we take it as simply a highbrow term for similarity. He thinks that first-order isomorphism, for example, would require mental states to be square or green in order to represent something as being square or green, and this is why he disavows first-order isomorphism in favor of second-order isomorphism:

> isomorphism should be sought—not in the first-order relation between (a) an individual object, and (b) its corresponding internal representation—but in the second-order relation between (a) the relations among alternative external objects, and (b) the relations among their corresponding internal representations. Thus, although the internal representation for a square need not itself be square, it should (whatever it is) at least have a closer functional relation to the internal representation for a rectangle than to that, say, for a green flash or the taste of persimmon (Shepard & Chipman, 1970, 2).

This makes sense if the only options under consideration are what I'm calling strong iconicity and quasi-iconicity.

The Arabic numeral scheme and a hash mark scheme are both scheme isomorphic to the domain they represent. The Arabic numeral scheme is not co-similar with its domain, though the hash mark scheme is as follows: "3" isn't any more similar to "4" than it is to "8," even though 3 is more similar to 4 than it is to 8; yet | | | is more similar to | | | | than it is to | | | | | | | |. The difference is that the representations that make up the Arabic scheme aren't dimensionable. There's an ordering to them, but it's an arbitrary one, in the sense that it's imposed from without; it's not intrinsic to the scheme, as it is for the hash mark system. Co-similar representational schemes are powerful, because they make it possible for representations of things to really serve as proxies for those things in a way that symbolic schemes don't allow. (Stronger forms of iconicity make for even better, more robust, proxy relations.) A system that exhibits co-similarity will automatically tend to handle error and novelty quite well. Mistaking one representation for a similar one will have relatively inconsequential effects, since the representanda will be similar as well. Representations of novel stimuli can be processed like the representations of familiar stimuli, normally with only minimal distortion.

For these sorts of reasons, co-similarity plays an important role in distributed vector representation in connectionist networks. Other vector representation schemes are similar: colors, for example, can be represented as triples of values along the three dimensions of hue, saturation, and value (or along red, blue, and green if you rather). In this scheme, nearby vectors will represent similar colors. We can talk here about points or locations in color "space," but these aren't normally thought to involve literal spaces in the way that strongly and weakly iconic representations sometimes do.[13]

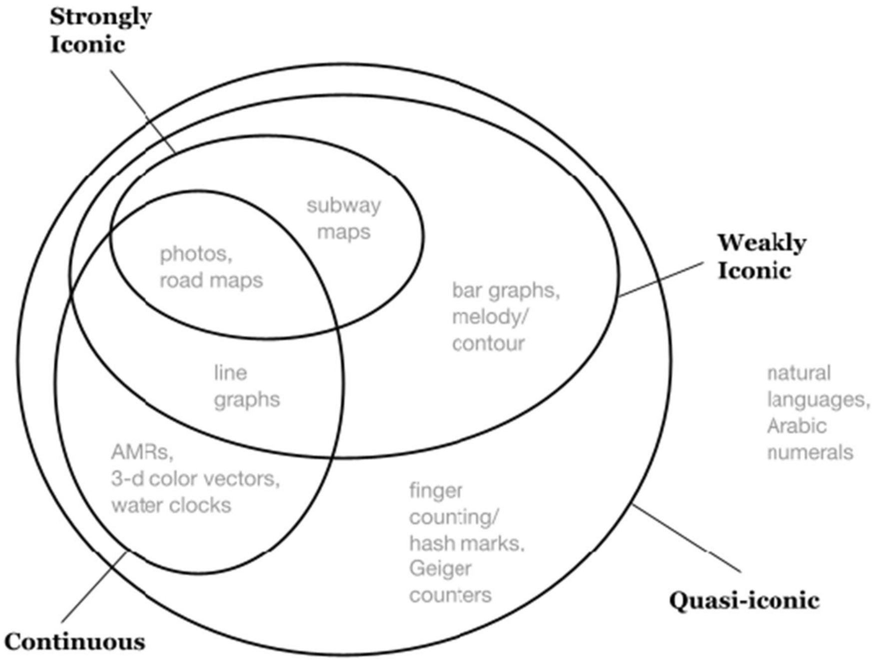
Although the nomenclature might suggest exclusivity, the definitions make it clear that at least the higher grades of iconicity are nested: all strongly iconic representations are weakly iconic, though not vice versa. It's somewhat less obvious that both of these are quasi-iconic. It's plausible enough that pictures with similar contents are themselves similar, that similarities among budgets will be reflected by similarities in their pie charts, and so on.

I think it's possible to give more of an argument than this. Strong and weak iconicity require that representations be similar to their representanda, in the sense that the representations be order-preservingly mappable onto their representanda in respect to their dimensionable properties. Part of what makes two representations elements of a common scheme is that they have a common semantics, which means that they employ the same order preserving mapping. But if the mapping is order preserving, then two representations that are similar in respect to their dimensionable properties will be mapped to representanda that are similar in respect to their dimensionable properties; similarities among the representations will be met with similarities in the properties attributed by those representations. Thus, the representational scheme will be isomorphic

[13] Although we don't normally, we could use literally spatial representations to represent colors. Most talk about vector representations of color, however, are meant to implicate quasi-iconicity rather than a stronger grade.

under an order preserving mapping, to the representational domain, in respect to the dimensionable properties of each. That is, any representational system whose elements exhibit vehicle isomorphism of the relevant type will exhibit scheme isomorphism of the relevant type.

The following Euler diagram indicates the relations among the three grades of iconicity, and continuity, along with some representative examples. Weakly iconic representations are, again, meant to include strongly iconic representations, and quasi-iconic include both. As I see it, the question of continuity and discreteness is largely orthogonal to the question of iconicity. Each of the three grades of iconicity admits of continuous as well as discrete instances. Strongly iconic representations include such continuous representations as photos and road maps, but also discrete representations like subway maps and molecular models. Line graphs are typically continuous and weakly iconic; bar graphs can be discrete (although they're typically continuous on the *y* axis). The contrived example of using melodies to represent contours is discrete, assuming the melodies are played on the piano, but it could be continuous if played on a trombone or a theremin, for example. Quasi-iconic representations include hash mark schemes, which are essentially a version of finger counting that can go past 10, as well as a number of familiar continuous instances, like water clocks, vector representations, and the analog magnitude representations used in cognizing about numerosities (see Carey, 2009; Dehaene, 2011). Outside the largest circle are symbolic representations.



Various sorts of hybrids and combinations are possible. Maps, of course, typically have symbolic elements in addition to their strong iconicity. They can also have weakly iconic and quasi-iconic elements, as when using color to represent air pressure or wind speed. Weak iconicity requires vehicle isomorphism, which requires internal vehicle structure; so when colors quasi-iconically represent wind speeds, then *patterns* of color will weakly iconically represent *patterns* of wind speeds. The relation between particular colors and particular wind speeds is scheme-level and a matter of co-similarity. Iconic representations can also combine discrete and continuous elements. A budgetary pie chart, for example, usually represents budget categories discretely and symbolically and represents percentages continuously and (weakly) iconically.

The diagram claims that all continuous representations are at least quasi-iconic.[14] I don't have a decisive argument for this, but I think it's probably true. All the instances I can think of are pretty clearly so. Additionally, I think that the noise inherent in continuous representations—the ease of mistaking one representation for a similar one—would be devastating if similar representations weren't constrained to have similar contents; thus, as a practical matter at least, continuity requires quasi-iconicity.

Again, this isn't about the *word* "iconic," and I wouldn't mind much if someone wanted to use it only for something much narrower, perhaps just those strongly iconic representations that represent in virtue of their spatial properties. What really matters are the distinctions illustrated by the Euler diagram. There are important differences between systems that represent by sharing properties and those that represent by having analogous properties. There are important differences between systems whose individual vehicles are similar to their representanda, and systems where the similarities only occur at the system level. And yet there are important similarities among all of these; co-similarity distinguishes all of these from symbolic representation and makes sense of the intuitive claim that iconic representation is somehow less "arbitrary" than symbolic representation.

4 Iconic *mental* representations

So far, I've deliberately focused on tangible, physical, non-mental representations, on paper maps and photographs and the like. My hope, however, is to connect this up with mental representation in an effort to better understand the latter. The most interesting question in this neighborhood, I think, is whether the mind employs any of the essentially spatial strongly iconic representations discussed above. I'll focus on maps, but I think the same considerations all apply to pictures as well. The idea that there are literally maps in the head now seems problematic, in a way that the idea that there are literally sentences in the head is not. Sentences are individuated only by their semantic and syntactic properties, so the thought that minds or brains could literally have sentences in them doesn't pose any fundamental trouble. It's just

[14] I.e., everything that's "analog" in the sense I would prefer (i.e., not digital) is "analog" in the sense some other authors prefer (i.e., not symbolic). Though not vice versa.

the thought that mental or brain states (or events, etc.) could have the same syntactic and semantic properties as some sentences. But strongly iconic representations represent by sharing properties with their objects, which requires them to *have* the relevant properties—in this case, spatial properties.

Surely there's no obstacle to the mind or brain exhibiting the isomorphisms required for weak and quasi-iconicity. Maybe that's iconic enough, without having to worry about strong iconicity? Can't we just say that some mental representations are at least quasi-iconic, and leave it there? I want to defend a fairly articulated view about the relation between iconicity and perception: although I think it is implausible that all perception is iconic in even the weakest sense (quasi-iconicity), I think it is nevertheless plausible that some perceptual representations may be iconic in the strongest sense. It is thus worthwhile to explore the possibility of strongly iconic mental (perceptual) representation, rather than just dropping the issue and settling for quasi-iconicity.

There's nothing problematic about the idea that mental representations might have temporal properties, which it might share with representanda and thus represent temporal properties in the environment in a strongly iconic way. But I think an even stronger, less obvious claim is defensible; I think the idea of strongly iconic *spatial* representations in the brain is not hopeless. Making sense of spatial vehicles in the brain is a difficult project that I can only touch on here, but it's a worthwhile project, as I hope to show below.

This difficult project, however, should be distinguished from a number of easy solutions that present themselves, all of which, I think, are dead ends.

4.1 Some dead ends

We want to try to make sense of the idea that mental representations could literally have spatial properties. One obvious possibility is to invoke sense-data. Sense-data are hypothesized mental items that literally have the very properties that external objects perceptually appear to have. For the table to look white, or rectangular, is for it to present a sense-datum that really is white, or rectangular. I won't say much about sense-data here, except to say that anyone who believes in them has a straightforward route to strong iconicity.

Another easy way out of the problem would be to simply appeal to phenomenal spatial properties. Suppose there's no special problem for the idea of phenomenal rectangularity and the like. Then maybe for any spatial property *F*, a mental representation "with" *F* (as opposed to merely *of F*) could just be a state that's phenomenally *F*? One problem with this move is that the question of whether (any) perceptual representations are iconic was supposed to be an empirical, scientific question, not one for introspection. The question wasn't supposed to be about how things seem to us, but about why they seem that way. Furthermore, a good many of the mental representations thought to perhaps be iconic are not (phenomenally) conscious, so phenomenal spatiality wouldn't help us with these questions anyway.

There is a lot of talk among neuroscientists about "cognitive maps" (Tolman, 1948; O'keefe & Nadel, 1978; Derdikman & Moser, 2011), especially in the hippocampus and related structures. What's meant by virtually all such talk, however, is that certain brain states encode navigation-enabling information about what is where. These internal representations have the *content* of maps and are "map-like" in the important sense that they carry information about metric relations among items in the world, and not just which direction home is. But this doesn't speak to our question here, which is about the *format* of these representations. Just as a subway rider could use a verbal list of the stations in order, rather than an actual *map* with this information, the mere fact that, say, place cells in the hippocampus carry information about conjunctions of features found in particular environmental locations, doesn't begin to argue that this information is encoded in a spatial, map-like format. But the format matters, as the subway map example illustrates: the very same information (content) is much easier to use when encoded in a map format, rather than a verbal one.

One last easy way out, which seems to have tempted nearly everyone who's written on the subject, is to note that the brain employs a large number of topographic maps. There are retinotopic maps throughout the early visual cortex, where adjacent brain regions code for adjacent retinal areas. There are tonotopic maps in the cochlea and auditory cortical regions, where high pitches are coded at one end and low pitches at the other, with in between pitches in between. And so on. These are highly significant phenomena, of course, but they don't have anything to do with the iconicity of mental representations. Of course the brain has spatial properties and some of these might be shared with distal objects; the question is whether *mental representations* might have spatial properties, and topographic maps do nothing to answer this question.

Take retinotopic maps. In a seminal study (Tootell et al., 1988), a monkey was given a radioactive dye while looking at a bullseye pattern; afterwards, the monkey's primary visual cortex was placed on a radiation-sensitive photographic film, where a (distorted) bullseye pattern was clearly visible. There were literally lines, curves, and wedges in the monkey's brain. (Or so let us concede.) Yet, there's no reason to think that at this very early stage of processing—V1—anything is being represented *as wedge-shaped*. In fact, there are reasons to think not, i.e., to think that lines, curves, wedges and the like are represented *as* lines, curves, etc. in different brain regions and later. The topographic map is often said to be a "representation" of the activity of retinal photoreceptors. Yet there's no reason to think that the brain *ever* forms a single unified mental representation of the retina. It's not needed, so long as a large number of highly coordinated representations of very local states of retinal activity exist. The contents at this level aren't things like "curve," "square," "wedge," but rather "edge at retinal location *x*," etc. Their topographic organization presumably facilitates the coordination of these local representations (as well as contrast enhancement by way of lateral inhibition, etc.), but the representations are still local.

Second, it is well known that, due to cortical magnification, the topographic maps are distorted: there are more cortical neurons with foveal receptive fields than with peripheral receptive fields. Consequently, a square presented to one side of the fixation point will produce a trapezoid in the cortex, the side nearer to the fixation point receiving a larger share of cortex. But if strongly iconic representations represent by having the very properties they attribute, this raises a puzzle: why shouldn't this represent the object as being trapezoidal, or projecting trapezoidally? There's nothing physically rectilinear in the cortex, so no strongly iconic representation as of rectilinearity.

The solution, of course, is what we knew anyway, that it's functional spatiality, rather than physical spatiality that's doing the representational work. If we could rearrange and jumble these cortical neurons while keeping their connectivity and timing the same, the topographic organization would disappear, but all the semantic and syntactic properties of the mental representations would stay the same, including the iconicity or not of these representations. Topographical realization is simply irrelevant to the question; we have to seek iconicity at the representational level, not the implementational level. All this means we can't avoid the task of coming to a better understanding of what functional spatiality—which I equate with spatiality at the level of the representation, rather than the realization—might mean.

4.2 Functional implementation

Let a *representational scheme* consist of (a) a set of representations and (b) a semantics for these representations. In compositional schemes, (a) is specified by listing primitive representational elements and providing a set of combination rules for generating well-formed representations out of these, and (b) is given by specifying contents for the primitives along with semantic composition rules to generate contents for the complex representations. Schemes will be individuated rather finely. Two schemes could differ even while having the very same "syntactic" elements (the same primitives and same possible combinations among them), if they had different semantics.

Much of cognitive psychology is concerned with discovering what kind of representational scheme a given system is using/implementing. This research, I think, is guided by the following considerations. To implement a representational scheme of a given type (e.g., Venn diagram, subway map, Roman numerals, etc.) the implementation needs to match the scheme in at least the following, functional, ways:

(i) It needs to have the same *expressive power* as the scheme. Standard Venn diagrams, for example, can express relations among three predicates, but no more; any system that can express more is not implementing a (standard) Venn diagram scheme. If pictures do have distinctive skeletal contents, then a system can't be implementing a pictorial scheme if its states don't have those contents.

(ii) It needs to exhibit the same *systematicities*. As Fodor and Pylyshyn (1988) famously noted, cognitive capacities tend to come in clusters. Among other things, this means that as certain representational capacities are gained or lost, they do so in groups. Gaining a new primitive automatically brings with it a host of new complexes; losing a combination rule means losing all the complex representations that rule made possible, etc. As Cummins (1996b) points out, however, it's not that symbolic systems are systematic and iconic systems aren't; it's that different kinds of systems exhibit different *kinds* of systematicities.

(iii) The "analyticities" of the scheme are preserved. Some information is encoded directly in the representations, while other information has to be calculated or stored separately as background knowledge, all in a way that varies from scheme to scheme (Lyons, 2005; cf Shin, 1994; Camp, 2018). A Venn diagram can't represent "no As are Bs" without at the same time representing "no Bs are As," although natural languages can. A map that shows *x* to the east of *y* and *y* to the east of *z* shows *x* to the east of *z* (and shows *z* to the west of *x*, etc.), although the latter information isn't strictly contained in the verbal description. Some facts (e.g., if no As are Bs, then no Bs are As; these two shades of blue are more similar to each other than the second is to this third, etc.) are encoded in a way that's intrinsic to some representational schemes.[15]

(iv) It makes the same *algorithms available*. It's well known that Arabic numerals are more conducive to long multiplication than Roman numerals, because there's a simple algorithm that can be used with Arabic but not with Roman numerals. Similarly, map-like representations of the environment would be amenable to different algorithms than sentence-like representations, or the kinds of vectors involved in dead reckoning. Pictorially encoded mental images would be susceptible to rotation and translation algorithms, which symbolically encodings of images would not; conversely, the latter would allow direct comparison of structural descriptions while the former wouldn't (Cooper & Shepard, 1975; Kosslyn et al., 1978; Shepard & Metzler, 1971).

(v) It will be susceptible to the same *errors*. The scheme in conjunction with the algorithms determine what kinds of mistakes will be rare or common or perhaps sometimes even inevitable. Noise-related errors connected with co-similar representations will tend to be approximately correct, while noise-related errors connected with symbolic representations are predicted to be wildly false as often as nearly correct (Beck, 2019).

In short, the representational format will determine what information can be represented and what computations will be easy, difficult, fast, slow, error-prone, impossible, or unnecessary.

Some iconic representations are essentially spatial, which means that the kinds of transformations available for processing them would be themselves mainly spatial: beyond the obvious introduction and elimination of elements, we'd have rotating, stretching, compressing, scanning, transposing, and similar operations. Transposition in iconic representations has an interesting and distinctive feature (Camp, 2007, 2018): when you move an element in, say, a map, you automatically update all the information about the spatial relations that item bears to the rest of the array. In a

[15] This might be connected to the above idea of representational entanglement: certain kinds of scheme can't represent something as shaped without representing it as colored. Thus, it might be not that representational entanglement per se is indicative of (strong) iconicity, but that iconicity implies certain kinds of representational entanglement, while other schemes imply other kinds.

language-like scheme, "*x* has moved four blocks south" would tell you with minimal computation where *x* is now, but to know *x*'s spatial relations to most of the other landmarks would require a lot more computation, all of which would take time (increasingly so the more other elements there are) and introduce new possibilities of error.

Functional sameness along the lines of (i) through (v) is at least partially constitutive of implementing a given type of representational scheme. This might be an incomplete list. A great deal more work in this area is needed. But finding that these conditions were satisfied by some computational black box would go a long way toward providing evidence that that box employs a particular type of representational scheme.[16] Although the evidential situation is more complicated than the easy ways from Sect. 5.1 suggest, there's no reason in principle why brains couldn't thereby implement literally spatial schemes and why we couldn't have evidence of such. To implement a representational scheme that's essentially spatial in nature, there's no particular spatial *physical* organization the brain must have; the spatiality occurs at the functional level, not the implementational level. Instead, the brain has to realize a system of representational states with the same expressive power as, the same systematicities as, the same analyticities as, and is subject to the same algorithms and same mistakes as, some essentially spatial representational scheme.

Note that functional implementation is first and foremost a scheme-level phenomenon. All of the previous discussion has been aimed at the implementation of such-and-such a *scheme*, with little attention to the individual representations. How does this address the question of vehicle isomorphism, over and above scheme isomorphism? The answer is that if a physical system has the relevant expressive power, systematicies, analyticities, etc., then it implements a Venn diagram scheme, or a Roman numeral scheme, or a subway map scheme, or a seating chart scheme, as the case may be. If that scheme then includes vehicle-isomorphic representations (no for Roman numerals, yes for subway maps, etc.), then we have at least weak iconicity.

The fact that it's only *functionally* spatial doesn't mean that the thing implemented is not *literally* spatial. It means, rather, that the phenomenon of interest resides at the representational level, not the implementational level. Thus, the idea of even strongly iconic mental representations in the brain seems to be back on a par with language-like mental representations: expressive power, systematicity, analyticity, algorithm, and error are as implementation-neutral as syntax and semantics. In particular, none of these impose any spatial constraints on implementation. And yet, functional spatiality is far from trivial; it is clear that a system could very well represent spatial information—i.e., have spatial content—without exhibiting the relevant systematicities, analyticities, etc.

Again, this is all very preliminary, but compare "cognitive maps" as briefly discussed above with "feature maps" as commonly discussed in the vision literature (Treisman, 1988, etc.—which in turn is to be contrasted with the way the term "feature maps" is used in the convolutional neural network literature, e.g., Kohonen,

[16] I'm gratified to see that Elisabeth Camp (2018, note 13) independently espouses something quite similar for the implementation of maps.

1982). Cognitive maps have geographic *content*—they say what's where—but that's not enough to have map-like *format*, since sentential representations can say what's where too. Feature maps, at least as I'll understand them, don't normally have geographic content—they represent structural relations among parts of objects, for example—but they do have map-like *format* in at least the following sense: it's a structural requirement imposed by the representational scheme that a feature can't just be represented as being present; it has to be *placed* somewhere, in some spatial relation to other elements of the array. Feature maps are in this way similar to seating charts (see Camp, 2007), in which symbolic representations (students' names) are tokened, but they can only be tokened *at some position* or other in the (strongly) iconic array. Seating charts are hybrid, although the present concern is with the iconic aspect. Feature maps are at least hybrid, and perhaps fully iconic, depending on the nature of the feature representations. In this way, feature maps really are maps, while cognitive maps are not. (More modestly: there's some reason to think the former are maps and no reason to think the latter are.)

Insofar as perceptual representation involves feature maps thus construed, in that respect at least is perceptual representation (strongly) iconic. But this sort of iconicity is likely considerably more widespread. The striking fact about seating charts noted above is a "placement requirement": you can't represent someone without placing them at some particular location. This requirement might govern a lot of perceptual representations that aren't normally associated with feature maps. Many of the neurons that subserve visual representation (perhaps all) have receptive fields: regions of space (often retinal, but sometimes egocentric and sometimes allocentric space) to which they are sensitive, should one of their preferred stimuli (the class of objects or properties to which they respond) occur in that region.[17] So, the firing of a particular *F*-detector in IT or V4 carries information not just about the presence of *F*, but also of the *location* of *F*. This seems sufficient for the implementation of a seating-chart-type scheme, depending on whether/how this mandatorily represented information is exploited downstream. Mandatory location representation is not all there is to being a seating chart scheme—the information also has to be susceptible to distinctive algorithms, certain relational information has to be directly encoded and thus extractable without reliance on stored, external, information, and so on. But it's suggestive enough, I think, to warrant further study.

Earlier, I dismissed topographic organization as mostly irrelevant to questions of iconicity. It is, however, relevant in the following way: topographic organization implies receptive field structure (trivially, from the definition of topographic organization), and it is *this receptive field structure* that is suggestive evidence in favor of a kind of iconicity, even if neural *adjacency* relations are entirely irrelevant to that question.

What emerges is a very speculative—and rather Kantian—suggestion about the iconicity of perceptual representation: that perceptual representation, at least for some modalities, is subject to a placement requirement, i.e., that in order to

[17] The term 'receptive field' is sometimes used to cover both what I'm here calling "receptive fields" and what I'm calling "preferred stimuli". Throughout, I am using the term in its narrower, spatial, sense.

perceptually represent something, that something needs to be represented as occupying a particular location on a perceptual map—where this map is a strongly iconic representation of some retinotopic or egocentric or allocentric, etc. space. In Kantian terms: "space is the form of outer intuition" (Kant, 1781/1998). As for the items that get placed on that map—the "matter" of outer intuition—these might also be iconic, to some degree or other, or they might be fully symbolic, as in a seating chart.

I offer this Kantian suggestion in a conjectural spirit. I don't know whether it's true. Likely it's only a subset of perception that has space as its "form". But the conjecture is worthy of investigation.

5 How iconic is perception?

There are several different ways to read the claim that perception is iconic. To get it out of the way, I don't think proponents of this view could mean that *only* perception is iconic; if there are iconic representations in perception, there are almost certainly iconic representations in cognition too. Analog magnitude representations might figure in perceptual representations (Beck, 2019), but they are also clearly used outside of perception, as when judging whether the number 87 is greater or less than 84 (Dehaene et al., 1990). The picture-like representations of visual imagery are used, cognitively, to decide whether item A can be rotated to match item B (Shepard & Metzler, 1971). Anyone who thinks that cognitive maps are really *maps* (i.e., iconically formatted) must surely think that they're really *cognitive*, as they're used in navigation to plan routes and determine directions and distances.

I don't think there's much of a chance that *all* perceptual representation is iconic either, for reasons to be made clear below. The more interesting questions here are as follows: which perceptual representations are iconic? which aspects of those representations are iconic (taking hybridity, for example, into account)? and which grade of iconicity do they exhibit?

First, however, why do researchers think that *any* perceptual representation is iconic, in any sense of the term? Setting aside the dead end arguments of Sect. 4.1, what I take to be the standard argument goes as follows. The classic mental imagery studies show that visual *imagery* utilizes representations with strongly iconic format, because the processing algorithms involve scanning, translation, and rotation operations of the sort that wouldn't be available for other formats. Parsimony considerations and brain imaging studies (Farah, 1989; Kosslyn et al., 2001, though see also Winlove et al. 2018) suggest that visual imagery uses the very same representations as normal vision. Therefore, vision uses (strongly) iconic representations. Therefore, at least some perception uses iconic representations. It's clear that this alone gives us little reason to expect iconicity in non-visual perceptual modalities. Furthermore, the argument might establish that (some) visual representations are *partly* strongly iconic, but it doesn't rule out hybrid representations. We noted above, for instance, that weather maps are strongly iconic with respect to spatial properties but symbolic or only quasi-iconic with respect to precipitation, using colors to represent the latter (the system of colors may be scheme isomorphic with the system of pressures in a way that exhibits co-similarity, but individual colors aren't interestingly isomorphic to individual pressure or precipitation values).

In addition to this argument, we have good empirical reason to believe in feature maps (e.g., Koch & Ullman, 1987; Treisman, 1988, 1998; Chan & Hayward, 2013); Clarke (2022) argues that this evidence supports genuinely map-like structures in pre-attentive vision.

I mentioned in Sect. 4.2 some suggestive partial evidence in favor of a Kantian iconicity hypothesis. Insofar as the presence of receptive fields argues for iconicity, this argument (unlike the standard visual imagery argument, or the feature map argument) readily generalizes to some other modalities; auditory and somatosensory modalities at least are known to use receptive fields. Furthermore, should this Kantian hypothesis turn out to be true, it would offer a statement of the iconicity of perception that distinguishes perception from cognition: space is the "form" of perception but not of cognition.[18] This should be a welcome possibility for those who think that perception is distinctively iconic (Beck, 2019; Burge, 2010, 2018; Block, 2014, forthcoming; Carey, 2009; Dretske, 1981; Raftopoulos, 2019), although they never frame their claim in such Kantian terms. Maybe this suggestion would satisfy these authors, but it isn't the thesis they've been endorsing.[19]

But even if perceptual representation involves a seating-chart-like format, the status of the filler items (corresponding to students' names) is up for grabs. Could it all be iconic? Which grades of iconicity?

Strong iconicity is only possible where properties are shared. The best candidates for shared properties are temporal and spatial properties, though there may be others. But this means that only these shareable properties—whatever they may turn out to be—could be strongly iconically represented. Similarly, weak iconicity requires representanda that are analogous, in the sense articulated above, to spatial, temporal, etc., properties. Co-similarity (which includes strong and weak iconicity) is only non-trivially possible for representandum domains that are dimensionable, as noted in Sect. 3: the determinate values can be laid out in a nontrivial and nonarbitrary similarity space. Binary domains, where a single feature is marked as either present or absent, are dimensionable only in a trivial and uninteresting sense.

Even if perceptual representation consists of placing features in a Kantian map, consider the information that would have to go into the map. The kinds of properties presumably represented in perception fall into several categories.

[18] Or at least, not of cognition in general. Again, this kind of Kantian iconicity could at most be necessary for perceptual representation, not sufficient, since the representations used in mental imagery, spatial navigation, etc., might well include this Kantian formal constraint and yet remain on the cognitive side of the border. Thanks to an anonymous referee for pressing this.

[19] For what it's worth, I think that should this speculative Kantian proposal turn out to be true, it would offer *a* cognition/perception distinction, not necessarily *the* cognition/perception distinction. That is, I don't think it would obviously entail that this is the only theoretically interesting distinction between perception and cognition. This is because category membership and other late, or high-level information does not seem to be subject to a placement constraint and yet is very plausibly perceptual in some important sense nevertheless.

(i) Shape (both perspectival and objective), size, motion, etc. Such spatiotemporal properties could easily be represented strongly iconically. Of course this is not an argument for thinking that they actually are, and further empirical investigation is needed here.
(ii) Color, pitch, timbre, etc. Though it's possible to represent the *space* of colors, pitches, etc. weakly iconically (as the italicized term suggests), perceptual representations don't obviously represent color spaces; they represent things as having specific colors. Such representations are therefore more plausibly thought of as involving quasi-iconicity, rather than vehicle isomorphism. (It's not that this rate of neuron firing resembles red; it's that this rate of firing figures into the system of neural firings in a way that resembles how red figures into the system of colors.) There is pretty good evidence (often neurological) for co-similarity here, but none to my knowledge for any stronger grade of iconicity.
(iii) Depth, numerosity, etc. There's very good evidence for analog magnitude representations of numerosity (see Beck, 2015 for an overview). These are quasi-iconic representations but unlike color and pitch, for example, don't seem in-principle amenable to any stronger iconicity.
(iv) Low-level features like edges or blobs, solidity, object boundaries, (Spelke-) objecthood, Gestalt features, like grouping, figure/ground, etc. These seem not to be amenable even to quasi-iconic representation, at least not in any non-trivial sense. These features are either present or absent; they don't come in degrees and are not dimensionable in any nontrivial sense.
(v) FINST representations (Pylyshyn, 2003). The aforementioned representations were all in some sense predicative, while FINST representations are supposed to serve a purely referential function. Like these predicative representations, however, FINST representations are not non-trivially dimensionable and thus not quasi-iconically represented.
(vi) Category membership. If perception involves the representation of items as falling under commonsense categories (e.g., *dog*, *chair*), this will be another instance of perceptual representation that doesn't involve co-similarity.[20] Obviously, it is highly controversial whether category membership is represented in perception or post-perceptually. Note, however, that anyone who denies that such representation occurs in perception, had better not do so on the grounds that perception is (fully) iconic, since it's not at all controversial that the kinds of representations mentioned in (ii)–(v) are genuinely perceptual. And we've seen that some of these (ii)–(iii) are at most only quasi-iconic and others (iv)–(v) not even that.

The best hope for the claim that perception is distinctively iconic might lie in the Kantian proposal sketched above. It's in need of a great deal of empirical, as well as more theoretical, work; but it potentially might cover all of perception (not just vision, for example), in a way that partially distinguishes perception from cognition, and—interestingly—it attributes to perceptual representation a very high

[20] Unless, maybe, concepts are just regions in multidimensional similarity space (Churchland, 1989). I take this to be a non-standard view of concepts.

grade of iconicity, not mere quasi-iconicity. Even if this bold proposal turns out to be true, however, it wouldn't show that perceptual representation is iconic through and through. It would be at most hybrid, with a strongly iconic form, into which are placed fillers using a combination of strongly iconic, weakly iconic, quasi-iconic, and symbolic elements. Even this hybrid proposal isn't at all obvious. I hope that by clarifying the nature of iconic representation in the abstract and the nature of implementation of iconic representation, this paper gets us a little closer to finding out whether it is true.

Acknowledgements Earlier versions of this paper were presented at Cambridge University, Radboud University, University of Glasgow, and workshops at the University of Barcelona and University of California San Diego. Thanks to the audiences there, especially Jonathan Cohen, Frances Egan, Joseph Levine, Louise Antony, Janet Levin, Harmen Ghijsen, Keith Wilson, Fiona Macpherson, Ori Beck, Maarten Steenhagen, Jake Quilty-Dunn, Christopher Gauker, and Indrek Reiland. Thanks to Sam Clarke for comments on a written draft and special thanks to Jacob Beck for helpful discussion.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Beck, J. (2015). Analogue magnitude representations: A philosophical introduction. *British Journal for the Philosophy of Science, 66*(4), 829–855.

Beck, J. (2019). Perception is analog: The argument from Weber's Law. *Journal of Philosophy, 116*, 319–349.

Berkeley, G. (1710/1975). *A treatise concerning the principles of human knowledge*. In M.R. Ayers, ed. *George Berkeley: Philosophical works, including the works on vision*. Dent.

Block, N. (1995). The mind as the software of the brain. In E.E. Smith and D.N. Osherson, eds., *An invitation to cognitive science, Vol. 3: Thinking*, 2nd ed. MIT Press. 377–425.

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

Block, N. (forthcoming) *The border between seeing and thinking*. Oxford University Press.

Burge, T. (2010). *Origins of objectivity*. Oxford University Press.

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

Burge, T. (2018). Iconic representation: Maps, pictures, and perception. In W. Shyam & F. A. Dorio (Eds.), *The Map and the Territory* (pp. 79–100). Springer.

Camp, E. (2007). Thinking with Maps. *Philosophical Perspectives*, *21*(1), 145–182.

Camp, E. (2018). Why maps are not propositional. In A. Grzankowski & M. Montague (Eds.), *Non-propositional intentionality.* Oxford University Press.

Carey, S. (2009). *The Origin of Concepts*. Oxford University Press.

Casati, R., & Varzi, A. C. (1999). *Parts and places: The structures of spatial representation*. MIT Press.

Chan, L. K., & Hayward, W. G. (2013). Visual search. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*(4), 415–429.

Clarke, S. (2022). *Mapping the visual icon. Philosophical Quarterly, 72*(3), 552–577.

Cooper, L. A., & Shepard, R. N. (1975). Mental transformation in the identification of left and right hands. *Journal of Experimental Psychology: Human Perception and Performance, 1*(1), 48.
Cummins, R. (1989). *Meaning and mental representation*. MIT Press.
Cummins, R. (1992). Conceptual role semantics and the explanatory role of content. *Philosophical Studies, 65*(1–2), 103–127.
Cummins, R. (1996a). *Representations, targets, and attitudes*. MIT Press.
Cummins, R. (1996b). Systematicity. *The Journal of Philosophy, 93*(12), 591–614.
Dehaene, S. (2011). *The number sense* (2nd ed.). Oxford University Press.
Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital: Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 626–641.
Derdikman, D., & Moser, E. I. (2011). A manifold of spatial maps in the brain. In S. Dehaene & E. Brannon (Eds.), *Space, time and number in the brain* (pp. 41–57). Elsevier.
Dretske, F. (1981). *Knowledge and the flow of information*. MIT Press.
Farah, M. J. (1989). The neural basis of mental imagery. *Trends in Neurosciences, 12*(10), 395–399.
Fodor, J. (2007). The revenge of the given. In B. McLaughlin & J. Cohen (Eds.), *Contemporary debates in philosophy of mind* (pp. 105–116). Blackwell.
Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*(1–2), 3–71.
Haugeland, J. (1991). Representational genera. In W. Ramsey, et al. (Eds.), *Philosophy and connectionist theory* (pp. 1–90). Lawrence Erlbaum.
Hill, C. S. (2016). Perceptual relativity. *Philosophical Topics, 44*, 179–200.
Hill, C. S. (2021). Perceptual existentialism sustained. *Erkenntnis, 86*, 1391–1410.
Hume, D. (1739/1978). *A treatise of human nature*. In L. A. Selby-Bigge & P. H. Nidditch (Eds.). Oxford University Press.
Kant, I. (1781/1998). *Critique of pure reason*. In P. Guyer & A. Wood (Trans. and Eds.). Cambridge University Press.
Koch, C., & Ullman, S. (1987). Shifts in selective visual attention: Towards the underlying neural circuitry. In L. Vaina (Ed.), *Matters of intelligence* (pp. 115–141). Springer.
Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics, 43*(1), 59–69.
Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance, 4*(1), 47.
Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience, 2*(9), 635.
Kulvicki, J. (2015). Maps, pictures, and predication. *Ergo, an Open Access Journal of Philosophy*, *2*.
Lyons, J. C. (2005). Representational analyticity. *Mind & Language, 20*(4), 392–422.
Maley, C. J. (2011). Analog and digital, continuous and discrete. *Philosophical Studies, 155*(1), 117–131.
Millikan, R. G. (2004). *Varieties of meaning: The 2002 Jean Nicod Lectures*, T. Roeper, ed. MIT Press.
Millikan, R. G. (2017). *Beyond concepts*. Oxford University Press.
O'keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Clarendon Press.
Pylyshyn, Z. W. (2003). *Seeing and visualizing: It's not what you think*. MIT Press.
Quilty-Dunn, J. (2020). Perceptual pluralism. *Noûs, 54*(4), 807–838.
Raftopoulos, A. (2019). *Cognitive penetrability and the epistemic role of perception*. Palgrave Macmillan.
Rescorla, M. (2009). Predication and cartographic representation. *Synthese, 169*(1), 175–200.
Schellenberg, S. (2008). The situation-dependency of perception. *The Journal of Philosophy, 105*(2), 55–84.
Shepard, R. (1968). Review of Ulrich Niesser, *Cognitive Psychology*. *The American Journal of Psychology, 81*(2), 285–289.
Shepard, R. N. (1975). Form, formation, and transformation of internal representations. In R. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 87–122). Lawrence Erlbaum.
Shepard, R. N. (1978). The Mental Image. *American Psychologist*, *33*(2), 125.
Shepard, R. N., & Chipman, S. (1970). Second-order isomorphism of internal representations: Shapes of states. *Cognitive Psychology, 1*(1), 1–17.
Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(3972), 701–703.
Shin, S. J. (1994). *The logical status of diagrams*. Cambridge University Press.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 2684*, 677–680.
Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55*(4), 189.
Tootell, R. B., Switkes, E., Silverman, M. S., & Hamilton, S. L. (1988). Functional anatomy of macaque striate cortex II. Retinotopic Organization. *Journal of Neuroscience*, *8*(5), 1531–1568.
Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology, 40*(2), 201–237.
Treisman, A. (1998). The perception of features and objects. *Visual Attention, 8*, 26–54.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.