THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Digital images

Citation for published version:
Isaac, A 2017, 'Digital images: Content and compositionality', *Journal of the American Philosophical Association*, vol. 3, no. 1, 106-126. https://doi.org/10.1017/apa.2017.13

Digital Object Identifier (DOI):
10.1017/apa.2017.13

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Journal of the American Philosophical Association

Publisher Rights Statement:
© American Philosophical Association 2017

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 11. May. 2020

Digital Images: Content and Compositionality

Alistair M. C. Isaac
University of Edinburgh
a.m.c.isaac@ed.ac.uk

§1. *Introduction*

Is there a distinctively *imagistic* form of *content*, with semantic characteristics different from those typically associated with linguistic content? The view that there is has played an important role in debates about the nature of perceptual content, mental representation, aesthetics, and scientific realism. Yet just what is it that is distinctive about imagistic content? One popular line of inquiry into these questions has focussed on the apparently *analog*, *continuously variable*, or *arbitrarily fine-grained* nature of imagistic representation, contrasting it with the discreteness of typical accounts of language-like representation. Different strands of this approach locate this fineness-of-grain in the vehicles of imagistic content, the content itself, or both.[1] A second influential strategy takes pictorial representations to be paradigmatic images and concludes that imagistic content is *essentially perspectival*.[2]

The aim of this paper is to motivate an alternative account of the distinctive features of imagistic content by taking as primary data the synonymy relations that (we take to) obtain between *digital* imagistic representations. The most familiar examples of digital images are those displayed and manipulated on our computer screens. These images are (i) composed of a finite set of discrete elements (the "pixels"); and (ii) each of these elements may take on one of a finite set of discrete values (the "colors"). Call any representation with these features and imagistic content a *digital image*. Intuitively, two digital images are "synonymous" if we ascribe the same content to them.

[1] Goodman (1976) gives the canonical account of analog symbol systems as syntactically and semantically dense, i.e. between any two symbols exists a distinct third symbol, and between any two represented classes is a distinct additional class. Peacocke (1986) and Haugeland (1991) identify analog content as content specified within dimensions of variance. The literature on non-conceptual content has emphasized the apparent fineness-of-grain of perceptual / imagistic experience (Gunther 2003).

[2] Since pictorial representation typically involves 2-dimensional projections of 3-dimensional scenes, and these are only defined with reference to a perspective point, some have argued it is essentially perspectival (e.g. Hopkins 1998: 36), or, abstracting from the restriction to a single perspective point, at least *essentially aspectual* (Lopes 1996: 124), and these conclusions have been extended to imagistic representation in general (van Fraassen 2008: Part 1).

I demonstrate that synonymous digital images are related by a homomorphism that does not preserve exact color values, and only preserves color boundaries at a granularity coarser than that of the pixels themselves; this motivates the central claim of the paper, that images exhibit a distinctive compositional structure. I argue that parts of an image always contribute in a systematic way to its overall content, i.e. images are always non-trivially compositional. However, I also argue that there is a fineness-of-grain at which apparently content-bearing parts of the image do not uniquely contribute to overall content, i.e. images are never "inverse compositional." This conclusion differs markedly from the (typical) analog account of image content, on which details of an image bear content at an arbitrarily fine grain. It also indicates a deep structural difference between language-like representations and images: languages may or may not be compositional, and they may or may not be inverse compositional; in contrast, images are always compositional, but never inverse compositional.

My method for defending this claim differs substantively from those of previous investigations in that it does not presuppose a particular theory of imagistic content. Since compositionality is a relation obtaining between meaningful parts of a representation and its meaning as a whole, rigorous analysis of compositional structure does not require a full syntax or semantics, but merely a system of synonymy relations.[3] The intuition is: if we can find a distinctive characteristic of images purely at this structural level, it will be completely general, and thus independent from any contingent features of *particular* systems of imagistic representation (though such systems may themselves exhibit analog or perspectival characteristics).

There are compelling methodological reasons to focus attention on digital images. Perhaps surprisingly, it turns out that most images are digital in the above sense—not just images on our computer screens, but traditional photographs as well, and arguably even paintings. "Analog" photography captures images in a discrete array of light-sensitive silver halide crystals, either turned opaque or washed away during the process of development, and thereby acting as "pixels." Likewise, paintings and other real world images may be pixelated in the limit by their molecular structure. Thus, the intuitively "analog" may in fact be "digital" in the technical sense! If most (arguably all) actual images are digital, then it would seem mistaken to take analog images as paradigmatic.

[3] For discussion, see Hodges (1998: 11); Pagin and Westerståhl (2010a: 253).

These claims are contentious, yet there are more mundane reasons to take digital images as the starting point for a theory of imagistic content. We engage in a rich practice of content-preserving, mechanical manipulations of digital images: we resize images on our screens, print them to paper, and change their file type routinely, all in the expectation that content is preserved. Every time we print or resize an image, we implicitly accept that the input and output of the algorithm our computer employs are synonymous, that they are in some suitable sense "the same" image. My strategy is to begin from this *de facto* trust of algorithmic image manipulation, treat the outputs of such manipulations as strictly synonymous, and extract the distinctive features of imagistic representation implicit in these synonymy relations. Thus, my target is not some Platonic concept of content, but rather the understanding of content implicit in our everyday interactions with actual images.

I begin with some methodological preliminaries that qualify and motivate the notions of content and synonymy employed in the sequel. Next, I survey the basic features of synonymous images presupposed by algorithmic image manipulation. This motivates a more formal analysis of imagistic content as structure preserved across probabilistic homomorphisms, which in turn leads to the central argument of the paper, that imagistic representations are distinctive in that they are always compositional, but never inverse compositional. I conclude by revisiting the claim that imagistic content is essentially perspectival, acknowledging a qualified role for perspective in delimiting, but not determining imagistic content.

§2. *Synonymy, transformations, and skeletal content*

How can study of synonymy between image vehicles inform a theory of image content?

In the abstract, synonymy relations partition a set of representations into equivalence classes, such that all members of any given class bear the same content. If we know the equivalence classes, we can ask how representational vehicles in the same class are related to each other, and thereby hope to learn something about their content-bearing features. One influential strategy has been to look at the *transformations* that take one member of a class into another member of the same—whatever stays *invariant* across all such transformations is a clue to the content and how it is borne by its vehicle.

The idea that content identity, or at least meaningfulness, should be understood as invariance across transformations can be traced back to Felix Klein's Erlanger Program. Klein unified disparate geometric systems by distinguishing and systematizing the transformations across which their objects remained invariant—for instance, in Euclidean geometry (unlike, say, projective geometry), the values of interior angles are part of the identity conditions for any triangle, and these only remain invariant across rigid rotations and translations of the plane. This same basic idea has reappeared in the perceptual constancy literature (Cassirer 1944); the transformational grammars of Zellig Harris and Noam Chomsky; the formal theory of measurement (Luce, at al. 1990: Ch. 22); and Nozick's final proposal to unify philosophy (2001). Patrick Suppes explicitly argued that propositional content should be investigated by Erlanger-style analysis of invariance across transformation in his 1973 presidential address to the American Philosophical Association.

The strategy advocated by Cassirer, Nozick, and Suppes is especially apt for the study of the imagistic content of digital images for two reasons. First, due to the practical demand for reproduction of imagistic content on different devices, at different sizes, and in different formats, content-preserving transformations over digital images have been rigorously honed and optimized. Second, these transformations provide *prima facie* reason to doubt the equivalence of linguistic and imagistic content. For digital images are stored as sentences in a language, namely the language governing the syntax of their file formats; in the simplest case, pixels may be named, a predicate assigned to each color, and the image specified by the conjunction of the formulae that attribute the correct color to each pixel. Yet the everyday manipulations of a digital image that we take to preserve its imagistic content—resizing it on a screen, printing it to paper—do not preserve the linguistic content of these sentences. Consequently, the semantics of digital images cannot be derived directly from the semantics of the programming languages in which they participate.

The synonymy relation over images implicit in our practices of mechanical resizing and reproduction is synonymy of a minimal kind of content, content considered independent of the use to which the image is put. Haugeland (1991) calls this "skeletal" or "bare bones" content: "the strict content of a representation, that not augmented or mediated by any other." Haugeland contrasts this with the "full-blooded contents of

everyday representations[, which] are shaped and supported by their skeletal contents, but . . . fleshed out and enlivened through other influences" (185). Haugeland's distinction is suggestive, but hard to pin down.[4] For our purposes, it will be helpful to take skeletal content as that which determines what an image is *apt* to depict, while fleshing out through background knowledge, context, or other nearby representations is necessary to specify what the image depicts *simpliciter*. For instance, a particular pattern of colored pixels may convey skeletal content apt to depict a group of 15th century soldiers. However, this skeletal content must be supplemented with conceptual knowledge and attention to the context of use if it is to be fleshed out. Depending on the context and surrounding representations, such as a caption or nearby text, the same skeletal content may be fleshed out in a variety of ways, for instance as "generic scene from the Hundred Years' War," "Henry V addressing the troops at Agincourt," or even "Kenneth Branagh as Henry V." Nevertheless, the skeletal content remains unchanged as the image is duplicated, resized, and captioned; it is this skeletal content that ensures the image is apt to depict either a generic battle scene or a Shakespearean performance.

So, the remainder of this paper concerns imagistic skeletal content: that aspect of an image, preserved by our practices of mechanical reproduction and resizing, that determines what it is *apt* to depict. In the interests of minimizing baroque circumlocutions, I often suppress the expression "apt to," with the understanding that "image of *x*" is always shorthand for "image apt to depict *x*" unless otherwise specified.

By restricting attention to skeletal content, we avoid an influential argument that imagistic content is inherently perspectival because it depends on *use*. Van Fraassen (2008) gives the example of a 1961 Doisneau photo of the Eiffel Tower: if it is used as a postcard, sent through the mail with writing on the back, it represents the Eiffel Tower, but reprinted in a book on photography, with suitable caption, it represents Doisneau's famous photo *Au Pont de l'Alma* (20–1). Van Fraassen concludes: "*what it is an image of* depends on the use, on *what I use it to represent*" (21). But the postcard and the image in

[4] Kulvicki (2006) discusses Haugeland's suggestive notion in depth. His "bare bones content" (BBC) shares many features with the interpretation offered here; in particular, he takes BBC to "articulate what pictures can be about" and holds that the BBC of a picture includes "a plane that is colored and shaped exactly like the [picture] itself" (59, c.f. Ch. 6). The primary contrast between the present approach and Kulvicki's is methodological: he defines BBC in terms of the representational system in which the picture participates (for instance, "linear perspective"); whereas here, skeletal content is derived from invariants across content-preserving transformations, and thus does not need to be indexed antecedently to a particular representational scheme.

the photography book both comprise the same pattern of black and white parts; one may be transformed into the other through mechanical reproduction, and *something* about this pattern, some context-independent feature, renders it apt for representing a holiday view of Paris or Doisneau's photo—that something is the skeletal content.

§3. *How synonymous images differ*

Everyday interactions with computers and printers illustrate how the imagistic content of a digital image is preserved across *resampling*. Resampling is the process by which the same imagistic content is presented across different numbers of pixels or colors. We resample to change image format, resolution, or "quality," yet take these changes to preserve content, implying that there is some granularity at which the details of a digital image are not content-bearing.

Content-preserving transformations of images between format or display type generally involve changes to color space, and thus to exact color values, without thereby compromising skeletal content. For instance, due to the physical differences between the additive color mixing of lights and the subtractive color mixing of pigments, printing a digital image to paper necessitates a change from an RGB (red-green-blue) color space to a CMYK (cyan-magenta-yellow-black) one. Likewise, the same image file may realize different images if presented on different monitors, as monitors may differ in their RGB hardware, or the calibration that determines how image information is realized by that hardware. Differently calibrated monitors may present the same image content as more or less saturated, lightened, or even shifted slightly across the color spectrum, without thereby changing what the image is apt to depict.

Changes in format also typically involve changes in resolution, the number of pixels that determine an image. Upsampling enlarges an image to fill more pixels, while downsampling shrinks it into fewer pixels. Both transformations employ sophisticated algorithms to ensure image content remains invariant. In the case of upsampling, pixel color values must be interpolated carefully to prevent "pixelation," the introduction of steps into smooth curves. In downsampling, color values are determined for the smaller array of pixels through statistical analysis of corresponding larger patches in the source image. This again is a complex procedure, which must be tuned to minimize the danger

of "aliasing," the introduction of artifactual patterns into the image—for instance, the moiré pattern that distorts striped neckties on TV.

Crucially, downsampling may *improve* image quality, and thus fineness-of-grain *simpliciter* is no decisive measure of the accuracy with which imagistic content is conveyed. For instance, if there is noise in an image, i.e. stray pixels with random values that do not contribute to content presentation, this noise will be washed out by averaging over nearby pixel values during downsampling. More generally, various algorithmic procedures for sharpening or blurring an image, i.e. systematically sampling regions and slightly reapportioning color values within them, may be taken to improve quality without altering content. Unlike changes to color or resolution induced by a change in display format, these latter changes are typically instigated intentionally by an expert user—in the extreme case, applied only to select areas of the image, as in photo-retouching to "clean up" noisy or blemished photos. Nevertheless, we typically take such photo-retouching to clean, improve, or reveal already present content, rather than impose or create new content, an attitude dramatically illustrated by the popular trope of police procedural fiction whereby algorithmic "sharpening" of blurry security camera footage "reveals" a perpetrator's face or license plate (rather than "creating" or "introducing" it). Such examples illustrate how we routinely accept images that differ in detail to nevertheless bear the very same content.

Does this conclusion go too far—may such radically different images bear *precisely* the same content? The kind of image manipulation depicted in the police procedural is unrealistic: one may not in general recover fine-grained details from a coarsely pixelated source image, no matter how powerful one's image processing software. Likewise, in practice, we do not accept two images that differ radically in degree of blurriness or sharpness to bear precisely the same content. Nevertheless, this very consideration cuts both ways, indicating, in the case of digital images at least, why we *do* treat images that differ in detail as bearing precisely the same content. For consider again upsampling—the reason we have developed algorithms that interpolate pixel color values when increasing pixel number is to ensure that curves or gradations of color are presented with equal smoothness across any pixel number. If a higher resolution image realizes precisely the same color regions down to the pixel level as a lower resolution one, then this will not be the case: smooth curves in the lower resolution image will be stepped, or pixelated, in the higher one. Thus, equal smoothness ("blurriness") over the curve as a whole demands a

change in "sharpness" near the pixel level. More generally, since the complex challenges for up- or downsampling may be resolved in multiple ways by different algorithms when converting an image from one presentation format (or file type, or resolution, or "quality") to another, and these strategies characteristically involve the statistical blurring or interpolating of color values across different arrangements of pixels, color boundaries are not deterministically preserved by these transformations; consequently, the correspondence between synonymous image parts is at best *probabilistic* or *fuzzy*.

This conclusion is inconsistent with a Goodman-style theory of imagistic content, which takes arbitrarily fine-grained details to be content bearing. A defender of this approach will insist that the slightest change in image "quality," number of pixels, saturation of colors, etc. corresponds also to a change in content. Yet this view cannot be correct in the limit. For, if we follow Goodman, then two prints from the same negative will not bear the same content, as their constituent silver particles will be arrayed in slightly different positions; two imprints of Hokusai's *Great Wave off Kanagawa* will not bear the same content, as there are slight differences in the density and boundaries of their colors; and two high-quality digitizations of the same imprint of Hokusai's *Great Wave* will not bear the same content if they are of slightly different resolutions, viewed on different monitors, or printed on different paper stock or with different inks.[5]

Goodman's view provides an important metaphysical foil against which other theories of imagistic content may position themselves. Nevertheless, it appears inadequate for explaining our *prima facie* content-preserving manipulations of images: we take resizing, printing, slight changes in color space, etc. to leave content unchanged, yet they all alter Goodman-style imagistic content. Once we embark on the project outlined in the previous section, and take the algorithmic transformations of digital images as primary

[5] As an anonymous reviewer has stressed, Goodman's canonical (1976) view may be supplemented to account for the intuition that images that differ only in small ways bear contents that also differ only in small ways—although at the cost of reintroducing Goodman's bugbear: *similarity*, a point he himself seems to recognize: "ironically . . . a ghost of likeness, as nondifferentiation, sneaks back to haunt our distinction between pictures and predicates" (Goodman and Elgin 1988: 131). However, in order to say that images that differ in small ways (e.g. two prints from the same negative) bear precisely the same content, a Goodman-style view must either take the images as participating in a syntactically disjoint symbol scheme, or as differing with respect to features that are not themselves bearers of content. The former strategy would contradict Goodman's own preferred account of digital images, on which they are images precisely in virtue of participating in a dense symbol scheme, only some members of which are pixelated (126–7). The latter strategy is pursued by Bach (1970) and has close affinities to that developed here; the difference is that Bach takes some *marks* to be representational and others not, whereas here we restrict attention to *pixels*, and take the same arrangement of pixels to exhibit representational features at a coarse granularity, yet non-representational ones at some finer grain.

data on synonymy relations, we are forced to accept a more coarse-grained account of image content. Yet the difference between image features that bear content and those that do not is a difference of scale, not of kind. The very same techniques that may be unquestionably employed to "clean up" blemished photos, improving the accuracy with which content is conveyed, become controversial when they are employed at a coarser grain, and thereby judged to alter the content of an image—as in fashion photography, where subtle changes in color boundaries may be used to alter the apparent shape of a model's body to conform to perceived norms of beauty. Such examples imply there is some granularity at which a boundary is crossed between content-bearing and non-content-bearing aspects of a digital image. It is this apparent boundary that is the key *explanandum* for an adequate theory of digital imagistic content.

[Figure 1 about here; caption:

> Three synonymous images, differing at pixel level. (a) and (b) are different resolution digitizations of the same imprint of Hokusai's "Great Wave off Kanagawa" (circa 1830), from the Metropolitan Museum of Art, cataloged JP1847. (c) is a digitization of a different imprint from the original printrun, also in the collection of the Met, catalogued JP2569.

]

§4. *Synonymy and structural similarity*

If preservation of content across digital image manipulation does not require the precise reproduction of arbitrarily fine-grained details of the image, what does it require? A rough, intuitive idea is that the spatial *structure* of an image must be preserved across resampling if its content is to be preserved. A manipulation of Hokusai's *Great Wave* that inverts the relative positions of the wave and Mount Fuji, or deletes the fishing boats, clearly changes the content of the image by failing to preserve large-scale structural elements. Detailed aspects of the image that bear content—the alternation of dark and light blues on the wave's underside, the topknots on the passengers in the fore of the forward boat—likewise comprise more fine-grained structural relations between patches of color that must be reproduced if content is to be preserved. This section aims to make these intuitions more precise, articulating a formal condition on image synonymy that

combines the demand for relatively fine structural similarity with the above insight that at some fineness-of-grain, detail no longer bears content.

Call a strict linear order over pixels an "axis." Axes arrange pixels in "space," whether space be understood as physical and extended or mathematical and abstract. While resampling doesn't preserve exact pixel values, it does preserve the spatial arrangement of patches of color, i.e. axes are preserved. Furthermore, not only pixel space, but color "space" as well exhibits determinate structure that is preserved across resampling. While uniform lightening changes absolute color values, it preserves the relations between colors. Likewise, relative color values are preserved across transformations between different color spaces. Color spaces may differ in relatively minor ways, for instance the RGB space used on Apple monitors and that used on NTSC television screens are defined by slightly different values for the hues of the component red, green, and blue dimensions; but color spaces may also exhibit large-scale structural differences, for instance RGB color space is 3-dimensional, while CMYK color space is 4-dimensional. When we recognize these spaces as digital, i.e. divided into discrete-valued increments, we have no guarantee of exact correspondence between elements of one space and elements of another, no matter how similar their gross structure—at best colors from one space may be reasonably *approximated* in another. Since color spaces are arranged such that nearby colors are close in appearance, however, we do have a guarantee that nearby colors in one space will be approximated by nearby colors in another. Thus, the content-preserving relationship between color spaces is analogous to that between spatial arrangements of pixels: "axes" of elements ordered by similarity in one space correspond to axes of elements likewise ordered in another.

These considerations motivate the following characterization of the relationship between two digital images with the same imagistic content. For any image with set of pixels X, call its color space C_X, and the function which partitions it, assigning a unique color to each pixel, c_X: $X \longrightarrow C_X$.

H If A and B are different digital representations, with the same imagistic content, then there exist homomorphisms f: $A \longrightarrow B$ and g: $C_A \longrightarrow C_B$, such that

i.) f maps the axes of A into axes of B

ii.) g maps the axes of C_A into axes of C_B

iii.) there exist non-zero bounds δ and ε, such that for any pixel a in A, $c_B(f(a))$ is less than distance δ from $g(c_A(a))$ with probability $1-\varepsilon$.[6]

Intuitively, clauses (i) and (ii) say that the pixels and colors of synonymous images may be placed in structural correspondence. Yet (iii) says that this correspondence is not exact, but rather *fuzzy*, or *probabilistic*. (iii) is needed to account for the examples discussed above, where a uniform patch of color in one image may correspond to a patch with slight variations in another, or a pixel of image "noise" may be present in one, but not the other, of a pair of synonymous images. **H** is of interest not so much for what it says, but for what it doesn't say; in particular, if we take everyday manipulations of digital images to determine synonymy, we cannot assert that content depends on precise details of the vehicle, i.e. we cannot assert that f and g are isomorphisms, nor that $\delta = \varepsilon = 0$.

Note that **H** applies also to 3-dimensional digital representations. In that case, **H** will preserve structure in a spatial volume defined by three axes rather than the pictorial two, and the "colors" may be understood as a partition over this volume into those regions that are solid and those that are not (more generally into regions of different densities, or other distinguishing properties). Consider the same design printed on two different 3-D printers with different resolutions; these two objects will be related homomorphically just as two pictures with the same content at different resolutions are. Similar considerations apply to the scanning of parts of the body or brain by means of magnetic resonance imaging (MRI) or x-ray computed tomography (CT scanning)—the output of such scans is an array of 3-D pixels, or "voxels," and content will typically be preserved across slight differences in scanning resolution or algorithmic filtering.

In general, **H** characterizes the relationship between synonymous digital images of arbitrarily high dimensionality.[7] There are good reasons to develop a theory of imagistic content adequate for representations of any dimensionality. For instance, perceptual

---

[6] I've suppressed formal details in the interests of space. More formal versions of clauses (i) and (ii) will follow the standard definition of homomorphism (i.e. for $a_1, a_2 \in A$, $a_1 < a_2$ implies $f(a_1) \prec f(a_2)$, see Isaac (2013) for further details). Clause (iii) requires a distance metric, which may be obtained by simply counting pixels or colors, since these are discrete, and a probability measure, which may be interpreted as relative frequency of color correspondence across all paired pixels in a set of synonymous images.

[7] But also arbitrarily *low* dimensionality! For instance, digital audio recordings satisfy the definition of digital image and when synonymous (e.g. of the same performance) are related by **H**. They convey "imagistic" content of a single dimension, with the (say) 44,100 samples per second understood as the "pixels" and the discrete-valued amplitudes assigned to each sample the "colors."

content is often interpreted as imagistic, yet our perceptual experience is of a world extended in three dimensions. The color spaces described above are in fact motivated by research on phenomenal color space, which, while frequently depicted in three dimensions, is actually at least 4-dimensional, probably more (Niederée 2010). Russian director Andrei Tarkovsky characterized motion pictures as something like 4-dimensional images when he called filmmaking "sculpting in time." On a grander scale, Churchland (2012) argues mental representations are many thousand-dimensional "maps"—plausibly understood as bearers of imagistic content. Each of these examples exhibits synonymy phenomena analogous to those uncovered for 2-dimensional digital images. For instance, a movie displayed on a cinema screen at 24 fps, each frame lasting 1/48 of a second, is related by **H** to the "same" movie displayed on a 1080p HD monitor—we interpret them as bearing the same content even though the screen is dark half the time in the former case, but not the latter, and the exact pixelation changes with each frame in the former case and never in the latter. Churchland's view faces the difficult task of explaining how the high-dimensional mental maps defined by neural activation may bear the same (or even similar) content across different brains, despite involving different numbers of neurons, and thus dimensions (110*ff*)—observing they are related by **H** suggests a start at answering this question.

§5. *The basics of (inverse) compositionality*

Just like images, synonymous linguistic representations are homomorphic. However, the structure preserved by the relevant homomorphism differs: intuitively, the (syntactic) surface structure of synonymous linguistic representations may be quite different, yet the semantic roles filled by morphemes still correspond precisely. Conversely, **H** dictates that synonymous images will share gross (spatial) surface structure, yet the correspondence between their color boundaries may be fuzzy rather than precise. I argue that these disanalogies indicate a general structural difference between imagistic and linguistic representations, with respect to the contribution the parts of a vehicle make to overall meaning. This section introduces the relevant concepts for articulating this difference: compositionality and inverse compositionality.

Compositionality is a property that characterizes the contribution of the parts of a representation to its meaning:

> A representation is *compositional* if and only if its overall content (meaning) is a function of the contents of its parts and the structural relations between them.[8]

It is easy to construct language-like representations that are trivially compositional. Consider for instance a system of maximally precise symbols, with no internal syntactic structure; such a system would convey linguistic content, and do so in a characteristically linguistic way. For instance, a propositional "language" with atomic proposition letters, but no logical connectives, would convey content of essentially the same form as that of a full propositional language, yet be "compositional" only in the trivial sense that none of its representations have meaningful parts. The content so conveyed would still be characteristically linguistic in that it is assigned to elements of the language in the same way as for a full-fledged propositional language, i.e. by an interpretation function from symbols to propositions.

What about languages that contain complex expressions with meaningful parts? The consensus is that such languages need not be compositional, and for any particular language, it is a contingent fact whether that language is compositional or not.[9] Nevertheless, it is widely accepted that some form of compositionality obtains for the case of greatest interest, natural language. If natural language were not compositional, then the ease with which we learn it, successfully interpret novel expressions, and produce our own novel utterances would be mysterious. There are difficulties for a compositional account of natural language, however. If a language is non-compositional, then it contains "a complex expression that changes its meaning when some of its constituents are replaced by synonymous ones" (Westerståhl 2015: 9). There are well-known instances of such expressions in natural language, for instance opaque contexts such as "Jo believes *x*": *x* may be replaced with a synonymous expression, and yet the truth conditions (overall meaning) of the sentence change. There is an active research program to find compositional

[8] See Hodges (1998) for a canonical statement of the formal notion of compositionality and Pagin and Westerståhl (2010a; 2010b) or Szabó (2013) for surveys.
[9] A minority holds that it is trivially true that all languages are compositional, see Westerståhl (1998) and Pagin and Westerståhl (2010b) for surveys and rebuttals. This claim is more convincing for special languages subject to *a priori* constraints, e.g. the "language of thought."

analyses of examples such as this, with a host of technical strategies available for generating compositional interpretations of purported counterexamples, for instance by adding hidden variables or "lifting" semantic types (Zimmermann 2012; Johnson *forthcoming*).

If a representation is compositional, then substituting parts of the representation synonymously won't change its meaning. An even stronger notion of compositionality applies if substitution of parts non-synonymously *always* changes meaning. This stronger condition obtains only when the function from the meanings of parts to overall meaning is injective, i.e. never maps different combinations of meaningful parts to the same overall meaning.

> A representation is *inverse compositional* if and only if its overall content (meaning) is an *injective* function of the contents of its (atomic) meaningful parts, and the structural relations between them.[10]

Inverse compositionality says of overall meaning that it is *unique* modulo the synonymy classes of atomic meaningful parts, where the atomic meaningful parts of a representation are just the smallest component bearers of meaning. In the linguistic case, these are morphemes—so, while "color" and "colour" differ in atomic orthographic components (letters), they are nevertheless completely synonymous as atomic bearers of meaning, both presenting the content [*color*] by different conventions.

Inverse compositionality was introduced to explain aspects of communication and thought that appear to require a stronger condition than compositionality. For instance, Pagin (2003) argues that, while compositionality may explain how a hearer is able to comprehend a complex expression, it doesn't help explain how a speaker is able to generate complex expressions in the first place. Beginning from a (complex) meaning she intends to express, the speaker must choose words to compose a suitable expression. How she does this is a mystery unless there is a function from overall intended meaning to the

[10] This notion is also known as "reverse compositionality" (Fodor and Lepore 2001). I follow here the terminology of Pagin (2003; Pagin and Westerstähl 2010a) as formally this is the inverse of the regular notion of compositionality. Compositionality says there is a function from meanings of parts to meanings of wholes, whereas inverse compositionality says there is a function from meanings of wholes to meanings of parts. This is equivalent to the statement in the text since a function is injective just in case its inverse is also a function.

meanings of parts, i.e. unless the inverse of compositionality holds. Similar considerations motivate Fodor and Lepore, who worry that, while compositionality explains how anyone who understands "dog" and "bark" understands "dogs bark," inverse compositionality is needed to explain how anyone who understands "dogs bark" also understands "dog" and "bark." They gloss the required principle as "each constituent expression contributes *the whole of* its meaning to its complex hosts" and assert that "as far as anybody knows, compositionality and [inverse] compositionality *always go together*" (2001: 366, emphases in original). Nevertheless, inverse compositionality places a strong requirement on a semantic system, and just as the compositionality of natural language is a matter of debate, so also is its inverse compositionality (e.g. Johnson 2006).

So, language-like representations may or may not be compositional, and they may or may not be inverse compositional. It is worth closing this section with a reminder that we are concerned here with skeletal content, content a representation bears independent of other representations or contextual background. One of the most compelling lines of argument against compositionality turns on the apparent context-sensitivity of familiar conversational examples—if meaning is determined by context of utterance, then it does not appear to be a function of the meanings of the parts of the expression. However, such context-sensitive meaning will be part of the full-blooded content of the expression, not its skeletal content. The skeletal content of an expression, sometimes thought of as "what is said," may be "impoverished," perhaps not even characterizing a full proposition (Szabó 2013: 4.2.3), yet still provide scaffolding to support whatever full-blooded contents context may impose (Bach 1994). One might consistently maintain both that full-blooded content is non-compositional, and that skeletal content is (inverse) compositional, determined by / determining the skeletal meaning of the expression's parts.

§6. *Are digital images (inverse) compositional?*

I argue that, in contrast to language-like representations, images are both always (non-trivially) compositional and never inverse compositional.

Images are always non-trivially compositional. This follows because images always have internal structure, and content-bearing properties of this structure contribute to overall content. More specifically, clause (i) of **H** implies that synonymous images have

the same structure, i.e. that structure is somehow critical for determining overall content. Clause (ii) implies that parts of synonymous images must be assigned analogous colors. Insofar as color values in some sense determine the content of image parts, (i) and (ii) together imply that overall content is a function of the contents of parts and their arrangement, i.e. images are compositional.

More intuitively, the status of an image *as an image* depends upon some structured "depiction" of contents—whether those contents are real or imagined, concrete or abstract. An image apt to depict a cat on a mat must involve some component corresponding to *the cat* and some component corresponding to *the mat*, and these components must stand in a relation corresponding to *on*-ness, or it is simply not an *image* of a cat on a mat. While the state of affairs that the cat is on the mat may by stipulation be represented with an unstructured proposition letter *p*, stipulation cannot ensure *p* bears the *imagistic* content that a cat is on a mat, and we can see this because *p* does not stand in relation **H** to images of cats on mats. Likewise, an abstract Mondrian grid of colors, such as *Tableau I*, while not (obviously) depicting any (recognizable) subject, nevertheless depends for its identity *qua image* on a particular spatial arrangement of color patches and lines, and this same imagistic content may be represented by many distinct replicas of *Tableau I*, synonymously related by **H**.[11] At stake again is skeletal content: the phrase "Great Wave off Kanagawa" may indexically refer to the same content as Figure 1a under the right circumstances, but its skeletal content is determined by the semantics of English, while the skeletal content of Figure 1a is determined by a spatial arrangement of color patches— "Great Wave off Kanagawa" is not *imagistically* synonymous with Figure 1.

The claim that the contents of image parts plus their spatial arrangement compositionally determine the content of the image as a whole has been defended many times for the special case of maps, although opinions differ on whether the compositional structure of maps is the same sort as that of language (Casati and Varzi 1999; Blumson 2012) or qualitatively different (Camp 2007; Rescorla 2009). The stronger claim that all images exhibit compositional structure has been defended less frequently, although an

[11] Although abstract examples such as this are "non-representational" in the sense of not depicting recognizable objects (although even that claim may be contentious or indeterminate—consider for instance Mondrian's own *Broadway Boogie Woogie*), this does not mean they are not "representations" in the more general sense of objects that convey meaning and thus demand a vehicle/content distinction. In this more general sense, we must treat abstract paintings (and musical recordings, footnote 7) as representations in order to make sense of the cultural practices in which they participate, especially our practice of reproducing them and treating distinct reproductions as in some sense "the same"—i.e. *the same in content*.

exception is Blumson (2014). Blumson argues that images are compositional "since one may proceed by rational inductive means from knowing what some pictures represent to knowing what other pictures composed of the same parts represent" (115).

Two key issues distinguish the position advanced here from other defenses of image compositionality. First, the target is different—the focus here is the compositionality of skeletal content, whereas others have addressed full-blooded content. Blumson's argument, for instance, invokes content attributions that rely on knowledge of the content of other representations, i.e. the kind of knowledge that fleshes out content. Second, the methodology is different, since my starting point is a given set of synonymy relations, rather than the principles of a particular representational scheme. This difference is nicely illustrated by the different role played by homomorphism in this approach and that of, say, Camp: Camp derives image compositionality from the homomorphism that obtains between an image vehicle and its content (2007: 156–9); in contrast, I derive image compositionality from the homomorphism that obtains between any two image vehicles bearing the same content.

Images are never inverse compositional. This is a corollary of clause (iii) of **H**. A representation violates inverse compositionality if non-synonymous replacement of an atomic meaningful constituent may leave overall meaning unchanged. Clause (iii) says that fine-grained parts of two digital images may differ in (relative) color value, even when images are synonymous. I argue that color patches are plausible atomic meaningful constituents of digital images, and that their synonymy conditions depend on color value, since in some contexts a change in the color of a patch produces a change in content. Nevertheless, in other contexts, the very same change in color of an equivalent patch (i.e. non-synonymous replacement of an atomic constituent) may leave overall content unchanged.

In the literature on image compositionality, a classic worry has been that there is no canonical decomposition of an image into parts (for survey and rebuttal, see Blumson 2014: Ch. 6; c.f. Abell 2005: 192–4). In the case of digital images, however, this worry seems ill-founded, as digital images are defined in terms of their atomic parts, namely pixels. However, pixels may not be atomic *meaningful* image parts, i.e. they may be analogous to letters rather than morphemes. If pixels are not atomic bearers of content, then some of the examples considered above do not count as counterexamples to inverse

compositionality. For instance, suppose a single pixel of an image is changed to a random value, introducing noise while leaving content unaltered; the two images, before and after this change, will be related by **H**, yet if that single pixel is not itself a bearer of content, this example would not violate inverse compositionality—the change would be analogous to a mere variation in spelling rather than non-synonymous morpheme replacement. Consequently, the main challenge here is not to show that images may differ yet still be synonymous, this was demonstrated in §3, but rather to show that these differences occur at the level of content-bearing image parts.

A plausible first pass at an atomic meaningful image constituent is the uniform patch of color. In the case of maps, for instance, regions of uniform color are treated as atomic constituents by Casati and Varzi (1999: 191–2). Crucially, color patches contribute substantively to content: whether an image is *of a red ball*, or *of a blue ball*, depends upon the hue value of a (set of) color patch(es). Likewise, color patches themselves are prima facie bearers of content, attributing, for instance, a particular surface property to a depicted surface. A red patch in an image vehicle bears the content that some surface, or region, is *red*.[12] However, we must be careful not to assume that a color patch always bears as content its own color. In the case of fleshed out content, it is clear that colors may convey more elaborate meaning, they "can become correlated not with colours but with feelings and moods" (Bach 1970: 121). Yet even the skeletal content of color patches may not be identified with precise color value; this follows already from clause (ii) of **H**, which indicates that it is *relative* color value, the relationship a color stands in to other possible colors, that is critical for image synonymy. For instance, the dark stripes of blue on the underside of the wave in Fig. 1a and those in Fig. 1c bear the same content, despite differing in absolute value, since they stand in the same relative relationship to other colors in their respective images. The color space as a whole for 1c is merely slightly desaturated relative to that for 1a.

However, the content of a color patch is not determined solely by the value it is assigned within the image color space. Its skeletal content is also determined in part by its spatial relationship to other color patches in the image. Consider, for instance, Edward

[12] Abell (2005) calls such regions "sub-pictorial parts," defending the view that they bear content, and that this content is not determined by convention, since they are not "salient" to the "makers or interpreters" of images; consequently, the semantics of sub-pictorial parts is disanalogous to that of ambiguous words. I take her arguments to complement and support those offered here.

Adelson's "checkershadow illusion," in which a cylinder casts a shadow across a checkerboard, yet the board has been cleverly arranged such that a "black" square outside the shadow and a "white" one within it have exactly the same reflective properties. Here patches assigned precisely the same value in color space differ in what they are apt to depict: in one spatial context, a black square, and in another, a white square. Thus, the skeletal content of a color patch is doubly contextual: it is determined on the one hand by its spatial position within the image, and on the other by the position of its color value within color space.

The double contextuality of skeletal content ensures that, in some instances, color patches may be replaced non-synonymously, and yet overall content remain unchanged. This is because in one spatial context a slight change in patch color may change what the image overall is apt to depict, while in a different spatial context, the very same slight change may have no effect whatsoever on image content. Consider, for instance, Figure 2a: the middle patch of the top stripe may be changed without any change in content (2b), while the middle patch in the bottom stripe may be changed in precisely the same way yet change content substantively (2c). Or consider again Hokusai's *Great Wave*: a white speckle located in one region may bear content, serving as a fleck of spray, yet a speckle of the very same color, shape, and size in another region may constitute mere noise or dirt. That a change in these small regions may change overall content indicates that they are indeed content-bearing. However, that they may in some instances change—be replaced non-synonymously—without changing overall content indicates that image content as a whole is not uniquely determined by such content-bearers, i.e. is not inverse compositional.

[Figure 2 about here, caption:

> The center square in both lines of (a) has exactly the same greyscale value (24%). In (b) the top one only has been darkened (to 35% grey), yet synonymy with (a) is preserved. In (c) the bottom one only has been darkened by exactly the same amount, yet this breaks synonymy with (a), since the bottom line is no longer apt to depict a smooth gradation of grey.

]

How may we reconcile the compositionality of images with the failure of inverse compositionality? When we consider the details of an image "from the bottom up" as it were, examining constituents in isolation, we see fine-grained details, even individual pixels, as bearing the very same properties (colors) that determine content at larger grains. If we consider the image as a whole "from the top down," however, as we "zoom in" from its holistic content to more and more fine-grained details, these at some point lose significance for the image as a whole, i.e. parts of the image no longer represent parts of the image's subject. A process of semantic decomposition that begins by revealing compositional structure eventually breaks down. When I look very closely at the reproduction of Hokusai's *Great Wave* on the printed page, I can see individual splashes of ink from the printing process, the "pixels" it comprises. These details are disanalogous to the orthography of linguistic representations, they are individually assessable as bearing distinctively imagistic content (whereas in most cases, a letter considered in isolation does not bear linguistic content). Nevertheless, I do not accept them as unique contributors to the overall content of the image *qua Great Wave off Kanagawa*, i.e. the content it shares with all other sufficiently high-resolution reproductions of Hokusai's masterpiece. So, from this perspective, the claim that images are compositional, but not inverse compositional, may be restated: images are compositional *at one level of granularity*, yet *non*-compositional *at a finer level of granularity*.

§7. *Thresholds and granularity*

The granularity at which synonymous digital images bear content is bounded in **H** by two parameters: δ and ε; but what sets these parameters? One strategy in the content literature pins them to the limits of human perceptual power: images that differ in their physical properties may nevertheless be indistinguishable to us.[13] I've argued here for a coarser

[13] This is the "lack of effective differentiation" that Goodman and Elgin identify as distinctively pictorial (1988: 131). Since perceptual indistinguishability is not transitive, representational systems that allow physical variance but perceptual indifference satisfy Goodman's notion of a dense symbol scheme (1976: 160). However, Zeimbekis (2012) argues that digital images may be physically distinct, yet type-identical with respect to their representational properties, effectively demonstrating a threshold at which indistinguishability, and thus (for him) synonymy of digital images is transitive. In contrast, I allow synonymous images to be perceptually distinguishable, yet still take synonymy to be strictly transitive. This introduces a tension, however, as the everyday practices that determine the thresholds for synonymy are not themselves transitive. Although this tension warrants further investigation, I set it aside here on the grounds that the algorithms discussed in the text, such as those for resizing images on screens or printing them to

granularity based on our everyday practice of treating images that differ slightly as nevertheless bearers of the very same content—Figures 1a and 1b differ in some respect, but that respect is not itself content-relevant. The bounds on this coarser granularity are encapsulated in the algorithms we trust to reproduce, resize, and otherwise transform our digital images. But under what circumstances do we trust these algorithms? I think the answer reintroduces a precise role for context or perspective, not in *determining* content, but in *delimiting* it.

Consider again Hokusai's *Great Wave* in Fig. 1a. The image you are looking at as you read this paper, and the image I am looking at while writing it, will be of different resolutions, displayed in different color spaces, and thus related by **H**. The algorithms that relate these two images will start from the high-resolution version available from the Metropolitan Museum of Art. The image that I view while writing this has been downsampled to fit my laptop screen; the image you view may have been resampled in printing the paper to pdf (a question for the typesetter); if you view it electronically, it will also be up- or downsampled to fit the resolution of your screen, by an algorithm dependent on the software you use to view the pdf; if you are viewing the paper as a physical printout, then the image will have been downsampled to fit the resolution of the printer, and possibly the colorspace downsampled to greyscale if you did not print in color. By what standard can we assess whether these various images are appropriately synonymous, i.e. what are the relevant values of δ and ε?

At the very least, you, the reader, need to be able to follow the discussion of Hokusai's *Great Wave* in the text—if the image is downsampled so severely that the passengers in the forward boat cannot be identified, or their topknots distinguished, then image content has clearly changed. Furthermore, you need to follow the textual discussion *non-inferentially*: if the image has been downsampled to greyscale, you can likely still identify the dark blue stripes on the underside of the wave, but the image itself only conveys the content that these stripes are darker than their neighbors; you must infer the stripes were originally blue from the text, your memory, or your inferential knowledge of waves. Generalizing, it seems that in order for imagistic content to be the same across the

---

paper, are treated as transitive, and it is this practical attitude that is the present *explanandum*. I thank Ben Blumson for emphasizing to me the importance of this question.

two images, any detail of the first image to which I might *potentially* have drawn your attention needs to be available to you in the second image (and *vice versa*).

The algorithms for downsampling digital images are designed to preserve as much imagistic content as possible, but whether they preserve enough content for the downsampled image to be synonymous with its parent depends on context—if we ask an algorithm to downsample an image past the threshold at which content is preserved, it will do so. But if algorithms for resampling digital images can be made to "break" like this, degrading or losing content, what ensures that in the typical case they do preserve content? I think the answer extends beyond a rough parity in human perceptual acuity to broad commonalities in our expectations and purposes—not only do our eyes operate at roughly the same granularity, but our attention to detail is roughly the same, and our expectations about what others can and do see in an image is broadly stable across standard contexts of image sharing. These common characteristics of our use of images, these facts about us and what we do, set the baseline threshold across which granularity of detail ceases to bear content. It is because of such commonalities in how we perceive and interpret images that we all accept two photos struck from the same negative as bearing the same content, two showings of a film on screens of different resolutions as the same, and our respective instances of Figure 1 as both Hokusai's *Great Wave*.

We can now see the limited sense in which imagistic content is perspectival. Human perspective and use sets the threshold between content-bearing and non-content-bearing aspects of an image. Perspective does not determine content itself—that is just determined by facts about the distribution of colors within the image—, rather it delimits content, inscribing a boundary between meaningful and artifactual details. Yes, this threshold may change with circumstances, interest, and ability; Figures 1a and 1c may be strictly synonymous for the collector of *Ukiyo-e*, but not for the historian of Japanese printmaking technology. Nevertheless, because the role of perspective is merely to set a limit on content, it does not block the possibility of an objective analysis of imagistic content modulo this perspectival threshold. This is good news for those who would develop a theory of imagistic content outside the realm of human artifacts, such as paintings and photographs, and apply it to features of the natural world, such as perceptual experience or mental content.[14]

[14] See Isaac (2013) for an extended discussion of this issue.

§8. *Conclusion*

Our everyday image-using practice accepts as synonymous images related by algorithmic transformations that alter their fine-grained details. This observation implies a structural constraint on digital imagistic representation, namely: digital images are always non-trivially compositional, but never inverse compositional. This formulation reveals a fundamental contrast between imagistic and linguistic representations, since the latter may or may not be compositional, and may or may not be inverse compositional. The semantic implication of this result is that, for any set of synonymous digital images, there is some granularity at which fine-grained details do not make unique contributions to content. The threshold at which this changeover from content-relevant to content-indifferent detail occurs is itself context dependent, determined by human perspective and use; nevertheless, a broad uniformity across all such contexts permits us to devolve synonymy assessments to algorithmic procedures, such as those from which our investigation began. Finally, the user-independent character of these procedures implies that perspective does not determine content, but merely delimits it—welcome news for would-be theories of imagistic content that extend beyond the arts.

Acknowledgements

This paper has benefitted from the comments of Ben Blumson and several anonymous referees, as well as conversations with Michael Johnson, Bryan Pickel, Brian Rabern, and J. E. Wolff.

References

Abell, Catharine (2005) 'Against Depictive Conventionalism'. *American Philosophical Quarterly*, 42, 185–197.

Bach, Kent (1970) 'Part of What a Picture Is'. *British Journal of Aesthetics*, 10, 119–137.

Bach, Kent (1994) "Semantic Slack: What is Said and More'. In Savas L. Tsohatzidis (ed.) *Foundations of Speech Act Theory* (London: Routledge), pp. 267–291.

Blumson, Ben (2012) 'Mental Maps'. *Philosophy and Phenomenological Research*, 85, 413–434.

Blumson, Ben (2014) *Resemblance and Representation: An Essay in the Philosophy of Pictures*. Open Book Publishers.

Camp, Elisabeth (2007) 'Thinking with Maps'. *Philosophical Perspectives 21: Philosophy of Mind*, 145–182.

Casati, Roberto and Achille C. Varzi (1999) *Parts and Places: The Structures of Spatial Representation*. Cambridge: MIT Press.

Cassirer, Ernst (1944) 'The Concept of Group and the Theory of Perception'. *Philosophy and Phenomenological Research*, 5, 1–35.

Churchland, Paul (2012) *Plato's Camera*. Cambridge: MIT Press.

Fodor, Jerry and Ernie Lepore (2001) 'Why Compositionality Won't Go Away: Reflections on Horwich's 'Deflationary' Theory'. *Ratio*, 14, 350–368.

Goodman, Nelson (1976) *Languages of Art*, 2nd ed. Indianapolis: Hackett.

Goodman, Nelson and Catherine Z. Elgin (1988) *Reconceptions in Philosophy and Other Arts and Sciences*. London: Routledge.

Gunther, Y. (ed.) (2003) *Essays on Nonconceptual Content*. Cambridge: MIT Press.

Haugeland, John (1991) 'Representational Genera'. In W. Ramsey, S. P. Stich, and D. E. Rumelhart (eds.) *Philosophy and Connectionist Theory*, pp. 61–90. Reprinted in *Having Thought*, 1998 (Cambridge: Harvard UP), pp. 171–206.

Hodges, W. (1998) 'Compositionality is not the Problem'. *Logic and Logical Philosophy*, 6, 7–33.

Hopkins, R. (1998) *Picture, Image and Experience*. Cambridge UP.

Isaac, Alistair M. C. (2013) 'Objective Similarity and Mental Representation'. *Australasian Journal of Philosophy*, 91, 683–704.

Johnson, Kent (2006) 'On the Nature of Reverse Compositionality'. *Erkenntnis*, 64, 37–60.

Johnson, Michael (*forthcoming*) 'Compositionality'. In L. Matthewson, C. Meier, H. Rullmann, and T. E. Zimmermann (eds.) *Companion to Semantics*, Wiley-Blackwell.

Lopes, D. (1996) *Understanding Pictures*. Clarendon Press.

Luce, R. Duncan, David H. Krantz, Patrick Suppes, and Amos Tversky (1990) *Foundations of Measurement*, vol. 3. San Diego: Academic Press, Inc.

Niederée, Reinhard (2010) "More than Three Dimensions: What Continuity Considerations Can Tell Us about Perceived Color'. In Jonathan Cohen and Mohan Matthew (eds.) *Color Ontology and Color Science* (Cambridge: MIT Press), pp. 91–122.

Nozick, Robert (2001) *Invariances*. Cambridge: Harvard UP.

Pagin, Peter (2003) 'Communication and Strong Compositionality'. *Journal of Philosophical Logic*, 32, 287–322.

Pagin, Peter and Dag Westerståhl (2010a) 'Compositionality I: Definitions and Variants'. *Philosophy Compass*, 5/3, 250–264.

Pagin, Peter and Dag Westerståhl (2010b) 'Compositionality II: Arguments and Problems'. *Philosophy Compass*, 5/3, 265–282.

Peacocke, Christopher (1986) 'Analogue Content'. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 60, 1–17.

Rescorla, Michael (2009) 'Predication and Cartographic Representation'. *Synthese*, 169, 175–200.

Suppes, Patrick (1973) 'Congruence of Meaning'. *Proceedings and Addresses of the American Philosophical Association*, 46, 21–38.

Szabó, Zoltán Gendler (2013) 'Compositionality'. *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2013/entries/compositionality/>.

van Fraassen, Bas (2008) *Scientific Representation*. Oxford UP.

Westerståhl, Dag (1998) 'On Mathematical Proofs of the Vacuity of Compositionality'. *Linguistics and Philosophy*, 21, 635–643.

Westerståhl, Dag (2015) 'Questions about Compositionality'. In P. Schroeder-Heister, G. Heinzmann, W. Hodges, and P.-E. Bour (eds.) *Logic, Methodology, and Philosophy of Science: Proceedings of the XIVth International Congress* (London: College Publications), pp. 123–147.

Zeimbekis, John (2012) 'Digital Pictures, Sampling, and Vagueness: The Ontology of Digital Pictures'. *The Journal of Aesthetics and Art Criticism*, 70, 43–53.

Zimmermann, Thomas Ede (2012) 'Compositionality Problems and How to Solve Them'. In Wolfram Hinzen, Edouard Machery, and Markus Werning (eds.) *The Oxford Handbook of Compositionality*, pp. 81–108.

a

b

c

a

b

c