Design & Syntax in Pictures

Mind and Language

Robert Hopkins

Abstract:

Many attempts to define depiction appeal to viewers' perceptual responses. Such accounts are liable to give a central role in determining depictive content to picture features responsible for the response, *design*. A different project is to give a compositional semantics for depictive content. Such attempts identify *syntax*: picture features systematically responsible for the content of the whole. Design and syntax are competitors. But syntax requires system, in how picture features contribute to content, that design does not. By examining John Kulvicki's semantics for basic depictive content, I argue that the relevant systematicity is absent from the pictorial realm.
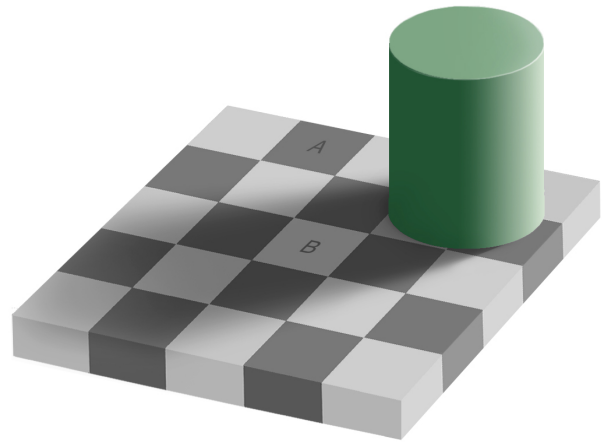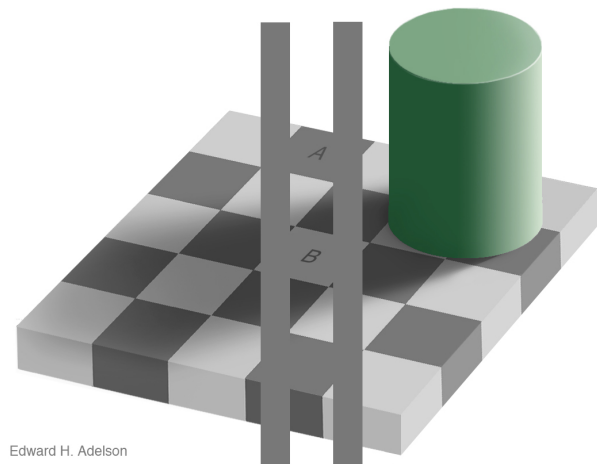
Figures

Figure 1

Figure 2

Edward H. Adelson

Figure 3

§1 Let's begin with a distinction some draw (e.g. Adelson 2000: 341) between three features of a surface.

> *Illuminance* is the amount of light falling on the surface.
> *Reflectance* is the amount of light the surface reflects.
> *Luminance* is the amount of light reaching the eye from the surface.

Reflectance is a capacity, the capacity of the surface to reflect a certain proportion of the incident light. Illuminance and luminance, in contrast, are categorical: the amount of light falling on, or reaching the eye from, the surface at a given time. Reflectance combines with illuminance to determine luminance. For instance, in figure 1 the two squares at bottom right match in illuminance, differ in reflectance and, as a result, differ in luminance. Not only does one square look to be a darker grey than the other (reflectance), and not only do we see them both to fall equally in the light (illuminance), but we'd expect that, if we viewed them through a viewing tube, thereby removing information about illuminance and reflectance, we'd see a darker tone (luminance) looking at one than at the other.

Figure 1 is Adelson's checkershadow illusion.[1] Some take it to show that we don't, in general, see luminance.[2] Consider squares A and B. Their luminances match, as demonstrated by figure 2. Were we looking directly at such a checkerboard, and viewed A and B through a viewing tube, we might see this. In each case we'd see that we are presented with the same amount of reflected light. But that's certainly not how things look without the tube. It's a shock to learn that the two squares match in luminance. Of course, we see the other properties above. We see that one square is darker than the other (reflectance) and that one is in light, the other in shadow (illuminance). What we fail to see is the factor these two others determine - that in respect of luminance A and B match.[3]

My interest is not in the argument's conclusion, but in two assumptions on which it rests. As it stands, it moves too fast. We're not looking at a checkerboard, but a picture of one. We must therefore distinguish features of the vehicle, the picture, from features of its content, the scene it represents. What figure 2 shows is that there is a match in vehicle properties: the part of the picture depicting square A is the same shade of grey (reflectance) as

---

[1] See http://persci.mit.edu/gallery/checkershadow.

[2] [Redacted] presented this argument in a class we taught together.

[3] 'See' here means 'consciously veridically perceive'. The visual system must be sensitive to luminance, since that is the basis on which it computes reflectance and illuminance (Burge 2010: 351-4). What the argument purports to show is that our conscious impressions of luminance can be quite wrong. We here perceive reflectance and illuminance correctly, but misperceive luminance. If the moral generalises, luminance is something our conscious states fail to track, even though it provides the data on the basis of which the visual system generates veridical experience of the other two values.

the part depicting square B. That is certainly a surprise, but how does it show anything about seeing (or not seeing) luminance?

The answer, if there is one, lies in the following assumption: pictures, or at least pictures of this kind, depict luminance by themselves exhibiting suitable reflectances.[4] The grey of the marks determines the luminance depicted, so that, if two parts of the picture match in reflectance, they match in the luminance they depict. Figure 2 shows the antecedent of this conditional to hold, thereby securing the consequent. A and B are depicted as matching in luminance, but that is something we fail to see when looking at the picture. And from this claim about the perception of luminance in pictures we can (somehow) draw a conclusion about our perception of luminance more generally.

How might this last step be justified? Here we encounter the second assumption. When we look at figures 1 and 2, we see, not checkerboards but representations of them. That is what opens up the gap between features of the vehicle (the reflectance of the picture parts) and features of its content (the luminances they purportedly depict). In general interpreting representations shows us little about perceiving objects. We learn little about shape perception, for example, by studying our *terms* for shapes. Why should things be different here? The answer must lie in some distinctive connection between understanding pictures and perceiving their objects. Understanding a picture, like reading words, involves seeing representations; but only in the picture case is our grasp on meaning itself in some key way perceptual.

The idea that pictures are distinctively perceptual representations is of course an old one. A good part of the philosophical literature on pictures has been devoted to working it out. Indeed, in the anglophone tradition, serious philosophical discussion of pictures in effect began with the debates between those, like Goodman, who emphasized pictures' status as representations (or symbols), and those, like Gombrich and Wollheim, who emphasized their perceptual character.[5]

There is no need to see these two aspects of pictures as in conflict. The past fifty years saw the development of many views that aim to reconcile them. Nonetheless, it is possible to find pockets of tension between appealing positions that take as central one aspect or the other. My interest is in one such tension. It lies between positions that seek to advance two rather different projects: defining depiction in perceptual terms (§2), and giving a compositional semantics for depicted content (§3). Each project foregrounds a role for features of the vehicle in

---

[4] Here it matters that reflectance is a capacity of the surface, not a matter of the amount of light it actually reflects on a given occasion. Only the latter varies with the illuminance. This squares the assumption with the fact that a picture depicts what it does regardless of how brightly lit it is.

[5] See Goodman 1968, Gombrich 1960, Wollheim 1987.

determining content. However, they conceive that role in rather different terms. A compositional semantics requires the determination to be systematic. It presupposes constant relations such as that enshrined in the first assumption above, that picture reflectance determines depicted luminance (§4). Perceptualist definitions of depiction, in contrast, make no such demands. The checkershadow pictures offer a way to test the issue between them. Those images cast doubt on the first assumption, and with it on the prospects for a compositional semantics of pictures (§5). The nature of our engagement with pictures, in particular the central role in it played by experience, at the very least complicates any compositional semantics they involve (§6). The rival perceptualist project, with its more relaxed conception of the relation between content and vehicle, avoids these problems. So much the better, then, for the perceptualist approach (§7).

§2 If we seek to distinguish pictures from other representations, we do well to begin by narrowing focus. Pictures represent in many ways. Some of those forms of representation are also exhibited by representations of other kinds, such as language. (Think of exemplification (Goodman 1968: ch.2) or iconography: e.g. representing a saint by her 'attributes' (Kulvicki 2020a: ch.5).) But one, it is supposed, is found only in pictures (along perhaps with other closely related symbols, such as sculpture, mime and even auditory imitations). Define this, and we will have defined the class of pictures (and their near neighbours). Thus much inquiry has sought to complete the following biconditional:

P depicts O iff……

Where 'depiction' names the form of representation we're trying to define, and 'O' can be replaced by any content that any picture can represent in that way.[6]

Now, any completion of the biconditional might be taken to answer either of two questions. One, the *boundary question*, is this: what makes (some of) P's content depictive? This is the question which led us to the biconditional in the first place. But many answers will also simultaneously respond to a different question: in virtue of what does P depict O, rather than anything else? This isn't a question about boundaries, but about what it is that endows P with the depictive contents it has. If semantics is the task of describing the contents of representations, and meta-semantics the task of saying in virtue of what they have those contents, then this second question is meta-semantic.

As an example of a view that answers both questions, consider the following:

---

[6] See, for instance, Wollheim 1987, Schier 1986, Peacocke 1987, Walton 1990: ch.VIII, Lopes 1996, Hopkins 1998.

Illusionism: P depicts O iff (1) suitable viewers, on looking at P, would have visual experience as of O; & (2) something about P's context of production renders it appropriate to experience P that way.[7]

Illusionism, at least in this blunt form, is not plausible. Very few pictures sustain the illusion that we are seeing what they depict. The view does, however, nicely illustrate the present point. It answers the boundary question, however implausibly: what marks out pictures from other representations is that only the former generate the perceptual illusion of seeing their objects. But it also promises a meta-semantics of pictures: any picture depicts whatever contents figure in the illusory experiences it appropriately sustains, and it is in virtue of sustaining those experiences that it has those contents. The biconditional alone does not, of course, secure this 'in virtue of' claim. Hence it is possible for views to complete the biconditional without addressing meta-semantic issues. (For an example, see Walton 1990 ch.VIII.) But those offering completions may well take themselves thereby to answer the meta-semantic question. That, indeed, has been the attitude of many.[8]

Illusionism completes the biconditional by appeal to some response on the part of picture viewers. In this it is far from alone. The general form of such views is:

Viewer Response Accounts: P depicts O iff (1) viewers S respond to P in some O-specific way R; & (2) something about P's context of production renders that response appropriate.

R might conceivably be specified in such a way that it does not admit of variation across different substitutes for O. (If it were, the account would tell us what makes P depictive, but not what makes it depict O in particular.) But most specifications are likely to be of a genus that has species specific to the various possible substitutes. This is certainly true of Illusionism: the visual experience as of a horse, say, is distinct from visual experience as of a dog; and (arguably, at least) visual experience as of Wellington is different from visual experience as of Napoleon; and so on. Provided R comes in O-specific flavours in this way, Viewer Response Accounts will be in a position to attempt to answer the meta-semantic question.

Examples include accounts that are more or less distant relatives of Illusionism (Gombrich 1960; Wollheim 1987; Lopes 2006; Newall 2009). But they also include other accounts which make R some other kind of experience, such as experienced resemblance views (Peacocke 1987, Budd 1993; Hopkins 1998). And they include views that frame R in non-experiential terms, such as those on which pictures trigger our capacities to recognise their objects in the flesh (Schier 1986; Lopes 1996). Provided the key response is itself in some way

---

[7] On the need for the second condition, see Wollheim 1987: 48-51.

[8] Walton aside, it is true, for instance, of all those cited in note 6.

perceptual, Viewer Response Accounts can be seen as different ways to endorse the second assumption identified above (§1), that pictures are representations that are distinctively perceptual.

Viewer Response Accounts will think that when a given picture does elicit the relevant O-specific response, it does so in virtue of visible features of its surface. It is because the canvas is marked a certain way that when we look at it, seeing those marks, we are led to respond in the relevant way. Following Lopes (2006: ch.1), we can give these features a name:

> *design*: those visible features of P responsible for its eliciting the O-specific response R.

Where Viewer Response Accounts aspire to answer the meta-semantic question, the picture's design is thus a set of features responsible for its depicting what it does. That responsibility is indirect: design matters only because it sustains the key response R, and that is what directly determines the picture's depictive content. But, indirectly or not, design is responsible for content.

§3 Let's turn to compositional semantics. A semantics of pictures, or more precisely of the depictive content of pictures, might make claims of various kinds. It might identify the contents of particular pictures. It might draw broad distinctions between kinds of depictive content: e.g. singular vs predicative (Goodman 1968: ch.1; Hyman 2013). It might assert, or deny, that depictive content can include content of certain kinds, for instance limiting it to the kinds of content that can figure in (non-pictorial) visual experience (Hopkins 1998: ch.1). It might describe the logical form of depictive content in general, or depictive content of certain kinds (Gregory 2020). Or it might identify a hierarchy within depictive content, e.g. describing some layer which is basic to all the rest (Kulvicki 2006: ch.6).

In addition to describing depictive content, however, a semantics for pictures might attempt the rather different task of explaining how the content of a whole picture is generated from that of its parts, along with the ways in which they've been combined.[9] The ambition of offering such a compositional semantics requires us to bridge the divide between aspects of the representation and aspects of its content. We need to identify the parts the combination of which determine the 'meaning' of the whole. Those parts must themselves have meanings, or semantic values, if we are to make sense of the meaning/value of the whole as arising out of their combination. But to have meanings they cannot themselves simply be elements at the level of meaning. They must be parts, in

---

[9] Is the resulting theory meta-semantic? Not in the sense in which the theories discussed in §2 are. For, while they promise to explain how any pictorial element has the content it does, a compositional semantics takes for granted the semantic value of its basic units, the syntactic parts.

some sense or other, of the representation itself. They must be aspects of its syntax: identifiable parts or aspects of the representation that contribute distinct parts or aspects of content.

What would such a theory look like? John Kulvicki (2015a) has shown us, at least for a certain kind of picture, and a certain, basic layer of (depictive) content. His inspiration comes from maps.

Viewed from the syntactic point of view, maps combine places on the map (what Kulvicki calls *locations*) with features placed at those locations (*locatables*). What are the semantic values of these elements of syntax? The locations represent places in the world. That is, some system of projection maps the locations to places in the mapped territory. The locatables predicate properties of the places those locations represent. Placing blue (a locatable) at a location may, for instance, represent the corresponding part of the territory as sea; placing yellow might represent it as land.

Kulvicki suggests we treat locatables as coming in degrees of freedom, bundles such that, in a well-formed map, every location displays some locatable from the bundle, and where the locatables within the bundle represent incompatible properties. (For instance, it may be that every place on our map must be either blue or yellow, thereby representing the relevant parts of the territory as bearing one of two incompatible properties, being sea or being land.) Where this holds, maps will obey the absence intution: absence of a syntactic value from a location suggests absence in the corresponding place in the world of whatever property F that locatable predicates. But this will be true because, in a well-formed map, absence of one value entails presence of a rival value from the degree of freedom, representing a property incompatible with F. (Kulvicki 2015a: §4)

Kulvicki's proposal is that we can treat some of the content of some pictures in much the same way. His focus is colour photography. In a colour photograph there are again two (and only two) kinds of syntactic feature. There are places on the photograph's surface - locations. Here, as in the case of maps, what locations represent is places in some slice of the world, the slice the photograph as a whole depicts. The locations are mapped to the places they represent by some projection system, in particular by linear perspective. (In maps, in contrast, a wide range of projection systems are possible.)[10] What of the locatables, the syntactic features that can occupy those locations? Well, surely the key determinant of what a part of a colour photo represents is its colour. And what those colours represent is itself some sort of colour property. A red picture part represents something red, a green part represents something green, and so on. Every location on the picture must be some colour or other. And the properties those colours ascribe to the corresponding places are incompatible: no part of the world can be both

---

[10] Since photography can exploit distorting lenses, perhaps linear perspective is not the only projection system it can involve. Other forms of picture are certainly not so bound. Even so, the range of acceptable projection systems for pictures in general, and photographs in particular, is far narrower than for maps.

red and green, or any other two colours. Thus the colour locatables form a degree of freedom. Colour photographs are like simple maps: maps with only one acceptable system of projection relating locations to the places they represent, and where the locatables that can occupy those locations form a single degree of freedom. (Kulvicki 2015a: §6)

It should be clear (as it certainly is to Kulvicki) that what is on offer here is at most a semantics for a particularly basic tier of pictorial content. Colour photographs can depict pretty much anything we can see, but the above limits its claims about content only to colours and places: the photo represents one colour at one place, another at another, and so on. Moreover, as Kulvicki also stresses, the space represented is two-dimensional. Linear perspective maps the 3D world onto the photograph's surface in such a way as to be wholly insensitive to differences in depth. (That is why a photograph of a scene may be indistinguishable from a photograph of a photograph of that scene.) And, since places in a 2D array are often not themselves coloured, if that means having the very same property that surfaces bear, Kulvicki also refines his account of the property the locatables predicate of those places. It is what he elsewhere (e.g. 2020b: 96) calls 'collumination', the property a pink thing in white light shares with a white thing in pink light. Thus this, the 'bare bones' content of colour photographs, is much less than their depictive content as a whole. It doesn't include 3D shapes, or surface colours, let alone many of the other properties we might take photographs to capture. Nonetheless, the bare bones content grounds the rest. Pictures represent rich 3D scenes in virtue of representing the bare bones contents those scenes involve. A photograph depicting a beach at sunset, from a certain point of view, does so in part in virtue of having as its bare bones content a 2D array of colluminations, the particular 2D array that the beach itself would present, viewed from the relevant position.

From all this we can extract a simple compositional semantics for the bare bones content of colour photographs. The semantic value of each colour locatable is some collumination. The semantic value of each location is the place in the 2D array that, given the projection system (linear perspective) corresponds to it in the scene photographed. Each colour locatable predicates the relevant collumination of the place in the 2D array represented by the location it occupies. These pairs of locatables-at-locations form the basic syntactic parts that determine the content of the whole. The semantic value of the whole photograph (as far as bare bones content goes) is the 2D array formed by conjoining the semantic values of each of these parts.

§4 Though Kulvicki takes colour photographs as his central example, his account will apply to a wider range of pictures. Much realistic painting, for instance, might be understood in the same terms, whether or not it depicts real scenes. (A projection system can relate pictures to 2D arrays that are merely possible.) John Willats calls such pictures, in which the basic representational tool is the use of colours to represent (something like) colours *optical* pictures (Willats 1997: ch.6). But, as Willats notes, not all pictures are optical. Line drawings, for

instance, seem to have as their key locatables, not points of colour representing colours, but lines representing edges (see Willats 1997: ch.5). Such pictures, if they admit of a compositional semantics at all, will require a different treatment from Kulvicki's.

One easy extension of Kulvicki's scheme is to black and white photographs, and pictures like them. Exploring how this extension can be made will allow us to apply his claims to the checkershadow pictures with which we began.

Colluminations are the chromatic analogues of luminances. Indeed to our original triplet:

> *Illuminance* is the amount of light falling on the surface.
> *Reflectance* is the amount of light the surface reflects.
> *Luminance* is the amount of light reaching the eye from the surface.

We can offer a parallel trio of chromatic features:

> Illuminant colour is the colour of light falling on the surface.
> Surface colour is the colour of light the surface reflects.
> Collumination is the colour of light reaching the eye from the surface.

Just as the amount of light reaching the eye is determined by the amount falling on the surface combined with the surface's capacity to reflect light; so the spectral composition (colour) of light reaching the eye is determined by the spectral composition of the light falling on the surface combined with the surface's capacity selectively to reflect light of varying wavelengths. Surface colour is a way of changing the spectral composition of incident light.[11] Indeed, if colour is the combination of brightness (apparent luminance), saturation and hue, the three values with which we began may be thought of, without much distortion, as these latest three with hue and saturation stripped out.

Given this, Kulvicki's semantics for colour photographs and other optical pictures easily yields a semantics for black and white photographs and their ilk. The claims at the close of the previous section can all stand, with only

---

[11] Compare Broakes 1992. As a definition of surface colour, this claim is controversial. It suffices for present purposes that surface colour at least correlates, by and large, with a surface's capacity to alter the incident light.

two adjustments. First, the relevant locatables are not surface colours but reflectances.[12] Second, what those locatables represent are not colluminations but luminances. Black and white photos and their kin are just achromatic optical pictures.

§5 I noted at the start two assumptions that must be made by those who appeal to Adelson's checkershadow illusion to argue that we don't in general see luminance. One of those assumptions was that pictures such as figure 1 depict luminance by themselves exhibiting a reflectance: the reflectance of a picture part determines the luminance it depicts. Kulvicki's semantics for achromatic optical pictures turns that assumption into an axiom. In such pictures the only locatable is reflectance. The semantic value of that locatable is luminance. If we are to build a compositional semantics on this basis, we need this relation to be invariant. A given reflectance must always represent one and the same luminance. If not, how can we derive the meaning of a whole picture from that of its parts? [13]

Is the assumption true? Is the theory that takes it as axiomatic built on solid foundations?

Consider again figure 1. The assumption entails that the A- and B-depicting parts depict their respective objects as matching in luminance. What are those objects? Ultimately, these picture parts depict squares on the checkerboard. But they do so in virtue of bearing bare bones content, and that is confined to sections of a 2D luminance array. As I'll say, *patches* on the picture (features of the vehicle) depict *areas* in the array (bare bones content) that correspond to *squares* on the checkerboard (part of its complete depictive content).[14]

Thus what the assumption entails is that the A- and B-depicting patches depict their respective areas as matching in luminance. But this purported match is not merely one it is easy to overlook. Even when one goes looking for

---

[12] Presumably Kulvicki thinks we interpret pictures by seeing the relevant locatables. Since reflectance is a capacity, to extend his account to achromatic pictures he must accept that capacities can be seen. But surface colours are certainly visible, and many think they are capacities, or at least powers of some kind. Indeed, Kulvicki may be among them (2020b: 96).

[13] Some pictures do not exhibit significant variations in reflectance. If you are reading this paper on your computer, the figures it contains are examples. Being composed of patches of achromatic light, they represent what they do in virtue of the illumination they emit. (This was true of the images in Adelson's experiments (Adelson 1993: 2042). Both the assumption and Kulvicki's compositional semantics should be framed disjunctively, so as to allow for these pictures, and not just printed images in which the key syntactic feature is reflectance. In what follows I ignore this complication.

[14] For ease of presentation, I'll use 'A' and 'B' to name either patches, areas or squares, as appropriate. Since each patch corresponds to just one area and just one square, provided the relevant term is supplied, there should be no confusion.

it, it is hard to see. Indeed, the match may well be impossible to see. For now consider again figure 2. The ladder bars make manifest to us the match in reflectance between the A- and B-depicting patches. If anything can help us to see the match in luminance which that reflectance match determines, this should. But I, at least, find the luminance match remains elusive. Indeed, I find my attempts to see it simply spread the rot. If I try to see area B as sharing the luminance of area A, by concentrating on the match in reflectance between the patches that depict them, using the bridge provided by the match in reflectance captured by the 'ladder', the result is not the match I seek, but that the ladder around patch B is lightened by the way I see area B. Rather than seeing the luminance that is purportedly depicted, I begin to misperceive the reflectance of the relevant parts of the picture.

So what? What do these facts about picture perception tell us about picture content? It's a tempting thought about many (non-mental) representations that their contents must be recoverable by suitable users. It's a tempting thought about pictures that their contents are recovered in our perceptions of them. (Any Viewer Response Account of the nature of depiction is likely to find this last thought attractive, but one need not be a theorist of that stripe to defend it.) Together the two tempting thoughts suggest the

Recovery Principle: pictures depict only what suitable viewers can see in them.

But if so, and if we are suitable viewers of figure 1 (and why should we not be?), and if we can't see in it the match in depicted luminance that the assumption entails, the assumption is false.

We might be tempted to argue that the luminances must match on the grounds of claims about how things would be were the depicted scene real. Figure 1 depicts square A as having a certain reflectance and illuminance, and square B as having another reflectance and another illuminance. In reality, any squares exhibiting those two reflectance-illuminance pairs would exhibit the same luminance. Doesn't this show that the match is depicted? It does not. The contents of pictures are not closed under entailment, any more than the content of other representations are.[15] Nor does it help to appeal to facts about luminance in the scene photographed. For one thing, figure 1 is not a photograph, but a computer-generated image. But even if it were a photograph, it hardly follows from the fact that a photo depicts a scene that is F that it depicts the scene as F. The camera can lie, or at least stay silent. No one would be photogenic, and no snaps of majestic mountains would disappoint, if photos were destined accurately to depict the features of their objects.

---

[15] In fact, Kulvicki does think that there are certain kinds of entailment under which pictures and other analog representations are closed. If a picture deploys a determinate property F to represent a determinate property G, then it also represents any determinables of G, in virtue of instantiating determinables of F (Kulvicki 2015b). The entailments above are quite different.

§6 If figure 1 doesn't depict areas A and B as matching in luminance, what does it depict? One option is that it does not depict luminance at all. That would render it a fatal counter-example to Kulvicki's compositional semantics, at least in its extension to achromatic optical pictures. That holds that luminance is the only feature predicated of places in the 2D array. If the picture doesn't depict luminance, then, according to the theory, it doesn't depict those places as bearing any feature at all. And since all other content is built on this base, the picture would lack any predicative content, not only at the bare bones level, but across the board. That is too much to swallow.

Kulvicki does better to accept that the picture does depict luminance, and to find some other way to respond to the threat thus posed. It is axiomatic in the compositional semantics that [1] picture reflectance determines depicted luminance. [2] Patches A and B in figure 1 match in reflectance. Yet [3] the areas A and B in the 2D array those patches depict do not match in luminance. Can these claims be reconciled, either by revision or clarification? If not, can Kulvicki reject [3]? Below I consider various strategies, some for reconciliation, others for rejection. The costs of these strategies vary, and not all are obviously unsustainable. However, every cost offers some motivation to adopt an alternative approach, if available. In §7 I show how the notion of design offers precisely that.

A. Blur semantic values.
Perhaps reflectance syntax does determine luminance content, only the content it determines is less precise than assumed. Square A and square B are depicted as differing in luminance. The parts of the picture that depict them match in reflectance. Reflectance may still determine luminance, but only by determining some luminance value that areas A and B are depicted as sharing: the determinable of luminance the two squares have in common. That common determinable is the limit of the picture's bare bones content. Ex hypothesi it is not the limit of the picture's luminance content *tout court*. But that just emphasises a feature of Kulvicki's view: his semantics determines only bare bones content. What determines the rest is another matter. The matching determinable luminances attributed to the areas are refined to distinct determinate luminances attributed to the squares that figure in the picture's full content. [3] is thus false.

Taking this line would, I think, amount to a significant deviation from Kuvicki's stated position. Both in offering his compositional semantics and in his other discussions of bare bones content he assumes that the distinction between bare bones content and the rest is a matter of which properties are represented, not how determinate a given represented property is. Moreover, the thinner bare bones content becomes, the harder will be the work of explaining how everything else is grounded in it.[16]

---

[16] For more of Kulvicki's treatment of bare bones content, see his 2006: ch.6 and 2020a: ch.2.

B. Subjectivise syntax.

If the A- and B-depicting parts of the picture match in reflectance but differ in the luminances they depict, luminance content cannot be determined by mere reflectance. But while the two patches match in objective reflectance, it is another question whether they look to do so. Indeed, part of the moral of figure 1 is that we don't see the two picture parts as the same shade of grey - that the two match in that respect is the first shocking revelation of figure 2, whatever other revelations might (or might not) follow for depicted luminance. So Kulvicki should simply subjectivise the syntax: what determines depicted luminance is not actual but apparent reflectance (what is sometimes called lightness (Adelson 2000: 342)). To this claim figure 1 is no counter-example at all. So revised, [1] no longer combines with [2] to entail the negation of [3].

This strategy in effect aligns the checkershadow picture with cases such as figure 3. To see the dog in this picture one must see the outline that divides the dog-depicting part from the rest. But that outline is not present objectively - it is a 'subjective contour'. Again then, the features key to determining content are not objective, but merely apparent features of the surface.

Now, figure 3 is perhaps best treated as another achromatic optical picture - rather than, say, as a line drawing. If so, the subjective contour is not part of its basic syntax, the syntax that Kulvicki has offered as determining bare bones content. Thus the two cases are not precisely on a par. The key content-determining feature of the dalmatian picture is not part of the syntax Kulvicki has described, and we should be wary of assuming it can be treated as syntax, in the sense required by a compositional semantics, at all.

Even so, the comparison illuminates a danger in the current strategy. When we see the dog in figure 3, we see the subjective contour. But it is very implausible that we can see that contour without seeing the dog (Lopes 2005: 41-2). The subjectivised 'syntax' can be grasped only by someone who understands the picture's content. And this raises at least a prima facie obstacle to treating the contour as syntax, in the sense required by a compositional semantics. For how can the meaning of the whole be understood as arising from that of its parts, if a grasp of the meaning of the parts requires a grasp of the meaning of the whole?

The worry is that similar issues confront the current treatment of figure 1. Why do the patches depicting area A and area B look so different in reflectance? No doubt part of the answer can be given by appeal to the features of surrounding parts of the picture. But it seems unlikely that the context provided by other properties of the vehicle does all the work. (For evidence, see Adelson 1993; Adelson & Pentland 1996.) Insofar as the disparity in apparent reflectance is down to features of the depicted scene, such as the fact that one square is depicted as in shadow, the other as in the light, the issue just raised reappears. The syntactic feature, apparent reflectance, does determine semantics, luminance, but only in the context of other depicted content (illuminance). (And note the

14

relevant context is not even confined to other bare bones content.) The promise of a clean route to a semantic whole from the purely syntactic character of the parts has to be abandoned.

Of course, often in interpreting a given part of language, such as a word, we must appeal to the context in which it occurs. That context is often semantic: I know which reading of an ambiguous term to apply because I grasp what the rest of the sentence means. Since that hardly shows that a compositional semantics for language is not available, why think any parallel conclusion threatens here?

The point is well taken. However, a sceptical conclusion survives it. As it stands, Kulvicki's compositional semantics is seriously incomplete. We cannot generate the bare bones content of figure 1 entirely out of any purely syntactic elements and the ways in which they are combined, let alone out of those purely syntactic elements Kulvicki identifies. Appeal to aspects of the wider semantic context will, it seems, be unavoidable.

C. Complicate syntax.

The role of semantic context in determining content has more devastating implications for another strategy. That is, not to subjectivise syntax, but to complicate it. If the A- and B-depicting parts of the picture match in reflectance but differ in the luminances they depict, luminance content cannot be determined by mere reflectance. But perhaps the lesson is just that we need to enrich our sense of what the relevant syntactic features are. Perhaps they are perfectly objective but include more than the reflectance of one picture part alone. For instance, perhaps what is key is something like reflectance relative to the surrounding context. This offers another way to revise [1] to avoid entailing the negation of [3].

This strategy really is ruled out by the considerations just given. If what determines the luminance depicted by the two parts of figure 1 is not merely the context of other vehicle properties, but the semantic context, then we cannot complicate objective syntax sufficiently unless we incorporate into its specification claims about the picture's content. That is plainly incoherent.

D. Reject Recovery.

Rather than tinkering with his compositional semantics, perhaps Kulvicki should instead challenge [3] by rejecting its theoretical underpinning, the Recovery Principle. The tenor of his whole account may seem to encourage such a move. His bare bones content is certainly distinctive. In general, we do not dwell on the 2D array of reflectances depicted in black and white photographs and other achromatic optical pictures, any more than we do on the 2D array of colluminations in coloured ones. If Kulvicki is right that such content is the building block on which other depictive content sits, then our interest is almost always confined to the latter. The bare bones content can certainly be overlooked, and *prima facie* need not be seen at all by many who engage with these images. Yet, seen or not, according to the Principle it must at least be visible. Why accept this?

One reason lies in Kulvicki's own theory of how the rest of pictorial content is generated. His idea is that a picture depicts whatever 3D scene meets various conditions (Kulvicki 2020: 29). One condition is that suitable viewers must have the capacity to recognize that scene, when encountered face-to-face. And Kulvicki makes clear that this recognition must be on the basis of the presence of the relevant 2D array. Standing before the scene, and presented with the array, they would recognize their environment as involving the former, in virtue of being presented with the latter.

It is very hard to see how the rest of pictorial content could be subject to this constraint unless the depicted 2D array plays some role in subjects' grasping that content. The condition requires that, face-to-face with the 3D scene, they recognize it because they are sensitive to the presence of the 2D array. How so, unless it's similarly the case that, when interpreting the picture, they grasp the depicted 3D scene because they are sensitive to the bare bones content depicting that 2D array? (See also Kulvicki 2006: 173.) But then what form does this sensitivity take? Since it sustains visually recognizing the scene (either as really present, or as depicted), it is overwhelmingly natural to assume it consists in *seeing* the array. Kulvicki's own account of how bare bones content grounds the rest appears to entail that bare bones content is subject to the Recovery Principle.

The point is not merely *ad hominem*. To make any progress in developing a compositional semantics for pictures we must first restrict our attention to an elementary level of content, such as Kulvicki's 2D arrays of luminances or colluminations. Since the depictive content of pictures obviously extends far beyond that, we need an account of how our elementary layer and the rest relate. How are we to build that bridge, if not by appeal to something like the link through recognition that Kulvicki proposes?[17] At this point, if not earlier, the need to acknowledge the perceptual character of pictures is liable to bite. In Kulvicki's development, the bite then goes deep enough to impose the Recovery Principle even on the bare bones content that grounds the content we are currently trying to accommodate. The challenge to anyone who thinks this consequence a mere accident of his theory is to develop an alternative that avoids it.

E. Unconscious seeing.
Still, seeing comes in various forms. Perhaps we are visually sensitive to the relevant bare bones content (the matching luminances of areas A and B), but unconsciously. Perhaps the same is true of the syntax that determines that content, the matching reflectances of the relevant patches in figure 1. Objective reflectance of the picture parts determines depicted luminance, just as [1] claims. We see the matching reflectances, and the matching luminances they depict. It's just that none of this shows up in conscious experience. [3] is false, despite

---

[17] He is not the only one who attempts to solve the problem this way. Compare Hyman 2013.

our failure to see the match consciously, and despite the fact that a version of the Recovery Principle is true. Pictures do depict only what suitable viewers can see in them, only the seeing need not be conscious.

It is another question, however, whether revising the Principle in this way respects the considerations that motivate it. The Principle is driven by a desire to acknowledge that pictures are symbols we use to communicate, and that what they mean is constrained by the communicative uses to which we put them. To suggest that they may have meanings that never figure in our conscious engagement with them seems to open up a shadow realm, of meanings for our own symbolic tools, of which in principle we users might know nothing. For this suggestion not to seem extravagant, we need to be given good reason to postulate depictive contents that lie beyond the reach of our conscious engagement with their bearers. What difference does it make that figure 1 depicts the two squares as matching in luminance, if no viewer of the picture is ever consciously aware of that fact? Until we are told, strategy E seems ad hoc.

Kulvicki may view matters differently. The pressures discussed under strategy D force him to accept the Recovery Principle in some form or other. Now we see what form it must take to give him his best hope of meeting the condition thus imposed. All this is done in the name of preserving a theory that offers a compositional semantics of bare bones content, and at least the start of an account of how that content might provide the foundation for the rest. If some of the intuitive motivation for the Principle must be sacrificed to make all this work, is that not a price worth paying?

Perhaps. Whether it is depends on what alternative view of the terrain is on offer, and how its merits compare. In a moment, I outline that alternative. But first let's consider a final strategy.

F. Look for appropriate viewers

If we are not to revise the Recovery Principle, one strategy remains. This is to hypothesize that some viewers do consciously see what the rest of us cannot. Consider the training artists traditionally undergo, in which they learn to focus on elusive aspects of the visual world. As it is sometimes put, they must learn to see what is there, not what they believe to be. Just as they acquire an improved eye for visible figure and visual size, perhaps they also acquire an improved visual sensitivity to luminance and reflectance. Perhaps they learn, not only to see subtle differences in these values that elude the rest of us, but how to disentangle them from the other visual elements in which, in ordinary experience, they are submerged. Might such a person not look at figure 1 and see both the matching reflectances of patches A and B, and the matching luminances they purportedly depict? [18]

---

[18] I owe this line of thought, like strategy D, to anonymous referees.

If there are such viewers, perhaps that is enough to square Kulvicki's axiom with the Recovery Principle. For the principle requires only that, for every aspect of depicted content, there is some suitable viewer able to see it. Just as a picture can depict some tiny detail that only the sharpest eyes are able to make out, so perhaps it can depict the luminances as matching even if most of us cannot see them that way.[19]

However, resting the fate of Kulvicki's compositional semantics on the existence of these visual experts offers a hostage to empirical fortune. Perhaps sufficient training leads to these results, perhaps it does not. Moreover, even if there are such characters, they hardly help Kulvicki's wider theory. We saw, in discussing option D, that his account of how bare bones content sustains further depictive content appears to require that everyone able to grasp the latter must be visually sensitive to the former. If only a select few see the match in depicted luminance, where does that leave the rest of us? We too certainly make out the checkerboard, so we appear to be reaching the rest of depictive content by a different route.

§7 We've seen various ways to adapt Kulvicki's view, refining or altering its claims, or disputing the theoretical background, so as to negotiate the counter-example provided by the checkershadow picture. Perhaps one, or some combination of them, will work. However, given the difficulties confronting them, and the costs of each, it is perhaps worth exploring an alternative. That is to abandon the project of compositional semantics, and with it the notion of syntax that is the source of the troubles above, and to substitute appeal to an idea close to the heart of Viewer Response Accounts, the notion of design.

Design looks like a competitor to syntax. Prima facie, they do the same job. Both are visible features of the marks in virtue of which they possess the particular content they do.[20] But it is not very likely that the two will divide features of pictures in the same way. The project of compositional semantics, at least at its most ambitious, requires the relations between syntax and content to be systematic in ways that appeal to the notion of design does not. Compositional semantics runs most smoothly when, within a given representational system, each syntactic feature always contributes the same elements to content; and when that contribution is secured by the presence of that syntactic feature alone, unsupported by contextual factors. That is certainly true of Kulvicki's semantics for optical pictures, as he presents it. The notion of design, in contrast, allows for much more flexibility. If, as Viewer Response Accounts claim, the picture depicts O only in virtue of engaging some

---

[19] 'Perhaps' because the cases differ in one important respect. While most of us simply miss the detail, our experience conflicts with that we are supposing the artists to enjoy: they see matching luminance in patches A and B, we see difference. Where one viewer's experience is not merely less precise than another's, but in conflict with it, the question which viewer is 'suitable', the one to whose experience pictorial content should be keyed, becomes harder to settle.

[20] If the Kulvickian opts for strategy E, syntax need be visible only in the sense that the visual system is sensitive to it.

key response R, in a form specific to O; then no doubt there must be something about the picture that enables it do to that. There is, however, no reason to think that what does it on one occasion must do it on all others. Nor is there any reason to think that what a given feature contributes, in its role as design on one occasion, it must also contribute on another. If we seek to depict a given content, what design must our picture bear? Anything that triggers the O-specific response will do. What does that job in one context might be very different from what does it in some other; and the response a given feature triggers in one context may be very different from what it triggers elsewhere. Perhaps the very same feature that leads us to see plants in one part of a picture leads us to see shadow in another.[21] The contribution of any element of design depends on context in a holistic way - varying both with the rest of design and with other aspects depicted. (Compare Langer 1957: 95.)

Now, of course, compositional semantics does not have to run in the smoothest way possible in order to run at all. Some of the amendments to Kulvicki's system explored above amount precisely to attempts to run it across rougher territory. And it is also true that there will be system of some sort in the contribution of design to depicted scene. Whatever form the key response R is purported to take, its triggering will presumably depend on psychological processes, and those processes will display causal order, at least at some level of resolution. Together, these two points might lead us to wonder whether syntax and design must compete after all. Perhaps wherever there is design, with whatever systematicity it displays in the relations between vehicle features and content, that very system could be tapped to redescribe design as syntax, to plug it into a compositional semantics, albeit of a more complex kind.

Perhaps. Whether this is so depends in key part on exactly what the project of a compositional semantics for pictures is supposed to be. Is it sufficient to execute it to describe relations between picture-features and content that are computable? Or must those relations meet further conditions - for instance, that they be articulable in such a way as to yield rules for interpretation that a person could, in principle, consciously follow? The fact that design operates by exploiting psychological processing may reassure us that the weaker of these two conditions is met, but it hardly grounds any confidence with respect to the stronger.

I have neither the space nor the expertise to settle what the ambitions of a compositional semantics should properly be. I will rest content, therefore, with drawing two modest conclusions. One is that there is some reason to think that even the bare bones content of optical pictures will not yield to treatment by the most straightforward compositional semantics. And the other is that the notion of design, favoured by Viewer Response Accounts of depiction, offers to do some of the same work as the notion of syntax required by a

---

[21] For an example, see Smith 2013: 309 (and figure 4). Smith's paper argues for points somewhat similar to those in this section, though from a rather different theoretical background. For an account of depiction that puts at its core the flexibility here illustrated, see Abell 2009.

compositional semantics; that the two may be genuine competitors, and that, if they are, some of the challenges to syntax do not arise for design.

References

Catharine Abell. 2009. "Canny Resemblance", *Philosophical Review* 118 (2), pp.183-223.

Edward Adelson. 1993. "Perceptual Organization and the Perception of Brightness", *Science* Volume 262, pp.2042-2044.

----. 2000. "Lightness perception and lightness illusions" in The New Cognitive Neurosciences, 2nd ed., M. Gazzaniga, ed. Cambridge, MA: MIT Press, pp. 339-351.

E.H.Adelson & M.Pentland. 1996. "The Perception of Shading and Reflectance" in *Perception as Bayesian Inference*, edited by David C. Knill, and Whitman Richards, Cambridge University Press, 1996, pp.409-423.

Justin Broackes. 1992. "The Autonomy of Color", in David Charles and Kathleen Lennon (eds.), *Reduction, Explanation and Realism*, Oxford: Clarendon Press, pp. 421–66.

Malcolm Budd. 1993. "The Look of a Picture" in Virtue and Taste ed.s D.Knowles & J.Skorupski Oxford: Blackwell, pp.154-175.

Tyler Burge. 2010. The Origins of Objectivity, Oxford: Oxford University Press.

Ernst Gombrich. 1960. Art and Illusion, Pantheon.

Nelson Goodman. 1968. Languages of Art, Oxford: Oxford University Press.

Dominic Gregory. 2020. "Pictures, Propositions and Predicates", American Philosophical Quarterly 57 (2), pp.155-170.

Robert Hopkins. 1998. Picture, Image and Experience. Cambridge: Cambridge University Press.

John Hyman. 2013. "Depiction" Royal Institute of Philosophy Supplementary Volume 71, pp.129-150.

John Kulvicki. 2006. On Images: their structure and content, Oxford: Oxford University Press.

2015a. "Pictures, Maps and Predication" Ergo 2:7, pp.149-174.

2015b. "Analog Representation and the Parts Principle", Review of Philosophy and Psychology 6:1, pp.165-180.

—— 2020a. Modeling the Meanings of Pictures, Oxford: Oxford University Press.

—— 2020b. "Colour and the arts: Chromatic perspectives", Routledge Companion to the Philosophy of Color ed.s D. Brown and F. MacPherson, London: Routledge, pp.91 – 106

Susanne K. Langer. 1957. *Philosophy in a New Key*, 3rd edition. Harvard: Harvard University Press.

Dominic Lopes. 1996. Understanding Pictures, Oxford: Oxford University Press.

—— 2005. Sight and Sensibility, Oxford: Oxford University Press.

Michael Newall. 2009. "Pictorial Experience and Seeing", British Journal of Aesthetics 49:2, pp.129-141.

Christopher Peacocke. 1987. "Depiction", Philosophical Review 96, pp.384-410.

Flint Schier. 1986. Deeper into Pictures, Cambridge: Cambridge University Press.

Paul Smith. 2013. "'Between a thing and a thought': syntax, twofoldness,and the wisdom of the Ancients", Word and Image 29:3, pp.304-333.

Kendall Walton. 1990. Mimesis as Make-Believe, Cambridge, Mass.: Harvard University Press.

John Willats. 1997. Art and Representation: New Principles in the Analysis of Pictures, Princeton: Princeton University Press.

Richard Wollheim. 1987. Painting as an Art, Thames & Hudson.