Philosophical Psychology
Vol. 18, No. 2, April 2005, pp. 187–203

Routledge
Taylor & Francis Group

The Twofold Orientational Structure of Perception

John Dilworth

I argue that perceptual content involves representations both of aspects of objects, and of objects themselves, whether at the level of conscious perception, or of low-level perceptual processing—a double content structure. I present an 'orientational' theory of the relations of the two kinds of perceptual content, which can accommodate both the general semantic possibility of perceptual misrepresentation, and also species of it involving characteristic perceptual confusions of aspectual and intrinsic content. The resulting theoretical structure is argued to be a broadly methodological or logical one, rather than a substantive theory that is open to empirical refutation.

1. Introduction

Several recent writers on perception have emphasized that our perceptions are not simply an objective recording of the world as it actually is, but that instead the content and structure of perception reflects our own activities and purposes with respect to the world (for an overview, see Noë & Thompson, 2002). However, there is a fundamental pair of factors in the mismatch between the structure of perception, and of the world, that yet remains to be adequately investigated or understood. I shall argue that perceptual contents of any sensory kind have a *twofold structure*—involving aspects of objects versus objects themselves—that has no literal worldly counterpart, but which arguably is an unavoidable result of any successful cognitive processing of complex, real-world perceptual data.

To begin, perception clearly involves some kind of representation of the perceiver's environment, traditionally assumed to be of some single, unitary kind. But a neglected possibility is that there might be *more than one* fundamental and general way in which we perceptually represent objects and events. I shall argue that there

Correspondence to: John Dilworth, Department of Philosophy, Western Michigan University, Kalamazoo, MI, 49008 USA. Tel.: 269-343-4009; Email: dilworth@wmich.edu

ISSN 0951-5089 (print)/ISSN 1465-394X (online)/05/020187-17 © 2005 Taylor & Francis Group Ltd
DOI: 10.1080/09515080500169256

are in fact at least *two complementary* kinds of representation in perception—or, otherwise put, that there are at least two different kinds of *perceptual content*—and provide an 'orientational' theory concerning their relations.

As an initial example, consider a chair, which may be visually perceived from many different spatial orientations or angles as it is rotated in front of a stationary observer. On the face of it, there is not one but *two different* kinds of perceptual content present to an observer during such a survey: first, content concerning the various *aspects* of the chair, as seen during each different orientation of the chair relative to the observer, and second, content concerning the *chair itself*—since an experienced observer normally regards all such perspectival views of a chair as being views of the *same* chair in each case. Thus, the relevant visual contents of an observer will normally include both the many variable orientations of the relevant object (to be called *aspectual* content) and a single invariant content that represents the object itself (to be called *intrinsic* content). While at the same time, any other sensory inputs, such as tactile data, will also have their own characteristic kinds of aspectual and intrinsic content.

To be sure, this initial description of such a perceptual survey could theoretically be handled in several ways: a perceptual theorist is not *forced* to accept that there are two fundamentally different kinds of content in such a case. For example, aspectual or orientational perceptual content could be taken as primary, with intrinsic object perceptual content being reduced to it in some fashion (Tarr & Vuong, 2002). Nevertheless, I shall show that it is possible to construct a plausible theory in which the above *twofold or double content* (DC) construal of the perceptual situation is accepted at face value. Also, the concepts of spatial aspectual and intrinsic content will be generalized to cover any kinds of aspectual versus intrinsic object factors.

2. Observer versus Object Frames of Reference

It is common to conceptualize perceived objects and events as existing within *orientational frames of reference*, which are either observer or viewer centered—in which a rotating object, such as a chair, is successively seen as having distinct aspects, each perceived from a single observer point of view—or object centered, from which perspective the object features remain unchanged during such a rotation (Edelman & Bulthoff, 1992). This contrast of orientational perspectives or frameworks provides a useful initial rationale for there being two kinds of perceptual content: variable *aspectual* content is content based on an observer centered or observer relative framework, while invariant *intrinsic* content is content based on an object relative framework.

To be sure, frameworks of reference are merely analytical or mathematical tools, which make no difference to the concrete objects and events that make up the real world. However, it must not be forgotten that an individual's perceptual contents are the results of *representations of* that framework-independent world from *a particular observer point of view*, and hence the *contents* of such perceptual

representations will inevitably involve, in the first place, an *observer centered* framework of reference.

Thus, an observer of a rotating chair will initially—in the chain of perceptual processing—acquire information about the many distinctive, observer relative *aspects* of the chair. Further perceptual processing would be required to arrive at an appropriate *object* relative (intrinsic) frame of reference for the relevant object, since the initial perceptual data or content is instead structured as aspectual, observer relative content.

This last point is an important one, because it establishes a *prima facie* relationship between aspectual and intrinsic content: the aspectual content *represents* the intrinsic content, in that the initial aspectual content provides only a transformed version of the object related content, rather than its directly providing the actual object related content itself. Or in formal terms, an aspectual reference frame requires an inverse mathematical projective transformation in order for it to be mapped onto an object relative reference frame, which projective relation establishes a precise sense in which the aspectual content represents the intrinsic content (for a more detailed account, see section 10).

Thus, my hypothesis is that perceptual processing involves both aspectual content and intrinsic content, with the aspectual content providing a representation of the intrinsic content, from which aspectual representation the intrinsic content is perceptually computed, presumably at some initial or lower level stage of processing that is prior to conscious perceptual awareness, but with both aspectual and intrinsic content elements remaining as part of a perceiver's resultant conscious perceptual experience.

3. Misrepresentation and Duality Constraints

So far we have some potentially useful initial features for a double content (DC) theory of perceptual content, but more constraints on an adequate DC perceptual theory are needed to cope with two significant issues.

3.1. The Possibility of Perceptual Mis*representation of the World*

To begin, since perceptual representation of any object having any perceptible qualities must be possible, such as a red book, the relevant perceptual content *X* must count as content that *represents* the relevant red book (or, more precisely, the content *X* explains why a person's *perceptual system* represents the red book). Yet, at the same time, it must also be possible for a person's perceptual system to *misrepresent* the red book as a green book, or as a red brick, and so on. Or in other words, there is a semantic dimension to perceptual content, which may be correct or incorrect with respect to the worldly objects and events that it represents (see Fodor, 1992). However, we have not yet taken into account the previous conclusion that there are *two* kinds of content to be considered, both aspectual and intrinsic, and not just one. A particular orientational perspective on a red book gives rise in the first place to *aspectual* red-book-related content—involving, in intuitive terms,

how the relevant object looks from a certain spatial perspective A′, or equivalently, involving a perceived aspect A of the object.

Now, it might be thought that misrepresentation should also be theoretically recognized at this initial *aspectual* stage, for it might seem to an observer that she is seeing aspect A of an object, when in fact she is seeing some other aspect B of an object. And this could take place prior to, and independently of, a procedure that determines the intrinsic content that is represented by the aspectual content, and which content, such as red book content, could also be correct or incorrect in its own intrinsic way. Thus, one possible kind of DC theory of perceptual content would allow for *two* levels or kinds of misrepresentation, or of correctness versus incorrectness of content—a *double misrepresentation* constraint on an adequate theory.

However, though such an approach is theoretically possible, arguably it is unnecessarily complex, for the following reasons. First, semantic structures should not be imputed to perceptual contents independently of the broad epistemic *functions* that perceptual contents serve for their users. In general terms, the object or function of perception is usually to extract useful information about the *intrinsic* features of objects and events, etc., in one's environment, and issues of misrepresentation therefore arise primarily with respect to such intrinsic contents, rather than with respect to their aspects. Second, the main function of having two different kinds or levels of perceptual content is to enable an organism to reliably *identify* those intrinsic features of objects, *independently* of the specific aspectual conditions under which those intrinsic features are found. Aspectual content is content that *makes allowances* for various potentially discordant or interfering contextual conditions, such as colored light that may distort the apparent color of an object. A perceiver whose aspectual content embodies correct *assumptions* about the prevailing aspectual conditions will thereby be able to correctly identify the intrinsic color of an object, whereas aspectual content embodying incorrect assumptions would typically lead to incorrect identification of intrinsic color.

Nevertheless, though the aspectual assumptions might help to explain *why* a perceiver was able to correctly or incorrectly identify intrinsic content, they are not themselves the focus of her practical perceptual activities of intrinsic content identification. Or in other words, contextual or aspectual conditions have a specific function to play in perception, *as* aspectual conditions of objects. To be sure, those same conditions, such as the color of the light impinging on an object, could, in distinct acts of perception, themselves be perceived as intrinsic features in their own right, such as when one examines a light source directly for its own *intrinsic* color value. But under those conditions the ambient light color is no longer perceptually functioning as *aspectual* color with respect to some other object.[1]

Consequently, it is theoretically simpler to postulate only a single, intrinsic kind of correct or incorrect representation in DC perceptual cases,[2] with issues about the implied or assumed correctness of aspectual conditions being handled as issues about the intrinsic content of *appropriately related* perceptual cases, in which aspectual perceptual conditions are examined as intrinsic object features

in their own right. Thus, the *misrepresentation constraint* on an adequate DC theory involves only a single, intrinsic kind of misrepresentation.

3.2. Duality Constraints

Another issue requiring one or more constraints on an adequate DC perceptual theory is that it must be able to explain the possibility of perceptual *ambiguity* and *interpretation.* For if a single stream of perceptual data is to be resolved into two different kinds and levels of content, one would expect that, in the absence of any explicit labeling or typing in the data input stream, in some cases there would be *more than one* way in which to split up the data between the two kinds of content.

For example, an apparently wedge-shaped object in the visual field might either actually be wedge-shaped, and viewed in a frontal view, or instead it might be rectangular, but viewed so that it is skewed in space, with one side being closer to the observer than the opposite side. Thus, the non-rectangular visual appearance of the object could be interpreted either as a fact about *intrinsic* content—the actual shape of the object—or as a fact about the relevant *aspectual* content, namely its spatial orientation relative to the observer. In an actual case, at most one of these interpretations could result in a correct intrinsic content.

Another ambiguity example would make use of a broader sense of the observer relative content of perception, that will be introduced in due course, and which will cover any kinds of observational conditions, such as viewing objects through red-tinted glasses. A person perceiving a scene through glasses, but who did not know whether they were tinted, would be in an *ambiguous* perceptual situation: the perceived colors of objects might either be their *intrinsic* colors, as seen through clear glasses, or they might be *aspectual* colors due to tinted lenses, which therefore would not show the intrinsic colors of the relevant objects. And again, at most one of these interpretations could be correct in an actual case.[3] Thus, in sum, there is also a *duality* constraint—of aspectual versus intrinsic factors—on an adequate DC theory of perception.

As to how the duality constraint relates to the misrepresentation constraint, a white object seen as being red would involve misrepresentation of its intrinsic color, perhaps because the perceiver wrongly assumes that there was no distinctive aspectual color, when in fact there was an aspectual redness that was caused by red-tinted glasses. On the other hand, a similar visual appearance might instead involve a case where the object really is red, but it is perceptually misrepresented as being intrinsically colorless, because the perceiver had assumed that the red color was due to the aspectual effects of red glasses, when in fact the relevant glasses had clear lenses, and hence added no distinctive aspectual color. In order to satisfy both the misrepresentation and duality constraints, it will be necessary to further develop the initial two frame (observer versus object frames) explanation of aspectual and intrinsic content, along the following lines.

4. Orientational Foundations

The following orientational concepts will be useful during the development of the theory. The concepts will be introduced in a form appropriate for discussing spatial examples, and then they will be generalized so as to apply to any possible perceptual situation. There will both be similarities and differences from the initial, simpler observer versus object frame of reference approach to defining aspectual and intrinsic content.

The concept of a *transformation field* (hereafter: a field) is a central one. In the spatial case, a field *F* is some three-dimensional environmental or contextual spatial structure, relative to which the spatial orientation of an object may be specified, so that an object has a *field orientation* relative to a field *F* in which the object is spatially contained. In addition, both fields and objects also have an *intrinsic orientation*, which consists in their having one or more characteristic or distinguished sides or values. For example, a spatial field has an intrinsic orientation because it has a unique top, bottom, front and rear—relative to an observer reference frame—while a chair also has an intrinsic orientation for the same reasons, except that its values are relative to an object rather than observer frame of reference.

Part of the importance of the concept of *intrinsic orientation* as applied to objects is that it provides a concept of certain features of an object that remain *invariant* through orientational *change*. For example, the top, bottom, front and back of a chair retain their identities as such, no matter to which three-dimensional *field orientation* the chair may be rotated, relative to its spatial field. The concept of intrinsic orientation is also important because it allows the definition of a concept of *uprightness*: an object *O* having an intrinsic orientation is *upright* just in case its field orientation in field *F* is such that the top of *O* is aligned with the top of *F* (plus that *O*'s front is aligned with the front of field *F*).

Now, so far there might seem to have been no significant departure from the original observer versus object frame of reference view. To be sure, the concepts of a *field*, of *field orientation*, of *intrinsic orientation* and of *uprightness* are new, but these concepts might be viewed as being in some way implicit in the 'two frames' approach—a field is like a frame, the field orientation of an object is the way in which its intrinsic reference frame relates to an observer frame, intrinsic orientation is closely allied with an object frame, and an object is upright when its observer and object frames of reference have the same values. Indeed, these similarities and putative identifications could be conceded for actual physical situations involving issues specifically of *spatial orientation*. But the new concepts will show their strength in two ways, first when generalized to apply to any kinds of aspectual versus intrinsic factors, and second in the analysis of *perceptual content*.

5. Generalizing Aspectual and Intrinsic Frameworks

The contrast of an aspectual, observer relative *spatial* framework versus an intrinsic, object relative spatial framework is only one species of a family of such distinctions. The most general contrast is between contextual or environmental

aspects of an object—whether or not they are defined in an observer relative way—versus *intrinsic* features of an object.

The relevant general concepts of an *aspect* of an object *Z*, and of an *intrinsic* feature of that object, could be initially explained as follows. First, the distinction is not an ontological one, but instead it pertains to the *informational structure* of low-level perceptual data. An aspect is a low level, compound structure of information, which requires perceptual or conceptual processing to extract its components. Thus, the *aspects* of an object *Z* are those perceptual factors such that the information they provide is the primary perceptual source for information about *Z* and its features, while the *intrinsic* features of *Z* are those that may be appropriately inferred or computed from those aspects.

For example, in the case of the perceived *color* of an object *Z*, the relevant aspect would normally consist of compound data giving information both about the color *D* involved in the ambient lighting of object *Z*, and of the color *C* of *Z* itself. Since the color aspect of *Z* contains information both about *Z* and about its ambient lighting, further processing is required to compute the estimated *intrinsic* color *C* of *Z*. Presumably this further processing would also produce, or assume, an estimate of the likely color involved in the ambient lighting also, though the isolation of such a contextual factor would usually be only of subsidiary interest during perceptual processing of intrinsic object-related information.

Hence, strictly speaking, colored ambient lighting is not *itself* an aspect of the color of an object *Z* that is thus lit, but instead it is a relevant contextual or aspectual *factor*, information about which is combined with intrinsic color information to make up a low level perceptual aspect of *Z*. And similarly with other perspectival or contextual kinds of factors for an object *Z*: their aspectival role depends on their informational contribution to relevant kinds of low level aspectual perceptual data, from which data the relevant kinds of intrinsic *Z*-related features may be computed or inferred. As for the current concept of an intrinsic feature of an object, this has no specific metaphysical implications concerning matters such as the necessity of possession of the feature by an object, or of the properties an object would have if it were not influenced by other objects,[4] or even of the objective status of the feature;[5] the term functions primarily as a relative term to distinguish features that may initially be assigned to an object itself, in cases of correct perception, from other cases where correct perception would rather require assigning a feature to some aspect-related contextual condition or factor.

6. The Double Content Theory Outlined

Given the previous preliminaries, a more complete form of the double content (DC) perceptual theory can now be outlined, using the concepts of a field, field orientation, intrinsic orientation and uprightness as presented in section 4—so that the theory could more fully be described as an *orientational* DC perceptual theory.

To begin, in any given actual situation involving one or more objects, each possible *intrinsic* content factor, such as shape or color, is paired with a relevant

aspectual factor that could affect the appearance of, or perceptual data concerning, the intrinsic content. Thus, intrinsic color may be paired with ambient color, and intrinsic frontal shape with perspectival spatial shape. Then in each case a single field *F* is defined, whose top element is the actual value of the relevant intrinsic content factor (such as the actual color of an object, or its actual shape), and whose other elements define other possible values of either the relevant intrinsic or aspectual factor, such as other colors or shapes. In the theory, appropriate elements of a field *F* may be coordinated either with actual aspectual or intrinsic factors, or with items in a low-level perceptual data stream that represent such actual or external factors.

A field may be regarded as a kind of *transformation* matrix or array, in which any non-top element *Y* is a *transformed version* of the top element *X*, and which hence may be used to *represent X*. As used with aspectual perceptual data, aspectual element *Y* of field *F* represents, or stands in for, element *X* in the low-level aspectual data stream. Or, in epistemic terms, the presence of item *Y* in a data stream indicates, or provides evidence for, the presence of an actual intrinsic factor *X* in the external environment, and hence it internally represents or indicates an intrinsic perceptual content factor *X*, which in its turn would externally represent the actual intrinsic factor *X*.

A field *F* is used in two complementary ways in the theory. First, it provides a standard of correctness for perceptual *intrinsic* content: one has correctly perceived a color, for example, if one's perceptual intrinsic color content is *upright* in the relevant color field, i.e. if it is aligned with, or has the same value as, the top element of the field; while *non-upright* (or 'inverted' in a broad sense) intrinsic color content would instead be incorrect, or a misrepresentation of the actual intrinsic color of the object, in that it would have some different field orientation with respect to the field *F*.

The second use of the given field *F* is to define the orientation of the *aspectual* perceptual content, with the orientational change from its upright position providing a metric for the degree of transformational change in the low-level perceptual data from an optimum purely intrinsic informational structure. In a case when the relevant aspectual factor has no effect on the low-level aspectual perceptual data, such as when the ambient lighting is neutral in color, or when a shape is in a frontal perspective, the aspectual content would be *upright* in the field, and hence perceptually 'transparent' or unnoticed by the viewer. Indeed, if all aspectual factors were always upright in their relevant fields, a single content theory of perception would be completely adequate, and there would be no need to introduce discussion of aspectual factors into a theory of perception.

However, the strength of a DC theory shows itself when the relevant aspectual content is *not* upright in its field. In the real world, the low-level perceptual data available with respect to a given object is typically *variable*, in ways that depend on the typically *non*-zero values of its aspectual factors. The DC theory models such non-zero factors in terms of aspectual content that has a *non*-upright, or broadly *inverted*, field orientation with respect to the relevant field *F*.

As a simple example, if one looks at the front of an upright book, and gradually rotates it to an inverted position, one's perceptual contents must simultaneously keep track both of its changing *aspectual* content, namely the varying, non-upright spatial orientation of the book in the relevant spatial field, and also of the unchanging *intrinsic* spatial orientation of the book itself—with respect to an identical, object-centered version of the relevant field. Or, in the case of the intrinsic color of an object, perceptual contents must be able to allow for changing aspectual lighting conditions in a similarly structured way. Thus, normally, perception is not simply of intrinsic object features, but instead of those features *as they are invariantly represented* by changing aspectual content features—or at least, invariant to the extent that perceivers are able to correctly represent the intrinsic features under changing conditions, since, for example, some extreme changes in lighting conditions could easily result in incorrect perception of intrinsic color, resulting in a non-upright intrinsic color content.

The theory as presented so far clearly satisfies the *misrepresentation* constraint of section 3, in that misrepresentation is explained as non-uprightness of intrinsic content in field *F*. In order to show that the *duality* constraint (similarly discussed in section 3) is also satisfied, some terminological abbreviations will be useful. Since a single field (or two structurally identical copies of the same field) is involved with both aspectual and intrinsic content, one may use the following simplified terminology. Using 'UA' and 'UI', 'respectively', for upright aspectual and intrinsic content, and 'NA' and 'NI', 'respectively', for non-upright aspectual and intrinsic content, the following characteristic cases need to be considered.

As a preliminary, actual aspectual factors and their objects may themselves be modeled in the terminology. They come in only two kinds, UA–UI or NA–UI. For any actual *intrinsic* factors are by definition upright in their corresponding relevant field (hence they are always UI), while their relevant *aspectual* factors may either be neutral or upright (a UA–UI case) or non-zero and hence non-upright (a NA–UI case).

As for cases of perceptual content, my claim is not that all cases exhibit duality phenomena, but only that some such cases may be found, which are best explained by a DC theory. There are three main cases of interest, to be presented using as an example an object Z whose intrinsic color is red, as perceived in various conditions of ambient lighting. First, a UA–UI case of perceptual content is one in which the intrinsic red color of object Z is both perceived correctly (a UI case), and also perceived to be in optimum, neutral lighting conditions (a UA case)—which joint UA–UI perception would normally be the result of the actual situation itself being a UA-UI case.

However, if the actual situation was instead one in which the lighting was *non*-optimal, such as one in which the red object Z was observed in a strong yellow light—an actual NA–UI case—a potential perceptual *ambiguity* or *duality* results. Viewers who were experimentally induced to concentrate on perceiving object Z itself, in isolation from its environment, might fail to notice contextual cues indicating that the ambient lighting is yellow, and so instead perceive the lighting

as being neutral (a UA case), and hence incorrectly perceive object Z's intrinsic color as being *orange*—an NI case, and hence a combined UA–NI perceptual case.[6] On the other hand, viewers who were induced to notice the contextual cues, giving evidence that the ambient lighting was yellow, would instead be able both to perceive the case as an NA case—since they perceive the non-optimal lighting conditions—and then presumably often be able to *compensate* for that condition in their perception of the intrinsic color of object Z—as being *red* rather than orange—resulting in a correct (UI) perception, so that this is a combined NA-UI perceptual case.

Thus, overall, these two possible UA–NI versus NA–UI interpretations of the same actual situation provide a *perceptual duality* case, of the sort that one would expect to occur, given that a single orange aspectual data stream has to be perceptually interpreted in terms of two distinct kinds of perceptual content.

Another duality example would be of an apparently wedge-shaped object in the visual field (as briefly discussed in section 3), which might either actually be wedge shaped, and viewed in a frontal view, or instead it might be rectangular, but viewed so that it is skewed in space, with one side being closer to the observer than the opposite side. Suppose that the latter interpretation is correct. Then a correct perception of it would be an NA–UI case—as a skewed or non-optimal view of a rectangle—while an incorrect perception of it would be a UA–NI case, as a non-skewed or optimal view of a wedge-shaped object.

To be sure, this orientational DC analysis of such duality examples does not provide a purely *phenomenological* account of a person's perceptual contents, in that perceivers normally take a perceived intrinsic content factor (such as the color orange, or a wedge shape) to be the *actual* intrinsic content of the object. So no one ever perceptually experiences their own perceptual contents as being a UA–NI case.

Nevertheless, this divergence is to be expected, given that an adequate analysis must account for the actual semantic correctness or incorrectness of intrinsic perceptual content, independent of a perceiver's own beliefs about their perceptual contents. Thus, phenomenologically perceived contents are logically similar to believed propositions—which are believed to be true, but which may nevertheless be false. Hence, just as an adequate theory of true propositional content must be independent of such propositional beliefs, so also must an adequate DC theory of correct *perceptual* content be independent of such perceptual beliefs—as is the case with the present DC theory.

7. More on Fields, and Aspectual versus Intrinsic Factors

Some additional considerations need to be discussed with respect to how a field relates to aspectual and intrinsic factors, and also to a plausible account of actual perceptual processing. The main point is that a given intrinsic factor, such as a particular shape, may be associated both with different aspectual factors and with different fields, in ways that depend on the specific nature of an actual perceptual situation, or on facts about the perceiver in question. For example, the apparent

intrinsic shape of an object may be altered by factors other than changes in spatial orientation—a straight stick will look bent if it is partially immersed in water, and a perceiver suffering from astigmatism or other eye defects may experience systematic aspectual shape distortions of her perceptual content. Each such kind of case might involve a distinctive aspectual factor, along with a resultant range of potential divergences from a given intrinsic shape, and hence it might require its own specialized field.

To be sure, one could postulate a single, multidimensional field for any given intrinsic factor, with an additional dimension for every possible aspectual factor that could distort that factor in characteristic ways, and in that way ensure that only a single field is required for each intrinsic factor, no matter what the perceptual situation may be. But actual perceptual processing is likely based on more selective and economical processing methods, such as a dedicated field whose top element is the relevant intrinsic shape, with its other elements being the characteristic alterations in aspectual shape that result purely from spatial orientational changes in a relevant object. Or, a perceptual situation recognized as involving the presence of water might in theory trigger the use of a specialized refractive distortion field, for perception of the characteristically changed shapes of objects that are partially immersed in the water. Nevertheless, in practice it seems likely that normal human perception of such relatively unusual and unimportant phenomena would not involve such a field, hence explaining why such cases are thought to involve perceptual *illusions*, i.e. perception of intrinsic factors that is usually incorrect, because there is no adequate aspectual compensation for the refractive phenomena.

In the terminology of section 6, these would be UA–NI (upright aspectual, non-upright intrinsic) cases, because the relevant aspectual refractive factors are perceptually unrecognized as such, and hence treated as being optimal or transparent, i.e. as being upright in what is in fact a degenerate field that lacks any compensating refractive elements—which also explains why the represented intrinsic content factor ends up being a *mis*representation (a NI case). However, it would not be hard to train oneself in such cases so that, for instance, one could come to perceptually recognize which sticks are actually (intrinsically) straight, on the basis of their characteristic aspectual bent appearance, which cases would involve at least an ad hoc, trained use of a refractive distortion field, so as to produce an NA–UI (non-upright aspectual, upright intrinsic) perceptual double content structure.

8. A DC Integration of Competing Psychological Perceptual Theories

There are two common psychological theories of how perceptual recognition of objects and their intrinsic features works, which theories are often considered to be in competition with each other. Initially they could be described as *viewpoint-independent* versus *viewpoint-dependent* theories (for an overview, see Biederman, 1995).

In outline form, on the viewpoint-independent approach, perceptual processing involves the extraction of invariant, object related intrinsic factors from an

environmental stream of 'noisy' perceptual data, whose other, non-intrinsic elements are *unrelated* to the desired intrinsic elements. This view could be dubbed the *irrelevant aspectual factors* (IAF) view of perceptual recognition, and it naturally leads to a view of perceptual content as involving only a single level of *intrinsic* content factors (as in the geon theory of Biederman, 1987).

On the other competing viewpoint-dependent approach, the primary task of object recognition is viewed as being one of recognizing *aspects* of objects, since, on this view, intrinsic features of objects are normally not directly available in low-level perceptual data streams, which instead consist primarily of aspectual data with respect to a given object. Thus, this view could be dubbed the *unavailable intrinsic factors* (UIF) view of perceptual recognition, and it naturally leads to a view of perceptual content as involving only a single level of *aspectual* content factors (see Tarr & Vuong, 2002).

From the perspective of the double content (DC) view, each of these competing theories is sorely in need of integration into a more structured theory that gives due place to both aspectual and intrinsic perceptual factors, as does the DC theory. Briefly, the DC view agrees with the first IAF view, to the extent that it agrees that the main purpose of perceptual recognition is to identify intrinsic factors. But on the DC view, this identification can usually only be achieved by making integral use of aspectual data, since on the DC view low-level perceptual content is primarily made up of such aspectual data items, which typically only *represent*, or are a *transformation* of, the desired intrinsic factors, so that those intrinsic factors are usually not directly available in the data stream.

This representational approach to intrinsic content also has a significant advantage over the IAF view, in that it need not require that there be some stored, well-defined intrinsic factors, precise matching with which is required for object intrinsic factor recognition. Instead, the task of object recognition is reconceptualized as being one of *field* recognition, i.e. of recognizing that a given aspect is an element in a given object-related field, whose elements *represent* the object, or intrinsic factors, in question. Thus, characteristic aspects of an object provide substitutes or stand-ins for intrinsic object features, which aspects may be recognized as being representations of those intrinsic features.

Thus, the DC theory agrees with the second UIF view, to the extent that it agrees that low-level perceptual data consist primarily of aspectual data. But it also, as just noted, embodies the insight that there is a systematic *transformation* or *relation* between a given kind of aspectual data and a given kind of intrinsic data, as captured in the concept of a field, and that a relation of representation exists between any actual aspectual element of the field and its hypothetical top element, which provides the desired intrinsic factor.

9. Double Content Improvements on the Two-Frame Approach

An initial rationale for a DC theory of perception was provided in section 2, in terms of the differing spatial perspectives provided by observer (aspectual content)

versus object (intrinsic content) frames of view. It was also pointed out there that the introduced additional concepts of a *field*, of *field orientation*, of *intrinsic orientation* and of *uprightness* could be viewed as being in some way implicit in, or a natural development of, that 'two frames' approach. Thus, on an initial view, a field is like a frame, the field orientation of an object is the way in which its intrinsic reference frame relates to an observer frame, intrinsic orientation is closely allied with an object frame, and an object is upright when its observer and object frames of reference have the same values.

However, the DC theory as it has now been developed goes beyond that initial two-frame conceptualization in several ways. First, the concept of a *field* generalizes the concept of a spatial frame so as to apply to any possible range of empirical or perceptual data. Second, the concept of a field, unlike the original concept of a frame, now has two complementary functions: for intrinsic contents, upright orientation in a field *F* indicates correctness; non-upright orientation indicates incorrectness. While for aspectual contents, upright orientation in field *F* indicates optimal or transparent aspectual factors, and non-upright orientation indicates non-optimal factors that would change the values in the low-level aspectual data stream from their optimal values for detection of intrinsic object features. Third, a field, unlike a frame, is not an arbitrary mathematical framework of coordinates, that could be applied in the same way to any objects or situations, but instead it is an ordered structure that has a top, or most salient element, that is used to indicate the actual value of some intrinsic feature of an object, such as its actual shape, color or position, and which field is such that its other elements define various degrees and kinds of *aspectual transformation* from its actual top value. Also, arguably the concept of a field is primarily a psychological—or psychophysical—rather than a purely physical or mathematical concept, in that the concept is geared to understanding how changing aspectual perceptual content relates to unchanging intrinsic perceptual content.

Nevertheless, these developments do retain a key feature in the initial two spatial frame view, namely that it is fruitful to regard intrinsic object factors in perception as having *the same structure* as aspectual factors, in that each can be modeled using the *same* field, with any given intrinsic factor being the result of a transformation from a given aspectual factor. As to how it is possible for a *single* field to be thus used, in spite of its two complementary aspectual/intrinsic functions, the fundamental reason is that any divergences from optimality with respect to intrinsic factors in an aspectual data stream initially only indicate *some* change in field orientation from an upright position, but without specifying whether that non-upright orientation is due to aspectual or intrinsic factors.

It is at this point that the *duality constraint*—that an adequate double content view must regard a single perceptual data stream as being potentially ambiguous, and only resolvable via alternative interpretations such as the dual NA–UI versus UA–NI forms—plays its part: cognitively only a *single* field structure *F* could be at work in such cases, because *any* element in a field *F* potentially could be interpreted as either an aspectual or an intrinsic content factor.

The resulting DC view also has much greater explanatory value than the initial two-frame view, in that, for instance, intrinsic perceptual mistakes or misrepresentations often could be explained as being due to related misjudgments about aspectual perceptual factors—which misjudgments need not be entirely or even partly perceptually based, since they could be based on incorrect information of other kinds. (This last point provides another reason not to analyze double content perceptual cases as involving two distinctive kinds of *perceptual* misrepresentation, which issue was discussed in section 3).

In addition, as pointed out in section 8, the DC theory can potentially incorporate the valid theoretical insights of both of the main mutually conflicting psychological perceptual theories—the viewpoint-independent or 'irrelevant aspectual factors' (IAF) approach, and the viewpoint-dependent or 'unavailable intrinsic factors' (UIF) approach.

10. The DC Structure of Perceptual Processing

Here is an overview of the primary salient factors that would be involved in veridical perceptual processing, according to a DC perceptual theory. Consider a perception of an object X, with relevant aspectual factors Y. To begin, since the theory claims that raw perceptual data have to be resolved into two related content factors before genuine perception is achieved, a clear distinction must be made between that raw data Z, and its aspectual and intrinsic content.

The incoming raw data Z is the complex result of the influences both of the object X and of the aspectual factors Y on the perceptual organs, and more specifically, the aspectual or contextual factors Y may be considered to have performed a *transformation* on the X-related data available to perception, which modifies the data in some way. Thus, the aspectual content may appropriately be symbolized as $Y'(X')$—as the functional result of a transformation of X-related data X' by Y-related data Y'.

Thus, in representational terms, the representation *vehicle* Z—the raw data stream itself, i.e. the concrete excitation or stimulation events or states in the perceptual organs produced by incoming physical stimuli—must be distinguished both from its aspectual content $Y'(X')$ and from the intrinsic content X' that is in turn represented by that aspectual content $Y'(X')$, which content X' provides information about intrinsic X-related factors.

Next, a DC theory must account for how a perceptual system is able to handle the inevitable ambiguity or duality issues resulting from an attempt to extract two distinct content factors from a single data stream. Initially, the extraction might seem problematic, because the data by hypothesis involve, or require for their correct interpretation, *two* kinds of factors—both Y-related and X-related information—whereas the incoming perceptual data Z only provide a single complex sum of those factors.

What is needed to resolve the problem is some related data, whether perceptual or known on other grounds, which can be used to identify the relevant aspectual

factors Y, independently of the data stream Z. Once those aspectual factors Y are known or assumed, the data Z can be interpreted as having aspectual content $Y'(X')$ and intrinsic content X': the effects of the aspectual factors Y in producing data Z can be subtracted from data Z, so as to calculate the represented data X', which calculation will also enable the relevant aspectual content $Y'(X')$ to be identified as such.

For example, a perceiver who knows that it is evening, or that the sun has set, may be able in this manner to compensate for the aspectual lower brightness and intensity in the apparent colors of objects, so as to correctly perceive their true colors.

To be sure, if all knowledge is gained through perception, there must be some optimal perceptual situations in which aspectual factors play no significant part, so that intrinsic factors in such cases can be correctly perceived without any need of independent knowledge of aspectual factors. Thus, initial learning of object-related information by organisms could be explained by the fact that under optimal conditions, only trivial inverse transformations (see section 2) are needed to calculate such object-related information from the initial aspectual perceptual information. But a DC theory is still needed in general, because not all perceptual cases are optimal, and hence presumably general-purpose perceptual mechanisms would maintain the DC structure even in optimal cases.

11. A Basic Justification of the DC Theory

First, here is a *G.E. Moore*-style epistemic argument for perception having a double content, at least at the higher, conscious level. If I see my hand, I know that it is *my hand* that I am seeing, i.e. that at least part of my current perceptual content includes my hand. But at the same time, as well as seeing my hand, I also see it in a particular perspective or aspect, that depends on my particular viewpoint, and hence I also know that I concurrently see a particular *aspect* of my hand. Hence, I know that the contents of my perception include both the relevant aspect of my hand, and my hand itself. But similar points would apply to any object that I see: no object can be seen without seeing a particular aspect of it, and similarly no aspect of an object can be seen without seeing the relevant object. Hence, perception in general must have a double content structure.

Next, here is a somewhat broader functional argument to the same effect. In general perceivers are capable of identifying and re-identifying an object X under a wide range of aspectual conditions. Hence, such a person's X-related perceptual content X' must in some way remain *the same* through such aspectual changes, since the object is in each case identified as being object X. However, at the same time P's perceptual contents must also adequately reflect the relevant aspectual changes as object X is rotated, moved, differently lighted and so on. Thus, P must be able to perceive different *aspects* $A_1, A_2, \ldots, A_n$ of object X, while nevertheless continuing to perceive them as aspects of *the same* object X.

But a single-level account of perceptual content would not be able to adequately account for this complex structure of re-identifiable perceptual contents, hence

showing the need for a two-level view. A person's perceptual content with respect to object *X* is neither simply object *X* itself—which would mean that the different perceptual aspects of *X* were not adequately perceived by her—nor simply the different *aspects* of *X*, which would result in a lack of perception of them as all being aspects of the *same* object *X*. Instead, in some way *P* must perceive *both* the aspects, *and* the fact that they are aspects of the same object *X*, and hence her perceptual contents must in some way include both components during her perception of any particular aspect of *X*.

The above arguments for perception having a high-level (conscious) double content structure do not depend on any particular scientific results about perception. Arguably a similar point applies to the hypothesized low level or pre-conscious processing as well. Since any low-level perceptual data impinging on sense organs will inevitably be a mixture of aspectual and object-related factors, low-level perceptual processing *must* have mechanisms for separating out aspectual from object-related data, on pain of failure to identify intrinsic object features, and hence it also must involve appropriate kinds of low-level double content structure.

Also, the basic point that aspectual content *represents* intrinsic content follows, as argued in section 2, from the fact that the initial aspectual content provides only a transformed version of the object-related content, rather than its directly providing the actual object-related content itself.

The additional contribution of this paper to the basic, broadly logical points summarized above is to be found in its specific orientational double content (DC) theory, as developed in the body of the paper. But this theory too is probably best regarded as providing no more than a minimum logical, mathematical and semantic framework of concepts, within which the basic structure of the relations of aspectual and intrinsic contents may be investigated. Thus, the DC theory is not a theory that could be confirmed or disconfirmed by any empirical investigations, but instead it provides a methodological framework, within which empirical results about the relations of aspectual and intrinsic contents, plus their relations to their real-world counterparts, may be articulated.

Thus, it would be a misunderstanding of the present paper if it were thought to be incomplete without more specific connections to actual scientific perceptual experiments and theories—or to standard philosophical problems about perception either. Of course the potential implications of both kinds are of great interest, and they will be pursued elsewhere, but the DC theory presented here is a very basic one—in articulating what are arguably necessary logical features of *any* adequate theory of perception—so that it deserves to be judged on its own merits, independently of such broader application issues.[7]

Acknowledgment

My thanks to an anonymous referee for the journal for very helpful comments.

Notes

[1] Clearly there is much more to be said about the resulting duality of function of aspectual versus intrinsic contents, but another paper will be required to further investigate the issue, since it would be too distracting currently.

[2] Such an approach should also be congenial to eliminativists or instrumentalists concerning the reality of perceptual content, since it does not require that there be some fact of the matter as to whether aspectual content is correct or incorrect, independently of its functional connection with the determination of correct intrinsic perceptual content.

[3] To be sure, in actual cases involving colored glasses, the resultant colors seen would presumably be due partly to both aspectual and intrinsic factors, so the example is a simplified one.

[4] However, this simple definition of the term has been disputed recently, as a result of a paper by Langton and Lewis (1998).

[5] For example, some colors could still be intrinsic features on this account, even if they are no more than dispositional properties to cause appropriate color qualia in humans.

[6] Presumably such an incorrect perception could be explained by the fact that the result of the yellow light being reflected from a red surface would usually be an orange aspectual data stream. There is a large psychological literature on such contextual effects, though for reasons of space, connections with it will have to be explored elsewhere.

[7] I develop a related thesis for artworks in my book (2005).

References

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147.

Biederman, I. (1995). Visual object recognition. In S. F. Kosslyn & D. N. Osherson (Eds.), *An invitation to cognitive science: Vol. 2: Visual cognition* (pp. 121–165). Cambridge, MA: MIT Press.

Dilworth, J. (2005). *The double content of art*. New York: Prometheus Books.

Edelman, S., & Bulthoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Research*, *32*, 2385–2400.

Fodor, J. (1992). *A theory of content.* Cambridge, MA: MIT Press.

Langton, R., & Lewis, D. (1998). Defining 'intrinsic'. *Philosophy and Phenomenological Research*, *58*, 333–345.

Noë, A., & Thompson, E. (Eds.). (2002). *Vision and mind*. Cambridge, MA: MIT Press.

Tarr, M. J., & Vuong, Q. C. (2002). Visual object recognition. In H. Pashler & S. Yantis (Eds.), *Stevens' handbook of experimental psychology: Vol. 1: Sensation and perception* (pp. 287–314). New York: John Wiley & Sons.