Illusion in the commonplace: reinterpreting Ernst Gombrich's concept of illusion

Jonathan Auyer

University at Albany, State University of New York, jonathanauyer@hotmail.com

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Aesthetics Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive. Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
ILLUSION IN THE COMMONPLACE:
REINTERPRETING ERNST GOMBRICH’S
CONCEPT OF ILLUSION

by

Jonathan P. Auyer

A Dissertation
Submitted to the University at Albany, State University of New York
in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy

College of Arts & Sciences
Department of Philosophy
2013
Illusion in the Commonplace:

Reinterpreting Ernst Gombrich’s

Concept of Illusion

by

Jonathan P. Auyer

Copyright 2013
Abstract

In the dissertation I analyze and interpret Ernst Gombrich’s book *Art and Illusion*, focusing on his view that illusion is involved in pictorial representation. Since Gombrich never gave a concise, systematic account of illusion, my goal will be to fill this void by using the text of *Art and Illusion* as well as Gombrich’s subsequent writings in order to present a coherent account of how illusion might play a role in a picture’s representing an object.

My goal is not to present an unassailable account of pictorial representation. Instead, I offer a version of Gombrich’s theory that pushes readers towards a better comprehension of what a Gombrichian theory of illusion involves. In the process I introduce and defend a number of terms and concepts in the service of filling in those places where Gombrich is silent. Among other things, in response to Gombrich’s notion of visual substitution I elaborate upon the claims that representational pictures function as relational models and afford recognition of the objects they represent; I reply to Richard Wollheim's "twofoldness" objection to Gombrich; and I contend that Gombrich's use of the notion of illusion is not open to the objections commonly made against it (e.g., that normal picture perception does not involve illusion because “illusion” is synonymous with “delusion”).
Acknowledgements

The pages of this dissertation hold the culmination of years of research and study, but they also document the discovery of a *trouvaille*—a chance encounter that has resulted in the dedication of a significant portion of my life towards something I did not even realize I was passionate about. The author Norton Juster once wrote that “the most important reason for going from one place to another is to see what's in between”. A great many kind and passionate people have helped and guided me as I set about seeing “what’s in between” (which might better be described as my graduate studies). Some of these people I have known my entire life; others I did not meet until quite recently. These written thanks do not do justice to the contribution each has given to me:

To my parents. Their love, compassion, and willingness to allow me to explore and find the truth for myself have afforded me the ability to develop the skills needed to endure years of tedious research and investigation. I owe them not only my work ethic but also my all-encompassing love of all things “learning.” Thank you for putting the locks on the outside of the doors.

To professors David Basinger and Andrew Koehl, both of whom sowed the seeds that have blossomed into a true love of philosophy, as well as the desire to question and to inquire deeper and deeper into the world around me.

To Richard Woodfield the lengthy correspondence with whom proved to be integral in ever so slightly dissecting the tangled web that is Gombrich’s thoughts and ideas.
To my committee members Ron McClamrock and Jason D’Cruz, who have both
dedicated significant time and energy to read and comment on my writings throughout
busy semesters, bouts of sickness, and the birth of a new baby. I am tremendously
grateful for the hours each of them gave in service to the completion of my degree.

To my advisor Robert Howell, in whose class it was that I stumbled upon the
trouvaille that was Ernst Gombrich’s book, *Art and Illusion*. Navigating the dissertation
process requires a patient and meticulous guide, and Professor Howell was more than
equal to that task. My continual, steady progress could not have been achieved without
his advice and the countless hours he committed to this project. I was, and continue to be,
honored to have him as my advisor, and I cannot express enough thanks for everything
that he done for me.

And finally, to my wife Katie. She has been my confidante, my partner, my editor,
my sounding-board, my supporter, and often-times the great optimist throughout this
journey. Too many car rides consisted of my ramblings; too many coffee dates involved
talk of minute details and esoteric problems. Perhaps now I can finally stop talking and
reading about illusions (for a little while anyway), and try to become absorbed in
something a bit lighter. If the most important reason for going someplace is to see what’s
in between, then there is no one whom I would rather see the in-between with.

To those mentioned here, and to those I may have forgotten, I can say, with the
utmost thanks:

“Hats off!”
Table of Contents

Abstract ................................................................................................................. iii

Acknowledgements .............................................................................................. iv

Contents .............................................................................................................. vi

Introduction: Why revisit *Art and Illusion*? ......................................................... 1

1 A Working Interpretation of Gombrich’s Account of Depiction ......................... 3
   1.1 The goal of Chapter 1 ...................................................................................... 3
   1.2 Two styles of depiction .................................................................................. 4
   1.3 Illusion and relevant similarity ....................................................................... 7
   1.4 Illusion exists on a continuum ....................................................................... 14
   1.5 The core necessary condition for depiction ................................................ 16
   1.6 Gombrich’s “theory” of perception ............................................................... 21
   1.7 Perception and the schema .......................................................................... 22
   1.8 Artists create equivalences ............................................................................ 24
   1.9 Images are always incomplete ........................................................................ 27
   1.10 Substitutes, schemata, and substitution ...................................................... 29
   1.11 The mental set .............................................................................................. 32
   1.12 The artist and the beholder .......................................................................... 34
   Recapitulation and further details ...................................................................... 35
   1.13 A change in the function of pictures .......................................................... 35
   1.14 The beholder, the mental set, and the act of projection ............................... 39
   1.15 Completing the incomplete ......................................................................... 41
   1.16 Gombrich and twofoldness ......................................................................... 42
   1.17 Revisiting illusion ......................................................................................... 45
   1.18 Conclusion .................................................................................................. 47

2 Visual Substitution and Relevant Similarity: The Relational Model .................. 50
   2.1 Presenting relevant similarity in two parts ................................................. 50
   2.2 The layout of Chapter 2 ................................................................................ 50
   2.3 Relevant similarity is more than what is given via a descriptive linguistic tool... 52
   2.3.1 Why “as of”? ............................................................................................ 53
   2.4 Relevant similarity: neither iconic or mimetic ............................................. 56
   2.5 The respects of similarity ............................................................................. 62
   2.5.1 Which respects? Ernst Gombrich, Karl Bühler and the notion of relationstreue . . . 71
   2.6 Final thoughts on relevant similarity as a relational model ....................... 79

3 Visual Substitution and Relevant Similarity: Affording Recognition ............... 82
   3.1 The goal of Chapter 3 ................................................................................... 82
   3.2 The importance of recognition ....................................................................... 84
   3.2.1 Recognizing objects and recognizing their pictures ............................. 86
3.2.2 Other theorists who advocate similarity between ordinary and pictorial perception......................87
3.2.3 More thoughts on recognizing objects and their pictures.................................91
3.2.4 Looking at pictures at-a-glance, piece-by-piece, or both?.................98
3.2.5 Recognition and the role of perceptual hypotheses..........................102
3.2.6 Recognitional duality.........................................................104
3.3 Perceptual constancy: an empirical hypothesis.................................108
3.3.1 Perceptual constancy in ordinary perception.......................................109
3.3.2 Constancy in pictures does not require identical visual information........111
3.3.3 Non-arbitrariness and the correspondence between picture and object.....120
3.3.4 Final thoughts on perceptual constancy and recognition..................123
3.4 Concluding relevant similarity..................................................126

4 Gombrich and the Concept of Twofoldness .......................................130
4.1 The goal of Chapter 4..............................................................130
4.2 Wollheim and Twofoldness.........................................................131
4.3 Why twofoldness?..............................................................134
4.3.1 Wollheim and a single conception of twofoldness..........................136
4.4 Gombrich and Twofoldness: A case of switching interpretations...........139
4.4.1 Symmetrical switching: switching between two meaningful interpretations....140
4.4.2 Can we really switch between two meaningful interpretations?..............143
4.4.3 Asymmetrical switching: switching between a meaningful and a meaningless interpretation........................................146
4.5 A way forward..............................................................149
4.5.1 The distinction between “being aware” and “attending to”.............150
4.5.2 Some Clarifications..............................................................153
4.5.3 Applications for Gombrich........................................................155
4.6 Gombrich and twofoldness: final thoughts.................................160

5 The Gombrichian Concept of Illusion..............................................162
5.1 The goal of Chapter 5..............................................................162
5.2 What illusion is not: distinguishing illusion from delusion...............165
5.2.1 Why there remains confusion over Gombrich’s position on illusion.........165
5.2.2 An illusion of what? The uniqueness of the commonplace...............169
5.2.3 The possibility of “epistemically benign” illusions.........................173
5.2.4 Illusion, trompe l’oeil, and medium awareness............................178
5.2.5 Perceiving a picture’s “true information”....................................187
5.2.6 Concluding thoughts on the claim that “illusion is delusion”..........190
5.3 Affordances, illusion, and visual substitution.................................193
5.3.1 A mediated phantom: the emergence of a substitute.........................193
5.3.2 A substitute and its affordances..................................................195
5.3.3 Pictures as visual substitutes......................................................200
5.3.4 Pictorial affordances..............................................................202
5.3.5 The visual substitute as a non-delusory artifact..............................203
5.3.6 A substitute for what? Visual substitution and the test of consistency....208
5.4 Final thoughts..................................................................211
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Conclusion</td>
<td>215</td>
</tr>
<tr>
<td>6.1</td>
<td>Where we have come from: Gombrich’s basic theory of pictorial</td>
<td>215</td>
</tr>
<tr>
<td></td>
<td>representation</td>
<td></td>
</tr>
<tr>
<td>6.2</td>
<td>Core necessary conditions revisited</td>
<td>222</td>
</tr>
<tr>
<td>6.3</td>
<td>The end is the beginning</td>
<td>225</td>
</tr>
</tbody>
</table>

References ........................................................................... 227

Notes....................................................................................... 235
The best books in the world are not the ones that give all the answers. These are the books you read, agree with right away or are convinced by, and then that is that, and you do not have any use for them any more. The best books in the world are those that are full of provocative points of view and loose ends, rich examples and inconclusive discussions, and maybe even inconsequences and confusions, for to these books you must return again and again, and their very inconclusiveness makes them take on new aspects every time you return to them with your own new background, and they will always be generous in giving stuff for new considerations and new ideas.¹

—Soren Kjorp

Introduction: Why revisit Art and Illusion?

I have set before myself a difficult task: recapitulating a text that is both influential in its impact on numerous and varied areas of study, as well as widely (although I will argue not conclusively) renounced regarding the thesis its author espoused. The thesis, put bluntly, is this: pictorial representation (i.e., depiction) involves an illusion—the illusion of seeing the subject that is depicted in the picture but which is not really there in front of the viewer. This kind of illusion is common; it is experienced everyday and I take it to constitute the basic thesis of Art and Illusion (Gombrich, 2000), the work that Ernst Gombrich is most often remembered for.² As Michael Podro put it:

“[Art and Illusion] has remained, for 40 years, central to the discussion of the visual arts by philosophers, art historians, and critics. It retained this position despite radical criticisms of parts of [Gombrich’s] argument because at its core it focused, as no art historian before had done, on the role of illusion, on the fact that in depiction, without our being deluded, we are caught up by the represented subject that we recognise within it - the expression of a face, the gesture of a figure, the spaces of a landscape.” (Podro, 2001)

Art and Illusion has been discussed, interpreted and re-interpreted, analyzed and critiqued, and it might be questioned what I hope to gain by returning to this much worked-over text. That Gombrich’s account of representation was never fully accepted is due not only to the controversial nature of the thesis, but also to misunderstandings and
misreadings on the part of the readers coupled with the author’s own vagueness in presenting his position.

My goal is to put forward as plausible an interpretation of Gombrich’s theory of representation as possible, and then to defend that interpretation against objections. My primary focus will be the position Gombrich presents in *Art and Illusion*, and so it will not be my task to address every criticism facing Gombrich’s whole cannon of work. Instead, I will seek to point out those issues that will allow me to assess the strengths and weaknesses of my interpretation of that work. It could be argued (and I am sympathetic to the claim) that the theory found in *Art and Illusion* is the same, or at least very similar to the one Gombrich defended throughout his lifetime. *Art and Illusion* may be a self-contained whole whose interest and influence justify investigating it on its own terms; yet, to bolster the case for my interpretation, I will often cite material written both prior to and after *Art and Illusion*.³ It is my belief that Gombrich consistently stood behind the majority of claims he put forward in *Art and Illusion*—using them, developing them, but never completely rejecting them. With that in mind, let us take to battle against those ghosts that have haunted Gombrich’s theory in the hope that we might also tie-up any loose ends Gombrich left for us.
1 A working interpretation of Gombrich’s account of depiction

1.1 The goal of Chapter 1

When we look at a painting such as Degas’ *Self portrait with soft hat* we often say (when questioned about what it is we are looking at) something like, “I see a man in a hat.” By this we are not only describing how the painting looks, we are also saying, by implication, that the picture depicts something, that it is *of* something—in this case that it is of a man wearing a hat (more specifically, the painter Edward Degas). But how is it that pictures are capable of this remarkable feat? How is a flat, marked surface able to be seen as being of some three-dimensional reality? Furthermore, how are we able to recognize a person or place or state of affairs in the forms, lines and colors of a picture? In short, what is involved in depiction?

In this first chapter I will offer a working interpretation of Ernst Gombrich’s answer to this question as it is found in his book *Art and Illusion*. I will outline the main points of Gombrich’s account of depiction, discussing those that I take to be essential to a coherent reading of his position, and I will highlight various important issues concerning Gombrich’s account, some that I’ll return to (along with other problems) in later chapters.

The following points will be discussed in Chapter 1. These include: whether Gombrich’s account distinguishes between the depiction of particulars and the depiction of kinds (section 1.2); the nature of depictive illusion and its connection to the idea of relevant similarity (sections 1.3 and 1.17); the notion that “illusion,” as Gombrich understands it, exists on a continuum or spectrum (section 1.4); the claim that Gombrich’s account of depiction consists of a “core” necessary condition rather than a set...
of strict necessary and sufficient conditions (section 1.5); the roles of the schema and the mental set in ordinary and pictorial perception (sections 1.6, 1.7 and 1.11); the claim that the features of a picture are an equivalence, in two-dimensions, to the three-dimensional features of the actual object or scene (section 1.8); the fact that all images are incomplete and require supplementation by the beholder’s mental set (sections 1.9 and 1.15); the claim that all images, and by extension all visual substitutes, are created with the use of a schema (section 1.10); the existence of a collaborative act between the artist and the beholder (sections 1.12 and 1.14); the idea that the history of art witnessed a gradual change in the function of pictures (section 1.13); and finally, the issue of twofoldness as it relates to Gombrich’s account (section 16).

I plan on returning to many of these topics in later chapters to shed a more critical light on the role they play in Gombrich’s account. For the time being I hope that the reader will find the presence of the working interpretation offered in this chapter to be somewhat useful when encountering the problems and concerns explicitly discussed in the later chapters.

1.2 Two styles of depiction

Although Gombrich is not explicit, Art and Illusion is concerned with explaining how it is that a picture P depicts an object O.5 To be clear, depiction can be of at least two kinds: (1) a picture can depict a kind of thing (a picture of an apple or a man), or (2) a picture can depict a particular thing (this apple or that man). So is Gombrich concerned with (1) or (2), or both? Many of his central examples involve (2)-style depiction: for instance, his extended account of Constable’s landscape painting Wivenhoe Park6, which
is not of just any park but of the particular Wivenhoe park in Essex; there is also
Cézanne’s *Mont Sainte-Victorie*, Inness’ *The Lackawanna Valley* and the portrait of Mrs. John Taylor by Gainsborough. He also, however, makes use of numerous examples of (1)-style depiction. Some of these include Daumier’s *Advice to a Young Artist*, Manet’s *At the Races*, and Fantin-Latour’s *Still Life.* So it seems that Gombrich must mean for his theory to somehow deal with (2) as well as (1).

A problem facing this conclusion is that nowhere in his work does one find an explicit account of how a picture represents something particular—how do we know that a picture P is of A rather then B? It does not seem as if resemblance alone is enough for the viewer to draw this conclusion. For instance, I (along with a many of the readers) have never been to Essex, and have therefore never seen the park that is there; so it seems that I (along with those who have never seen the park) cannot know that the painting *Wivenhoe Park* is of the particular part in Essex just by looking at the image. Yet Gombrich makes no specific suggestions about how to handle this problem in any other way. As it stands, *Art and Illusion* is simply unclear about this point.

As far as I can see, however, there is no reason to think that Gombrich’s theory cannot be interpreted or supplemented so as to account for (2)-style as well as (1)-style depiction. There are a number of things one might appeal to to make this supplementation: for instance, we might point to the intentions of the painter or some sort of causal relation that holds between P and some object A rather than P and some object B. Or one might argue that P is of A and not B on the grounds that not only is P in some way relevantly similar to A (perhaps the visual features in P are in some way like the visual features in A), but that there is also something about our shared, social knowledge
of the background or facts about \( P \) itself, something that various individuals can agree upon that makes it such that \( P \) is of \( A \) and not \( B \).

At this point I am not advocating any specific approach, nor am I claiming that any such way of treating (2)-style cases within Gombrich’s theory will work. But it seems that if some appeal along these lines will work for other theories of depiction, then there is nothing in Gombrich’s positive theory that will prevent him from using such an appeal in his own theory, too. I will therefore assume, for the present, that Gombrich can ultimately provide some way of dealing with (2)-style as well as with (1)-style depiction, and I will feel free to make use of his own discussion of paintings such as *Wivenhoe Park* without worrying unduly about how Gombrich can explain their (2)-style as well as their (1)-style depiction.

That being said, the reader will no doubt come to see that I do, despite my plea in the previous paragraph, tip my hat towards one of the approaches mentioned above: specifically, I think that Gombrich integrates some understanding of relevant similarity into his account. He does this by connecting relevant similarity with his idea of visual substitution, which I will discuss in section 1.5 of this chapter. In discussing how relevant similarity accompanies the idea of illusion (as he understands it), I will claim that Gombrich places it (i.e., relevant similarity) at the very heart of his account in *Art and Illusion*. In so doing he employs the idea of relevant similarity (coupled with the idea of visual substitution) into the development of what I will call, below, the core necessary condition for depiction. Furthermore, according to Gombrich, it appears as though the core necessary condition is satisfied as much by (2)-style as (1)-style paintings; and as far as I can see, nothing in his use of *Wivenhoe Park* and other (2)-style examples, used to
justify his illusionistic account of (1), really involves, in any crucial way, the fact that *Wivenhoe Park* is in fact a (2)-style picture of a particular park and not just a (1)-style picture of some park. For the present I will focus just on (1) and assume, as I mentioned above, that Gombrich can also deal with (2).

### 1.3 Relevant similarity

How does Gombrich understand depiction? The short answer is this:

> when we see a picture P, which is a depiction of some object O, P creates for us *the illusion* of seeing something that is not there, namely, O.

This answer certainly does Gombrich a disservice, since the question we really want answered concerns the nature of illusion. Does “illusion” mean that we are tricked or deceived into thinking that we are seeing real objects when we see pictures of them? No — this is not what Gombrich means at all, though this is certainly what many have attributed to him (something I will discuss further in Chapter 5). When he says that P creates for the viewer the illusion of seeing O he appeals to the idea that P functions for the viewer as a visual substitute; and for P to function as a visual substitute means that P is *relevantly similar*, in some way not yet specified, to O. The artist has found a way to arrange the flat marked surface of P to function in such a way that it suggests O by triggering off a similar recognitional response. Furthermore, for P to function in this way means that P is fulfilling what I will call the core necessary condition for something’s being a picture (which I will talk about in section 1.5).

One question to immediately ask is: in what way is seeing a picture “relevantly similar to” seeing the object that the picture is of? Can we really say that a picture, a flat two-dimensional artifact, is similar in any way to a three-dimensional object? There are a
few different ways that Gombrich might understand relevant similarity in this context; or, to put it slightly differently, there are a few possible sources of relevant similarity. The first source locates relevant similarity in the viewer’s visual experience of the picture and of the depicted object; the second involves the features or elements of the picture and the features or elements of the depicted object; a third way to understand the source of relevant similarity focuses on the visual recognitional abilities engaged or triggered off by the picture and the object.\textsuperscript{10}

In this chapter I will touch on the first reading of relevant similarity, that it is in terms of the beholder’s visual experiences of the picture and the object depicted in or by it. I will give a more detailed discussion of the other two sources of relevant similarity in Chapters 2 and 3, respectively, in the end arguing that each of the three sources plays a role in determining whether a picture is relevantly similar to an object. According to the first understanding of relevant similarity, when looking at a picture an individual has a visual experience that is similar, in a relevant way or ways, to the visual experience of looking at the object the picture is of—for instance, a picture of an apple, according to this understanding, is able to generate a relevantly similar visual experience to the visual experience of seeing an actual apple. But this raises a further question, namely, what counts as the two experiences being similar? After all, on the face of it, we could say that two things are similar just in case they both have a property in common. But this would make any two things alike in some respect because one can always find a property, however odd, that two things share (e.g., being looked at before 10 p.m., or being viewed by a person standing next to an oak tree, or being viewed by a person wearing the color blue). But properties such as these are arbitrary—they are not relevant to two things
being similar. If such properties were relevant, a picture of an apple could represent an
orange if the picture possessed any of these rather absurd properties: being viewed before
10 p.m., by a person standing next to an oak tree, while wearing the color blue.

The issue thus becomes this: in virtue of what common property of my experience
of seeing an actual apple (s-a) and my experience of seeing a picture of an apple (s-pa)
are the two experiences relevantly similar? That is, how are s-a and s-pa similar in such a
way that it is plausible to hold that someone that has s-pa is, in Gombrich’s sense, having
an illusion of s-a?

Let me clarify on a few points. Firstly, we very often use the same word or words
to describe both what we see when we see an actual object and what we see when we see
a picture of that object. For instance, we use the word “apple” to describe both our really
seeing an actual apple (“I see an apple”) and our seeing a picture of an apple (“I see a
picture of an apple”). More specifically, we very often use the same word or words to
describe both the visual experience of seeing the actual object and the visual experience
of seeing a picture of the object. Looking at a picture of an apple, when we are asked,
“what does the picture show?” or “what do you see in the picture”, we often say, “I see an
apple” or “I see an apple in the picture.” So it seems that we use the same word “apple”
in the course of describing the two visual experiences. What’s more, there is a way
Gombrich is able to say that we can describe the two visual experiences using the very
same whole sentence and the very same words “see” and “visual experience”—we are
able to describe the two experiences in similar visual terms. When looking at an apple
and then back at a picture of an apple, we can say, both times, “I’m having a visual
experience as of seeing an apple” or “I’m having an experience that in many ways is,
visually, like the visual experience of seeing an apple.” Gombrich can thus say that seeing an apple and seeing a picture of an apple are relevantly similar when we are able to describe both experiences in similar visual terms by saying both are “as of seeing an apple.”

So in this chapter the working interpretation of relevant similarity is in terms of the property of being expressible, linguistically, in such “as of” visual terms of a property that is common to both the visual experience of seeing an object and the visual experience of seeing a picture of that object. The focus on “as of” is merely to point out a common way we often describe our visual experiences of pictures and objects, since we are able to describe seeing a picture of an apple in similar visual terms to seeing an actual apple. That I take this to be an uncontroversial common occurrence only precipitates the need to return to the idea of relevant similarity in Chapters 2 and 3. There I will not only clarify how the linguistic expression “as of” fits into a more philosophically dense account of relevant similarity, but more importantly I will establish what exactly this more philosophically dense account of relevant similarity consists of.

As I mentioned above, Gombrich clearly wishes to distance his idea of illusion from the idea of delusion, that it is only in extreme cases that pictorial “illusion” should be considered pictorial “delusion,” and that only in such extreme cases are we actually fooled into taking the picture for the object. Our visual encounter with a picture of an apple triggers off the release of a visual experience that affords an “as of” description—“as of seeing an apple” or “as of seeing an apple in the painting.” While there may be respects in which the visual experience of the picture is relevantly similar to the visual experience of the object they certainly won’t be identical, and under “normal” viewing
conditions any number of factors will help keep us from confusing the painting with the object (e.g., the flatness of the page, the frame, the presence of the surrounding wall, the glare of the glass enclosure, and so on).

The kind of “triggering off” I have been describing is not exclusive to humans, and Gombrich was perceptive enough to point this out. For instance, not only will dummies of stickleback fish release a violent reaction in the male fish if the bottoms of the dummies have been painted red, but the fish “always postured in their aquarium when red mail trucks passed the window at some distance, for to their brains red stands for danger and rivalry.” (Gombrich, 2000, p. 102) There are a number of respects in which the dummy stickleback is quite different from the mail truck, but there is at least one respect in which the dummy and the truck are relevantly similar—color.\textsuperscript{11} The presence of a red mail truck triggers off a reaction that is similar to that of a red fish. This ethological fact is important to keep in mind because despite Gombrich’s insistence upon the role that conventions play in depiction, the role of biology plays an equally important role in his account of the creation and perception of pictures. Furthermore, much of \textit{Art and Illusion} involves Gombrich illuminating places in the history of art where artists sought to discover which respects were relevant for triggering certain responses in the audience, and where the artists discovered visual facts about how audiences could respond to the images the artists produced.

Our having a pictorial illusion, says Gombrich, amounts to our ability to describe our seeing of P in terms of “seeing O.” One reason we can do this is because our perception of the relationally arranged visual elements in P affords the possibility of seeing aspects of O (e.g., depth of space, perspective, color, texture, lighting, and so on)
—it affords us seeing P as of O. The relevantly similar visual experiences of P and O are triggered off by the relevantly similar sets of relationally arranged visual elements in P and O, leading to Gombrich’s claim that “the shapes add up to pictures of objects” (Gombrich, 1994, 155). When we look at Velázquez’s Las Meninas we are able to see the receding room; in Constable’s Wivenhoe Park we are able to see the cows grazing in the pastures, and see a manor house in the distance. For Gombrich, in the usual non-trompe l’oeil cases (of a picture of a room or a picture of a hillside) we are able to agree that we aren’t having a delusory experience of seeing an actual room or an actual hillside even though the two experiences are relevantly similar—we are simply able to describe our experience of seeing the picture in terms of seeing the object.

I want to end this section on relevant similarity by briefly elaborating upon the topic of visual elements. The notion of a “visual element” is one that I will utilize throughout the rest of the dissertation, so it would be wise to offer some classificatory remarks before continuing on. Offering such remarks is difficult due to the fact that Gombrich gives us very little detail concerning how exactly we should understand the nature of visual elements. Gombrich implicitly describes visual elements through the example of the three-tone modeling system. This is an effective artistic tool for creating perceived texture and depth of space through the use of a neutral tone and two modifications towards light and dark. Of this system Gombrich writes that it is the “relationships rather than the individual elements” that are significant (Gombrich, 2000, p. 49). Since it seems that we must infer what constitutes an individual element, I find it helpful if we begin by working backwards. Gombrich sees relationships as being significant for our visual system (and for pictorial representation); in order for a
relationship to exist there must be two or more things between which such a connection exists; according to Gombrich, it is the relationships between the light, neutral, and dark tones (in the above example) that prove to be significant; therefore, it stands to reason that it is the individual tones that count as the individual visual elements. In ordinary perception, it is not that we fail to see the individual visual elements—we can certainly point out the individual areas of light—but rather that individual visual elements become “set-off” as the intended object of our focus because of the way it relates to the non-intended background (the neutral and dark areas). (In this regard the idea is very much like the Gestalt notion of figure/ground.) The same is then said of the artist’s task of transposing light into paint: the brightly light area is significant because it is related to the dark and/or neutral areas around it.

We may very be able to expand the kinds of examples of visual elements to include things like lines segments, corner intersections, and possibly even outline shape. Of course, such examples are merely extrapolations as Gombrich offers no details about the nature of visual elements; my comments are thus not meant to be conclusive. Gombrich simply assumes that the reader will understand what he is talking about, especially given the types of non-technical visual examples that he offers. Additionally, I believe that his understanding of the significance of relationships (in the manner discussed) fits with some of the biological data concerning the way that certain cells in the eye respond to certain light/dark boundaries as opposed to individual points of light or dark. Again, there is no conclusive evidence to support these claims. Throughout the rest of this dissertation I will attempt to make clearer (often in an implicit way) what the idea of relationships between visual elements amounts to. As such, my goal is simply to
offer a reasonable interpretation of Gombrich’s position on the matter, and not to resolve all the issues that may surround it.

1.4 Illusion exists on a continuum

Readers of Art and Illusion will no doubt know that Gombrich applies “illusion” to a variety of cases in a number of different contexts. It is used to describe the experience of seeing a naturalistically rendered picture just as much as the experience of seeing a line drawing, and possibly also to describe the child’s experience of riding a hobbyhorse. Of course the toy and the picture function differently depending upon the context of action in which each exists (playful for the toy or perceptual for the picture); but, both the toy and picture will trigger off experiences (for the individuals using them) as of the respective object’s each represents—of riding a horse (in the case of the child) and of seeing some object (in the case of the viewer). In the former example, the experience of riding the hobbyhorse is thought to be relevantly similar to the experience of riding an actual horse (although not necessarily a particular horse); in the latter example, the visual experience of seeing the picture is thought to be relevantly similar to the visual experience of seeing the object depicted in the picture (although not necessarily a particular instance of the object).

A pertinent question to ask is how it is that illusion can be present across the spectrum of cases just mention? Is Gombrich talking about one type of illusion in the case of the painting and a different type of illusion in the case of the line drawing and the trompe l’œil (and then possibly another in the case of the hobbyhorse)? The hobbyhorse does not look all that much like a real horse; and while many pictures may
include meticulous detail and thus appear to resemble what each depicts, recognition of a picture’s subject need not depend on the existence of point-for-point similarities between the depiction and the depicted. Often we are able to recognize the subject of a painting from just a few simple, well-placed lines. If illusion is about creating a relevantly similar visual experience between picture and object, this can be achieved in a variety of ways, and it can be done using a variety of artifacts and a variety of mediums (which highlights the conventional elements in Gombrich’s account). As the artist creates, he or she engages in an ongoing sequence of trial-and-error experiments to determine how best to achieve the desired results, to determine which arrangements trigger off a relevantly similar experience to “an apple,” “a horse,” or “Jacque Cousteau.” As Gombrich put it: “The history of art…may be described as the forging of master keys for opening the mysterious locks of our senses to which nature herself originally held the key”, and what these “keys” are made out of will differ depending on the style, the medium, and the intentions of the artist (Gombrich, 2000, p. 359).

Perhaps what will count as an illusion, for the child or the viewer, will wholly depend on convention or culture, on our socially agreeing what should or should not count as an illusion. But for Gombrich this would credit the artist with too much power: the degree to which different arrangements of marks or forms do trigger off a certain experience or recognitional response in the viewer is, to a certain extent, biological and thus out of the artist’s hands. The artistic experiments that succeed will “fulfill certain demands on the organism. [They will be] keys which happen to fit into biological [italics added] locks.” (Gombrich, 1994, p. 4) The keys can be made of different material, but whether the lock opens will depend upon whether we turn the “right key”—on whether
the artist’s experiments hit on a correct arrangement of pictorial features. Furthermore, whether an image is more or less naturalistic will depend on the context of action in which it is created and used. There are times when a more naturalistic image may “contribute to its potency” but at other times the artist may use more schematic forms to achieve his or her intended response (Gombrich, 2000, p. 110). Regardless of the image’s degree of naturalism, once the arrangement of pictorial features “fit into our biological locks” the artist will have found “a correct arrangement,” and thus will have found how to create a certain visual experience. What the artist discovers, then, is how to create a substitute for something, and thus how to afford the possibility of a substitute experience. It does not matter whether the context of action is a child at play or a Dutch still-life painting; for Gombrich, artifacts ranging from toys to pictures are able to create illusions, if only in the sense they can serve as substitutes for other objects.

1.5 The core necessary condition for depiction

I want to use this section to address the platitude that in Art and Illusion Gombrich develops strict necessary and sufficient conditions for depiction. It is my contention that this platitude is false, and we should instead interpret Gombrich as offering what I will refer to as a “central” or “core” necessary condition for depiction. As I understand it, this core necessary condition involves Gombrich’s idea of being-a-visual-substitute, which is just to say that a picture can stand in, visually, for the object(s) it depicts. While I think that Gombrich’s account of visual substitution may extend to a wider range of artifacts, here I will only be talking about its application to pictures.
Let me begin by first investigating whether Gombrich offers strict sufficient conditions for depiction. Let us assume that some picture P is a picture of Winston (W), and that when we see P we have an experience of seeing a picture of Winston (s-pW). Now let me make the claim that P is a visual substitute for Winston. If what we mean by visual substitute is “sharing some property or set of properties,” then it seems that P must also be a visual substitute for Winston’s twin, let’s call him Edward (E). We can say this because the experience of seeing Winston (s-W) shares many similarities with the experience of seeing Edward (s-E). In fact, one might draw the further conclusion that the experience of s-W and s-E are more alike than the experience of s-W is like the experience of s-pW. If the experience of s-W shares more similarities with the experience of s-E than it does with the experience of s-pW, it seems that we cannot say that P is a depiction of W rather than E. Further, we must conclude that P’s being a visual substitute is not a sufficient condition for P’s being a depiction of W rather than E. Thus, we are left with the conclusion that being-a-visual-substitute is not a sufficient condition for depiction.

Perhaps, however, visual substitution might count as a strictly necessary condition for something’s counting as a depiction. It goes without saying that we often use one thing to represent another—a Coke bottle can be used to represent a person; some squiggly lines can be used to represent a mountain range on a map; an image of a lion can be used to represent courage. Can we describe these as examples of visual substitution? One might conclude that this claim, on the face of it, seems dubious, and so choose to argue that these are not examples of visual substitution and thus that the idea of being-a-visual-substitute can’t be a strictly necessary condition for depiction.
To this sort of objection Gombrich might point to the stickleback fish dummy example\textsuperscript{21}: the dummy is not, nor does it need to be, \textit{exactly like} an actual stickleback fish; rather, it need only be similar \textit{enough} (given the context) in order to trigger off a response similar to that of an actual stickleback fish. Analogously, the experience of seeing the Coke bottle is not exactly like the experience of seeing a person (it would certainly be ludicrous to claim that it is); rather, Gombrich could claim that the experience is only \textit{relevantly} similar—relevant, that is, \textit{given the context} in which the Coke bottle is being used.

One response to this claim might be to question whether Gombrich really thinks that the experience of seeing the Coke bottle is similar \textit{in any relevant way} to the experience of seeing a person. Isn’t the Coke bottle being used in a way akin to that of symbols on a map? Any arrangement of lines \textit{could} represent a range of mountains so long as we agree that it does. Of course, mapmakers have agreed to use a certain symbol (a pictogram) to represent a range of mountains, another to represent a river, and yet another for a rest-stop. So long as we are in agreement, we can use the Coke bottle to represent a person; however, it seems that in such a context we are \textit{not} using the Coke bottle to represent a person in the same way we would use a picture to represent a person.

It might still be the case, then, that the Coke bottle serves as a substitute for a person (in the same way that an idol serves as a substitute for a deity) because Gombrich understands the idea of substitution to exist on a continuum (something he talks about in \textit{Meditations on a Hobbyhorse}). At its simplest the substitute is an artifact that stands in for something else; it takes its place. I can use the Coke bottle to stand in for my friend Ed, but I could also use a rock that I thought had some Ed-like-features. In either case the
chosen artifact serves as a substitute: in the former case because it was readily available, in the latter because I happened to think it bore some minimal Ed-like features. The rock and the Coke bottle serve as physical substitutes, as stand-ins; but as we move along the substitute-continuum we can actually generate substitutes that bear greater visual similarity to the represented object. That is, a feature or set of features included in the substitute could be described, by a viewer, as relevantly similar to those of the represented object. It is these cases, those further along the substitute-continuum, that we should describe as *visual* substitutes.

Thus, a short answer to the above line of thought is that there is indeed a sense in which the Coke bottle represents my friend Ed, the squiggly lines represent a mountain range, and the image of a lion represents courage; but if we wish to count these examples as *depictions* (and not as representations in general), the idea of being-a-visual-substitute cannot be a necessary condition. On the face of it, however, these are not the kind of examples Gombrich focuses on in *Art and Illusion*. They might still be described as representations, but not as depictions; thus they won’t show that being a visual substitute is not necessary for depiction because they really are cases of something else entirely (let’s call them cases of non-depictive visual symbolism). This, however, still leaves open the problem of how to separate these cases from those genuine cases of depiction that concern Gombrich. How do we separate the Chinese ideogram for “dog” from a Thomas Eakins painting of a dog? (Or even from an abstract Giacometti sculpture of a dog?) As far as I can tell Gombrich does not provide a systematic way of doing this, nor do I think that Gombrich would try to do so; but for my purposes this is not of primary importance. Instead of offering a way to separate depiction from cases of symbolism Gombrich
implicitly focuses on the kinds of certain central or core cases of depiction I have mentioned. He picks out those examples he believes are uncontroversial examples of depiction—paintings like Constable’s *Wivenhoe Park*, Manet’s *A Day at the Races*, and Velázquez’s *Las Meninas*.

This way of proceeding, by interpreting Gombrich as offering a core necessary condition for depiction, has the following benefit, which can best be seen by returning to the discussion of Winston, his twin, Edward, and a picture of Winston. It *seems* true that the experience of s-W is more relevantly similar to the experience of s-E than the experience of s-W is relevantly similar to the experience of s-pW; and, it *seems* true that Edward is more of a visual substitute for Winston than a picture of Winston is a visual substitute for Winston. And if Gombrich claimed that relevant similarity or being-a-visual-substitute were, *by itself, sufficient for depiction* than he should conclude that Edward depicts Winston. Furthermore, Gombrich shouldn’t conclude that the picture of Winston depicts Winston (given the claims of sufficiency above). But of course he does claim that the picture of Winston depicts Winston, and he will reject the claim that Edward depicts Winston. As this is the case we can conclude that Gombrich is clearly *not* claiming that relevant similarity or being-a-visual-substitute is, *by itself, sufficient for depiction*, and thus he is not trying to give a strict sufficient condition for all cases of depiction.

Instead, and as I have previously mentioned, Gombrich is appealing to certain central cases of depiction and giving a condition *necessary for those central or core cases*. Thus, my way of interpreting the core necessary condition for depiction is that the relevant similarity or being-a-visual-substitute condition is necessary for, but limited to
all cases that Gombrich and others identify as central or core cases of depiction (we
might also think of these as “paradigm” cases). Thus I take the core necessary condition
to be a “core” condition for two reasons. First is that it holds in all such central cases of
depiction. If something is a central or core case of depiction (i.e., an example that most
would agree is a case of depiction) then it will always involve relevant similarity and
being-a-visual-substitute. Second, this condition is at the heart of Gombrich’s account of
depiction: we will have a better understanding of the facts of depiction if we attend to this
relevant similarity or being-a-visual-substitute necessary condition.

My interpretation of the core necessary condition for Gombrich’s account of
depiction certainly requires further discussion. In the rest of Chapter 1 I will try to fill in
the details of why and how the components of the core necessary condition hold; and in
later chapters, as I discuss specific problems and objections to Gombrich’s account, I will
discuss whether Gombrich is right in accepting the core necessary condition.

1.6 Gombrich’s “theory” of perception

In this section and the next I will attend briefly to Gombrich’s understanding of
perception. As Paul Richter (1976) rightly points out, Gombrich does not actually present
a fully developed theory of perception in *Art and Illusion*. We can, however, determine
Gombrich’s general understanding of perception. For him, perception depends upon a
kind of “the searchlight theory,” 23 which involves constantly probing and testing the
environment and then sorting and categorizing the information found there. The act of
perception involves the acquiring of information about the environment through the
senses, but the information we acquire is often incomplete—so what we perceive will
always require an interpretation. For instance, I cannot see the entire bottle sitting on the table in front of me, I can only see the part facing me—the backside, the bottom, and the inside are all hidden from my view. I am therefore forced to offer an interpretation of what I see based on the incomplete information available. One interpretation could be that the bottle lacks a back. But this is not likely. Barring the presence of a fake bottle that was placed in my room to deceive me, a better interpretation, and probably the correct one, is that I just can’t see the back or the bottom of the bottle. This is simply a fact about the limitations of human vision: we cannot perceive all of an object’s features at once. Of course, perception does not only rely on immediate visual evidence; it also relies on past knowledge and expectation. I know that the bottle (under normal conditions) does have a back because of facts I know about what I am capable of seeing, and thus I do not expect to see the back of the bottle when I look at it from a single, stationary viewpoint.

1.7 Perception and the schema

As we encounter the world around us we will always need an interpretation ready in advance of the visual data. Gombrich refers to this prepared interpretation as a “schema.” The evidence for this is drawn from certain biological facts. For instance, we don’t want to be attacked (perhaps by some wild animal) while we are working out how to read the ambiguous shadow moving towards us in the woods. To help us make sense of the ongoing stream of sense data we develop schemata that allow us to pigeonhole or categorize what we perceive; these will then act as a kind of guess or initial interpretation upon which we can base future decisions. The schema we form is not perfect and can sometimes go wrong, so we must be able to alter and adapt it as we encounter new visual
data that yields new information. Our inability to take in everything within our visual field means that we will instead focus on those elements that are deemed to be significant. And what counts as significant in perception is not, in general, the perception of individual visual elements, but rather *relationships* between elements (Gombrich, 2000).\(^{25}\) If it were otherwise, a “passing cloud over the sun…even a tilt of the head, or an approach from a different angle” would affect what we are seeing (Gombrich, 2000, p. 50). By focusing on the relationships between elements we are able to keep the world around us in “constancy,” even while the varying stream of sense data changes. The sets of visual elements we recognize in our environment will be those that fall within (i.e., are filtered or sorted by) the sets of categories that we are using (i.e., the schema); the schemata, according to Gombrich, is the set of categories we have developed as a means for quickly interpreting and classifying the diverse relations of visual elements we perceive in our environment.

Having just noted the importance Gombrich places on perceiving relationships in nature, we can now turn to the importance of relationships of visual elements in representation. The artist, Gombrich claims, will assign an interpretation to the relationships of pictorial elements of those features in the world that those relationships are meant to signify. Of course, the interpretation assigned to a specific arrangement of elements is not wholly determined by the artist; it also depends on how we actually see those arrangements of elements, that is, it also depends on the biological facts of which arrangements work to trigger off a response in the viewer. One example of this (which I have already mentioned) is the three-tone modeling system.\(^{26}\) The three tones give us a set of categories (which sort perceived sense data), and the relations between the tones


\(^{26}\) The three-tone modeling system is a technique used in painting to create the illusion of three-dimensional space on a two-dimensional surface.
are meant to signify relations of features in the world: white next to dark means intense light; white next to grey, not so intense, and so on. So, the contrasts of light, neutral, and dark in the tones of a drawing can be used to signify textural features in the world.

For Gombrich perception is not the passive process whereby the retina simply registers sense data like a photographic plate (Gombrich, 2000); instead it is an active ongoing process of probing and testing the environment in terms of the categories we are using. Perception involves the “inquiring mind” not the “innocent eye.” (Gombrich, 2000, p. 395) The very act of perception is conditioned by expectation and adapted to different situations. Understanding perception in this way means that we cannot hope for a childish perception of the world as flat stains of color and shape devoid of any meaning because all visual perception, ordinary and pictorial, is colored by interpretation.27 Put differently, all perception, according to Gombrich, will be influenced by what we bring to it.

1.8 Artists create equivalences

Earlier I said that when we look at a picture we can often times describe our visual experience as of some object (e.g. “my experience of P is as of seeing O” or “my experience of P is as of seeing O in P”). But what exactly does this mean? To get a better understanding of this we should remember that for Gombrich a picture serves as a visual equivalent. “Equivalent” here, is different from “identical” or “copy,” and to help clarify this difference we can look at Gombrich’s discussion of caricature. Citing Filippo Baldinucci, Gombrich writes that portrait caricatures “aim at the greatest resemblance of the whole of the person portrayed, [and] yet…they disproportionately increase and
emphasize the defects of the features they copy, so that the portrait as a whole appears to be [italics added] the sitter himself, while its components are changed [italics added].” (Gombrich, 2000, p. 344) So it is possible that we can have a global impression of P as of O, while at the same time noticing that parts of P (i.e., constitutive areas of the picture) are dissimilar to or possibly quite unlike those of O. For instance, we have a global impression of likeness in the stages of Charles Philipon’s Les Poires even while we perceive some of the individual elements of the picture as being very unlike those of King Louis Philippe. Gombrich goes on to say that,

All artistic discoveries are discoveries not of likeness but of equivalences [italics added] which enable us to see reality in terms of an image and an image in terms of reality. And this equivalence never rests on the likeness of elements so much as on the identity of responses to certain relationships [italics added]. We respond to a white blob on the black silhouette of a jug as if it were [italics added] a highlight; we respond to the pear with these crisscross lines as if it were Louis Philippe’s head. (Gombrich, 2000, p. 345)

This is exactly how Daniel Dennett (2002) describes his experience of looking at Bellotto’s painting Dresden from the Right Bank of the Elbe above the Augustus Bridge.

He writes how he “marveled at the gloriously rendered details of all the various people walking in bright sunlight,” but, upon closer examination,

I was astonished to find that all the little people were merely artfully positioned single blobs and daubs of paint—not a hand or hat or foot or shoulder to be discerned. Nothing shaped like a tiny person appears on the canvas, but there is no question that my brain represented those blobs as persons. Bellotto’s deft brushwork “suggests” people crossing the bridge, and my brain certainly took the “suggestion” to heart.28

An equivalence, then, is the fact (discovered by the artist) that certain dissimilar visual elements, when arranged in relation to each other, suggest a global impression of similarity. The relations of lines and blobs of paint in a picture can create the kind of as-
if-it-were-ness Gombrich talks about, such that we can see the whole picture as if it were the object; put differently, we can see the picture in terms of the object, and we are able to describe our visual experience of seeing the picture using the “as of” expression.

I think that we can now fit the idea of equivalence with the previously mentioned notion of triggering off: the arrangement of relations of two-dimensional visual elements in P, which give P its features, can lead the beholder to see P in terms of O (we can see P “as of…”, or say of P “this is a picture of…” because the relationally arranged visual elements in P are perceived as being visually equivalent to those in O. Artists experiment to find which relations of elements will trigger off which responses in the audience. As I have said before, there is a certain amount of convention in the artist’s choice of which relations of elements to include in a picture. The artist is certainly free to work in any medium he or she chooses, and the sketches, doodles, and under-paintings that lead to the completed work no doubt show the range of choice that the artist has in constructing the picture’s composition. But if I am trying to draw a picture of a hand, the relations of elements I choose to include in my drawing will be different than the relations of elements I would choose to include in a drawing of an apple. Certainly some individual elements might be similar (e.g., some intensities of light might happen to be the same, or certain lines or tones of color might be similar); and yet, Gombrich appears to be saying that the textures, lines, and shapes used to construct a picture (e.g., of a hand) make it such that the picture actually affords being read a certain way (e.g., as of a hand). The picture possess a certain “affordance”—it offers a visual “use” for a viewer—and thus the picture offers a viewer a way or ways of being being interpreted. The picture does this through its constitutive pictorial designs. For example, the pictorial designs that
Constable included in *Wivenhoe Park* delimit or constrain the possible ways in which the painting can be interpreted, making it much more likely that different people, seeing the painting at different times and under different viewing conditions, will almost always arrive at the correct interpretation.

To say that the features in P are equivalent to the features in O means that the former trigger off the release of a visual experience that is similar to the visual experience triggered off by the latter. Even though the features in P must be arranged in a certain way, and even though the viewer must come to the picture with a certain mental set (which I will talk more about later), the features in P will nevertheless allow the viewer to have the illusion of seeing the depicted object (be it an apple, Louis Philippe, or the city of Dresden) because they have been discovered to be equivalent to the features they are depicting. This is what Gombrich means when he claims that even though the world does not look like a flat picture, some flat pictures can look like the world. (Gombrich, 2000)

1.9 Images are always incomplete

Although it might seem redundant, it is important to reassert that under ordinary picture viewing conditions (which I take to involve binocular vision along with the ability of the viewer to move around the picture) viewers can tell that a picture is not identical with the represented object. A picture can be made to look more or less naturalistic, but there is no way to transcribe all of the features of an actual object into a picture. It may be impossible to match the color of an actual apple within the artist’s medium, therefore
the artist must reconcile the “local color” of the apple itself with the range of tonal
gradations available within his or her medium. So, the features in the painted apple won’t
be identical to those features perceived in the actual apple. The task of the artist is to find
those relations of tonal gradation in her medium that are the best equivalent to the
relations of tonal gradations found in the apple—hence the claim that a picture is “a
transposition, not a copy” (Gombrich, 2000, p. 48). The artist is tasked with finding “the
patch on the window that might be mistaken for a house in the distance viewed from a
given spot.” (Gombrich, 2000, p. 301)31 This is why, for Gombrich, the artist is not after
identity and exactness, but rather equivalence and suggestion.32

While the complete transcription of visual elements is impossible, the history of
art has shown the ingenious ways artists have discovered for capturing more and more
aspects of an object. Thus, for Gombrich, art history consists of developments, through
trial-and-error, of creating on the canvas visual equivalents to the types of perceptual
cues33 found in nature. Again, what our visual system takes as significant in nature are
relationally arranged visual elements that are perceived as having a certain visual
meaning (e.g., “rounded surface”). Given this fact about the perception of nature, artists
discovered that certain arrangements of relations of visual elements in a given medium
(e.g., paint or charcoal) could lead a viewer to take them to visually approximate the
meaning of those in nature. Having figured this out, artists can then assign the two-
dimensional relations of elements interpretations based on their representing natural
relations of elements because the interpretation fits with a category we use and recognize
in ordinary perception.
1.10 Substitutes, schemata and suggestion

“All image-making,” writes Gombrich in *Meditations on a Hobbyhorse*, “is rooted in the creation of substitutes” (Gombrich, 1994, p. 9). In this well-known essay Gombrich begins his discussion of substitutes by arguing that a child’s hobbyhorse will serve as a substitute for a real horse if it (the hobbyhorse) fulfills a relevant function within a context of action for the child—the hobbyhorse stands-in for the actual horse because it fulfills the central function of being “ridable.” Analogously, the function that a depictive picture fulfills is a visual one, such that our seeing the picture is roughly like seeing the object it represents. In writing about the hobbyhorse Gombrich says that “the common factor,” between a stick (which qualifies, for the child, as a horse) and an actual horse is “function rather than form; or, more precisely, that formal aspect which fulfilled the minimum requirement for the performance of the function” (Gombrich, 1994, p. 4). When Gombrich turns to pictures he will say that the picture will appear to the beholder as the object it depicts because the picture fulfills a *visual* function that is similar to the visual function that the object fulfills in ordinary perception.

How does this relate to the idea of the schema? Gombrich is saying that the schemata used in ordinary perception (made up of the sets of categories plus the interpretation of the relations of elements) becomes a part of the artist’s artistic vocabulary and is utilized in both the making and the perception of pictures. So, given the schemata used in ordinary perception, the artist will set out to discover which sets of categories are best suited for representing which objects in a given medium. The artistic vocabulary is passed from artist to artist and generation to generation, and it embodies
certain formulas and “pictorial devices” (e.g., foreshortening and sfumatto) discovered to work best for arranging the relations of pictorial elements to produce specific results.

All picture-making (what Gombrich refers as “image-making”) involves the creation of substitutes, and when creating a picture an artist begins by choosing how to use his or her vocabulary, which is to say, his or her schemata. The knowledge of which schema works best serves the artist as an established rule or modus operandi, which can be recalled easily and applied to various situations. As artists employed the schemata in different ways and to varying degrees (by arranging the relations of elements differently or placing greater or lesser importance on following and implementing established traditions), they inevitably altered their artistic style; and as artists sought to change or break with familiar schemata they would then establish new schemata, new vocabularies, with new sets of categories that required viewers who were familiar with the previous schemata to alter their interpretations and re-discover what schemata the artist was using.

It is important to note that, as the above discussion already implies, while the inherited schema forms the starting point for the picture-making process it does not have to necessarily rely too heavily on the resemblance of external features. A smiley face does not look all that life-like—the merest outline suffices for it to be recognizable as a smiling face. So, the context of action in which a picture is created will influence the artist’s choice of which features to include in it, and thus whether it will be more or less schematic, or, conversely, more or less naturalistic.

Above I said that for a picture to function as a visual substitute means that it must trigger off the release of a relevantly similar visual experience (in the way I’ve described) to the object it is depicting. I also said that when any picture is created it will rely upon
the artist’s choice of schema. Now, no matter how naturalistic the picture is intended to 
be the schema (the artistic vocabulary) should “preserve and repeat” those significant or 
distinctive features which can “release a similar response” in the beholder to the one 
released by similar features found in the object (Gombrich, 2000, pp. 110-11). Since 
pictures must be developed through the use of a schema, and since all pictures-making 
involves the creation of substitutes, we can conclude that for Gombrich a substitute will 
embody certain schematically distinctive features (or some set of them) that are capable 
of triggering off, in the beholder, a relevantly similar response to that triggered off by the 
depicted object.

The artist, says Gombrich, through his or her choice of schema, is performing an 
act of suggestion when he or she creates a substitute; and when we “mistake” a dab of 
paint for a house the artist has succeeded in finding a way to use his or her materials in 
such a way that they suggest something beyond their base materiality, something beyond 
their being perceived as merely flat marks. Again, nature does not literally look like the 
picture, but the features of a picture can, in certain respects, look like nature. Our visual 
system can respond to pictorial features in a similar way to the features of the object 
because of the inclusion, in the picture, of features deemed to be significant for 
recognition.

The artistic quest to determine which features count as distinctive, and thus which 
features will help serve the act of artistic suggestion as well as the process of recognition, 
is also important in that it will help the artist create a more potent visual substitute. The 
choice of which features to include in an image, however, is not a totally random act on 
the part of the artist. This is something I have already hinted at. The ability of an artist to
pick out an object’s distinctive features is really no different than the ability of an ordinary individual to pick out the distinctive features of an object during the normal course of a day. The primary difference lies in the artist’s task: the artist understands which of the distinctive features of an object or scene can be used to construct a successful depiction of that object or scene and, similarly, which will aid in successful recognition by the beholder. Of course, it is possible for us to be fooled or “taken in” by the picture and believe that the picture is the object it depicts—but this is not really the point. Gombrich is primarily concerned with those cases where the picture is not confused with the actual object and yet the features of the picture suggest a visual reading in terms of some object (cases I have described as the “ordinary” or “usual” cases). When we say that a picture suggests an object what we are really saying (to put it in Gombrichian terms) is that a picture is capable of being a visual substitute for an object; and for something to be a visual substitute means that it must be created using an artistic schemata that, in the hands of the artist, can be wielded to suggest any number of things beyond the picture itself.

1.11 The mental set

In the next two sections I will sketch some remarks on the mental set and the relationship between the artist and the beholder. I will then return to these topics in section 1.14 to give a fuller, more detailed description of how Gombrich understands their interconnectivity.

For the visual features in a picture to be taken as equivalent to features in the object the artist and the beholder must collaborate in what Gombrich calls a “game of
equivalence” (Gombrich, 2000, p. 326). In creating a picture the artist must base the choice of schema (at least partly) on facts of human perception. Artists will selectively include features in the picture that correspond to features of the object, and while the choice of which features to include in the picture is up to the artist (e.g., should the picture be colored or monochromatic? Is the size of the object important? Should there be a pronounced outline?), those that are included will be those that are best capable, given the medium and the intentions of the artist, of representing the features of the object. The creation of specific pictorial features depends on the arrangement of relations of pictorial elements in the picture, and the artist can construct and arrange the chosen relations of elements on the canvas in any number of ways. However, if he wants the picture to be a successful representation of, say, an apple (such that the beholder is able to recognize the picture as a picture of an apple), he must be conscious of the expectations and limitations of the beholder. This means that he must be aware of the proactive role of the beholder’s mental set in influencing how the picture will be interpreted.

Looking at any picture means looking at it with a certain mental set—we will have certain expectations of how the picture should be read and what the picture is meant to “do.” Just as the artist is selective about which features to include in the picture, the beholder will also be selective about which mental set to bring to the picture. For instance, the mental set we bring to the reading of the *Bayeux Tapestry* is different from the one we bring to a Rembrandt painting, which is different still from the one we bring to a magazine ad. Knowing that the mental set associated with the *Bayeux Tapestry* is different from the one associated with a Rembrandt painting requires the beholder to be aware of any number of facts about pictures (e.g., the stylistic, cultural, and social
background; the artist’s intention; the conditions under which the image is meant to be viewed, and so on). Of course there is always the possibility that we can get it wrong. There is no guarantee that we will always have the right interpretation. We could read the Bayeux Tapestry like a Rembrandt (or worse, like a Picasso). But this just means that we must back up, figure out what has gone wrong, and attempt to “re-attune” our mental set.

1.12 The artist and the beholder

No matter how detailed an artist’s schema may be there will always remain a pictorial deficiency: the marks on the canvas will only be able to convey a limited amount of information about the depicted object, and “even the most meticulous realist can accommodate only a limited number of marks on his panel” (Gombrich, 2000, pp. 219-220). What’s more, pictures not only fail at representing the infinite, they fail at representing the finite as well. And yet, it is because a picture is deficient that we are all the more in awe of its transformative power. We are in awe of the way that flat swaths of color and shape can become a horse or a hillside, a person or a unicorn. What is just as amazing, Gombrich will say, is that as beholders we have an active role in this act of transformation—as beholders we help supplement the deficiencies on the page by affixing our mental set to the picture. Gombrich understands this collaborative act in the following way: the artist chooses to arrange the pictorial marks in such a way that they afford the possibility of being read in a certain way; when we read the picture we will “mobilize our memories and our expectations of the visible world [i.e., our mental set] and …test [the artist’s] image through tentative projections.” (Gombrich, 2000, p. 314)

The artist knows that the pictorial marks are just that—flat marks on a page; the artist also
knows is that this fact requires the help of the beholder to ensure that the picture is seen correctly.

When the beholder perceives arrangements of two-dimensional relationships of elements he or she will offer an interpretation of what is being seen, and the perception of the relationships as interpreted yields a visual experience that, Gombrich claims, is relevantly similar to the experience of seeing the depicted object. In this way the beholder, in looking at a picture, is seeing something that is not really there. To understand Gombrich’s conception of illusion involves recognizing that a combination of things are working together at the same time: one is the artist’s ability to create a two-dimensional artifact capable of suggesting something beyond its base materiality, and consisting of features that can be recognized as representing features in the world (I will label this as a picture’s capacity to “afford” recognition of what it represents); another is the beholder’s capacity to respond to the features of a picture, recognize them as being of something else, and yet refrain (in the ordinary cases) from being deluded and confusing the picture for the object it is representing.

Recapitulation and further details

1.13 A change in the function of pictures

I have already alluded to the relationship between the substitute and the schematic image, but let me briefly recapitulate Gombrich’s position. Recall that according to Gombrich it is through the use of a schema that all artists begin the picture-making process, and in making a picture the artist makes a visual substitute. That a picture is a
visual substitute means that it embodies certain features that suggest, to some viewer, the substituted object. As I have said before, a substitute can be recognized as visually “being” the depicted object (i.e., the picture can be seen as relevantly similar to the object it depicts) despite the fact that the substitute’s individual constitutive parts do not resemble those of the depicted object all that much. Ancient craftsmen who constructed idols with the intention that they should stand for some deity only needed to include those distinctive features that fulfilled the intended function such that the idol became (in the viewer’s dealings with it) the deity it was standing in for. We can see this same sort of thing happening in the case of the hobbyhorse. While it doesn’t look all that much like an actual horse, Gombrich holds that within a context of action (for the child) it becomes a horse by fulfilling a relevant function of what it means for something to be a horse—it is “ridable.” Eventually artists found that they could add more naturalistic features to the simple schematic shapes and forms of their substitutes, giving them a more lifelike presence that was more dramatic and believable (as Gombrich put it: the hobbyhorse must have eyes or else how can it see?).

As the history of image-making progressed, though, a change in function occurred: the substitute (the picture) no longer had to exist in its own right as the idol did; instead it could refer to something outside itself (like a landscape or an apple), and therefore it could be a record of a visual experience and not merely a physical stand-in. We might refer to these types of substitutes as being more “holistically” depictive because artists chose to allow more parts of the surface of the picture to play a relevant role in depiction (e.g., the empty space surrounding the forms in a landscape painting is interpreted as atmosphere or air surrounding the depicted foliage). Thus, the artifact as a
mere substitute (i.e., the artifact consisting of overly schematic shapes and forms) became endowed with a presence, with a perceived sense of depth and space. I think that this is what Gombrich means by the phrase “imagined reality”\textsuperscript{39}—we are presented with the presence of a world existing within the framed boundaries of the picture. For example, a schematic line drawing of an apple can function as a visual substitute because it embodies visual features that allow it to be recognizable as an apple. (Of course it will only contain a limited number of visual features, making it easier to interpret the image as just meaningless marks on the page.) Equally, a painting of an apple can function as a visual substitute, and by embodying more visual features it will create a presence of space around the depicted apple thus making it more difficult to see the painting as just meaningless marks on the page. The painted apple is not seen as floating in a white void on the page, as is the case with the line drawn apple; instead the painted apple is seen as existing in a depicted space (the “imagined reality”). This is what I meant by saying that the change in the function of a picture can be described as the progression from mere substitute (or schematic image) towards an artifact that is more holistically representational. By making more of the picture’s features (and thus more of its relations of visual elements) representationally relevant, pictures progressed from artifacts functioning (within a context of action) as mere stand-ins for objects, to artifacts that were endowed with the presence of an imagined reality, and which point to objects, scenes, and states of affairs outside of themselves.

When I refer to an image as being “schematic” I am referring to both the “simplistic” nature of the image as well as the fact that not all of its parts are relevant to the image \textit{qua} depiction. Let me give an example to illustrate this. In a line-drawing of an
apple, the apple that is depicted may appear to be rounded and have a sense of depth and space; but any sense of space is of such a limited nature that it can easily be broken by a host of contravening features (the flatness of the paper, the edge of the page and the surface underneath it, the uniform whiteness of the paper, and so on). So not every part of the line-drawn apple is representationally relevant to its being as a depiction of an apple. In this way the drawn apple is said to be “schematic.” When it comes to a painted apple, the sense of space and depth surrounding the depicted forms is much more pronounced and, I think Gombrich would say, disregarding such a visual presence becomes much more difficult. This is because more parts of the picture are representationally relevant (e.g., the whiteness surrounding the line-drawn apple is usually not relevant while a painted white background usually is). Many pictures need only contain a minimal set of visual features that allow the picture to be recognized as (for it to be seen as of) the depicted object. And as the artist includes more and more features (e.g., shadows and highlights, gradations of color tones, textural elements, and so on) the transformative power of the image is expanded, allowing more parts of the image to function depictively. Schematic forms are thus the building blocks for all picture-making according to Gombrich; all artists begin from a schemata that can be added to, adapted, or possibly rejected, but which nevertheless must embody a minimal set of features that allows the picture to be recognizable.

1.14 The beholder, the mental set, and the act of projection

Let me now return to the topics I picked up in sections 1.11 and 1.12. At its simplest, the concept of the mental set is a form of selective attention—it consists of
expectations and assumptions, the idea that what worked in the past will work in the future. We make use of these sorts of expectations everyday. When I walk out my front door I don’t expect to find comedian John Stewart waiting to bid me “Good morning.” Why? Because based on the mental set associated with the act of walking out my front door, I can expect to perceive some things (trees, my car, the paved street) but not others (a lion, a pint of Guinness, comedian John Stewart). Similarly, my mental set can influence what I actually observe. If I look out my window and see a dark object slowly moving across the driveway I might assume that it is one of the local geese waddling back to the pond. My expectation of seeing the goose actually changes what it is I believe I am seeing, influencing my interpretation of the ambiguous shape. As the shape moves under the streetlight, though, I find that my interpretation was wrong: it is not a goose, but a dog—my expectation did not lineup with my observation. In perceiving the ambiguous information in the world we must bring a certain level of preparedness or “horizon of expectation” to what we are seeing (Gombrich, 2000, p. 60). In *Art and Illusion* Gombrich quotes a passage from M.D. Vernon’s *Visual Perception* that is meant to illustrate this point—that often times the familiar can induce the expected.40

Gombrich’s use of the mental set applies most explicitly to our perception of pictures. No one thinks the trees in the *Bayeux Tapestry* actually looked like how they are depicted; we see them through the filter of the style used by the artist. As we look from one picture to another we often find the need to re-attune or alter our mental set based on the different expectations of the picture’s style or what it was intended for, and therefore how it should be read. Our mental set is not only influenced by our expectations, but also by our habits, beliefs, memories, culture, and social upbringing, as well as by facts about
the artist, the image itself, and our own visual system. While the schema filters sets of relations of elements based on whether or not they fit into the sets of categories we are using, our mental set forms a hypothesis based on our preparedness for, and expectations of, perception, which will not only influence what we are seeing but also what we are *likely to see*. Given that we bring our mental set to the perception of pictures, the idea that we are capable of directly influencing what we perceive (“of seeing what we believe”), and thus influencing how a picture is read, does not seem all that impossible (Gombrich, 2000, p. 210).

With these facts in mind we can understand the sense in which, for Gombrich, the mental set is coupled with the psychological act of “projection.” Projection is the ability of a viewer to read an interpretation into something (e.g. moss on a wall, a cloud or a picture), which in turn acts as a “screen” onto which the interpretation is projected. The simplest and most well known example of this can be seen in the Rorschach inkblot test. When we perceive an ambiguous figure, like the inkblot or a cloud or moss on a wall, we offer an interpretation of what we are perceiving, and in a similar way a picture will afford different interpretations to be projected into it. The interpretation that is ultimately settled upon will depend (in part, though not completely) on the beholder’s being keyed up in a certain way (due to the expectations and assumptions that make up his or her mental set). The beholder compensates for the incompleteness of an image by projecting his or her mental set into it—beholders have the ability to see things within things, to see whole worlds within indeterminate and ambiguous forms, and thus are able to fill in the gaps (i.e., supplement the lacking information) left by the artist. According to Gombrich, as representational art developed, artists increasingly relied on the knowledge
that viewers had this perceptual capacity. By gradually shifting the “load of creation” to the beholder, artists could represent objects by doing less and less (Gombrich, 1994, p. 10). As Gombrich puts it:

“...The willing beholder responds to the artist’s suggestion because he enjoys the transformation that occurs in front of his eyes... The artist gives the beholder increasingly “more to do,” he draws him into the magic circle of creation and allows him to experience something of the thrill of “making” which had once been the privilege of the artist. (Gombrich, 2000, p. 202)

1.15 Completing the incomplete

As the beholder looks at the image he or she will test it for consistency, that is the beholder will read the image and attempt to fit his or her interpretation with the information present in the picture—the beholder attempts to classify the whole image “within a possible category of experience.” (Gombrich, 2000, p. 238) This is a harder task for some pictures than for others. We often take this ability for granted because pictures have become so commonplace—we read them with such ease, and often with the slightest glance despite their visual incompleteness, and despite the fact that their individual parts, upon a second-look, do not look all that much like what they are supposed to represent. So, I think that there are really two ways to compensate for the pictorial deficiency of incompleteness, both of which I have already mentioned: one is the artist’s clever ability to suggest a depicted object by arranging specific relations of elements on the page in such a way that they create an equivalence to relations of elements in the world. Constable is able, quite beautifully, to suggest Wivenhoe park in Essex because the features he chose to include in the painting, features which are constituted by specific arrangements of relations of visual elements on the canvas, make it relatively easy for us to classify the entire image within a possible category of
experience, say “a landscape” or “manor house on a hill.” This is what the portrait
caricaturist does when she includes the “characteristic invariants” of, say, a politician or
famous actress (Gombrich, 1982b, p 29). The second way to compensate for a picture’s
visual incompleteness draws on the “beholder’s share,” on the ability of a viewer to
project an interpretation into the image given the specific sets of relations of elements
chosen by the artist. In regards to the latter, Richard Woodfield writes that, “The natural
incompleteness of all pictures leads to an appeal to our capacity for visual
supplementation.” (Woodfield, 1996, p. 35) Throughout Art and Illusion Gombrich tries
to remind us of the amazing act we perform when we look at a flat arrangement of shapes
on a page and, seemingly out of thin air, interpret them as a three-dimensional reality
within the frame of the picture.

If I understand Gombrich correctly, then, the incompleteness of all images is
supplemented through the harmonious combination of these two things: the artist’s ability
to suggest through the creation of a two-dimensional equivalent, and the beholder’s
ability to project an interpretation into the picture.

1.16 Gombrich and twofoldness

Before ending this chapter I want to discuss the notion of twofoldness, which will
be integral to understanding the discussion of illusion in Chapter 5. Twofoldness is the
idea that our experience of viewing a picture involves two different experiences that
occur at the same time: the experience of seeing what the picture is made of (its medium),
and the experience of seeing what the picture is a picture of (its content). Richard
Wollheim (1980) developed this idea as a reply to what he took to be Gombrich’s position
in *Art and Illusion*. Wollheim writes that “if I look at a representation as a representation, then it is not just permitted to, but required of, me that I attend simultaneously to object and medium.” (Wollheim, 1980, p. 213) Gombrich, says Wollheim, denies that we can, at the same time, see a picture’s medium and the pictorial content. Seeing the flatness of the canvas or the painted brushstrokes requires me to switch my interpretation in order to then see the subject of the picture. For example, when I look at Degas’ *Self portrait with soft hat* it is impossible for me to see both the canvas and the manor house simultaneously. To understand the content means that I must “momentarily forget the canvas.” (Gombrich, 2000, p. 279) According to Wollheim, then, the pictorial experience is one of “Now canvas and nature,” or “Now medium and object”; whereas for Gombrich, the pictorial experience is said to be one of “Now canvas, Now nature,” or “Now medium, Now object.” (Wollheim, 1998, p. 221)

The significant of this is that since Wollheim put forward the account of twofoldness it has been taken as a “fact” of pictorial perception, the consequence for Gombrich being that his claimed rejection of it means that his account has been labeled as “mistaken” (Lopes, 1996, p. 42).

In Chapter 4 I will argue that contrary to the current position, Gombrich’s account actually accommodates a form of Wollheimian twofoldness (resulting in the two accounts actually being much closer than was once thought). Let me briefly offer two reasons why I believe this to be the case. First, I believe that one of Wollheim’s major failings is that in describing our dual experience of seeing the medium and the content Wollheim runs together the ideas of “attending to…” and “being aware of…” This is problematic because being aware of something is markedly different than attending to it, and saying
that we can simultaneously attend to two things is a different claim than saying that we
can simultaneously remain aware of those same two things. Applied to the perception of
pictures, I will claim that we can see a picture’s content and concurrently remain aware of
its medium without thereby giving the latter our full attention. What is more, it might not
even make sense to describe our experience of looking at a picture as being “twofold”
because we simply cannot, contrary to Wollheim, simultaneously attend to both the
medium and the pictorial content.

Secondly, if we now adhere to this distinction between “attending to” and
remaining “aware of” medium and content, we will see that Gombrich’s claim that we
must switch visual interpretations makes perfect sense if we read him as talking about
“attending to” the medium and the represented object. We must switch interpretations
because we cannot see, at the same time, the picture as having a meaningful content (e.g.,
as being of Degas) and as being a meaningless arrangement of marks on a page. To
simultaneously attend to both the medium and the content would be to do the impossible
—to simultaneously interpret the picture as being meaningless and meaningful. Because
we cannot do this, we must switch our attention from medium to content (that is, switch
our interpretation of the picture).

Finally, this interpretation of twofoldness is related Gombrich’s account of
illusion in the following way: if our experience of looking at a picture were not to involve
the simultaneous awareness of the medium and the pictorial content but rather our
attending to both the medium and the pictorial content, then there would be a real worry
that switching interpretations from “the medium” to “the content” would result in the lose
of the awareness that what we are looking at is a picture and not the object itself. This is
problematic for Gombrich’s account of illusion because it would mean that in such a case our experience of the picture would be one of a delusory nature. I will argue (in Chapter 5) that because Gombrich’s account can maintain this distinction between “attending to” and being “aware of,” the experience of seeing an ordinary picture can create a pictorial illusion that is not delusion—the creation of a delusory experience will only occur in cases of trompe l’oeils and not in cases of ordinary pictures.

1.17 Revisiting Illusion

One of the major stumbling blocks when reading *Art and Illusion* has been what Gombrich means by the central term “illusion.” Many critics of Gombrich have misunderstood his claims concerning illusion, and while I will return to the concept of illusion in Chapter 5 it would be prudent to discuss it just a bit more before concluding this chapter.

In the cases in which he’s interested, I take Gombrich to be claiming that illusion is simply a way of describing what is going on when we recognize a picture as being of some object. Illusion occurs when a picture triggers off, in our visual perception, the release of an experience similar, in a relevant way, to the experience of perceiving the object depicted in the picture. The picture cannot be of just anything since the specific features of the picture will constrain the possible visual experience(s) that it will trigger off in a beholder. The important thing to remember is that in these ordinary cases we know that what we are seeing is not the actual object, and therefore in such cases we don’t suffer a perceptual delusion nor do we have a false belief about the picture. Also, and as I have stated above, Gombrich’s description of illusion means that a picture can be
less than completely naturalistic-looking (and in fact might be quite dissimilar to the object it is depicting), and yet retain the capacity for creating the illusion (in his sense) of seeing the picture in terms of the object it is depicting. Put differently, Gombrich takes illusion to exist on a continuum, the line drawing being just as capable of generating an illusion as the ordinary painting or the trompe l’oeil. There is no progression from non-illusionary to illusionary; illusion is present at all stages of image-making (including in the trompe l’oeil). Gombrich will say that, beginning with the line drawing and moving along the spectrum towards the naturalistic painting, the sense of illusion the picture creates will be heightened or more pronounced. It is only once we reach the trompe l’oeil that the worry of illusion becoming delusory or deceptive is manifest—it is there that we can lose an awareness of the picture’s medium and those features of the picture that make us aware that the picture is a picture. When this occurs we are left unable to distinguish the picture from the actual object.

A trompe l’oeil is a special case because it is the point where illusion becomes about our real environment, “where our eye is actually deceived”—it is only there that illusion turns into delusion (Gombrich, 2000, p. 253). Importantly, in claiming that depictive art involves illusion Gombrich is not claiming that all images are like the trompe l’oeil. We don’t see Wivenhoe Park as an actual landscape spreading out in front of us, nor do we think that Self portrait with soft hat is Degas peering back at us through a frame-sized window. The image’s capacity to suggest and the beholder’s ability to supplement the lack of information are the master keys of illusion for Gombrich. Illusion arises from these two collaborative acts: the artist’s experimenting to find the arrangement of relations of pictorial elements that best work to suggest the subject, and
the beholder meeting the artist half way by supplementing what was incapable of being translated into two-dimensions. No artist has ever been able to create an illusion because she could replicate what she saw. This has been shown to be impossible. Instead, Gombrich is saying that something even more amazing is taking place: artists experiment with lines, shapes, tonal gradations and colors to create a visual substitute, in two-dimensions, that is able to suggest the intended motif—they then leave the rest to be supplemented by the roving eyes and imaginations of their audiences.

1.18 Conclusion

So, where do we stand regarding Gombrich’s account of depiction? I have tried to make it clear that in Art and Illusion Gombrich is not trying to give strict necessary and sufficient conditions for depiction. Instead, he gives what I have called a “core necessary condition” for an artifact’s counting as a depiction. The core necessary condition holds that a picture P will count as a depiction of an object O when P functions as a visual substitute for O in the following sense: the arrangement of relations of visual elements in P, being constitutive of P’s features, triggers off the release of a visual experience that is relevantly similar to our visual experience of seeing the relations of visual elements in O; additionally, P can be described as relevantly similar to O if P affords the viewer the ability to describe the experience of seeing P with the linguistic expression “as of.” Which is to say, P is relevantly similar to O if we can describe our seeing of P “in terms of” seeing O. If the visual experience of seeing P can be described in this way, and if P thus functions as a visual substitute in the way I have described, then the seeing of P creates the illusion of seeing something that is not really there, namely O. Furthermore,
Gombrich says that the artist and the beholder must be in sync with each other: the artist is aware of the limitations and the expectations of the beholder; the beholder is aware of the schemata employed by artist, and is able to supplement the limitations of the picture through the projection of his or her mental set.

Apropos of substitution, Gombrich thinks that any number of artifacts can serve as a substitute for something else, and he extrapolates from cases of physical substitution, where one thing physically stands-in for another (e.g. the hobbyhorse or the idol), to cases of visual substitution, where the relevant substitutive features are visual ones. A visual substitute is created through the use of a schema (in the artist’s medium), which picks up on and approximately reproduces relevant visual features that are read by the beholder as similar to “ordinary” visual features in the world; the pictorial and ordinary visual features are relevantly similar in that the perception of each triggers off the release of a relevantly similar visual experience—the visual features of a picture of an apple will trigger off a relevantly similar visual experience to the ordinary visual features of an apple. The choice of which features to include in P is up to the artist (and is thus conventional), but the degree to which certain features do trigger off a specific visual experience is out of the artist’s hands (and thus might be deemed natural or biological). We, the beholders, will read (i.e., interpret) the arrangements of relations of elements in nature and in pictures in specific ways due to the facts about how our visual system responds to perceiving specific relations of elements. Not just any set of relationally arranged visual elements in a picture will trigger off a visual experience like that of, say, an apple. In certain respects, then, the choice of which features to include in P delimit how P will be interpreted—our visual system will only respond to the relations of visual
elements in a limited number of ways—thus constraining the range of possible visual experiences triggered off by $P$.

Gombrich believes that pictures involve an illusion because they afford us the possibility of seeing in them a presence, a perceived reality that is beyond their flat, base materiality. Pictures can be of many different things (people, places, states of affairs), and the illusion a picture creates involves our ability to recognize the content of a picture. Gombrich is quite clear, however, that illusion (of his sort) should be kept distinct from delusion and false belief because the ordinary picture is not intended to deceive the viewer, nor will usual viewers have reason to take the ordinary picture to be the object it depicts. Gombrich’s understanding of depiction is that a picture is a visual representation of an object owing to the fact that the picture functions as a visual substitute for the object it is depicting. In functioning as a visual substitute the picture fulfills the core necessary condition for something’s counting as a depictive picture by being an artifact the seeing of which approximates the visual experience of seeing the object, thus allowing a viewer (who is properly situated and mentally-set) to see the object in the picture.
2 Visual Substitution and Relevant Similarity: The Relational Model

2.1 Presenting relevant similarity in two parts

In this chapter I will return to the concept of relevant similarity presented in the first chapter. As it stands, the working interpretation of this concept is inadequate and is in need of elaboration. Therefore, I will begin to fill in the details by using this chapter to introduce the first part of relevant similarity: the importance of the picture as a relational model; then I will continue and complete the investigation of relevant similarity in Chapter 3 by discussing the ability for a picture to afford recognition of its subject. Recall that I proposed the concept of relevant similarity as a way to shed light on a somewhat vague yet integral part of Gombrich’s theory of depiction—the idea of visual substitution (specifically, a pictorial substitute). Thus, the job of investigating relevant similarity in this chapter and the next will be to offer a better understanding of this part of Gombrich’s writing by demonstrating how relevant similarity might fit into Gombrich’s already existing theory.

2.2 The layout of Chapter 2

The working interpretation of relevant similar in the first chapter was in terms of the property of being expressible, linguistically, in “as of” visual terms. I suggested that a picture serves as a visual substitute for an object, and is thus relevantly similar to the object, just in case there is a similar way that we are able to describe the picture and the
object the picture depicts. I tried to express this in my claim that we can see a picture “as of,” or “in terms of,” an object. For example, our experience of seeing a picture of an apple can be described as an experience “as of seeing an apple,” just as can our experience of seeing an apple itself. So the picture is relevantly similar to the apple because both the picture and the apple have the property of being such that someone can have, with regard to each of them, a visual experience “as of seeing an apple”. This initial interpretation is inadequate because it does not explain why a picture and an object are relevantly similar beyond the similar linguistic description we offer. It fails to show whether there is a direct relation between the picture and the object and so will not get us any closer to understanding Gombrich’s claim that a picture can be a visual substitute for an object. If relevant similarity is to be more substantive than a mere descriptive label we must investigate whether there are any reliably relevant respects in which a picture P is similar to an object O, and this will mean looking at a number of possible sources for relevant similarity.

I will begin this more detailed account of relevant similarity by briefly revisiting the idea (discussed in Chapter 1) that we cannot hope to locate relevant similarity in a picture’s being visually identical with an object, which precludes the possibility of a picture’s capturing everything that is visually true about it. None of the designs and arrangements of marks on the canvas can be visually identical to those of the object itself, and thus a picture will fail to capture everything that is visually true of the object itself. As such, relevant similarity does not consist in the idea that the individual visual elements of a picture are identical to the individual of an object—point-for-point the picture fails to be similar to an object. Instead, I pursue Karl Bühler’s notion of
relationstreue (what Gombrich comes to call a “relational model”), as key to identifying which respects are the ones that make similarity relevant. In looking at the role Bühler plays in influencing Gombrich’s writings I claim that we should focus on whether a picture is relationally faithful to its subject. In other words, it is not the individual visual elements but rather the relationships among visual elements, the arrangements of similarities and dissimilarities between visual elements, that are important to a picture being relevantly similar to an object.

2.3 Relevant similarity is more than what is given via a descriptive linguistic tool

An initial worry that should be addressed is this: we should understand relevant similarity is understood as merely a descriptive linguistic tool. That is, the source of relevant similarity between a picture P and an object O resides wholly in the viewer’s ability to describe the seeing of P and the seeing of O using similar words or even whole sentences. Seeing an apple on a table triggers off a specific visual experience for a viewer, and, I wish to claim, seeing a picture of an apple on a table triggers off a visual experience that is, for the viewer, similar in certain respects. The source of the similarity between the picture and the object, so this initial worry goes, is found in the viewer’s ability to describe the visual experience of seeing the apple itself and the picture of the apple in similar “as of “ terms.

I have already said that we should not pursue this conception of relevant similarity. I would certainly agree with this conception up to a point, for it remains helpful to point out that there is truth in the claim (uncontroversial as it may be) that we
do describe a picture and an object using similar words and even whole sentences. What remains to be explained is why we describe picture and object using similar words and even whole sentences. For this reason I ask the reader to continue along the current discursive path a bit more, if for no other reason than to determine that the concept of relevant similarity is to be analyzed via something more than a mere descriptive linguistic tool, and, additionally, that such an initial account of relevant similarity does not fit with anything Gombrich says in *Art and Illusion*.

In the next section I will elaborate on the role of the linguistic expression “as of” introduced earlier on in the first chapter. This will allow me to show that the claim established at the beginning of this section (i.e., that we are able to describe seeing a picture and an object in similar visual terms) is really symptomatic of more substantive similarities between a picture and an object, and thus that the source of relevant similarity is more than the viewer’s descriptive label that a picture and an object are alike.

2.3.1 Why “as of”?

I am certainly not the only writer to use the phrase “as of” to describe the pictorial experience: Robert Hopkins48 (1997, 2003, and 2005), Dominic Lopes (1996, 2000, 2005), Michael Newall (2009), and Richard Wollheim (1998) all use “as of” in various contexts to describe the experience of seeing a picture. For instance, Dominic Lopes writes (although I don’t think that he intends the phrase to have any special status) that, “Admittedly we experience pictures as of their subjects” (1996, p. 3), which is how I have been using the phrase. So it would seem that Lopes and I are in agreement on this
rather general way of using “as of”. When we see a picture and recognize its subject (i.e.,
the object or scene the picture is depicting), we often say, “I see…” . By doing this we are
describing our seeing the picture in terms of seeing the object. This ability to describe a
picture in terms of an object can be summed up in the way that I, Lopes, and others have
done, by saying that we see (i.e., experience) the picture *as of* an object.

Nowhere in *Art and Illusion* do we find the phrase “as of,” a fact which is not
necessarily of concern. Gombrich certainly does acknowledge the common way in which
we use similar descriptions for both pictures and object,49 which seems to hint at the
possibility that he may also acknowledge that there are aspects of the picture and the
object that warrant the use of such similar descriptions. We might therefore assume that
when asked, “In what way is a picture of an apple similar to an actual apple?” , the answer
is that there exist certain respects, between picture and object, that warrant the response,
“The picture of the apple and actual apple are similar in ways X,Y, Z.” You could always
respond, rather sarcastically, by listing just *any* respects that come to mind in which the
pictured apple is like an actual apple; but most such sarcastic responses will be irrelevant
to the question of how the picture is similar to the object (e.g., “I saw both today”), not to
mention that there will be a number of obvious respects in which the actual apple and its
picture are dissimilar (e.g. the fact that the one is a three-dimensional object and the other
a flat, two-dimensional artifact). These facts alone, however, do not prove that there are
no respects in which picture and object are similar.

That being said, it is clear from the text that Gombrich does not hold the position
that a picture is similar to an object simply *because* the viewer can describe the picture
and the object in a similar manner using similar words or even whole sentences.50
Perhaps, to address a famous riddle, a raven really is similar to a writing desk, but even if this is true we don’t know how or in what way a raven is similar to a writing desk.

Likewise, perhaps I can say that a picture of an apple is similar to an actual apple, and perhaps I can even describe the many ways in which the similarity holds, but this conception of relevant similarity still does not tell us how or in what way a picture is similar to the object it represents; and furthermore, it cannot be a fundamental account of relevant similarity since such an account should not depend on the existence of someone who can describe both the experience of seeing the picture and the experience of seeing O itself in “as of seeing O” terms. It remains plausible to say that seeing a picture P may lead us to have an experience as of O, and this may lead us to “interpret P as of O” (Hopkins, 1997, p. 446) or, to have an “experience as of O” (Hopkins, 2003, p. 157); and, in turn, we might say that the picture can “trigger experiences as of what [it represents]” (Lopes, 2000, p. 230). But if the reader can assume with me that a picture of an apple really is similar to an actual apple in some more substantive way (e.g., in a way that is relevant to Gombrich’s visual substitute idea), then the source of relevant similarity is more than the result of what is given via a linguistic description.

If this account of relevant similarity was inadequate from the off, what was my reason for introducing the phrase “as of”\(^5\)? I have tried to show that using “as of” is a helpful way of describing the seeing of a picture, since we often will talk about a picture and its subject using similar words and even whole sentences. Thus, using the phrase “as of” (e.g., “I see the painting as of Wivenhoe Park” or “I experience the painting as of an apple”) is merely a means of illuminating both our ordinary pictorial experience and the common way we talk about pictures, and neither of these is at odds with how Gombrich
understands the act of pictorial perception or picture making. It is uncontentious that we
are able to describe a picture and an object in a similar manner, however it does not make
much sense to deduce from this that it is the use of a similar description that makes a
picture similar to an object (even if some degree of similarity does in fact exist)—which
is why I will not attempt to do so. Neither does this conception of relevant similarity fit
with the text of *Art and Illusion*. Instead we should be concerned with *why* we use similar
descriptions at all. Thus, if a picture is relevantly similar to an object we need to be
concerned with something of a more substantive nature, and this pushes us towards
considering whether the source of relevant similarity is something intrinsic to a picture
and its subject.

2.4 Relevant similarity: neither iconic or mimetic

What might a more substantive source of relevant similarity be? I have asked the
reader to entertain the claim (and to entertain the idea that Gombrich holds it as well) that
a picture is similar to an object in some relevant way; this, in turn, forces a further
question upon us: “How, or in what respects, is a picture similar to an object?” One
answer to this question is that there is something visually *in* the picture that is similar, in
some way, to something visually *in* the object. Choosing this route to explain relevant
similarity might push us towards the traditional theory of *mimesis*: a picture (as a
representation) is an imitation or a copy of an object, and the picture represents the object
because it resembles or looks like the object. For instance, in the *Sophist* Plato writes,
“And what shall we say of human art? Do we not make one house by the art of building,
and another by the art of drawing, which is a sort of dream created by man for those who
are awake?” (Plato, 2011b)53 And in the *Cratylus* Plato puts it this way: “Returning to the image of the picture, I would ask, How could any one ever compose a picture which would be like anything at all, if there were not pigments in nature which resembled the things imitated, and out of which the picture is composed? (Plato, 2011a)54 If this is the kind of explanation we want for relevant similarity it means that the relevant way in which a picture is similar to an object is in the picture’s resembling or imitating the visual appearance of the object.

This type of position is also echoed in Charles Peirce’s account of iconicity. Peirce writes that of the three kinds of signs that exist the first are “*likenesses, or icons*; which serve to convey ideas of the things they represent simply by imitating them.” (Peirce, 1894)55 An image serves as an icon for something because it is “similar in some respects to what it denotes”; it has something in common with its meaning or denotation, and so it is thought of as being “like that thing” (Peirce, 1955, p. 102).56 For example, an image of a horse is an icon of an actual horse (it is visually like an actual horse) because it refers to an actual horse “merely by virtue of characters of its own” (Peirce, 1955, p. 102). Peirce does not explicitly say (though I think we can safely infer) that what the picture and the object have in common is their shared visual likeness —the picture looks like the object.

The investigation thus far has left us with the following possibility: we can explain a picture’s being relevant similarity to an object by appealing to a shared visual resemblance. It is fairly obvious that Gombrich rejects this view: pictures cannot be truly mimetic in the way Plato describes nor iconic in the way Peirce describes, and this distances Gombrich from the notion that representation is explained by a picture’s merely
resembling the visual appearance (“the look”) of an object. (I still maintain that
Gombrich holds there to be some “connection” or “correspondence” between picture and
object, and I will talk about what this might be later in the chapter.) In his review of
Charles Morris’ Signs, language and behavior Gombrich (1949) criticizes the writer’s
insistence on the possible equivalence between pictures and language. (For instance,
Morris claims that the language of the arts [pictures] is translatable, unaided, into the
language of words; this would mean that the informational content of a picture could be
presented equally effectively, and without any informational lose, through the use of
words.) Gombrich further criticizes Morris’ claim that there exists “pure” or “objectively”
iconic images. As Gombrich writes, “[it may be demonstrated that] an image that passes
as highly iconic of reality is really less directly related to the object it denotes than it
would first appear.” (1949, p. 248) Gombrich is here separating himself from the view
that what ties a naturalistic picture to its subject is merely a shared visual appearance.

Gombrich goes on to say that the artist:

relies on the beholder’s capacity to read ‘iconicity’ into his sign. The contextual,
emotional, or formal means by which the type of interpretation is evoked or
facilitated—in other words, the relation between objective ‘iconicity’ [italics
added] and psychological projection—would have to form one of the main fields
of study of a descriptive semiotic of the image. Perhaps it will show that what has
been called the history of ‘seeing’ is really the history of a learning process
through which a socially coherent public was trained by the artist to respond in a
given manner to certain abbreviated signs. (Gombrich, 1949, p. 248)

One thing Gombrich will say is missing from such traditional mimetic or resemblance
accounts is what the viewer brings to the act of perception. These quotations show us the
origins of his insistence on the active role of viewer (the beholder’s share), which he later
develops in the essay “Meditations on a Hobby Horse or the Roots of Artistic
Form” (1996) and then in Art and Illusion (2000). In that later work Gombrich claims
that pictures consist of different “languages,” that they employ “cryptograms,” and that viewers must learn a picture’s “code” or “vocabulary” before being able to understand what the picture is of. From these claims it would appear that the picture is only “like” the object it represents to the extent that the viewer is able to read the painted code correctly (though what Gombrich means by these phrases is unclear). “Likeness,” it would seem, is not found in the picture itself but is instead projected into it by the viewer; it is only once we know the code, once we have learned to decipher the “cryptograms on canvas,” and thus once we have learned the pictorial language that we can know what the picture is of.

I believe Gombrich is correct in rejecting the simplistic notion of resemblance or iconicity (what he calls “objective iconicity”) as a means of explaining representation—no picture is capable of capturing everything a viewer perceives of the world. No picture bears perfect verisimilitude to reality, and thus no picture can look exactly like the object it represents. This is not merely because the artist lacks the skill to perform such an act of artistic magic, but because the very act of creating a picture involves certain limitations that make it impossible to attain such a goal. Gombrich refers to these limitations as the “limits of likeness.” Klaus Lepsky describes them this way: “The perfect representation of reality cannot be successful because these limits prevent the artist from including completely the complex optical information of the reality in his or her picture.” (Lepsky, 1996, p. 34) Pictures will always be deficient in capturing the range of rich information present in a visual motif. They will never be able to fully resemble their subjects and so, Gombrich will say, pictures will always need to contain conventional elements (certain pictorial devices) that compensate for the lack of a point-for-point match between an
object and its representation.

Gombrich outright dismisses the possibility of pictures being objectively iconic, and yet I can still interpret him as holding that a picture and object are similar in some relevant way. I can do this so long as my explication of “relevant similarity” diverges from the above naive conceptions of mimesis and resemblance held by Plato (and dismissed by Nelson Goodman); and so long as relevant similarity is separated from the kind of iconicity proposed by Peirce. Understanding relevant similarity to be different from these views does not mean that we spurn commonsense: it is true that we often feel a picture captures the likeness of its subject, for instance when we look at a Titian or a Leonardo and exclaim, “That looks so life-like” or, “The people look so realistic.” Our initial reaction is certainly understandable, yet we remain aware of the fact that the picture never looks exactly like its subject—the felt presence of an exact likeness can remain, and yet Gombrich will say that it is partly the result of our ability to read it into the picture. Gombrich’s idea of visual substitution (and by extension my idea of relevant similarity) is meant to excavate and explain how, in spite of a picture’s limitations, we are still able to see that a picture is of something or other (and, in seeing that a picture is of something or other affords us the ability to describe the picture in the “as of” terms mentioned previously).

I have already noted that the concept of relevant similarity does not rely on the traditional notions of mimesis or iconicity (here understood as identity or exactness of visual appearance), and I have avoided this path partly because of the kind of co-operative game played between the beholder and the artist just mentioned. “Any picture, by its very nature,” writes Gombrich “remains an appeal to the visual imagination; it
must be supplemented in order to be understood.” (Gombrich, 2000, p. 43) To create a picture is thus to create something that requires the input of the viewer, and yet the viewer does not give the picture the meaningful content that it has. I do not make the painting *Wivenhoe Park* a painting of Wivenhoe Park—it is a painting of Wivenhoe Park regardless of whether I project my interpretation into the paint, or whether I ever see the painting at all. So if the content of a picture exists independently of our visual interpretation of it we are left with the question of how, or in what respects, a picture is of the content it is of.

It is partly to address this problem that I have introduced the concept of relevant similarity. In the first chapter I said that in creating a picture an artist resorts to creating an equivalence. “Gombrich’s argument,” Klaus Lepsky asserts, “is that for the artist, the representation of reality is mainly a question of attaining an approximate adaptation [italics added] to the ideal of complete [verisimilitude].” (Lepsky, 1996, p. 34) And Gombrich, in the Preface to the Millennium Edition of *Art and Illusion*, states that there is “undeniable evidence that images can be approximated [italics added] to the experience of reality.” (Gombrich, 2000, p. xxvi) An image does not copy or replicate some subject, rather it serves as an equivalence or approximation for it. In order for the picture to attain this the artist must find those arrangements of visual elements or designs in the picture that the artist feels work best (given a particular medium) for depicting a certain subject. It seems that if I am to say in which respects a picture is similar to an object then those respects must be capable of approximating reality in the sense of being the kind of a visual equivalent Gombrich has in mind.
2.5 The respects of similarity

A critic may respond to the current investigation as follows: when we say of a portrait that it is “similar” to the person it depicts we cannot really mean what we are saying—we don’t mean that the pictures is similar to the object since no actual object exists that really looks like dabs of colored paint or smears of black charcoal. This criticism rings of Nelson Goodman’s denunciation of resemblance, which he presents straightaway in the opening pages of *Languages of Art* (1976). Recall that for Goodman representation involves the implementation of a symbol system, the core of which involves reference or denotation. There is “no degree of resemblance [that] is sufficient to establish the requisite relationship of reference”, writes Goodman (1976, p. 5); thus, as Tony Skillen describes Goodman’s position, there is no point in trying to fashion a “‘resemblance’, ‘copy’ or ‘likeness’ of some three-dimensional object” because when we claim that a picture “truly resembles” its object we are simply using “an idiom, a locution that should not mislead us into thinking that in a literal or serious sense the picture resembles the object.” (Skillen, 1996, p. 78) A picture does not resemble an object in any substantive way, and I think Goodman might extend his objection to cover my notion of relevant similarity: we cannot say that a picture is relevantly similar to an object in any substantive way, and my use of the phrase “reliably similar” is (like “resemblance”) merely jargon that fails to explain anything. Therefore, since a picture’s properties are always dissimilar to an object’s properties we should not assert that the picture is relevantly similar to the object unless we are prepared to seriously dilute the meaning of the word “similar.”

I agree that such objections have worked to scare individuals away from
considering resemblance theories of depiction, and it is tempting to think the same will work against the incorporation of relevant similarity into Gombrich’s theory. In truth, though, our fear is based on two mistaken assumptions. The first assumption holds that similarity must be *absolute*, and the second assumption holds that similarity must be *perfect*.

*Absolute similarity* between a picture P and an object O amounts to P and O being similar in some non-contextual, universal way. So saying that any P is absolutely similar to any O will thus depend on P and O always sharing some non-contextually specified respect or property F. That is, P and O will be absolutely similar just in case there is some respect or property F such that in a context C, P’s being absolutely similar to O always amounts to P and O both having F.

*Perfect similarity* between a picture P and an object O means that the respects in which P is similar to O (if such respects exist at all) do not admit of degrees—similarity is perfect in that it is either on or it is off. So, to say that P is perfectly similar to O is to say that P simply *is* similar to O with regards to some respect or property F, and that there is no sense in which P is or could be more or less similar to O. This idea of perfect similarity only assumes that there are *some* respects in which P is similar to O; it *does not* assume that P and O are similar in all respects. This latter assumption, the claim that any property that P has, O has too (and conversely), might be referred to as *perfect identity*, and I will mention it briefly below.

I find these assumptions problematic because they fail to describe the kinds of comparisons of similarity we normally make between objects in the real world. When we
say that a son looks like his mother we do not discount the claim as meaningless just
because the son is a male and the mother a female, or because the son is younger and the
mother older, or because the son is tall and the mother short. We can see that the son
looks like his mother despite the fact that some respects will be relevant to the question of
whether the son and the mother are alike while others will be irrelevant—similarity is not
absolute. Additionally, we often talk of things being more or less alike. “Doesn’t Billy
look just like his mother?” someone may ask, to which we might reply, “Perhaps a little
bit.” The son might be more like his mother with respect to shared facial features and less
like his mother with respect to his gait or demeanor; there can be similarity despite the
fact that those respects result in the son and the mother being alike to greater or lesser
degrees—similarity is not perfect. Gombrich astutely points out that “we are free to
categorise [sic] things in any number of ways and order things according to any quality
they may have in common…Moreover, in this ordering activity we can always specify in
which respect one thing is like another.” (Gombrich, 1972, p. 8) We can certainly choose
to specify in which respects the son is similar to the mother, focusing on these respects
rather than those, but the respects we do focus on will be neither absolute nor perfect. We
do not compare the son and the mother based on just any respects regardless of their
relevance to the question of how the son and the mother are similar (e.g., we may
discount gender or age but include eye color and demeanor). In addition, the similarity
that we see may be present in varying degrees, making the son and the mother more or
less similar depending on the respect(s) in question (e.g., the two may be more similar
with regards to their facial features or hair color than with regards to their posture or
gait).
Perhaps, however, pictures are different. Perhaps with pictures, to speak a bit colloquially, I am trying to compare apples to oranges. When we ask whether a picture is similar to an object we are, so the critic will say, comparing two things that are incommensurable—which, to be quite honest, is the real motivation behind these two assumptions. In other words, pictures have less in common with the objects they represents than they do with other pictures. This is Nelson Goodman’s position; but, if such a position invokes this kind of apples-to-oranges idiom (albeit in a more philosophical manner) then I believe it is being somewhat misleading. Obviously apples are not oranges, but that does not preclude us from finding a valid way to compare them—they are, in fact, both fruit. This means that there is a valid relevant respect in which we can compare apples and oranges despite there being any number of irrelevant respects for comparing them (e.g., asking whether they both exist). In the case of pictures it might be that a picture P is similar to an object O if P and O share some common respects, respects which are relevant to determining whether similarity exists. Yet, making the claim that a sketch of a horse is only vaguely similar to an actual horse, since horses are not flat and made of graphite, “seems to have violated background assumptions concerning which respects of the picture’s appearance are relevant [italics added].” (Neander, 1987, p. 214) The respects in which we say the son is similar to the mother or in which the apple is like the orange are limited to those relevant to a specific context (e.g., comparing the look of the one to the look of the other in the former case, and applying the label “fruit” in the latter case). Furthermore, I will argue that in an analogous way the respects in which a picture is similar to an object will be limited to those relevant to a specific context.
Let me rebuff the first assumption (absolute similarity) as it applies to pictures, and then turn to the second assumption (perfect similarity) in a moment. Our concern when comparing a picture and an object is with certain relevant respects, but we must still ask whether the relevant respect or respects are absolute or universal, and thus whether such a respect or respects can be found to apply in all cases and for all pictures. Assuming that such relevant respects exist at all it should be easy to see that the possibility of their being absolutely or universally relevant does not hold up to scrutiny, since the same respect can be relevant in one instance, and irrelevant in the next (Neander, 1987). For some pictures the relevant respect might be color; for others it might be lumination; for still others it might be shape. Even the very same subject matter can yield differing relevant respects when rendered in different mediums. An oil painting and a charcoal drawing can both depict a horse, but the respects in which the oil painting is relevantly similar to a horse can differ from the respects in which a charcoal drawing is relevantly similar to a horse: both the oil painting and the charcoal drawing can be similar to the depicted horse with respect to (e.g.) outline shape or the gradations of light (luminance), but the charcoal drawing will not be relevantly similar to a horse with respect to color in the way that the oil painting can be. It appears that a relevant respect, if one can be found, fails to be relevant for all pictures and in all cases: intensities of light, color value, shape, size—all of these may be relevant respects in which a picture can be similar to an object, but they are not universally relevant in all cases and for all pictures. So given the fact that any relevant respect can “shift about,” in this way, from context to context, we may wish to amend the statement “P is relevantly similar to O” to read something like “P is similar to O with regards to P and O sharing some non-universally or contextually
relevant respect or respects.”

The conclusion that we will not find any universally generalizable respects in which a picture is similar to an object should not prove fatal to my account of relevant similarity since the absence of any universally relevant respect or respects does not prove that there could not exist some relevant respect or respects indigenous to different pictures. Just as we can agree that there are certain respects relevant to confirming the conclusion that the son looks like his mother without assuming that those respects will always be the same respects for all cases of comparing mothers and sons, so too should we agree that there could be certain respects relevant to confirming the conclusion that a picture and an object are similar without assuming that those respects will always be the same respects in every case of comparing picture and object.

The second assumption is that similarity between a picture and an object is perfect, and thus the respects in which the picture and the object are similar do not admit of shades of similarity. This assumption states that while P and O may be similar in regards to certain respects those respects must be all-or-nothing, which means that P and O are not similar in degrees. There are a couple of points to make here: while it may prove effortless to see that a picture is of a woman, an elephant, or a rolling landscape, this fact certainly does not preclude us from admitting that some parts of the picture are not quite right. In fact, due to the nature of the limitations of likeness inherent in all pictorial representations, the relevant respects in which a picture and an object are similar to one another will only ever exist in a matter of degrees. We often find ourselves caught up in the presence of highly naturalistic pictures, for instance the portraits of Leonardo da Vinci or those of Chuck Close—pictures that offer us a heighten sense of reality—and it
is easy for us to assume that such pictures capture a higher degree of similarity with the subject than, say, a child’s finger painting. While Gombrich will say that the former are more “correct,” he will remind us that all pictures remain flat artifacts, and thus, the range of similar relevant respects between them and actual objects will always be quite small. 73

My claim, then, is that we are mistaken to think that the respects in which a picture is similar to an object are perfectly similar, if by this we mean that the respects of similarity do not admit of degrees. Let me briefly address the possible stronger assumption of perfect identity, which I mentioned above. This assumptions holds that to say a picture is similar to an object is to say that the picture and the object are similar in all respects, and thus any property that the picture has, so to will the object have; and conversely. This, I believe, is just an impossible assumption: no picture will be identical in all respects (relevant or not) to an object, and yet while “we know that pictures frequently do not look exactly like their subjects…to deny that perfect identity obtains everywhere and always is not to deny frequent and deep similarity [italics added].” (Gilman, 1988) Thus, I think we should reject the assumption of perfect identity in favor of the much more likely assumption that pictures can be more-or-less similar to their depicted objects with regards to certain relevant respects, a fact that we often verbalize when we say that something in the picture is “off” or “not quite right”: we see what the picture is of, and yet we recognize certain problems with it. (e.g., the artist may get the shape of the sitter’s head right but fail in the placement of the eyes; or we see that the outline shape and scale of the trees is correct but the lighting and shading is off).

Finally, to tie the issue of perfect identity back to the case of the mother and the son, even if the mother and the son are found to be similar to one another it is clearly false that they
are, or even could be, similar in every respect: for the mother and son to have all of the exact same properties would mean that the mother is the son and the son the mother, or that the one is a replica of the other (and conversely). Barring evidence of some bizarre scientific anomaly, this is clearly not the case; so just like with the picture and the object, mother and son do not bear perfect identity.

A further point to make regarding the rejection of perfect similarity is that the degrees of similarity between pictures and objects are not wholly arbitrary, which fits with Gombrich’s notion of approximation (also referred to as the notion of equivalence). In discussing this point Gombrich draws analogies between pictures and onomatopoeic words since such words share something in common with the sounds they stand for. 74 The words “quack-quack,” “bang-bang,” and “bzzzz” are not wholly arbitrary signs for the sounds made by a duck, a gun, and a bee, respectively. The individual phonemes of a word in a given language may be conventional (and thus might be described as noniconic), but this does not mean that the word itself has nothing in common with its referent. Onomatopoeic words are not “proper imitations,” if what we mean is that they are identical to or a copy of a certain sound—instead they are “approximations, within the given medium of language, to the sound heard [italics added].” (Gombrich, 2000, p. 362) So within the English language uttering “quack” approximates the sound a duck makes (the same with “bang” and a gun, and “bzzz” and a bee). For Gombrich, the idea of an approximation involves similarity by degree rather than similarity as all-or-nothing (i.e., perfect identity). Furthermore, it is possible that we can, in an analogous way, apply the case of onomatopoeic words to that of pictures: within the medium of language onomatopoeic words more-or-less approximate certain sounds, and within its medium75 a
picture can more-or-less approximate certain relevant visual respects of an object.

I mentioned the idea that the degrees of similarity are not arbitrary, and so it is with the visual aspects that an artist choses to include in the picture. For most representational pictures the included visual aspects will depend upon facts about how or what our visual system is able to ordinarily perceive. For instance, it is a fact that we cannot see around corners; it is also a fact that we can only see a certain range of light wavelength. These are facts about us as perceptual agents and about the kind of visual system we possess; as such, the idea of relevant similarity apropos of representational pictures that I have been developing—and which I am attributing to Gombrich—will be dependent upon a specific context, viz. the visual system that we, as humans, possess.

The account of relevant similarity I am proposing holds that a picture P and an object O can be similar in certain respects. Furthermore, I have tried to show that the respect or respects in which P and O are similar are neither absolute or universal (i.e., there is no single respect or respects in which all Ps are similar to all Os, since a respect or respects can relevant at one time and irrelevant at another), nor is it perfect (i.e., P and O can be more-or-less similar, similar to a degree). This account may yet prove to be acceptable, but we are again left with another question: which respect or respects are the relevant ones? To help answer this question let me introduce the work of Karl Bühler.

2.5.1 Which respects? Ernst Gombrich, Karl Bühler and the notion of relationstreue

To integrate relevant similarity adequately into Gombrich’s theory we need to
discuss Karl Bühler’s work. Gombrich does not cite Bühler in *Art and Illusion*, and there are only a few places where Gombrich mentions Bühler directly, yet, it is clear that Gombrich means to adopt Bühler’s notion of *relationstreue* (truth to relation) and apply it to his own account.\(^{76}\) For Gombrich a picture aims to be a relational model for the objects it represents, and it is my belief that by integrating the idea of the relational model into the account of relevant similarity I will be able to identify, at least partly, how an object and an picture are relevantly similar.

Above I said, quoting Neander, that the idea of resemblance is always relative to context, and from this I drew parallels to my conception of relevant similarity. We cannot hope to say how two things A and B are relevantly similar if we fail take into account certain facts about A and B that constrain which respects are relevant to the question of whether A is similar to B. Karl Bühler makes similar claims in his book *Theory of Language* (2011).\(^{77}\) According to him a particular symbol can only function successfully when it is seen against an appropriate background or within a specific context, what he calls “the appropriate symbolic field.” (Lepsky, 1996, p. 30) Take a symbol for a musical note: it will function properly when placed on a stave, but when placed on a map it leads to confusion or misinterpretation. What is important for Gombrich’s purposes is that Bühler sees a close relationship between the concept of the symbol and that of a pictorial representation. Gombrich writes that

Bühler…starts from the insight that there exists a spectrum, ranging from the extreme fidelity to nature exhibited by a waxwork, which (even so) resembles the model only relatively, to, for instance, a temperature chart, which merely records certain relationships in a given field. In between we find (if I may simplify and supplement his account a little), for instance, the notes of a musical score, the map, the landscape painting, and the illusionistic backdrop of the stage as different but equally valid systems of signs…What is at stake is the notion of
"relational fidelity," which is brilliantly explained in connection with black-and-white photography (Gombrich, 1984, p. 164).

Here Bühler’s representational spectrum involves two poles: at one end we find those representations that are more “symbolic” (exemplified by language or a temperature chart); and at the other we find those that are more “image-like” (exemplified by a completely realistic representation, e.g., the waxwork). 78 What might Bühler mean by an “image-like” representation? Above Gombrich describes these types of representations as exhibiting “extreme fidelity to nature,” and Lepsky points out that Bühler talks about them as having the “highest conceivable level of image-likeness” as well as “gradations of verisimilitude”. From these comments we can infer, to quote Lepsky, that Bühler “probably means nothing more than the fact that certain forms of representation correspond more closely to the represented object than others.” (1996, p. 31) While these descriptions do not settle the matter—what does it means for one representation to “correspond more closely to the represented object” or for it to exhibit “extreme fidelity to nature”?—it is safe to assume that Bühler is tying the idea of image-likeness to how things “look,” at least in a general sense.79

It is also quite telling that in describing the example of the temperature chart Gombrich mentions the role of relationships. This seems to imply that at least one important factor for something’s being a representation, whether it is a temperature chart or a picture, is that it be concerned with recording relationships; and in a very general way both a temperature chart and a picture do just this—both work to be relationally faithful to what they stand for. So while the pure symbol lies at one end of the spectrum and the perfect or complete representation lies at the other, both can be said to correspond to their represented objects (albeit in different ways). It is in the vast middle ground that
the majority of examples are to be found (including ordinary pictures), and they will exhibit the symbolic and the image-like to greater or lesser degrees. All representational forms along the spectrum will correspond to their respective objects, but as a representation moves towards the visual (the image-like end of the spectrum) it can be said to correspond more closely to an object because, as I hope to demonstrate, it will exhibit a higher degree of visual relational faithfulness.

If a depictive picture is of something, it is a truism to say that it should be like what it is of in some way: why, for instance, should any artist create a portrait if no one can recognize the end result? Why paint an elephant or an apple or Wivenhoe Park if no one can tell what the end result looks like? As already noted, theorists such as Goodman (1976) may wish to discount the idea of similarity (or resemblance or the idea of “looking like”) between picture and object as being a necessary condition for depiction, but it remains that to tell what a depictive picture is of we must look at it. This is simply the nature of pictorial representation: if P is supposed to be a representation of O but P looks nothing like O, and thus no one looking at P would say that P is a representation of O because no one would recognize P as a picture of O, then I believe P will fail as a representation of O. There is no controversy in saying that the intended result of creating a depictive picture P is that, at the very least, a viewer should be able to tell by looking at P that it is of something. Not every viewer will get it right—some will say that P is of Q rather than O—but often this just means backing up and trying another interpretation.

In applying Bühler’s ideas, I am suggesting that the fact that P can be similar to O is not because there is an identical, point-for-point matching of the individual visual elements of P and O, but instead because the relations of visual elements in P match the
relations of visual elements seen in O (e.g., variations in color value and tone, or arrangements of lines responsible for size and outline shape) in such a way that P can trigger a similar response from our visual system to that triggered off by O. The idea is that there is some set of visual elements \([g_1 \ldots g_n]\) in P that is in a certain relation \(R\). There is also a set of visual elements \([f_1 \ldots f_n]\) in O that is in a certain relation \(S\). \(R[g_1 \ldots g_n]\) in P “matches” \(S[f_1 \ldots f_n]\) in O in the sense that seeing \(R[g_1 \ldots g_n]\) in P triggers off a sufficiently similar, though not identical, visual response to seeing \(S[f_1 \ldots f_n]\) in O.\(^{82}\) The capturing of these kinds of relationships in a representation is what Bühler means by \textit{relationstreue}, and a representation that exhibits \textit{relationstreue} (and is thus a relational model) will render, in two-dimensions, a similar set of relations of similarities and dissimilarities between visual elements found in nature.

Yet we should remember that a picture “is not a mere transcript of nature but a \textit{transposition} [italics added] of light into paint” (Gombrich, 2000, p. 252), meaning that the “artist cannot copy… but he can suggest” or remind us (Gombrich, 2000, p. 38). As I mentioned before, the history of art shows us the lengths artists went to push beyond and expand the range of idealistic goals: creating an image that is not only a likeness of nature, but which might also rival creation itself.\(^{83}\) Recall the threatening curse Donatello hurled at his \textit{Zuccone}: “Speak, damn you, speak!”\(^{84}\) This aim for perfection, however, can never be achieved—the statute remains still; the waxwork is still made of wax; the portrait fails to capture all that is true of the sitter. The notion of \textit{relationstreue} is meant to explain how, despite these facts, the representation (e.g., the picture) is connected to the subject and allows the viewer to see it as of the subject: the individual visual elements in a picture can be quite visually unlike those in the subject, and yet should the picture
capture the relations of similarities and dissimilarities between the visual elements of the
subject the picture will produce a correct representation of it. The importance of
relationships is confirmed by Gombrich when he writes that, “What matters to us is that
the correct portrait…is not the faithful record of a visual experience but the faithful
construction of a relational model [italics added].” (Gombrich, 2000, p. 90) The
successful depictive picture will not capture the exact visual likeness or the identical
individual visual elements of object; nor can it trigger off the exact same visual
experience as one would have when seeing the object itself (we would think the picture is
the object). What a depictive picture can do is set out to preserve the relations of
similarities and dissimilarities between visual elements found in an object, and present
them in such a way that they serve, for our visual system, as a visual equivalent to the
depicted object.

If pictures are not concerned with copying or with “the faithful record of a visual
experience,” and if they can only go so far towards preserving the ideal of perfect
verisimilitude before the brush or the pencil is thrown down in frustration, what are the
reliable respects in which a picture is similar to an object? The answer that Gombrich
seems to be offering in the above quote is that the relevant respects are whichever ones
are responsible for a picture’s being a relational model. But what does it mean to be
relationally faithful, or to be a relational model, and which specific respects are
responsible for relational faithfulness? The best illustration is found in Bühler’s
application of his theory to black-and-white photography. As Karl Lepsky describes it:

Bühler impressively demonstrates that the reproduction of objects by transposition
into black and white is only accepted as ‘vivid’ when the relations of grey tones of
the photograph correspond to relations in the represented object. Naturally there
exist numerous representational variants, for example because of different exposure of processing of the photograph, but besides these differences it is true that these pictures are “correct representations of the objects in the sense we have determined: whenever one point on the picture is more white than another the same is true for the albedo values (grey values) of the object (but not with the same steps)” (Lepsky, 1996, p. 32).

Not only does this example lend credence to my claim that it is the relations between visual elements which are important in explaining how a picture is similar to an object, but it also offers us at least one concrete example of a relational model: the three-toned modeling (which Gombrich also mentions in Art and Illusion). More importantly, we have an illustration of which specific relations of visual elements might be important for a picture’s being similar to an object: namely, the relational changes in light (described as tones and values). (I will later discuss the rather cryptic assertion that light values in the picture and the object are the same “but not with the same steps.”)

Let me mention two final items as further evidence for the importance of relationships for our visual system: the first comes from one of Gombrich’s ethnological examples, the second from his example of representing light falling on a white handkerchief. Humans are not the only creatures that find their way around the world with a visual system attuned to relationships. Gombrich’s example of how chickens “respond to light intervals, to what have been called ‘gradients’” illustrates the point as well (2000, p. 46). When chickens are fed they can be taught to expect food on the brighter of two grey pieces of paper: they see the brighter paper and they come to get their food.

If you then remove the darker [paper] and replace it by one brighter than the other one, the deluded creatures will look for their dinner, not on the identical gray paper where they have always found it, but on the paper where they would expect it in terms of relationships—that is, on the brighter of the two (2000, p. 46).
Humans, it would seem, are subject to a similar dependency upon relationships in perception. Facts about “constancy” in perception demonstrate our ability to remain relatively impervious to the variations and changes that occur in the world around us. Despite changes in distance, illumination, angle of vision, etc., the shape, color, and brightness of things remain to us relatively constant. (I will talk more about the issue of perceptual constancy in Chapter 3.)

The second point I wish to mention involves Gombrich’s example of rendering a white handkerchief. The local color of the white handkerchief will need to be represented by the highest intensity of lightness available in the artist’s palette, and the artist’s task is to find a way of representing the relative intensity of light falling on the handkerchief. Thus, what matters on Gombrich’s account is, as Karen Neander puts it, not the “direct correspondence of local colour, but [rather] maintaining the same pattern of similarities and dissimilarities.” (1987, p. 220) The artist does not try to directly match the individual pigments of color to the individual local color of the subject. Beyond the limited range of colors available to the artist, there remains the fact that the “white handkerchief [seen] in the shade may be objectively darker than a lump of coal [seen] in the sunshine.” (Gombrich, 2000, p. 52) Instead, the artist must ensure that “the pattern of similarities and dissimilarities…be carefully contrived to achieve a matching of qualities along their respective scales.” (Neander, 1987, p. 220) This is made clearer in the case of the Adelson Checkerboard, which serves to highlight both the issue of constancy and the difference between focusing on individual visual elements (i.e., the local color) and focusing on the relationships between visual elements. In the image, the local values of the check in light (A) and the check in shadow (B) are identical, and so we might say that
they are identical *individual* visual elements. The context in which A and B are seen, however, that is the relationship each check bears to the surrounding color values, makes each appear different. So while each individual visual element (each individual color value) is constitutive of the visual experience of perceiving the scene as a whole, it is the relationships *between the* color values of the checks that is significant to our visual system because it is those relationships that our visual system focuses on.

What we see here is that when the artist preserves (in the picture) the patterns of relationships seen in the world he or she is undertaking a similar sort of process to our mind’s ability to maintain color constancy in ordinary perception. Again, the local color of, say, an orange changes when we see it in bright sunlight, in candlelight, or in flashlight. The lightness and darkness of the local color (the color value) of the orange constantly changes, but the mind’s ability to discount this continual changing allows us to not only recognize the object (e.g., an orange) but to specify the object’s color throughout these different viewing conditions. We understand that the immediate, local color of the orange will change depending up the relationships established across changing lumination levels—shift the direction of the light in the image of the checkerboard and the relationships between the individual checks will change as well. In picture-making, then, we see that the artist attempts to render certain relationships between visual elements that correspond to certain relationships between visual elements found in an object. In doing this, the rendered relationships in the picture are meant to preserve a similar pattern of similarities and dissimilarities to those relationships found in the represented object; by preserving these relationships the picture serves as a relational model for the object.
2.6 Final thoughts on relevant similarity as a relational model

According to Gombrich picture-making involves creating visual substitutes that are meant to stand-in, visually, for some object or scene. Relevant similarity was introduced to help better explain how this works. Initially, I stated that we might be tempted to think that a picture P is a visual substitute for an object O simply because P and O share some property or set of properties. I showed that this account was too vague to serve as a sufficient condition for P’s depicting O, but that it can serve as what I called a core necessary condition for P’s depicting O. If we begin the discussion by focusing on certain central or core examples of depictive picture we can see that P can be a visual substitute for O because P and O have in common certain relevant respects; and from there we can expand our search out to locate where the fringes of visual substitution lie.

But what are the respects in which a picture is similar to an object? In section 2.3, and previously in Chapter 1, I pointed out that we commonly describe an object and a picture of that object using similar words and even whole sentences. An easy way to explain this linguistic similarity is to say that we see a picture “as of” (or “in terms of”) its subject. This account was inadequate because thinking of relevant similarity in terms of what is given via a mere descriptive linguistic tool failed to capture the substantive nature of the concept. The painting Wivenhoe Park may be a visual substitute for the actual park in Essex, and it may be true that we describe the picture and the object using similar language (“I see a manor house”, “I am having an experience as of field of cows” and so on), but the fact that the picture is a visual substitute and the fact that we do use similar language does not itself make the picture similar in any relevant way to the actual
park, even if, as I think, that fact does ultimately turn on the presence of some sort of underlying similarities.

In section 2.4 I rejected the possibility that the respects in which picture and object are similar are found in a picture and an object sharing identical visual elements (or visual properties or visual aspects). No picture can adequately embody the range of rich information found in the depicted object, and point-for-point identity between picture and object is impossible. No amount of artistic skill can overcome the limitations of likeness that are inherent in all pictorial representations—no amount of effort can achieve perfect verisimilitude. What is more, I claimed that the respects of similarity are neither absolute nor perfect. They are not absolute since not only will certain respects be relevant in one case and irrelevant in another, but equally there are no single universal respects that hold in all cases and for all pictures in which a picture is similar to an object. Furthermore, it is possible for a picture to be more-or-less similar to and object, meaning that the respects of similarity are not perfect. Thus, similarity can be present to varying degrees, and we can make sense of the claim that a picture is similar to an object despite the fact that the picture and the object are not similar in every respect.

As such, we are still left with the question of which respects underlie any similarity between picture and object. In sections 2.5 and 2.5.1 I took up this question and proposed that the respects we should be concerned with are the relations among the visual elements found in the object. (In Chapter 3 I will show that these relations serve to afford recognition of the represented object.) In other words, we should look at whether the picture accurately captures an arrangement and pattern of similarities and dissimilarities among visual elements that is similar to that found in the object. Not being
identical to those found in the object, the relations of visual elements rendered on the page serve as an equivalent (within the limitations of the medium) to those relations of visual elements found in the object. A picture that successfully preserves these relations can be said to function as a *relational model* for the object it represents, and it is creating such a model that is the task of the artist wishing to produce a correct or successful depictive picture of some motif.

As I have already stated, this account of relevant similarity needs to be fleshed out further. Chapter 3 will remedy this by introducing the idea that a picture which preserves the relations among visual elements found in some actual object will also afford recognition of that object in the sense that it (the picture) engages our visual system in such a way that a similar recognitional experience is triggered off.
3 Visual Substitution and Relevant Similarity: Affording Recognition

3.1 The goal of Chapter 3

In this chapter I will complete the discussion of relevant similarity by turning to the idea that pictures afford recognition of its subject. The concept of affordance originally comes from J.J. Gibson (1986). My use of affordance breaks with Gibson’s in that I seek to extend the term to include visual affordances.87 In extending and adapting the term I hope to show how it can help describe the ways in which a picture is relevantly similar to some object O, and how it is that we are able to recognize a picture as of O. There is good textual evidence from Gombrich’s later works that he viewed recognition as a significant feature of depictive picturing, and I hope to show that such an account fits well with the idea of a picture as a relational model. The recognition of pictured scenes normally occurs without much effort—we often find that we experience instant and effortless recognition of a object represented by the picture88—and I believe this is because the visual processes used for recognizing objects in the real world do not wholly differ from those used for recognizing pictures of those objects. Pictures trigger off a recognitional experience that is similar to that triggered off by the object. There is always the possibility that we can mis-recognize what a picture is of, yet most ordinary depictive pictures will exhibit some degree of relational fidelity to a depicted object. If this is the case, and we can say that a picture serves as a relational model in the manner I have previously described, then that picture will afford (though it certainly will not guarantee) recognition of its subject matter.
Given my discussion of the relational model in the previous chapter and the exploration of the role of recognition in this one, in the final sections of this chapter (3.4 through 3.4.4) I will put forward an empirical hypothesis that one possible set of relations of visual elements—a picture’s possession of which constitutes or at least implies its relevant similarity to the object it represents—is the set of relations that are responsible for perceptual constancy in our perception of the external world. For instance, visual aspects such as color value, size and shape are examples of relations of visual elements, responsible for perceptual constancy in our ordinary perception, that the artist can choose to capture in the picture. My claim will be that the presence of these relations of visual elements in the picture triggers off a perceptual reaction (i.e., a recognitional response) similar to that of seeing corresponding relations of visual elements in the real world. We can recognize an object in the actual world under varied, changing conditions—especially when we encounter certain perceptually constant features that allow for a constant awareness of the object; similarly, we can recognize a picture of an object (which is just another varied, changing condition) when we encounter certain equivalent perceptually constant features present in the picture. While recognition may not require the perception of an object’s “usual” perceptually constant features (e.g., recognition can occur when we see an object from an unusual viewpoint, or when we see odd or uncharacteristic visual aspects of an object), should we encounter the sorts of visual features that are responsible for perceptual constancy (usual or not) they will serve to more easily afford us recognition of both actual and pictured objects. Furthermore, in the case of perceiving the pictured perceptually constant features we can say that the picture in question serves as a relational model for the depicted object. Thus, I hope to show that in that case the picture
functions as a visual substitute for the depicted object in the way that Gombrich intended.

In the end, Chapter 2 and Chapter 3 stand as an attempt to fit the concept of relevant similarity into an already existing theory, and specifically into a certain part of that already existing theory: the idea of substitution. My claim is that a depictive picture serves as a visual substitute for an object. As such, the evidence presented can, at best, point towards a convincing interpretation of the concept of relevant similarity. The goal is not to present an account that is unassailable. Instead, I wish to offer a version that pushes readers towards a better understanding of Gombrich’s theory while at the same time filling in those places in which Gombrich is silent. If in doing this new paths of exploration emerge and insightful questions can be asked then I take it the process of developing a coherent version of relevant similarity will have been successful.

3.2 The importance of recognition

So far I have said that the concept of relevant similarity does not rest on the idea of comparing a picture and an object in all respects, but instead depends on its capturing certain relevant respects, viz. the relations between visual elements found in an object. This account of relevant similarity fits the text of Art and Illusion and, I will claim, serves to help explain what Gombrich means by visual substitution. What needs to be added, however, is a discussion of the role that recognition plays in the making and perception of pictures. I will need to explain what it means for a picture to afford recognition of something else, and then explain how this fits with both the concept of relevant similarity and, more generally, a Gombrichian theory of depiction.
Gombrich does not explicitly discuss the importance of recognition in *Art and Illusion*, though there are places where he alludes to it. For instance, Gombrich writes that an artist must create a picture that “works” for the audience of viewers, which I take to mean that a viewer is able to see what the picture is of. If we read this in the light of Gombrich’s later explicit insistence on the importance of recognition we can say that a picture that “works” is a picture that affords recognition of its subject.

An affordance can be thought of as a part of an object that “offers,” “provides,” and “furnishes” something to an user (Gibson, 1986, p. 127). In this case the claim is that a picture offers or provides a viewer recognition as of it subject. I do not take it as a demerit to integrate the notion of recognition into Gombrich’s understanding of depiction. On the contrary, to neglect the importance of recognition is to have a deficient understanding of Gombrich’s broader understanding of pictures.

Let me begin by describing Gombrich’s understanding of recognition in ordinary perception. This will allow me to illustrate, again, Gombrich’s insistence that the perception of pictures draws upon many of the same processes as the perception of objects—how we recognize the object a picture represents is not wholly different from how we recognize the object itself. I will also demonstrate how Gombrich’s understanding of recognition fits with, and completes, the concept of relevant similarity that I have been developing as a means for explaining visual substitution. I will argue that an object O and a picture of O are similar in that the presence of certain reliably relevant respects rendered in P affords recognition of O. In seeing a picture P of O a viewer can recognize P as of O—the viewer can see P as, or in, O—all the while knowing that his or her perceptual experience is that of looking at a picture and not at the object itself.
Recognizing P as of O does not mean that I literally recognize P as the actual O; instead, we should understand it as describing the very common fact that we have the ability to recognize a flat, marked surface in terms of some robust reality.

3.2.1 Recognizing objects and recognizing their pictures

For Gombrich perception is an act of interpretation. It involves an active process of interpreting and classifying—a kind of “exploratory activity” (Nam-Gyoon, 2001, p. 988). Our perceptual system is tuned to respond to the environment, and we respond in an immediate and direct way to the stimuli we encounter. This means that perception, so understood, does not always involve the performance of a conscious or intentional act. As we perceive we classify and sort perceived stimuli; and once the stimuli become familiar and have been sorted into “slots” for easy classification, recognition can occur almost instantaneously. If I see a particular shape, recognizing it as an instance of O will seem to occur immediately and in direct response to seeing that particular shape in my field of vision; thus, to determine that what I am perceiving is indeed an O all I have to do is perceive it, and based on my perceiving I can tell whether it is an O or not. Recognition, then, is not the result of consciously piecing together the various parts of my perception like a jigsaw puzzle; nor does it involve building-up meaning from various clues in my field of vision. Recognition occurs directly and immediately as I perceive, and it is not decomposable into conscious parts. Recognition is an essential part of perceiving the external world, allowing us to quickly make sense of our surroundings and picking out those objects that are familiar. In many cases the “recognition of meaningful features” is
essential for our survival, something seen in other animals as well (Gombrich, 1982, p. 285). We have adapted to scan our surroundings, seeking things to avoid or pursue, and in recognizing something we see it as familiar. If we were never able to recognize the content of our perceptions then each time we looked around it would be as if we were starting from scratch—nothing would ever be familiar.

From this discussion of recognition in ordinary perception we can now turn to how Gombrich incorporates it into the activity of picturing and pictorial perception. I have asserted that for Gombrich, recognizing an object and recognizing a picture of that object do not result from wholly different perceptual abilities. That is, the perceptual processes and mechanism used in the perception of objects in the real world are also those perceptual processes and mechanisms used in the perception of pictures of those objects. While Gombrich is somewhat unclear on this position in Art and Illusion he does make a rather evident set of remarks in the article “The Evidence of Images.” (1969) Gombrich, nor I, wish to assert a direct identity between how we look at the world and how we look at pictures of it, but it would be wrong to overlook the fact that there is a great deal of overlap between the two. To show that Gombrich is not alone in presenting this position let me briefly mention a number of other theorists who make similar claims.

3.2.2 Other theorists who advocate similarity between ordinary and pictorial perception

The idea that looking at pictures and looking at objects are not wholly distinct perceptual activities is a position held, in various forms, by Flint Schier (1986), Daniel Gilman (1988, 1992), Dominic Lopes (1996, 2005) and, to some extent, Kendall Walton
The first three seem to explicitly hold to some form of a recognitional theory of depiction, which would mean that pictures “embody information from their objects on the basis of which these can be recognized: they have been designed to trigger roughly the same visual processing that allows a perceiver to recognize the object represented when seen face to face.” (Batinaki, 2009, p. 239) One of the theorists who best exemplifies this position is Flint Schier. For him, looking at pictures depends on mechanisms or abilities used for looking at objects in the world. He argues for his theory directly from facts about our ordinary visual processes, flatly asserting that, “if you can see it, chances are you can see it in pictures.” (Schier, 1986, p. 43) Significantly, according to Schier the recognition of pictures only requires our ordinary perceptual abilities. I believe this is important for my discussion of recognition because Schier’s account points to the possibility that the possession of ordinary perceptual capacities to recognize an object also plays a significant role in recognizing pictures of that object.

Currently, the most thorough going recognition account of depiction can be credited to Dominic Lopes, who at one point writes that “the ability to work out what pictures depict covaries with the ability to recognize their depicta in the flesh.” (2005, p. 170) The details of Lopes’ account are not important for us. What one should be aware of is that while ordinary recognition is not identical to pictorial recognition, “The recognition skills we bring to pictures depend on and extend the dynamic recognition skills exercised in ordinary perception.” (Lopes, 1996, p. 149) For Lopes, recognition is basic to all pictures and while we may not be able to recognize every picture’s subject Lopes believes that we can clearly see that the importance and significance of the skills used in recognizing picture’s comes, in part, from the skills used in recognizing objects in
the actual world. The process of recognizing what a picture represents “exploits perceptual recognition skills…[Furthermore] this is not simply that pictorial representation makes use of recognition abilities for objects, kinds, and properties. Rather, the ability to recognize pictures’ subjects is an extension of the dynamism of recognition.” (1996, p. 144) Thus, for Lopes acquiring the recognitional-capacity to recognize picture’s of O will depend, at some level, on having the recognitional-capacity to recognize actual O’s.

Another writer who draws comparisons between how we see pictures and how we see objects in the world is Daniel Gilman. At one point Gilman asserts that the flat and immobile pigments on a page can, in fact, reasonably imitate features found in the world, and this is so because our visual system is sensitive to a similarity that exists between the two. To understand what this similarity is, and to gain a fuller understanding of the workings of pictorial representation, Gilman says it is of the utmost importance that we understand the workings of ordinary vision (Gilman, 1988). Given that our visual system is sensitive to a picture’s designs in a way similar to how our visual system is sensitive to features found in the world, it is safe to assume that Gilman sees similarities between how we look at an object and how we look at a picture of that object. Much of the position in his doctoral dissertation (1988) rests on details of the biology and physiology at the retinal level, which leads him to conclude that, “a surface and a picture of that surface can generate retinal images that are identical in relevant respects.” (Gilman, 1988) This means that “we might fashion a perspective on a scene, and a picture of that scene, and a description of a collection of neural states at the level of the retina (constituting a retinal image) in such a way that the retinal image might just as well be
Gilman further affirms the possibility that looking at pictures and looking at objects are not dissimilar perceptual activities in a later article entitled “A new perspective on pictorial representation” (1992). There he defends a position on pictorial representation held by many a non-philosophical reader. His position states that “pictures simply resemble their subject and…we are able to see the world through pictures in whatever ever manner we are able to see at all [italics added].” In other words, the kind of “natural” seeing we do when we look at actual objects in the world is also used when we look at pictures of those objects. This position is, Gilman claims, more or less correct: “there are facts about us, as human beings, and facts about the world, that prompt us to recognize the world in certain sorts of pictures.” (1992, p. 174) We are able to do this because “we react in similar fashion to both,” and our reaction is due, he claims, to the fact that a painting and an object “impinge on our sensory apparatus in the same way.” (Gilman, 1992, p. 184)

The last theorist I wish to discuss is Kendall Walton. In Mimesis as Make-Believe (1990) Walton points to a number of “analogies” between looking at objects in the actual world and looking at pictures of them: both kinds of looking are visual; both can yield corresponding kinds of information about certain visual characteristics of “the world of the picture” in one case and of “the real world” in the other (e.g., the color and size of a mill’s roof, the presence of peasants or farm animals, and so on); and in both kinds of looking there is “a correspondence in the order” in which an observer acquires information, and as such the sequence is “realistic”—general information can be acquired at a glance (the roof is red, there is a peasant in the field), while more detailed
information requires greater, or closer scrutiny (the tool in the peasant’s hands is a hoe, there is a woman standing in the doorway) (Walton 1990, p. 305). At another point Walton asserts, in response to Richard Wollheim’s case of a scrambled picture, that if “reading” an ordinary picture were the type of inferential process required in order to understand such scrambled pictures we, “[would] not see what is represented in the representation.” (Walton, 1990, p. 311) This is because normally we just look and see that something is in the picture, and if “we could not do this more or less automatically, the canvas would not be a picture for us.” (Walton, 1990, p. 311)

This does not prove that how we “inspect” (to use Walton’s word) or look at a picture is in fact congruent with how we “inspect” or look at an object itself, but it does offer evidence that the two types of perceptual activities might bear similarities.

There are a number of details in the positions of these three theorists that I do not wish to adopt; nor do I think that Gombrich would agree with everything they have to say. My only goal in citing them has been to bolster the claim, which I believe Gombrich holds as well, that when we look at a picture we are making use of similar perceptual processes to those used when we look at an object itself.

3.2.3 More thoughts on recognizing objects and their pictures

Let me briefly return to the discussion of recognition by mentioning Gombrich’s article “Visual Discovery Through Art” (1982). There he tells us that Art and Illusion was a set of investigations that aimed at exploring “the relation between visual perception and pictorial representation.” (Gombrich, 1982, 12) He then goes on to assert that if he could
start *Art and Illusion* again he “would pivot the argument on the distinction between recall and recognition.” (Gombrich, 1982, p. 12) “Recognition,” he writes, “is easy, it is almost automatic”—we see a picture of a bicycle and it will often trigger effortless recognition (Gombrich, 1982, p. 12); on the other hand, trying to recall the exact placement of a bicycle’s chain in relation to its tires can prove challenging.  

One device that aids in the recognition of a picture is the artist’s use of the laws of perspective. A picture that includes perspective “will generally evoke instant and effortless recognition. It will do so to such an extent that it will in fact restore the feeling of reality, including—and this is most important—the constancies.” (Gombrich, 1982, p. 19) When we look at a building with a receding row of columns, those columns that are further away will look to be much smaller than the columns closer to us (should we hold up our finger and measure their perceptible size from where we are standing); yet we know that the columns are in fact the same size and height, and that they will only appear to diminish in size as they move farther away from us—they are all the same size.  

The pictured column-representing-lines have a similar relation to each other as do the real columns in the actual world. That is, the relations between the pictured columns-representing-lines are similar to the perceived relations between the visual aspects of the real columns in the actual world. The relations preserved in the picture by the column-representing-lines are those relations one would perceive the real columns as having if they were seen from some stationary point of view. We know that when seeing the real columns they remain the same size and height even though they appear to grow smaller and smaller. Using these facts about how perspective works in the actual world and then employing them in a picture can create a similar kind of effect on our visual system.
Since the pictured columns have been depicted from a single stationary point of view (much like how real columns would be seen in the actual world) the experience of seeing the pictured columns impinges on our visual system in such a way that a similar visual reaction is triggered off. We have a feeling, as Gombrich says above, that the picture restores a sense of reality, and so we see the picture as of columns receding into the distance.

To create a depictive picture is to create an artifact that is of something or other, and if a picture is to successfully depict its subject it must be recognized as being of its subject. It is easy to forget that the artist stands in the unique position of being both creator and viewer (first viewer, in fact) of a picture. The artist creates a picture through trial and error, and even though the artist “cannot foretell exactly what [pictorial] device may have the desired effect, he can judge whether he recognizes the effect in his picture.” (Gombrich, 1982, p. 162) If an artist fails to recognize the effect of, say, perspective in a painting of a building (perhaps the angles are wrong), and thus he fails to see the arrangement of paint as a building, he must choose to either re-work the picture so that it does afford him recognition of a building or else he must re-evaluate the task of attempting to depict a building at all. In using certain devices to construct the picture the artist attempts to arrange the relations of visual elements on the page in such a way that they are visually equivalent to the ways in which such relations of elements impact our visual system when looking at objects in the actual world. The artist plays around with these kinds of pictorial devices, trying to achieve different effects, and ultimately trying to create a picture that affords recognition—creating an artifact that can be recognized as of something or other beyond the constitutive marks on the page.
Let me offer a few more remarks about the type of recognition that occurs in the sorts of depictive pictures I have in mind. If a picture is to be a representation of something the artist will actively scrutinize it to determine whether a viewer (beginning with the artist himself or herself) can see the subject as being depicted in the picture. In other words, if the artist finds that the picture is “not working,” if she sees that aspects of O are wrong or if she simply does not see O in P and thus does not recognize P as O, she will start again.105 But if the picture passes the test of recognition then the picture will have been recognized as being of something or other and should (though there is no guarantee) be recognized by other viewers as being of that which it depicts.

Let us now imagine that an artist makes another picture (Q) that is simply a random tangled-arrangements of lines, shapes, and colors. The artist does not intend Q to be a representation of anything, and thus she does not recognize Q as being of anything in particular. A viewer or viewers may nevertheless recognize Q as of something (e.g., a sunset, a lion, or a nude woman) despite the fact that it was not intended to be recognized as of anything. In this second case we have recognition but not depiction, and this should demonstrate that recognition may not be a sufficient condition for depiction. Even so, this need not trouble my argument since I have not been arguing that there are any sufficient conditions for depiction. That being said, I remain hesitant to say that in this second case the viewer recognizes Q as of something or other O at all, at least not in the same way that he or she recognizes P as of something or other O.106 While we may be able to see a sunset, a lion, or a nude woman in a tangled-arrangement of lines, and thus we may wish to say that in seeing the picture we recognize a sunset, a lion, or a nude woman, such recognition should rather be understood as being indirect (or as involving a kind of
“symbolic recognition”). Whatever it consists in, my claim is that such recognition is different than the ordinary type of recognition that I have been discussing above. Much like how Walton describes Wollheim’s scrambled picture, we may need to learn how to see an object in pictures of this sort; we may even get very good at recognizing faces in scrambled or tangled-lines pictures, but this type of recognition is, I wish to claim, markedly different from how we recognize objects in ordinary depictive pictures. I believe this is because the arrangement of relations of visual elements in ordinary depictive pictures (the designs constitutive of the picture), as opposed to those arrangements of visual elements in tangled-line or scrambled pictures, afford us the type of recognition that is more similar to how we recognize actual objects. One possible reason for this is that the pictorial devices used in constructing the picture have been chosen for a specific reason, namely that they successfully activate our perceptual apparatus in a similar way to how it is activated in real world processing (Casati and Pignocchi, 2007). If this is so, then we might say that ordinary pictures afford viewers a similar kind of recognitional act to that of looking at objects in the actual world because the designs constitutive of the picture (i.e., the rendered relations of visual elements in the picture) are able to engage our visual system in a similar manner to how it is engaged by looking at objects in the actual world.

Of course, even if we assume that someone can recognize that a picture is of something or other this does not guarantee that someone will recognize it as something or other. There will be cases where recognition fails and we either do not see that the picture is of anything, or else “false recognition” occurs and we see the picture as being of something that it is not. This kind of failure is not limited to the fringes of figurative
pictures—recognitional failure can occur when looking at even the most realistic of images: are those birds flying out of the distant trees in *Wivenhoe Park*? Is that an orange resting at the base of the pedestal in Manet’s *Young Lady in 1886*? We may be in doubt as to whether we are able to recognize certain parts of a picture, and there is no doubt that the individual constitutive marks on the picture’s surface look quite unlike a swan or an orange—yet Gombrich’s claim is that when we look at an ordinary picture we have a “global impression” of the represented motif (i.e., we see the “overall” content or the primary focus of the picture or something like this). Having such a global impression of a picture’s content leads me to conclude that we are able to recognize what a picture is of (i.e., its subject matter) despite the elements of the picture being visually dissimilar to the subject and even if we fail to recognize every single constitutive part of the picture.

Having now discussed the type of recognition that occurs when we see pictures, let me make a brief remark on ordinary perception. I have said that Gombrich takes ordinary recognition to be a direct, immediate response to seeing some object in our visual field: we look and if what we perceive fits a classificatory pigeonhole then recognition occurs. Should we find ourselves unable to recognize something we may simply need to alter our viewing position (step to the left or right and we recognize the shape as a mailbox not a large child), or else construct a new or alter an existing slot by which to classify the unfamiliar object. Walking through the grocery store I am able to recognize apples and oranges, cucumbers and peppers, boxes of cereal and other various sundries. I might be mistaken and think that I recognize something as an orange when in fact it is a kumquat, but this just means that recognition, like perception, is not infallible. Recognition is always based on the available evidence within a certain viewing context,
and there is always the possibility that we can mis-recognize or mis-perceive—we just have to stop, look again, and then make the necessary changes.

It is obvious that things are a bit different for pictures. A picture is a uniquely different perceptual object from, say, a rowboat or an elephant. We can physically move around the rowboat or the elephant in a way that we cannot move around the objects in the pictured space. While I can alter my viewing position of a row boat or an elephant to get a better view or to see a different aspect, altering my viewing position of a picture of a row boat or a picture of an elephant will not change the position from which either one is depicted by the picture. All that will change is the position from which I see the picture and the space and objects in it: the depicted rowboat remains still and fixed in the middle of the lake; the depicted elephant’s gaze remains fixed on the approaching lion. This marks a clear, and fairly obvious, difference between looking at objects in the world and looking at pictures of those objects. And yet, we are able to see the objects of the actual world in a variety of pictures. In most cases we have no troubling recognizing the objects in these pictures. The marks on the page create a visual experience as of an object, such that, despite their deficiencies the pictorial marks allow us to see them as more than what they are—to see them as of some depicted object. If my claims are correct, a depictive picture can retain the sorts of visual features of an object that allow our visual system to be triggered off in a similar manner to seeing such visual features in the actual world.

3.2.4 Looking at pictures at-a-glance, piece-by-piece, or both?

Contradictory phenomena about pictures, however, raise issues that complicate
the above account. If, as Gombrich claims, pictures are incapable of being visually identical to objects in the world because the designs constitutive of a picture will always be visually dissimilar to the object itself, and yet Gombrich intends pictures to serve as visual substitutes for objects in the world because the designs constitutive of a picture can be visually equivalent to the object itself, then there will be a sense in which looking at a picture is looking at a visually contradictory artifact. A picture is both like and un-like the subject it depicts, and so looking at a picture will be both like and un-like looking at the subject it depicts. Gombrich hints at this distinction when he discusses the unique characteristics of M.C. Escher’s work. He writes that

When we look at a normal representation, there is nothing to prevent us from forming a hypothesis about the figure-ground relationship or about the way the shapes add up to pictures of objects. We therefore believe that we take in the picture more or less at one glance and recognize the motif. Our experience with Escher’s contradiction shows that this account is inadequate. We read a picture, as we read a printed line, by picking up letters or cues and fitting them together till we feel that we look across the signs on the page at the meaning behind them. (Gombrich, 1996, p. 155)

When we look at one of Escher’s impossible pictures or at Hogarth’s *Satire on False Perspective* we try and take the picture in at a glance but quickly realize that we cannot—we are seeing something that is impossible or contradictory. We try to read the picture piece-by-piece but find that the pieces do not add up to a consistent reading of the picture. If the parts of a picture makes the whole thing impossible, it seems that while the kind of immediate “at-a-glance” perception may be applicable to ordinary perception, Gombrich appears to rule it out for pictorial perception for the very reasons that arise when looking at impossible or contradictory pictures. It should also be pointed out that the perception of some cubist works seems to require a similar perceptual act: we find ourselves
attempting to create a coherent whole out of a patchwork of inconsistent parts. Given these cases, how are we to reconcile these seemly divergent claims?

I believe that cases of perceiving impossible or inconsistent pictures are unique cases, and that we can handle them within a Gombrichian framework by taking Gombrich’s theory to apply, in an at-a-glance way, to the perception of parts of the picture one-by-one and not (without further discussion) to our perception of the whole of the picture. If this is so, when perceiving a work like *The Belvedere* we take in the coherent constitutive parts of the picture in the immediate, at-a-glance way Gombrich describes, only to discover that we must backtrack and attempt to fit those parts into a coherent whole—which we cannot do. In this way we can understand Gombrich to be describing a spectrum of perceptual responsiveness to different pictures, a range of ways in which our perceptual system responds when we encounter a picture. Additionally, immediate perception could still involve the same sort of underlying process (of responding to relations) as do the point-by-point cases.\footnote{111}

Gombrich is certainly correct that our eyes roam the page in a sort of “saccadic” way when we read a picture (though he does not mention the term itself). This sort of visual hop-scotch is a normal and well-documented perceptual action that occurs in the perception of real-world objects and scenes\footnote{112}, and while it has been a much studied aspect of how we read words on a page\footnote{113}, Gombrich indirectly point out its existence in the perception of pictures.\footnote{114} When we look at a picture our eyes are not static, but are continually roaming around the page, focusing on different parts of the picture (sometimes for longer and sometimes for short amounts of time). While we can intentionally shift our attention (and therefore our eye movements) to certain areas of a
picture, most people are unaware that their eyes make these sorts of saccadic movements when they are looking at a picture (and that the region of the visual field that is in focus is surprisingly tiny). We can become conscious of it, however, when the pieces of the picture do not add up. When we become aware that our initial reading of *The Belvedere* is inconsistent we can become aware that our eyes are moving around the picture in an attempt to cohere the picture’s inconsistent parts into a consistent whole.\(^\text{115}\)

While our eyes jump around *Wivenhoe Park* just as they jump around *The Belvedere* this does not contradict the fact that the recognition of objects in pictures can occur incredibly fast and with a high degree of success. One study has shown that when subjects saw objects at different durations of time they “performed just as quickly and accurately on [a] categorization task [e.g., is the object a bird or a car?] as they did on a task requiring only object detection” (Grill-Spector and Kanwisher, 2005, p. 152). In other words, “By the time subjects knew an image contained an object at all, they already knew its category.” (Grill-Spector and Kanwisher, 2005, p. 152) Being flashed a picture of *Wivenhoe Park* or *The Belvedere* might result in a kind of initial, loosely “global,” recognition of the depicted motif (“landscape” or “building,” respectively). The fact that *The Belvedere* requires re-examination only helps to bolster Gombrich’s claim that “all perception is more or less conjectural” and that any interpretation can be over-ruled with new evidence (Gombrich, 1986).

Most ordinary pictures, however, are not meant to frustrate viewers, and in most cases such pictures offer an initial interpretation that is correct—meaning that we are able to correctly recognize a picture for what it represents. Creating a picture with the intention of visually tricking viewers or fooling them into thinking they are looking at the
real object is not usually an artist’s aim. Granted, trickery, visual frustration, and visual deception are all legitimate artistic intentions (and there are clear examples of each in the cannon of art history), yet I take them to be exceptions, cases that deviate from the customary function or role that pictures serve. Depictive pictures are to be recognized, and constructing a picture that intentionally thwarts recognition would, in most cases, undermine the picture’s very function as a representation.

We perform acts of recognition all the time, and we do it with relative ease and in an immediate way. I am not frustrated when I look at Degas’ *Self portrait with soft hat*, and it seems that I can, and do, take the picture in the way Gombrich has been describing for ordinary perception. In the case of Degas’ self portrait, should I encounter an inconsistency I can return to the picture’s “intricacies” (its constitutive parts) to examine the details of (e.g.) the scarf or the shading of the face (possibly altering my interpretation of them). What I will not do is offer a different interpretation of the picture *as a whole*. Gombrich’s example of impossible pictures seems to imply that when we encounter a picture with truly inconsistent or impossible visual elements, visual elements that stump our perceptual apparatus, we find the need to consciously re-read across the picture in search of a coherent meaning of the picture’s constitutive parts.

We can hold that the perception of a picture involves reading it piece-by-piece if what we mean by this is that our eyes make saccadic jumps from one spot of the picture to the next trying to fit the parts of the picture into a coherent whole. We are rarely conscious of these saccadic movements because in most cases a picture (like the scene it depicts) affords an immediately coherent and consistent reading. In those interesting cases where the picture is impossible or inconsistent (as in some of Escher’s pictures),
this process of picture viewing breaks down and we are forced to try and reconcile our perception of the picture as a whole with our perception of the constitutive parts of the picture. Thus, it might not be possible for us to arrive at one overall consistent perception of an impossible or inconsistent picture, but instead consistent perceptions of the areas within the overall picture.\textsuperscript{117}

3.2.5 Recognition and the role of perceptual hypotheses

At the very least the above remarks offer evidence for the possibility that there exists some congruence between pictorial and ordinary perception. A clear example of this congruence can be found in Gombrich’s insistence that perception involves the formation of hypotheses. Perception is not a passive process but instead involves an ongoing act of interpreting visual stimuli; it is the constant search for meaning. Perception will always involve the formation of hypotheses because perception is always more or less conjectural. These claims easily lead us to discuss the influential role of Karl Popper. Popper held the idea that confirmation of a hypothesis can never be more than provisional while its refutation is final. We assume there is regularity in the world—“the assumption that things are simple until they prove to be otherwise”\textsuperscript{118}—and we take this as something on which we can at least provisionally rely; not for the paradoxical reason that our assumption “is more probably right but because it is more easily refuted and modified.” (Gombrich, 2000, p. 272) Hypotheses always have the possibility of being rejected and replaced with new ones; confirmed hypotheses will remain, but they do so with the caveat that the existence of new evidence may eventually disprove them.\textsuperscript{119} As
we move around the world we are always categorizing and then adjusting and re-adjusting our interpretations of what we see. The wealth of knowledge gained from past experience has shown us that through practice our perceptual hypotheses are often correct, and that our perception of the world is relatively stable and reliably consistent.

To be clear, the formation of a hypothesis will occur immediately and be the result of directly looking at the world. I am not claiming some sort of delayed hypothesis-formation occurs when we perceive the world. The monkey doesn’t stop to form a hypothesis when it jumps to the next tree; neither do I stop to form a hypothesis before each step along an icy sidewalk. The monkey jumps; I take a step. It is in the action, I believe, that we are to find the hypothesis, for it is in that moment of action that an instantaneous (or nearly so) decision was made (that it was safe to jump; that the next step was not icy). Should the action prove unsuccessful (the monkey falls; I slip) then the perceptual information will be reassessed and the hypothesis rejected or adapted. So the hypothesis that we form will be based on the specific stimuli we encounter; should we be faced with an indistinct or ambiguous scene the assumption or interpretation that we form (i.e., how we see the indistinct or ambiguous scene) can be influenced by a range of non-perceptual factors (e.g., background and cultural expectations, or previous knowledge).

For instance, I look out my window and see a dark shape lumbering down the street. Given that I live in an urban area my guesses will probably not include lions, or tigers, or bears; let alone fictional possibilities such as thestrals, werewolves, or hippogriffs. Thus, even though we form perceptual hypotheses we do not form just any perceptual hypothesis, and this means that a viewer will not see an indistinct or ambiguous motif as just anything, but instead will offer an interpretation constrained by the context, evidence,
and, possibly, certain non-perceptual factors,\textsuperscript{120}

Given these various influences we are left with the following: perception is interpretation, which consists of a viewer forming hypotheses, and though these perceptual hypotheses will always have the possibility of being overthrown in light of new evidence, misperception is the exception and not the norm. Humans have become very good at getting perception right. Pictorial seeing involves similar perceptual processes; while “new evidence” is evidence that conflicts or is inconsistent with our initial interpretation of the picture, in most realistic pictures we settle on our initial interpretation because the evidence in the pictures affords us a consistent or coherent reading of the entire image. That is we subject the picture to a “test of consistency” and finding that it passes we offer the hypothesis that we are looking at a picture of some O and not something else. Gombrich describes the process this way: “If the [picture] is faulty the cycle slows down, [but] if the expected features include colour [sic] and atmospheric effects such as we experience in nature it will not only run smoothly, it will come more closely in a specifiable way to the experience” of actually seeing the object O (Gombrich, 1982, p. 297).

3.2.6 Recognitional duality

I have been claiming that for Gombrich pictures afford recognition of their subjects, but to be clear I believe that there are actually two ways of looking at a picture that are pertinent to my discussion: on the one hand we can recognize that P is a picture and not O itself (i.e., we can recognize that P is a flat, marked two-dimensional artifact);
on the other hand we can recognize that P represents something else (i.e., we can recognize P in terms of or as of O). This kind of “recognitional duality”\textsuperscript{121} means that a depictive picture can be seen as a flat artifact, or as the subject depicted. When I see Degas’ *Self portrait with soft hat* I can recognize that it is a picture: I can recognize the smooth washes of painted color, the peaks and outlines of the dabs of paint on the canvas, the fact that I cannot look behind it and that it is surrounded by a frame, and so on. These characteristics are part of what makes the picture a picture. I can also recognize what the picture is of. For instance, I can recognize that *Self portrait with soft hat* is of a man, or precisely, I can recognize that it is of a specific man (the painter Edgar Degas). Regardless of whether I recognize the particular or the general the point is that I recognize the picture as being of something beyond the base materials that make it up.

There will undoubtedly be cases when the viewer is fooled by P and he or she will think that P is O itself (e.g., when P is a trompe l’oeil or when P is seen in some impoverished viewing situation). It is only in these rare cases, however, that a viewer will confuse the second type of recognizing with the first and thereby be deceived by P. The vast majority of ordinary pictures do not result in confusing these two recognizings, and so the perception of the vast majority of ordinary pictures does not result in deception. This is so because a picture consists of numerous visual characteristics that hinder our perceptual system from making the leap to deception.\textsuperscript{122} As we look at a picture and recognize (and thus gain the awareness) that we are looking at a picture, we will continue to retain this awareness even when we then recognize the picture’s content. In other words, once we know that P is just a picture and not O itself we retain this awareness even when we recognize the content depicted by P and thus see P in terms of, or as of, O.
While I take it to be a natural feature of a picture to afford this kind of dual recognition it rarely leads to problems in our everyday encounters with pictures.

Saying that a viewer can recognize a picture “in terms of” or “as of” an object may sound a bit stilted: how can we recognize a flat artifact as an object with space and depth? Or, put differently, how can we recognize something flat as something that is not? If someone were to point to the wall in my living room (on which Lichtenstein’s print *Ball of Twine* is hanging) and ask what I saw, I could reply, “I see a wall,” or “I see a picture of a ball of twine,” or “I see a ball of twine.” Having recognized the picture (in the two ways I have described) the talk about “as of” can, or may, come in as a means of explaining that we “see the object in the picture” or that we “see the picture as the object.” Thus, introducing the phrases “in terms of” and “as of” is a means of highlighting these facts: the picture is as of the subject though it is not the subject itself.123 Saying that “I see O” will not lead to any confusion about what I mean, though to be even clearer about my perceptual experience I can either say “I see P as of O” or that “I see P in O” (the latter being a favored phrase).

It is rare for us to recognize an ordinary three-dimensional object as some other ordinary three-dimensional object124—it is rare, that is, unless we are mistaken, perhaps recognizing our friend as a stranger or the oncoming dump truck as a bus.125 Looking at pictures may be similar to looking at objects in the world in a number of ways (getting things wrong being one of them), but they are not exactly the same and part of what makes them different is that pictures afford us the kind of recognitional duality that I have been describing. A picture (specifically a depictive picture of the kind Gombrich is concerned with) is a unique perceptual object that affords viewers the ability to see, and
to recognize, the picture as something. We can see a picture as a flat, two-dimensional artifact—we can recognize a picture as a marked surface, an arrangement of pictorial design features; but we can also see a picture as the subject built-up from these design features—we can recognize that the picture is of something. When we look at a picture and recognize a cow we really do have a recognitional experience of a cow. The content of our perception really does consist of us looking at the picture and recognizing a cow, and this is a perfectly sensible thing to say even if what we mean is that we are not looking at an actual cow but instead we are looking at a picture of a cow.

A depictive picture will afford a viewer both types of recognizing, and it is only when one of these types of recognizing is absent that problems arise: if we only recognize the painting as a flat, marked surface we will simply see it as a configuration of lines, colors, and shapes and not as a representation of anything; on the other hand, if we only recognize the painting as its content we will lose the knowledge that we are looking at a picture, and we will become deceived and ultimately think we are in the presence of the actual object. Both types of recognizing are essential to successful picture viewing. An apt reminder of this fact can be found in Kenneth Clark’s colorful narrative (cited in Art and Illusion) of how he would “stalk” Velazquez’s Las Meninas in the hope of catching the picture in the act of creating its content out of “a salad of beautiful brushstrokes.” Try as he might, though, “[Clark] could never hold both visions at the same time, and therefore the answer to his problem of how it was done always seemed to elude him.” (Gombrich, 2000, p. 6) We can recognize the representational picture as a flat, marked surface (a “salad of beautiful brushstrokes”) and we can recognize it as a vivid representation of an imagined reality. The picture that “works” for the audience of
viewers in the way I believe Gombrich has in mind, then, will be the one that affords this kind of recognitional duality.

3.3 Perceptual constancy: an empirical hypothesis

I have spent a great deal of time elaborating on the claim that the activity of perceiving pictures involves many of the same perceptual processes as the activity of perceiving objects in the actual world. That is the perceptual abilities used to recognize the objects we see everyday are involved in the perceptual abilities used to recognize pictures of those object. As such, I take recognition to not only be central to an understanding Gombrich’s account of depiction but to my interpretation of his writings as well. That being the case, in this section I will present the notion of perceptual constancy as a means for fleshing out some of the details of recognition; in this way, the concept of perceptual constancy should be included within the concept of recognition described previously. Cases of ordinary perception that involve perceptual constancy (i.e., where we are able to recognize an object under changing or varied perceptual conditions) will also be cases in which a viewer is afforded, though not guaranteed, recognition of O. We see an object from far away and close up; we see it as it moves quickly or slowly across our field of vision; we see it from a variety of different angles and in different lighting conditions. Despite these varied perceptual conditions we find that we can recognize the object—we have a “perceptually constant” experience of it. Given these facts about ordinary perception, and given my claim that the processes involved in ordinary perception are not fundamentally different than those found in pictorial perception, I now wish to propose that artists are able to pick up on and render in a picture aspects of our
ordinary visual experience that exhibit perceptual constancy.\textsuperscript{128}

In Chapter 2 I said that the relevant respects in which picture and object are similar resides in the relations between the arrangements of visual elements, which in turn affords us recognition of the picture’s subject-matter. Yet, I have failed to point out exactly which shared relations are the ones responsible for the picture and the object being similar. Despite the claim that there does not exist any set of relations of visual elements that are universally relevant to all cases and for all pictures, there still remain some sets of relations of visual elements that are relevant in some cases and for some pictures (specifically, the types of depictive pictures I am concerned with). In this section I will offer just such an empirical hypothesis of a possible set of relations of visual elements that are responsible for a picture and an object being visually relevantly similar: specifically, these will be the visual elements that are the best equivalence to those involved in perceptual constancy in ordinary perception.\textsuperscript{129} I will then show how such an hypothesis might fit into an interpretation of Gombrich’s account\textsuperscript{130}

3.3.1 Perceptual constancy in ordinary perception

In ordinary perception, perceptual constancy is the idea that we are able to identify an object under changing perceptual conditions. Branco Mitrovic writes that perceptual constancy allows us to maintain an “awareness of the real size and colour [sic] that affects or is reported to affect our phenomenal experience.” (Mitrovic, 2010, p. 16) Gombrich describes it in this way: “The color, shape, and brightness of things remain to us relatively constant, even though we may notice some variation with the change of
As the visual world changes we must learn to focus on the meaningful changes of the properties of objects and not merely accidental perceived changes. In this way we are able to recognize an object despite proximal changes in the stimuli we perceive. When I walk around my living room the various intensities of light are always changing, and yet I am still able to recognize my living room in morning, afternoon, and evening light. When looking around my living room there will be certain aspects of my visual experience that afford constancy to my perception; these visual aspects allow for the awareness that my livingroom remains my living room despite changes in my immediate visual experience. The aspects that allow for the continued “constancy” can include relations among size and shape, but also certain surface properties such as contrasts and gradations of color value. If we now conjoin the idea of perceptual constancy with my previous claims about the significance of relations of visual elements to what constitutes relevant similarity we get the following set of statements: when we look at an actual object O our visual experience (consisting of certain sensory information) is constantly changing; our visual system focuses on the relations among the visual elements in O; and finally, it is the relations between the visual elements (i.e., the relations among the sensory information) that remain constant through the changes in the object’s appearance. When we look at O we see a number of visual elements (f1…fn) constitutive of the experience of perceiving O, and we see that (f1…fn) are in some relation S to each other. Thus, in perceiving O our visual system is attuned to focus on S(f1…fn), rather than the individual elements (f1…fn). In many cases our perception of O will involve the perception of distance, illumination, angle of vision, and so on.” (Gombrich, 2000, p. 52)
relations of visual elements that can afford a viewer a perceptually constancy visual experience of O. When we see that S(f1…fn) in O we see certain features of O that remain constant despite changes in O’s immediate appearance. We might refer to the relations S(f1…fn) in O as its “perceptually constant features.” Furthermore, it is the invariance of these relations, and thus the perception of the perceptually constant features, that affords us awareness of the actual nature of the objects we perceive. Familiar objects remain familiar—my living room remains my living room, my dog remains my dog, my wife remains my wife. I recognize each of these despite the fact that the perceived visual aspects of each can change from moment to moment.

3.3.2 Constancy in pictures does not require identical visual information

Though an artist cannot match exactly those aspects found in nature he or she can, as Gombrich has described it, transpose light into paint, meaning that an artist can create an equivalence or an approximation given the limitations of his or her medium. Since the important aspect to be captured in the picture is the relations among the visual elements, should an artist wish to depict relation S found in object O (in the actual world) she will experiment to find the best equivalent relation R in the picture P. That is, she will do the best (within her medium) to capture the relation as it is perceived in the actual world. So long as the relation R between the visual elements (g1…gn) seen in P corresponds to the relation S between the visual elements (f1…fn) seen in O, the artist can preserve in P an equivalence to those visual aspects that exhibit constancy in our ordinary perception of O. It is important to remember that while the artist is able to preserve relation S in O in
the picture P by means of relation R, the object depicted in P does not (in P itself) undergo changing perceptual conditions (e.g., changes in lighting). The depicted object O we see in P is “arrested,” and thus we recognize O depicted in a frozen moment without seeing O undergo the changing perceptual conditions that the actual O would go through. Furthermore, the depicted relation R exhibited in P is fixed. In the actual world O exhibits relation S in a variety of actual, different conditions under which O is perceived; in this way perceiving relation S allows us to recognize the object as O (in the actual world) despite these changing perceptual conditions. As long as the relation S—the relation that makes for perceptual constancy in our perception of O in the actual world—is there in the picture (in the form of the equivalent relation R) and can be grasped perceptually by a viewer, then no matter how much the experience of viewing P may change (e.g., changes in lighting or the angle at which P is viewed) R in P will afford the viewer recognition of O in a similar way to S in O.

This goes somewhat further towards explaining why when a viewer looks at P he or she sees that P is of O even though R(g1…gn) in P is not identical to S(f1…fn) in O. I earlier quoted Karen Neander as saying that the elements of the picture and the object need to match along “respective scales,” and that this might help to explain how the relations in the picture can be equivalent to the relations in the object: the relation R(g1…gn) in P will be, for our visual system, an equivalent to the relation S (f1…fn) in O if the response from our visual system is triggered off by the former in a sufficiently similar way to how it is by the latter; and this response can occur even though the individual visual elements (g1…gn) in P are not identical to the individual visual elements (f1…fn) in O. Thus, preserving the correspondence between R(g1…gn) in P and S (f1…fn) in O
means that P preserves those aspects responsible for perceptual constancy in O, and our visual system, in turn, picks up on this correspondence and allow us to see that P is a picture of O.

Let me speak a bit more about a specific set of the relations, namely the relations of contrasts (and gradients) of color value, since I have already mentioned that Gombrich discusses these in his example of representing a white handkerchief. In various places throughout the chapter “From Light to Paint” Gombrich points out that our visual system is attuned to respond to “light intervals, to what have been called ‘gradients,’ rather than to the measurable quantity of light reflected from any given object.” (Gombrich, 2000, p. 50) A wonderful example of this, which also helps to demonstrate the kind of arrested nature of the depicted relations of visual elements mentioned above, is the numerous studies of wheat stacks done by Monet. An actual haystack will exhibit a relation S between visual elements (f1…fn). This relationships is recognized throughout the day (and different seasons) despite the numerous changes in perceptual conditions. In an individual Monet painting of a wheat stack at, say, 6:00 p.m., the wheat stack, as it is depicted in the painting, is not changing, but is fixed. The painting stops the perceptual variation at a specific point (e.g., 6 p.m.) But the relation S among the elements of the actual wheat stack (that stays the same throughout the day and lets us recognize the haystack under the different viewing conditions) is also displayed in the picture of the haystack as it is at 6:00 p.m. (in the form of relation R). So even though our perception of the depicted haystack isn't itself varying in the way it does in the real world, the presence in the picture of R allows us to recognize the haystack in the picture.

Given this line of thought, our perception of the relation S [f1…fn] as it is
exhibited in the picture in the form of relation R \([g_1\ldots g_n]\) allows us to recognize \(O\) in the picture; and thus the presence of the equivalent relation \(R\) in the picture constitutes (at least in part) the relevant similarity of the picture \(P\) to the object \(O\).

Moreover, I also believe this is what Bühler was concerned with in the previously mentioned quote from Lepsky: where the contrasts of color value in the picture (what Bühler called the “intensities of light) changes from light to grey we will find a corresponding change in the contrasts of color value found in the object. This also helps to clarify Lepsky’s cryptic claim that the contrasts of color value are the same “but not with the same steps,” as well as Neander’s claim that the pattern of similarities and dissimilarities between visual elements in the picture and the object match qualities “along their respective scales.” The changes in contrasts of color value cannot be identical since the variations of light in nature are vastly greater than the limited range of variations the artist is able to render in his or her medium (\(R [g_1\ldots g_n]\) in \(P\) will never be identical to \(S [f_1\ldots f_n]\) in \(O\)). What the artist can do is ensure that the relationships of color values rendered in the picture follow a similar pattern to those relationships of color values seen in the object itself. If one area of the object gets lighter or darker then the artist will do his or her best to ensure that the pictured area also gets lighter or darker (relative to the limitations of the medium). Thus, a successful representation does not require the rendered contrasts of color value to be identical, only that the rendered relations faithfully follow the same pattern as those relations of color value found in the object—doing this is the best the artist can hope to achieve (though it would be an understatement to say that the results can nevertheless be stunning).

I have already talked about the possibility that the ways we perceive objects in the
world and the ways we perceive pictures are not fundamentally different, but let me make a few more brief comments on the topic. Above I claimed that perceptual constancy is preserved in a picture because the picture accurately renders similar relations of visual elements found in the represented object (e.g., the contrasts or gradients of color value), and that the relations of visual elements captured in the picture are those that exhibit perceptual constancy. If we continue with Bühler’s example, the idea is that the rendered gradients of color value in the picture \( (g_1 \ldots g_n) \) follow a similar pattern to those gradients of color value \( (f_1 \ldots f_n) \) found in the object; the relationships between \( (g_1 \ldots g_n) \) in the picture are similar to the relationships between \( (f_1 \ldots f_n) \) in the object. So when \( (f_1 \ldots f_n) \) in the object exhibits a change from light to grey then \( (g_1 \ldots g_n) \) in the picture will exhibit a similar pattern of change. In short, the artist exploits facts about how we see the actual world and attempts to render an equivalence in the picture. (I will say more on this below.)

Take also the example of an object's size: when I look up in the sky and see an airliner high over head, I know that it is actually much larger than my dog even though the immediate appearance of the airliner is smaller than the immediate appearance of my dog. The artist understands that large objects seen from a distance will appear smaller than objects seen up close. The artist will then set out to create a picture that employs such facts about our ordinary perception. So when we see a painting in which an airliner is depicted in the sky as being smaller than a dog we can recognize that the airliner is most likely depicted as being farther away than the dog. Of course which objects are being depicted, and the surrounding visual context in which those objects have been arranged and are viewed will determine whether the picture preserves perceptual
constancy. The significance of this is that the artist recognizes facts about how we normally see objects in the world and incorporates those facts into the creation of depictive pictures through the incorporation of those aspects of an object or motif responsible for perceptual constancy.

Faithfully representing the relations of visual elements found in the subject is important for Gombrich’s theory, but the idea implies more than just that relations of visual elements be included in a picture—it also implies that the included relations be arranged on the page as they are seen in the object. In other words, the relations of visual elements are not arranged arbitrarily in the picture. This is clear from Karl Bühler’s example of the photographic relational model: in the case of such non-linguistic methods of representation “there are rules and principles which are not based on an arbitrary allocation of the symbol [picture] and the symbolised [object].” (Lepsky, 1996, p. 34) So, given the significant role that relationships play for our visual system in general, when an artist attempts to represent O he or she will not only choose to include a specific set of relations of visual elements in P, but will also ensure that the relationships be arranged on the page as they are in the object. In Bühler’s example he tells us that the relationships of grey values in the picture correspond to the relationships of grey values in the represented object. But how? In the sense that the changes in the relations between visual elements in the picture (whether it is gradients of light and color, or arrangements of lines and shapes) follows the changes in the relations between visual elements in the object, such that where the visual elements in the object exhibit a change from (e.g.) light to grey so too will the visual elements in the picture exhibit a change from light to grey. If it is the following of a pattern of change that is important, then the placement and
arrangement of the relationships between the visual elements in the picture will, in a similar way, follow the placement and arrangement of relationships between visual elements seen in the object. A viewer looking at an actual apple on a table will see certain visual elements \((f_1...f_n)\) as being in some relation \(S\), and in seeing \(S(f_1...f_n)\) the viewer sees \(S(f_1...f_n)\) arranged in a certain way within in her visual field. Should an artist wish to produce a painting of an apple on a table seen from a similar viewpoint she must ensure that the picture includes an equivalent set of visual elements \((f_1...f_n)\) and that they are in a similar relation \(R\) and they are arranged on the page in a similar manner—in this way when a viewer sees \(R(g_1...g_n)\) in \(P\) it triggers a similar visual response to that of seeing \(S(f_1...f_n)\) in \(O\). Hence, the relation \(R(g_1...g_n)\) in \(P\) is not arranged randomly around the page if \(R(g_1...g_n)\) in \(P\) is meant to correspond to \(S(f_1...f_n)\) in \(O\), and if \(P\) is thus meant to be a representation of \(O\).

These sorts of claims fit with what Daniel Gilman says when we writes about being “impressed by…an almost universal appeal of certain sorts of pictures as a way of transmitting information about what something looks, or looked, like.” (Gilman, 1988) For Gilman, and Gombrich as well, pictures are vehicles for delivering information about the objects depicted in them, and according to my interpretation it is the sets of relations of visual elements in a picture that affords us this kind of information (e.g., information about what the object looks or looked like). Gilman says that the sorts of pictures that clearly present the kind of information about what something looks or looked like include clear photographs and drawings in perspective—I might expand the list of examples to include pictures of faces, given Gombrich’s numerous enthusiastic discussions about the significance of physiognomic features for our visual system. Gilman and Gombrich can
make these assertions because they see a real connection between picture and object. The
information presented to our visual system by a picture really does tell us something
about the depicted object (why else would we look at pictures of planned vacations spots
or of deceased relatives?); but more than this, an artist includes one set of relations of
visual elements in a picture rather than another set because the one will be better able to
trigger the visual response of the depicted object than the other. Seeing a drawn
courtyard in perspective or a painted face presents information in an equivalent way to
seeing the actual courtyard or the actual face; it is not exact, but it does need to be to
illicit a correct response. It is sufficient enough for our visual system that it encounter a
pictorial equivalent, which is what Gombrich is claiming pictures are capable of
providing.

There is a well-known criticism brought against the position that a picture
resembles an object, namely that such a theory entails that a picture deliver to the eye
identical bundles of light rays but that the conditions required to guarantee this are too
extraordinary. There is not enough space to adequately address this topic, and nor will I
introduce the locution “bundles of light rays” into my account. I will say, however, that
given what I have said thus far such a demand seems unreasonable: no picture, not even a
tromp l’oeil, delivers to us bundles of light rays (i.e., information) identical to those
encountered when seeing the object itself. Here I am in agreement with Gilman when he
claims that it should not be a problem that pictures “may not always provide exactly the
same spatial information as surfaces,” since “a permanent and thoroughgoing identity of
appearance between pictures and their subjects is not a fact at all, [and] consequently it is
not a fact that any sort of account of pictorial representation needs to explain.” (Gilman,
Since no picture can present us with identical bundles of light rays, and thus no picture is capable of presenting us with exactly the same information as the represented object, we should not criticize a theory for neglecting to include such a requirement. The information we get from a picture of O will never be identical to the information we get from O itself, for “so complex is the information that reaches us from the visible world that no picture will every embody it all.” (Gombrich, 2000, p. 90) And yet, the picture of O can be successful; we can effortlessly recognize that the picture is of O. A successful picture is, Gombrich would say, simply one that does not present us with false information. A picture will not present us with identical information, but why should we think that seeing truthful information in the picture entails information identical to that perceived from the actual object?

That a successful picture can present us with truthful information seems compatible with my claim that a picture can be relevantly similar to an object by preserving the relations of visual elements found in the object. This can mean that the picture need simply present us with visual information that is sufficiently similar to the visual information in the actual object for our visual system to recognize the picture as of the actual object. Presenting “sufficiently similar enough” visual information is plausible given the robustness of our visual system and the mind’s ability to extrapolate and “run ahead of the stimulus situation.” (Gombrich, 2000, p. 303) It is possible for us to easily recognize an object or scene correctly despite not perceiving the totality of an object’s visual features—we do “get” all of an object’s perceptible information. (In fact, we rarely get all of an actual object’s perceptible information as well.) In the case of pictorial perception the formation of hypotheses and our ability to extrapolate beyond the
immediately available information allows us to fill in the informational gaps in the picture. In short, we should not be concerned with the fact that a picture presents “incomplete” visual information about an object or scene: we manage to instantly and effortlessly recognize objects in the actual world despite not perceiving their complete visual information; when we then look at pictures of those objects we find it to often be the case that recognition occurs just as instantly and effortless despite the “bundles of light rays” presented to our visual system not being identical to those that would or could be seen when standing in front of the actual object. This being the case, I find it an impossible demand to claim that representational pictures need to present viewers with information identical to that presented by the represented object.

3.3.3 Non-arbitrariness and the correspondence between picture and object

In *The Image and the Eye* Gombrich tells us that the relationships of visual elements in a photograph need to be transformed and compressed, since, like any two-dimensional medium, a photograph is only capable of representing a limited range of relationships. He goes on to say that “the fact that [the relationships] are thus transformed does not entitle us to call them an arbitrary code. *They are not arbitrary* [italics added], because a gradation from dark to light observed in the motif will still appear as such a gradation [in the picture], even if reduced in span.” (Gombrich, 1982, p. 282) This is a significant point, and one that I have already mentioned: the picture captures the relations of visual elements found in the object and it does so in a way that is not random. A different way of saying this might be to say that there are “constraints on
the configuration of marks” in the picture (Newall, 2003, p. 94). This means is that artists cannot include just any set of designs or configurations of marks in the picture if they are to depict one object or scene rather than some other. As Gilman puts it, “there are not other aspects of the world that we might just as well choose for our focus if our focus is to be achieved by looking.” (Gilman, 1988) The idea that the designs or configurations of marks in the picture are not arbitrary also means that there will be certain designs or configurations of marks that we simply cannot use to depict features of the world. For instance, rectangles and triangles cannot be used to depict the circular rim of a vase. Painting a picture of the New York City skyline means that I cannot choose to include just any designs or configurations of marks in the picture, nor can I arrange those marks in any arbitrary way if I want the picture to be a picture of New York City. I need to maintain certain configurations and arrangements of relations of visual elements on the page such that they correspond to the relations of visual elements seen in the New York City skyline. Thus, the artist’s choice of which designs to use in the picture is not unlimited nor can it be arbitrary given what is to be represented.

It would seem, then, that a whole range of representations can rest on the same principle, which means that doodles and pencil sketches can correspond to some depicted object or scene no less than caricatures and naturalistic oil paintings. In all of these cases the designs and marks constitutive of the picture are not arranged arbitrarily on the page if the picture is to be a correct representation. Both the relationships between the visual elements and their arrangement and placement on the page must correspond to those found in the object itself. For instance, why are the marks constitutive of a stick figure arranged in the way they are? Certain marks stand for the head, arms, torso, legs and feet,
and these marks are meant to correspond to the arrangement of body parts in an actual person. The correspondence is certainly not precise, and one might even argue that such minimally detailed drawings have a more symbolic function. Yet, we can easily recognize such a picture even when it is constructed of only a few well-placed marks—in fact, we can get along quite well with pictures that represent very limited or even distorted aspects of a subject. Take for instance Charles Philipon's caricature of King Louis Philippe, for which Philipon was accused of satirizing the king as a poire (a pear or “fat head”). In looking at each stage of the picture it is easy to find aspects of the emperor's face that correspond to the shape of a pear (if this were not so, why else would Philipon have been brought up on charges!). We can see this even more clearly when we compare the caricature with a portrait of the emperor himself. By looking back and forth between the portrait and the caricature one can pick out those features that are “pear-like,” the ones Philipon chose to highlight and exaggerate—yet we are still able to see the emperor. That is, despite the exaggeration of certain visual aspects of Louis-Phillippe we can still see that the caricature is of Louis-Philippe. While this example does not conclusively show that every picture employs features responsible for perceptual constancy, I think that it does illustrate that there are cases in which a picture can preserve the kind of perceptual constancy found in ordinary perception (i.e., where we can see the subject in various situations in which the appearance features change).

3.3.4 Final thoughts on perceptual constancy and recognition

Let me conclude the discussion of perceptual constancy with a few final remarks
on the connection between perceptual constancy and recognition. My claim has been that a wide range of pictorial representations (e.g., a stick figure drawing, a picture of a row of columns or train tracks, a portrait) can all exhibit visual aspects of an object that are responsible for perceptual constancy. These visual aspects constitute the relationally arranged set of visual elements—the similarities and dissimilarities between the visual elements found in an object (not the exact individual visual elements)—and a picture that preserves these relations will afford our visual system the ability to recognize what the picture is of. I believe that all cases in which we perceive perceptually constant features will involve the perception of relations of visual elements, and I believe that the perception of relations of visual elements is necessary for the recognition of an object or a picture of that object; I am hesitant, however, to claim that every case of recognition (whether of an object or a picture of that object) will involve the perception of perceptually constant features. More specifically, recognition can occur despite the fact that we do not perceive an object’s usual perceptually constancy features: there is always the possibility that we can recognize an object or a picture of an object in odd ways, or as a result of seeing some uncharacteristic features. In other words, on my account perceptual constancy remains a sufficient but not a necessary feature for recognition. Let me clarify this a bit.

In most cases recognition will be easy and effortless, which is due to the fact that we perceive an object from a usual or “characteristic” viewpoint, but this fact does not preclude us from recognizing an object from some unusual or “non-characteristic” viewpoint. A characteristics view might be thought of as “the one (or sometimes two) which exhibits most of those distinctive features by which we classify and name the
things of our world.” (Gombrich, 2000, p. 302) While there is no single or universal characteristic viewpoint for viewing an object, it is fair to say that we tend to encounter (i.e., perceive) most of the objects in our environment from a roughly straight on view, rather from above or below. While this is certainly not always the case (e.g., we often see our shoes from above while we see chimneys from below), most pictures tend to depict objects from a straight on, side-to-side view, rather than from a vertical, up-and-down view; this is even the case for objects (like shoes or chimneys) that are normally seen from above or below. In depicting an object from a straight on, side-to-side view (or from whatever view is found to be the characteristic viewpoint), we are choosing to depict the object from the viewpoint that allows for the clearest classification and, I would argue, this will often result in depicting those distinctive features of an object that are responsible for the usual or characteristic perceptually constant experience of that object.

The human visual system, however, is robust enough to allow recognition of an object even when that object is perceived from some unusual viewing position (e.g., seeing a courtyard from directly above rather than from straight on); or when we perceive odd or unusual visual aspects of an object (e.g., seeing just the mouth or the top of the head of a spouse or child); we can even recognize an object by means of its non-visual characteristics (e.g., if an object has a distinctive aroma or sound). Given these facts, it appears a plausible conclusion that recognition can occur when our visual experience lacks the usual distinctive features responsible for perceptual constancy. In light of my claims about the heterogenous nature of ordinary and pictorial perceptual processes, I wish to make the extensional claim that we can recognize pictures of depicted objects.
even when such pictures fail to include an object’s usual perceptually constant features.

All of these facts about the sufficiency of perceptual constancy can be true, and
nevertheless there can remain obvious cases of picture that do exhibit relations of visual
elements responsible for perceptual constancy, relations of visual elements that are
similar to those relations of visual elements responsible for perceptual constancy in the
object. My claims amount to an empirical hypothesis that requires us to look at actual
pictures in the world. Whether a certain picture does include relations of visual elements
responsible for perceptual constancy is an empirical claim that must be taken up on a case
to case basis (though I think it safe to say that most ordinary naturalistic pictures,
specifically the ones Gombrich is primarily concerned with, will present us with some set
of such perceptually constant features). Pictures that include such relations of visual
elements will present visual aspects of a object that correspond to those relations of visual
elements found in the subject matter—the picture presents visual aspects that are
responsible for perceptual constancy. It matters not whether a picture includes relations of
visual elements that exhibit the usual perceptually constant features of an object or
whether such included relations exhibit some kind of unusual perceptually constant
features object (assuming this is possible): a picture that captures those relations of visual
elements of an object that do exhibit perceptual constancy will serve as a relational model
for that object; the picture can thus be described as relevantly similar to the object it
represents, and, furthermore, can be said to function as a visual substitute for the object.
3.4 Concluding relevant similarity

According to Gombrich picture-making involves creating visual substitutes that are meant to stand-in, visually, for an intended object or scene. In Chapter 1 I introduced the concept of relevant similarity to help better explain just how the notion of visual substitution works. I initially stated that we might be tempted to think that a picture $P$ is a visual substitute for an object $O$ simply because $P$ and $O$ share some property or set of properties. I showed that this account was too vague to serve as a sufficient condition for $P$’s depicting $O$, and in Chapter 2 I said that we should instead look at whether $P$ and $O$ have in common certain relevant respects, namely, the relations between visual elements. In other words, we should look at whether the picture accurately captures an arrangement of similarities and dissimilarities between visual elements similar to that found in the object. Not being identical to those found in the object, the relations of visual elements rendered on the page serve as an equivalent (within the limitations of the medium) to those relations of visual elements found in the object. A picture that successfully preserves these relations can be said to function as a relational model for the object it represents, and it is creating such a model that is the task of the artist wishing to produce a correct or successful depictive picture of some motif.

In this Chapter I sought to complete the investigation of relevant similarity by claiming that the relevant respects in which a picture is similar to an object are also those respects capable of affording us recognition of the depicted object. In this way, I said that we should be concerned with those aspects of a picture that allow it to be seen as something beyond a merely flat marked surface—we should be concerned with a picture’s representationally relevant aspects. To succeed as a representation means that
viewers must recognize what the picture is of—a depictive picture that no one can recognize is like a plate of food that no one can eat. I believe it will be rare, though certainly not impossible, for a viewer to experience recognitional failure when the picture functions as a relational model in the manner I have described. The kind of instant and effortless recognition afforded by most ordinary pictures is due to their accurately preserving the relations of visual elements found in the represented object. Thus, recognizing what a picture is of remains intimately tied to the preservation and inclusion of certain relations of visual elements in the picture, and it is my belief that by including the latter you greatly increase (though you do not guarantee) the chance of the former occurring.

In sections 3.2-3.2.6 I asserted that we are able to recognize what a picture is of because it affords us similar recognitional experiences to those elicited by the perception of the visible world; this is due to the fact that looking at objects in the actual world and looking at their pictures do not involve wholly different perceptual processes. There exists some congruence between how our perceptual system processes the world and how it processes pictures of the world. The key difference being that as flat artifacts, pictures (at least the types of ordinary depictive pictures I am concerned with) can be seen as both meaningless marked surfaces and as meaningful representations of something beyond those marks. Thus, pictures affords us a kind of “recognitional duality” that objects in the world does not.

To answer the question of exactly which relations (capable of affording recognition) are relevant for a picture being similar to an object, in sections 3.3-3.3.4 I put forward, and elaborated on, an empirical hypothesis that a possible set of relevant
relations of visual elements includes those that exhibit constancy in ordinary perception. By taking those relations of visual elements that exhibit perceptual constancy in ordinary perception (e.g., contrasts and gradations of color value, as well as size and shape) and transposing them into an equivalent two-dimensional medium, the artist creates an artifact that both serves as a relational model and affords recognition of some depicted object. It is important to remember that recognition is built into the idea of perceptual constancy, meaning that perceptual constancy is sufficient but not necessary for recognition. Pictures can afford (though they will not guarantee) recognition, and just as our visual system is robust enough for us to recognize an object in the absence of perceiving an object’s usual perceptually constant features, so too can we recognize a picture of an object in the absence of perceiving any usual perceptually constant features rendered in the picture.

This is the manner in which the notion of relevant similarity should be understood and incorporated into an interpretation of Gombrich’s account of visual substitution. Initially I said that a picture P is a visual substitute for an object O because P is similar to O with regards to some relevant respect or respects. We can now say that P is relevantly similar to O when: (1) P functions as a successful relational model for O, meaning that P captures certain relations among the of visual elements found in O (i.e., certain patterns of similarities and dissimilarities between the visual elements found in O). Furthermore, (2) in preserving this kind of relational fidelity P affords us recognition of O: in preserving the relations of visual elements, some of which are capable of exhibiting usual perceptual constancy, P presents a viewer with certain visual aspects capable of triggering off a recognitional experience that is similar to the recognitional experience triggered off
by O; this allows us to see the picture “as of” the represented object, or to see the
represented object “in” the picture. P is thus relevantly similar to O in that P serves as a
relational model for O, capturing certain relationships among visual elements that afford
us recognition of the depicted object O. Relevant similarity does not lead to mistaking
pictures for objects because we remain aware of the characteristics of a picture that make
the picture a picture, and thus the kind of recognition afforded by pictorial perception
remains of a dual nature: we can recognize a picture as being “a flat artifact” or as being
“a representation of…” Both ways of recognizing a picture are essential to our pictorial
experience, and the absence of one or the other means that we either fail to see the picture
as being a depiction or that we see the picture as the object itself, respectively.

In this chapter and the last I have tried to make clear both the complex nature of
relevant similarity and how it might clarify an interpretation of Gombrich’s position in
Art and Illusion. If my exposition has adequately captured the symbiotic relationship
between the idea of the relational model and the idea of affording recognition, then the
concept of relevant similarity should fit with and help to explain Gombrich’s claim that a
picture functions as a visual substitute for the object it represents.
4 Gombrich and the concept of twofoldness

4.1 The goal of Chapter 4

In this chapter I will turn to the notion of twofoldness, first introduced by Richard Wollheim (1963). In addressing this topic I will discuss, in part, a question I addressed in the first chapter: what kind of experience do we have when we look at a depictive picture? In other words, what exactly occurs when we encounter a surface consisting of an arrangement of marks that are meant to represent something beyond the base materiality of that surface? The idea of twofoldness holds that the experience of looking at a picture is actually of a simultaneous dual nature: one aspect of this experience is of the picture's medium (the surface features), while the other aspect is of the picture's subject (the representation, the content). As we consider this matter the reader should keep in mind the following two questions: (1) Do Wollheim’s criticisms accurately reflect the position he attributes to Gombrich? (2) Can Gombrich’s position accommodate the kind of twofold experience that Wollheim has in mind?

To begin the discussion of twofoldness I need to first mention the difference between Wollheim’s early and later development of nature of twofoldness, as he revised his account of the concept from an experience of two different entities to a singular experience involving two different aspects. The inception of twofoldness arose out of a reaction to Gombrich’s position in *Art and Illusion*, which makes use of the duck/rabbit image in conjunction with the idea of “switching” interpretations. Viewers cannot hold or see two conflicting interpretations at the same time and must instead switch between them in two different ways: (1) viewers can switch between seeing two meaningful
interpretations (duck versus rabbit); and (2) viewers can switch between seeing a meaningful interpretation (duck or rabbit) and a meaningless interpretation (the medium, surface, etc.). I plan to show that, Wollheim’s reaction to the contrary, a Gombrichian type account of twofoldness can be incorporated alongside this notion of switching. In doing this I will need to dig somewhat deeper into what Richard Wollheim meant by the idea of a “twofold” experience of a picture; this, in turn, will precipitate a brief discussion on whether Wollheim was all that consistent in his explanation of the concept in the first place.

It should be noted that I do not deny that Gombrich’s position (as it is found in *Art and Illusion*) is rather vague regarding the subject of twofoldness, specifically as to whether or not Gombrich rejects twofoldness completely. That said, it is my belief that Wollheim’s criticisms misrepresent Gombrich’s stance and, as such, fail to completely rule out the possibility of a Gombrichian account of twofoldness. Some theorists (myself included) believe that Wollheim equivocated on an important distinction between two different psychological attitudes: “being visually aware” of such-and-such and “attending to” it. If this distinction proves to be significant it will help to illuminate a possible answer to the second of the two questions I posed above: namely, that we can interpret Gombrich’s position in *Art and Illusion* as accommodating twofoldedness.

4.2 Wollheim and Twofoldness

Richard Wollheim has become Gombrich’s self-proclaimed antagonist concerning the issue of twofoldness. In fact, Wollheim developed the notion of twofoldness to directly argue against what he took to be Gombrich’s position in *Art and Illusion*.149
There, in discussing the now famous example of the duck/rabbit picture, Gombrich writes that:

We see the picture as either a rabbit or a duck [italics added]. It is easy to discover both readings. It is less easy to describe what happens when we switch from one interpretation to the other. Clearly we do not have the illusion that we are confronted with a ‘real’ duck or rabbit. The shape on the paper resembles neither animal very closely. And yet there is no doubt that the shape transforms itself in some subtle way when the duck’s beak becomes the rabbit’s ears and brings an otherwise neglected spot into prominence as the rabbit’s mouth. I say ‘neglected,’ but does it enter our experience at all when we switch back to reading ‘duck’? To answer the question, we are compelled to look for what is ‘really there,’ to see the shape apart from its interpretation, and this, we soon discover, is not really possible. True, we can switch from one reading to another with increasing rapidity; we will also “remember” the rabbit while we see the duck, but the more closely we watch ourselves, the more certainly we will discover that we cannot experience alternative readings at the same time [italics added]. (Gombrich, 2000, p. 5)

When we see the picture as a duck (or a rabbit), since “we cannot experience alternative readings at the same time” we must switch our interpretation if we then want to see the picture as a rabbit (or a duck). I will argue that Gombrich is actually concerned with two distinct types of switching, and I will return to discuss each of them in a later section.

With this in mind, all that the reader should be aware of is that while Gombrich’s idea of switching appears to result in the rejection of twofoldness, I will attempt to show that this need not be the case.

According to Wollheim, the problem with Gombrich’s account is that he tries to collapse two different disjunctions into each other: the seeing canvas/seeing nature (or “seeing medium/seeing object”) disjunction, which holds for picture perception generally; and the “seeing duck/seeing rabbit” disjunction, which holds for the special case of the duck/rabbit picture. The two interpretations in the seeing duck/seeing rabbit disjunction are two meaningful interpretations. That is, they are meaningful in the sense
that each is an interpretation with some representational content. Contrast this with the other disjunction, where the two interpretation are between a meaningless interpretation (“canvas” or “medium”) and a meaningful interpretation (“nature” or “object”). If our seeing of Degas’ painting *Self portrait with soft hat* is just a seeing of the canvas (medium), then we will simply see the painting as an inchoately marked surface—dabs of paint on a canvas. We do not see the marks as adding up to anything more than just marks. We can then, should we so choose, switch our interpretation and see the marks as adding up to some representational content: we can see the marked surface as having the meaningful interpretation “A man with a soft hat,” or “Edgar Degas,” or something like this. Regardless of what our interpretation is, when we see what the picture is of we have a meaningful interpretation of the picture.

Having a meaningless interpretation of the picture accords with what Wollheim calls the picture’s “configurational aspect” (i.e., seeing the medium); while having a meaningful interpretation of the picture accords with what he calls a picture’s “recognitional aspect” (i.e., seeing the picture’s representational content). (I will talk more about these two aspects later on.) For now the reader should know that both Wollheim and Gombrich hold the second disjunction to be exclusive—we cannot see both meaningful interpretations (duck and rabbit) at the same time—but that Gombrich is accused of attempting to prove that the first disjunction is exclusive as well—that we cannot see both the medium and the object of the representation at the same time. This claim, if it is true of Gombrich, seems to entail the rejection of a simultaneous experience involving both the configurational and the recognitional aspects. Put differently:
Gombrich seems to reject the possibility of having a meaningful and a meaningless experience of the same picture plane at the same time.

It will require more investigating to determine whether these claims are true of Gombrich. For now, let me take the next two sections to talk a bit more about what is involved in twofoldness.

4.3 Why twofoldness?

The introduction of twofoldness arose through Wollheim’s rejection of a seeing-as theory of representation. In the wake of this rejection Wollheim proposed a new theory: seeing-in. I will not take the time to mention all the details of this theory. Suffice it to say, readers should know that Wollheim offers three reasons for making the switch to seeing-in and it is his third reason that is most relevant here. Wollheim contends that the perception of pictorial representations involves a distinctive kind of seeing, seeing-in, and when engaging in this distinctive kind of seeing viewers have an experience of the picture that is of a twofold nature. Bence Nanay succinctly describes Wollheim’s understanding of this experience in this way: “If an agent sees x in y, she is visually aware of both x and y simultaneously.” (2004, p. 285) So seeing an object O in a picture P means that I am, at the same time, visually aware of both O and P—I am aware of the object O and the picture P in which I see O. This claim is meant to highlight the unique difference between seeing-in and seeing-as: seeing-in allows for, as Wollheim describes is, “unlimited simultaneous attention to what is seen and to the features of the medium. Seeing-as does not.” (Wollheim, 1980, p. 212) This is supposed to be due to the fact that there are
certain restrictions placed upon what we can see something as. If I see P as O there will be certain “sustaining features” that permit me to see it, or explain my seeing of it, as y. On the other hand, “I cannot simultaneously see [P] as [O] and be visually aware of the features of [P] sustaining this perception.” (Wollheim, 1980, pp. 212-213) Instead, in order to be aware of the features that sustain my perception of P I must switch my attention. In other words, I cannot be aware of seeing P as O and be aware of how I am able to see P as O. The advantage of seeing-in is that “I may very well be able to see [O] in [P] and yet there be no delimitable features of [P] that can be looked upon as sustaining features of my doing so.” (Wollheim, 1980, p. 213) Furthermore, in cases where there are sustaining features of my seeing O in P, “I can simultaneously be visually aware of the [O] that I see in [P] and the sustaining features of this perception.” (Wollheim, 1980, p. 213)

The idea, then, is that twofoldness (Wollheim refers to it as “the twofold thesis”) allows for “unlimited” simultaneous visual awareness of both what is seen in the picture (the object of representation) and the features of the medium (the materials and the surface of the picture). When we look at a picture we see, at the same time, (a) the marks constitutive of the picture as a flat surface, and (b) the representational content built up out of those marks. In actuality, though, Wollheim makes a stronger claim: “If I look at a representation as a representation it is not just permitted but required of me that I attend simultaneously to object and medium.” (Wollheim, 1980, p. 216) So anyone who looks at a picture as a representation must see both the object of representational content and the medium. Furthermore, Wollheim goes on to say that when looking at a picture, twofoldness acts as “a normative constraint” upon anyone who tries to appreciate
This means, as Katerina Bantinaki puts it, that should a viewer focus “exclusively” on either the medium or the object that is being depicted, the object of sight does not figure in her visual awareness as a picture.” (Bantinaki, 2010, p. 131) If this should prove to be the case and our seeing of only the medium or only the object of representation does not result in our seeing the picture “as a picture” at all, then this fact would be a serious problem for Gombrich’s description of the pictorial experience. It would mean, in effect, that when we switch from seeing the medium to seeing the represented object we will lose awareness of the fact that what we are seeing is a picture. (I will return to this challenge in a later section.)

4.3.1 Wollheim and a single conception of twofoldness

It should be noted that there may prove to be some difficulty in talking about Wollheim’s having a single notion of twofoldness, since Wollheim seems to have reworked the concept over the span of his lifetime. This is evident from the three primary texts in which Wollheim presents and develops the notion of twofoldness: “Seeing-as, Seeing-in, and Pictorial Representation” (1980); Painting as Art (1988); and “On Pictorial Representation” (1998). The conception of twofoldness that Wollheim introduces and develops in the first two texts is modified, but not wholly rejected, in the last. While I have already quoted extensively from the first text (since it is there that we find the inception of twofoldness), I have done so in such a way that allows for many of my remarks to apply to the later, more settled conception of twofoldness found in “On Pictorial Representation.” Briefly, Wollheim’s major shift can be understood as revolving
around whether the experience of looking at a picture involves multiple perceptual experiences, or whether it involves a single perceptual experience with multiple aspects.\textsuperscript{155} The earlier Wollheim (1980, 1988) adhered to the first description of twofoldness as involving “two simultaneous perceptions: one of the pictorial surface, the other of what it represents.” (Wollheim, 1998, p. 221) On this account, then, when we look at a picture we are actually having two simultaneous \textit{but separate} perceptions: one of the picture’s surface and the other of what the picture is of. In his later work, Wollheim came to see this as problematic and instead chose to adopt a conception of twofoldness as involving “a single experience \textit{with two aspects} [italics added], which I call configurational and recognitional.” (Wollheim, 1998, p. 221)

Wollheim description of these two aspects (and especially how they are meant to relate to each other) is not all that detailed; and this, in turn, has provoked countless discussion.\textsuperscript{156} That being the case, let me attempt to offer some clarificatory remarks.

The configurational aspect is intended to relate to the marked surface (the medium) while the recognitional aspect relates to what it is we recognize the picture as being of (the object). As regards the configurational aspect, Wollheim writes that when we see a picture we should be aware "of the sustaining features of representation—features of x that permit me to see it as y" (Wollheim, 1996, p. 212-213). In short, we retain an awareness of the picture’s medium. Conversely, the recognitional aspect relates to whatever it is that we see in the picture, the subject or content of the picture. Here we are aware of the object represented by the picture. While there remains some contention over what exactly Wollheim has in mind as to the content of the configurational aspect\textsuperscript{157}, the basic idea is that when we look at a picture we simultaneously take in the medium
(the features of the picture as a picture) and the depicted object (the features of the picture as a representation). In the first case we see things like the flatness of the picture surface; various arrangements of marks (e.g., lines, shapes, and colors); the picture frame, and so on. While some of these features can be what we might call “representationally relevant” (i.e., they are constitutive of “how” a picture represents something and thus might also be features of the picture as a representation), as applied to the configurational aspect all of these sorts of features are perceived as inchoate (i.e., they are perceived as having no representational meaning or relevance). In other words, when we perceive the configurational aspect we don’t perceive the arrangements of marks on the surface of the picture as “adding up” to anything—they are just marks. When we turn to the experience of the picture’s recognitional aspect we find that we can see the picture as being of something; we see the marks as having meaning, and thus that they “add up” to some representational content.

If we are now to describe what it is for the experience of viewing a picture to be “twofold” on Wollheim’s account, we can say that a viewer simultaneously experiences a configurational awareness (an awareness of the picture’s medium, its surface) and a recognitional awareness (an awareness of the picture’s subject, the object it represents). Furthermore, these two aspects, though distinct, are connected together, somehow, in the single perceptual experience of looking at a representational picture. So when Wollheim says that the experience of looking at a picture is simultaneously a twofold experience he means that it consists of these two aspects, occurring together in the same perceptual experience, without requiring the viewer to switch back and forth between
them. The viewer has “access” to both aspects, at the same time, in the same perceptual experience of looking at the picture.

I will not belabor the differences between Wollheim’s earlier and later conceptions of twofoldness. Suffice it to say, as I proceed in this discussion I will be relying on the later conception of twofoldness—understood as a single perceptual experience consisting of two aspects—since he sees that conception as the settled conception; additionally, most theorists take the later conception to be the conception of twofoldness. Furthermore, given that this is how we should understand twofoldness, I will attempt to offer a reply to the criticisms of Gombrich’s supposed rejection of twofoldness in light of this later conception.

4.4 Gombrich and Twofoldness: A case of switching interpretations

For most theorists, the issue of whether or not Gombrich rejects twofoldness has been a non-starter. Andrew Harrison states it quite bluntly: “It is central to Gombrich’s account of our response to pictures that attention to pictorial surfaces and to what they depict are mutually exclusive.” (2001, p. 40) It is no doubt easy to arrive at this conclusion since Gombrich’s own words in Art and Illusion seem to make him out to be the antagonist of any sort of twofold experience159. In fact, Gombrich’s very own discussion of the duck/rabbit picture is often used as evidence against him. As I have already pointed out, Wollheim accuses Gombrich of conflating two things: (a) the experience of seeing the medium and the represented object (i.e., a meaningless and a meaningful interpretation of P), and (b) the experience of seeing the duck and the rabbit
(i.e., two meaningful interpretations of P). So while Gombrich may have shown that we cannot experience the duck and the rabbit at the same time, and thus that we cannot simultaneously hold two meaningful interpretations, this does not prove that we cannot see the medium at the same time that we experience the represented object. If Gombrich fails to accept this idea, so Wollheim asserts, then we must reject his description of the pictorial experience.

But what exactly does Gombrich have to say on the matter of twofoldness? I contend that what Gombrich says on page 5 of *Art and Illusion* coherently fits with other passages throughout the text to form a more nuanced position on twofoldness then theorists give him credit for. To bolster my claim I will need to first contrast the two types of “switching” that Gombrich discusses. Section 4.4.1 will discuss the first type of switching, which involves switching between two meaningful interpretations; then, in section 4.4.3 I will discuss the second type, which involves switching between a meaningful and a meaningless interpretation.

### 4.4.1 Symmetrical switching: switching between two meaningful interpretations

The duck/rabbit example has become (for better or worse) intimately tied to the argument that Gombrich must surely reject any possibility of a twofold experience of looking at pictures. The fact that we must switch from one interpretation to another, that we cannot hold two interpretations at the same time, is fodder for the critics’ cannon. While I believe that Gombrich intends the conclusions of the duck/rabbit example to apply, in some analogous way, to all types of representational pictures, it remains a
strange device for achieving this end. This is due to the fact that, as an example that frequents psychological textbooks, it is a uniquely contrived image. It functions like a novel party trick, one specifically constructed to amaze and befuddle the audience (in much the same way as do many of the works of M.C. Escher). This does not mean that the duck/rabbit image fails as an example for the conclusion that we must switch between two meaningful interpretations; rather, it means that more discussion is necessary in order for the conclusion based on this example to apply to a wider range of pictures.

Take the following two sets of statements:

1. I see the picture as of a duck. I switch my interpretation and see the picture as of a rabbit.
2. I see the picture as of a landscape. I switch my interpretation and see the picture as of some cows.

When we look at the duck/rabbit picture we find that we can have two different meaningful interpretations, O (duck) and R (rabbit), but to “hold” O (i.e., to see the picture as of a duck) and then to “hold” R (i.e., to see the picture as of a rabbit) requires a viewer to switch interpretations. For Gombrich, this “switching” is a necessary part of representational seeing: we simply cannot experience two meaningful interpretations of the same marked surface at the same time. The type of switching that Gombrich is talking about in the first part of the quote from page 5 involves this type of switching between two different meaningful interpretations of the same picture plane. Gombrich goes on to reemphasize this point later on:

“What is interesting…is not so much the flexibility of our interpretations as their exclusiveness. It is easy to see the bull’s-eye as a head facing us, as a button, or as a letter. What is difficult—indeed impossible—is to see all these things at the same time [italics added]…It is through the act of “switching” that we find out that different shapes can be projected into the same outline. We can train ourselves to
switch more rapidly, indeed to oscillate between readings, *but we cannot hold conflicting interpretations* [italics added].” (2000, p. 236)

It might also be helpful to think in terms of the labels that Katerina Bantinaki (2007) uses: she refers to this first type of switching as “symmetrical”, since it involves switching between two subjects (duck and rabbit); in contrast, the other type of switching she labels as being “asymmetrical.” Moreover, it is only the second type of switching, and not the first, that involves holding a *meaningless* interpretation of a picture—seeing a picture as just an inchoately marked surface. (I will talk more about this second type of switching in the next section.) So in the first sort of cases we see the marks as meaningfully arranged on the page, and we interpret them in two different but meaningful ways. What we cannot do, however, is experience both of these interpretations at the same time.

What is odd about the duck/rabbit example is that, being uniquely contrived to afford a viewer two meaningful interpretations, it makes each interpretation fairly easy to discover. The image is used as a psychological experiment, one that plays with our ability to see a single image in multiple ways. And what it shows us, claims Gombrich, is that we cannot hold two meaningful interpretations of the same picture plane at the same time. This is no less true of our perception of objects in the world. If I interpret the large squarish shape outside my window as a house it will be nearly impossible to see that shape *at the same time* as a giant tank or an elephant. Even if the shape has some tank-like or elephant-like features I will need to dislodge the house-interpretation in order to hold a tank-interpretation or elephant-interpretation. The impossibility, here, is in perceiving one and the same object in two different ways *at the same time*. The present issue, then, is not whether I can see the house as both a house and as an arrangement of
inchoate colors and lines, nor is it whether I can see the house as something else (i.e., whether it is possible for me to offer a different interpretation of the same object in my visual field). Instead, the issue is whether I can see the shape out there as a house and as something else (tank, elephant) at the same time. And this, says Gombrich, is impossible. When he says that “we cannot experience alternative readings at the same time” and that we “cannot hold conflicting interpretations” I believe he means that our experience of the picture cannot involve consciously perceiving alternative meaningful interpretations at the same time.

4.4.2 Can we really switch between two meaningful interpretations?

Before I continue let me make a few critical comments about the idea of switching between two meaningful interpretations of the same picture plane. Gombrich claims that all pictures are ambiguous; that they require supplementation by a viewer; and that they can afford multiple interpretations that are just as equally probable but which might not be consistent. This is all well and good, but does a picture like Wivenhoe Park really afford us another equally probable interpretation? Can we really see the picture differently than the way we do? Of course we can fail to recognize this or that part of the picture, which will simply mean that the part we fail to recognize will be seen as inchoate marks; we can also mis-recognize this or that part of the picture, taking it to be a horse when it is really a cow. But even if this is the case, it does not entail that we mis-recognize the entire picture. Let me put this differently. In the case of the duck/rabbit we can switch between seeing two meaningful interpretations, and we can also switch
between seeing a meaningful interpretation (“duck” or “rabbit”) and a meaningless interpretation (inchoate marks) of the duck/rabbit image. When we look at *Wivenhoe Park* we can certainly see it as having a meaningful interpretation (“landscape” or “of some cows”) or as having a meaningless interpretation (dabs of paint on a canvas), but are there really multiple meaningful interpretations of the entire picture that we can switch between? Even if there are possible interpretations that one might have, will we ever have one, or do we simply either see the picture as having the content that it does or as being inchoate marks on a page?

Constable, writes Gombrich, “had learned to paint…a flat patch that allowed of any number of readings, including the correct one. Ambiguity cannot be seen, and so we rightly ignore the innumerable weird interpretations that must also lurk behind the serene surface of the painting. For as we scan the flat pigments for answers about the motif ‘out there,’ the consistent reading suggests itself and illusion takes over. Not, be it said, because the world really looks like a flat picture, but because some flat pictures really look like the world.” (Gombrich, 2000, p. 329) If my claims from the previous chapters are plausible, the “correct” interpretation, the one with a “consistent reading” that “suggests itself,” will be the one that both faithfully captures the relations of visual elements found in the subject and which then affords recognition of that subject. Furthermore, if the flat picture can “really look like the world” and, in functioning as a relational model, is able to capture the arrangement of relations of visual elements in some non-arbitrary way, how is it that we can interpret the picture in any other meaningful way?
Gombrich does not say anything about the other “innumerable weird interpretations” that a picture might afford, nor how, exactly, one arrives at the “correct” one. In the case of a depictive picture that is relationally faithful in the manner I have described (and which is able to “really look like the world”), my claim is that in being a relational model the picture affords us recognition of the subject and thus affords us a meaningful interpretation of the picture’s subject. And once we have a meaningful interpretation it can be very difficult to detach it (Gombrich, 2000), which means that in most ordinary cases of looking at depictive pictures arriving at a meaningful interpretation will mean retaining that meaningful interpretation. Humans have become very good at interpreting images, and in most ordinary cases the meaningful interpretation that we arrive at will most likely be the “correct” one. If this is the case, then in most ordinary cases of looking at depictive pictures the possibility of being able to switch between two meaningful (albeit possible) interpretations will be very small. This does not mean that it is impossible, and certainly in cases where the picture is intended to yield different meaningful interpretations such switching will be possible (e.g., Surrealist or Cubist works). But if Gombrich intends this type of switching (i.e., between two meaningful interpretations) to apply to cases of naturalistic pictures such as Wivenhoe Park, then I am unsure exactly how he can maintain the competing claims that I have been discussing.

There is no doubt that we cannot hold two meaningful conflicting interpretations simultaneously, and on this Gombrich and I are in agreement, but this leaves open the question of whether we can even switch to another meaningful interpretation in the great many cases that Gombrich has in mind. Even if Gombrich is correct in his claim that
all pictures require supplementation and that completing an image requires the beholder’s share (by means of the projection of an interpretation into the picture), it is unclear how the constitutive marks of a picture like *Wivenhoe Park* can be interpreted as anything other than “a landscape with cows and a lake and a manor house…”

While these are important questions to ask of Gombrich’s account of switching I do not find the absence of an answer to the question of whether we can switch between to meaningful interpretations to be detrimental to the argument that switching occurs between *a meaningful and a meaningless* interpretation of a picture. With that said I must postpone any further investigation of them until a later time.

4.4.3 Asymmetrical switching: switching between a meaningful and a meaningless interpretation

Thus far I have been skirting the real issue that concerns Wollheim and which precipitated his criticisms of Gombrich. “For it is a central thesis of *Art and Illusion*”, writes Wollheim, “that, in looking at representational pictures, I am incapable of this kind of twofold experience.” (Wollheim, 1980, p. 213-214) When he says “this kind of twofold experience” Wollheim is not concerned with our being aware of different meaningful interpretations; instead Wollheim means the kind of experience that involves a simultaneous visual awareness of both the picture’s medium and the object represented in the picture. Gombrich’s position in *Art and Illusion* is captured in a quote by the French writer and painter Maurice Denis who tells us to, “Remember that a picture, before being a battle horse, a nude woman, or some anecdote, is essentially a plane
surface covered with paint in a certain arrangement.” (Gombrich, 2000, p. 279) This is a fact that, while not forgotten, can nevertheless remain “unseen” by most viewers. By this I mean that while we know that the picture hanging on the wall is, say, a portrait of some charging battle horse, we do not forget that it remains a flat, marked surface—we can walk around it and hold the thing in our hands to clearly see this! And yet the question remains: can we see the picture in both ways at the same time? As Gombrich puts it, the question we must ask is whether “is it possible to ‘see’ both the plane surface and the battle horse at the same time?” He concludes that, “If we have been right so far, the demand is for the impossible. To understand [italics added] the battle horse is for a moment to disregard [italics added] the plane surface. We cannot have it both ways.” (Gombrich, 2000, p. 279)

Here it appears quite clear that for Gombrich we cannot see the object of representation (in this case a battle horse) and the medium (the flat surface of the picture) at the same time. And yet the above passage requires further investigation, for it seems to me that we must inquire into what it means for a viewer to “understand” one interpretation and “disregard” another. What is it to understand the battle horse? And then, where does the other interpretation go when we disregard it? While Gombrich does say that in the course of pictorial seeing we “momentarily forget the canvas,” (Gombrich, 2000, p. 279) I think that he intends “forget” to be synonymous with “disregard.” When we see the charging battle horse we don’t throw ourselves out of the way because we forget that it is only a flat, painted horse. Disregarding the picture surface when we recognize the picture’s subject does not mean that we forget that we are looking at a picture; neither does it mean that we are not having a visual experience of, being visually
aware of, a flat surface. This, however, is exactly the conclusion Wollheim arrives at in his critique of *Art and Illusion*. Gombrich’s illusion account, writes Wollheim, “implies or suggests” that “seeing someone’s representation is quite continuous with seeing that person face-to-face” (Wollheim, 1980, p. 215). As Katerina Bantinaki puts it, Gombrich’s supposed failure is that “in seeing a picture we have an experience as of really [italics added] seeing the objects represented therein” (Bantinaki, 2007, p. 269). If this were Gombrich’s position then we must surely reject it, for it is certainly not the case that the visual experience we have when looking at a picture of O is the exactly same visual experience we have when looking at O itself. Were Gombrich to hold such a preposterous conclusion it would lead to all kinds of deceptive problems for viewing even the most mundane pictures hanging in our houses.

Thankfully Gombrich avoids this conclusion. Bantinaki points out that the above claim can only be attributed to Gombrich’s theory in the special case of trompe l’oeil. Trompe l’oeil pictures are, as Gombrich puts it, the “height of visual ambiguity” (Gombrich, 2000, p ), and as such they are able to “mask both their two-dimensionality and their medium.” (Bantinaki, 2007, p. 271) When we look at an ordinary picture we are cognizant of at least two important visual facts about it: that it is a flat, two dimensional artifact, and that it is composed out of a given medium. This is not the case with the trompe l’oeil. Because the trompe l’oeil is meant to literally fool the eye—we are meant to perceive it as the actual object—we will not recognize those visual cues that indicate flatness nor will we recognize the marks constitutive of the medium.163 We are fooled, deceived, and we take the pictured object to be the actual object.164 The reason for this deception, I take it, stems from the very fact of Gombrich’s supposed
denial of the twofold nature of pictures: if we do not see the pictorial medium concurrently with seeing what the picture represents, when we then find ourselves able to see what the picture represents it must be that we lose awareness of the medium and thus lose awareness of the fact that we are looking at a picture in the first place.\textsuperscript{165}

Gombrich’s view of pictorial seeing requires a viewer to switch between the interpretations one has when looking at a picture. Remember that the types of switching are of two sorts: between symmetrical pairs of possible interpretations (e.g., duck versus rabbit) and between asymmetrical pairs of possible interpretations (e.g., duck versus design). I have shown that there is no real controversy in his holding the first type of switching; it is the second type of switching that has garnered controversy and resulted in Gombrich’s account being rejected on the grounds of implausibility. Even if we grant that switching between an asymmetrical pair of interpretations does not result in forgetting about the picture, it still appears that the evidence is damning enough to cast doubt on the possibility that Gombrich’s position is compatible with twofoldness. How can I claim that Gombrich holds any semblance of a pictorial theory involving the sort of simultaneous experience Wollheim has in mind?

4.5 A way forward

What sort of experience do we have when we look at a depictive picture? The position laid out thus far is that we can either see a picture as merely inchoate marks on a flat surface \textit{or} as some represented content—we cannot see both at the same time. It may appear, then, that we are at an impasse: twofoldness requires a simultaneous experience;
Gombrich seems to reject such a suggestion. Is there a way to make sense of these two claims?

I believe that there is. To do so, however, necessitates making a subtle but significant distinction between two phrases Wollheim took as interchangeable: “being aware of” and “attending to” a picture’s medium. My claim is this: just because Gombrich claims that we cannot simultaneously see (i.e., visually recognize) both the medium and the object represented does not imply that we cannot be aware of both simultaneously. More specifically, we can be aware of the fact that the picture is a flat marked surface and that it is of a represented object without attending, in some conscious way, to both the medium and the represented object. I will pursue this distinction by discussing a number of other theorists, notably: Jerrold Levinson (1998), Bence Nanay (2005), and Katerina Bantinaki (2007). In bringing their commentary to bear on the present topic I hope to demonstrate how Wollheim was wrong in criticizing Gombrich for rejecting twofoldness, and how we can construct an interpretation of Gombrich’s theory of pictorial seeing that involves at least some form of a twofold experience.

4.5.1 The distinction between “being aware” and “attending to”

Jerrold Levinson (1998), in a symposium paper on Richard Wollheim, distinguishes two ways in which we can be visually aware of a picture’s medium or surface. The inception for this distinction stems from a passage in Wollheim’s book, Painting as an Art. There Wollheim writes: “Looking at a suitably marked surface, we are visually aware at once of the marked surface and of something in front or behind
something else.” (Wollheim, 1987, p. 47) Continuing on Wollheim says that if I then see “a boy in a stained wall I may very well concentrate on the stains, and how they are formed, and the materials and colours they consist of…and I might in consequence lose all but a shadowy awareness [italics added] of the boy. Alternatively, I might concentrate on the boy…and thus have only the vaguest sense [italics added] of how the wall is marked.” (Wollheim, 1987, p. 47) In reply to this Levinson asks the following question: what, exactly, does it mean to be visually aware of the picture surface? If I can go from having a “shadowy awareness” of the object represented to having “only the vaguest sense” of the marked surface, what exactly is going on when we make this kind of switch? Levinson concludes that when we see a picture’s content, at least in the cases he labels as “simple seeing-in,”“it is far from clear that…you must in some measure be attending to, taking notice of, or consciously focusing on [italics added] the picture’s surface or patterning as such.” (Levinson, 1998, p. 229) So when I see O in P I am able to take in P’s medium but not necessarily in a conscious or intentional manner, meaning that “the seeing-in involved in grasping pictorial content [does not] always [entail] or [include conscious] visual awareness of the surface as well.” (Levinson, 1998, p. 229) The surface and the medium can fade into the background as I concentrate on (i.e., attend to) the depicted content of the picture. I remain aware of the picture’s surface and the fact that it is flat; I am also aware that the picture is constructed of a specific medium (e.g., paint or charcoal). However, seeing the depicted content in the picture does not entail that I am, at the same time, consciously attending to the medium, nor does it mean that in attending to the one that I forget the other.
This sort of distinction, between “being visually aware” and “attending to,” is picked up on and elaborated by Bence Nanay (2005). Nanay points out that being aware of X and attending to X involves two different psychological attitudes, which Wollheim, however, appears to use interchangeably. This can be seen in the following passages:

The seeing appropriate to representations permits simultaneous attention [italics added] to what is represented and to the representation. (Wollheim, 1980, p. 213)

The explanation offered of this constancy is that the spectator is, and remains, visually aware [italics added] not only of what is represented but also of the surface qualities of the representation. He engages, in other words, in twofold attention [italics added], and has to if he wants to see representations in the way that we have come to regard as standard. (Wollheim, 1980, p. 215-216)

It is this development that…is invoked by Leonardo when…he encourages [the aspirant painter] to look at damp-stained walls or at stones or broken colour and discern there scenes of battle or violent action and mysterious landscapes. Now in the grip of such experiences the spectator enjoys a rather special indifference or indetermination. On the one hand, he is free, if he wishes, to overlook all but the most general features of the thing present. On the other hand, there is nothing to prevent him from attending to any of its features that he selects; he may not, of course, give them his full attention [italics added], but certainly he can give them peripheral attention [italics added]. The source of this indifference, or the spectator’s ability to attend or not to attend to the features of the thing present, lies in the fact that his essential concern is with the further visual experience, with seeing the battle scenes or the landscape, and this is, except in broad outline, discrete from visual awareness [italics added] of the wall or the stones. (Wollheim, 1980, p. 218)

While Wollheim seems to use the two phrases interchangeably—not explicitly differentiating between “being visually aware” and “attending to”—he does, especially in the final passage, hint at the fact that they are two different attitudes. It would thus seems that Nanay is correct in his assessment of Wollheim: there is a difference between giving the features of a picture “full attention” and giving them “peripheral attention.”
Drawing on Levinson’s discussion of the difference between simple seeing-in and seeing a picture aesthetically\textsuperscript{170}, Nanay shows that “the notion of visual awareness Wollheim uses when describing the twofoldness of the experience of looking at pictures is not necessarily ‘attending to, taking notice of, or consciously focusing on’ something.” (Nanay, 2005, p. 255) In other words, seeing O in P does not necessarily involve a conscious activity of seeing the surface features of the picture\textsuperscript{171}; instead, “in order to see an object in a picture, one only needs to register (without attending to it) [italics added] the surface one sees this object in. And even little children looking at a photograph of their mother are capable of this.” (Nanay, 2005, p. 229) Nanay is saying that to “register” the surface of the picture is to engage in a subconscious perceptual activity, one that does not involve our explicitly attending to the features of that surface. To have a twofold experience of the picture we need only be aware of the medium (perhaps only giving it our “peripheral attention”) when we recognize the represented object, and this notion of “being aware” is distinct from that of “attending to, taking notice of, or consciously focusing on”.

4.5.2 Some Clarifications

Let me try to get a bit clearer on what is involved in this distinction between “being visually aware” as opposed to “attending to.” Let us say that I am talking to a friend at a coffee shop. As we are conversing, a customer walks into the coffee shop and orders a coffee. Now, I can carry on my conversation with my friend and simultaneously be visually aware that there is a person walking around the peripheries of my visual field.
In this case, what I cannot do is consciously attend to both my conversing friend and to the customer. That is, I cannot focus my visual attention on both these objects in my visual field, at the same time.\textsuperscript{172} Consciously attending to my friend requires me to disregard the person walking behind her. Of course “disregarding” the customer does not amount to some kind of visual blackout—it simply means shifting my (visual) attention from one object (or focal point) to another. So as I attend to my friend I visually disregard the customer, though I remain aware that he or she is there.

We can see a similar sort of thing occurring in cases of auditory stimuli. Imagine that as I am conversing with my friend there is a conversation going on behind me, to which I turn my attention. My friend is talking—I can hear sounds and make out words here and there—but I could not tell you exactly what is being said; I don’t know the content of what she is saying. I am aware of her speaking since I am aware of the sound of her voice, but I am attending to (i.e., consciously focusing on) the other conversation. I simply cannot devote my conscious attention to both. Even more, when I return my focus to the conversation I am having with my friend I will find it impossible to hear what she is saying as \emph{both} meaningful words and meaningless noise. That is, when I hear her saying, “Well, then I decided to simply walk out…” I cannot at the same time hear that utterance as merely meaningless audible noise. If I am attending to her and to what she is saying (i.e., I interpret what she is saying as having meaningful content) I cannot also hear it as mere inchoate noise. In the previous case, when I overheard the other conversation I registered that my friend was speaking but I could not recognize the content of what she was saying since I was attending to the other conversation; I could not interpret what my friend was saying at the same time as I interpreted the other
conversation because I am not able to attend to both conversations simultaneously—I was attending to one and merely aware of the other. One required a more conscious activity on my part while the other required me to merely register the audible stimuli.

This, then, is what I believe Levinson and Nanay are pointing out when they make the distinction between “being visually aware” and “attending to.” Seeing something in a representational picture does not require us to consciously attend to both the content and the medium simultaneously. We are able to remain visually aware of the medium (in a peripheral way) while we attend to the content of the picture.

4.5.3 Applications for Gombrich

It is my belief that, in building upon the material found in Levinson and Nanay, one can construct an interpretation of Gombrich’s position of pictorial seeing that allows a viewer to be visual aware of a picture’s medium concurrently with recognizing the picture’s content but which does not require the viewer to consciously attend to the picture’s medium. In this way we find that there is a place for twofoldness in Gombrich’s account. Recall that Gombrich says that when we see a picture P’s content O (i.e., when we see O in P) we “disregard” the medium, and so “forget” the surface of P. But in an analogous way to the case of the customer walking behind my friend, disregarding or forgetting the medium does not entail visually blacking it out all together. The medium is still there and I remain aware of it—I remain aware that the picture is made up of globs of paint or strips of newspaper, that it consists of flat lines and shapes—and yet my ability to
momentarily disregard it simply amounts to disengaging it from my conscious visual attention.

This is the sort of reply to Wollheim made Katerina Bantinaki (2007). Bantinaki argues that Gombrich “allows that when one sees what a picture represents the medium can be somehow part of one’s visual awareness: what it excludes is the possibility of seeing the picture’s designs as a meaningless design—as just marks on a surface—while seeing what the picture represents.” (Bantinaki, 2007, p. 272) Bantinaki is saying of Gombrich that we cannot, at the same time, see the picture as a meaningless marked surface and as the represented object. Bantinaki offers a wonderful example to help illustrate this point:

Think of a matchstick-man: it is an artefact [sic] made of matches arranged so that the whole roughly corresponds to the structure of the human body, and it represents a real-world entity—a man, maybe even a particular man. When we look at the artefact [sic], if it is successfully crafted, we will see a matchstickman: neither a meaningless set of matches, nor the real-world entity. Of course we do see the matches when we see the matchstick man—the ‘medium’ is not absorbed by our recognition of what it represents; but we do not see them as just matches, rather we see matches shaped in a familiar configuration, the form characteristic of a man. That the medium is thus part of our awareness is what makes the experience of seeing a matchstick-man phenomenologically different from the experience of seeing a real man. (Bantinaki, 2007, p. 272)

There are two conclusions we can take from this example and apply to my assessment of Gombrich. The first is that Bantianki offers credence to the idea (following Levinson and Nanay) that there exists an important difference between being our visually aware of and our visually attending to a picture’s medium (at least as the terms are being used as regards twofoldness). So when Gombrich says that we must switch interpretations between medium and object what he must mean is that we must switch our attention between the medium and the object. We can remain aware of the medium when we see
the represented object (in the same way that we remain aware of the matches when we see the matchstickman), but what we cannot do is consciously attend to both at the same time. We cannot hold both interpretations at the same time, and so we must switch from one to the other.

The second conclusion arises from Gombrich’s position on ordinary and pictorial perception. I have interpreted Gombrich as holding that the processes involved in pictorial seeing are very similar to those used in ordinary seeing—certainly not identical, but at the very least each engages similar sorts of visual activities. Wollheim takes issue with this position, chastising Gombrich for failing to “assign to the seeing appropriate to representations a distinctive phenomenology [which] impels him towards the view that there is nothing distinctive about the seeing of representations, or that seeing someone’s representation is quite continuous with seeing that person face to face” (Wollheim, 1980, pp. 214-215). For Wollheim, if we understand the processes involved in pictorial seeing to be similar to those involved in ordinary seeing we are in danger of losing what is “unique” about the experience of looking at pictorial representations—Wollheim calls this “the delights of representation.” (Wollheim, 1980, p. 219) Ordinary seeing is a capacity that humans and other animals have as a means for perceiving things in the external world; the seeing distinctive of perceiving pictures, which for Wollheim is seeing-in, remains something “which presupposes, but is something over and above” the capacity of ordinary seeing.” (Wollheim, 1980, p. 217).

Now, even if we grant that pictorial perception presupposes ordinary perception it certainly does not entail that the former is identical to the latter; at the same time, however, granting this does not deny the possibility that there exists some similarity
between the two. Just because ordinary perception is not identical to pictorial perception it does not follow that the one is wholly different from the other. Listening to my wife read me a recipe, listening to a recording by Thelonious Monk, and listening to a chickadee may all involve difference kinds of auditory processes, but there is no reason to conclude that each involves \textit{wholly different} processes. Each involves the processing of auditory information, and in that way listening to my wife, listening to Thelonious Monk, and listening to the chickadee each presuppose the ability to hear in the first place (i.e., the ability to hear and processes, in a consistent manner, the auditory information I receive).

Wollheim seems to think that for Gombrich to see similarities between ordinary and pictorial perception is to some how do a disservice to the seeing of pictorial representations. This is a rash conclusion to jump to, and I believe that it is the result of the underlying misconception by Wollheim and many other interpreters of Gombrich that Gombrich’s illusion theory leads to deception: because we take the picture \textit{to be} the object we find ourselves fooled; hence, Gombrich’s account offers us no reason to think that pictorial perception is a unique perceptual activity, different from that involved in perceiving the external world.\textsuperscript{176}

In the previous chapter I proposed a plausible account of how there might be some similarity between ordinary and pictorial perception, and, more importantly, I argued that Gombrich’s holding such a position does not entail our being deceived or fooled by a picture. In this way I believe that Bantinaki’s matchstickman example offers a correct assessment of Gombrich’s position on twofoldness. In looking at a picture we remain visually aware of the picture’s medium despite not consciously attending to it, and
because the medium remains part of our awareness the experience of seeing a picture of O is phenomenologically different than that of seeing O itself.

The awareness/attention distinction I have called attention to is equally applicable to our perception of objects in the world, and it is roughly akin to the phenomenologist’s distinction between thetic and non-thetic conscious and the Gestalt principles of figure/ground (see endnote 167). In ordinary perception the idea is that I make something in my visual field the object of my attention, while the rest of the objects in my visual field recede into a kind of peripheral awareness. Attention shifts and changes constantly as I focus on different objects within my visual field, making some objects the primary focus of my conscious perception and relegating others to serve as conduits for that perception (e.g., there is some object behind or above me; I can tell that there are bright colors over there behind the passing truck; there is something moving away or towards me). I remain aware of these peripheral items even though I am not attending to them. This analogy of a distinction of consciousness allows me to maintain an organized and coherent experience of the world. Applied to pictures such as Wivenhoe Park, Gombrich will tell us that when we shift our attention to the manor house and “disregard” or “forget” the medium we remain aware that the medium is there despite not consciously attending to it. The medium (when we are focusing on the manor house) remains in our perceptual awareness. So it would seem that Wollheim’s criticism can only be applied to certain unique cases where our experience of the picture involves a lose of the visual awareness of the picture qua picture; which is what can occur in the case of trompe l’oeils. It is only when such an awareness of the medium is lost, when it is, as Bantinaki puts is, “not
absorbed by our recognition,” that pictorial perception fails to be distinguished from ordinary perception and we become deceived or fooled by the picture.

4.6 Gombrich and twofoldness: final thoughts

I have tried to show that Gombrich has been wrongly accused of rejecting twofoldness for a few different reasons. There is no doubt that Wollheim brings to his critique certain underlying prejudices regarding Gombrich’s illusion theory in general. At the same time, Gombrich does not make things easy on himself: he does not explicitly distinguish between the two psychological attitudes of “being visually aware” and “attending to”; furthermore, his writing leaves the reader up to his or her own devices when attempting to piece together details of the kind of experience had when viewing pictures. Add to this the unique nature of the duck/rabbit example, the two different types of switching involved in perceiving it, and Gombrich’s failure to detail explicitly the sorts of experiences involved in each type of switching, and it is no wonder that readers and theorists have drawn the conclusion that Gombrich rejects twofoldness.

Despite these facts, I attempted to demonstrate a way to interpret Gombrich that does not exclude twofoldness. If we draw a distinction between our “being visually aware of” and our “attending to” a picture’s medium, Gombrich’s position in *Art and Illusion* can be interpreted as allowing for some form of a twofold experience. When we see P we can, at the same time, remain visually aware (in a kind of peripheral way) of P’s medium while seeing that P is of O (i.e., while seeing O in P). Should we wish to attend, in a conscious way, to P’s medium we must switch how we see P. In this way it is impossible
for us to see a picture as both a meaningless design and as representing some object. Should we wish to see the picture only as a meaningless design we must switch our interpretation, attending to the marked surface in a conscious manner. And, at the same time, we remain visually aware of the medium when we see attend to the object depicted in P: the medium remains in the background as a part of our awareness, it is not given our conscious attention.

Interpreting Gombrich in this way succeeds in two ways: first, it follows the textual evidence found in *Art and Illusion*. Any systematic account of Gombrich's position needs to be pieced together from the available textual material, and a Gombrichian account of twofoldness is no different. With that being said, I think that the account I have proposed fits with what Gombrich says and does well to fill in the gaps left in his writing. Secondly, the proposed account serves as a reply to the critical remarks made by Wollheim, and others, that Gombrich should be faulted for failing to integrate twofoldness into his theory of pictorial representation. While I have raised questions of Wollheim's own account of twofoldness, I have also shown that incorporating the distinction between "being visually aware" and "attending to" serves to complement the position that Wollheim lays out.
The Gombrichian concept of illusion

5.1 The goal of Chapter 5

In many ways the preceding chapters have been a preamble to this one. As any reader who is familiar with Gombrich’s writings will know, the major worry that critics have expressed over Gombrich’s writings concerns the nature of illusion: what exactly does Gombrich mean when he says that a picture creates an illusion? What is it for viewers to have the illusion of a represented object? Is illusion synonymous with delusion? This chapter will utilize the conclusions from the previous chapter to offer answers to each of these questions.

The first issue to be discussed is why there remains confusion over whether the Gombrichian concept of illusion is synonymous with delusion. In section 5.2.1 I offer two guesses as to why this position continues to exist despite the accessibility of Gombrich’s writing. Then, in sections 5.2.2-5.2.3, I offer more detailed evidence that aims to disentangle Gombrich’s position on illusion from that of delusion: because ordinary pictures are different from trompe l’oeil (in the visual experiences they are intended to produce), apropos of pictorial representation “illusion” cannot be used interchangeably with “delusion.”

I maintain that because ordinary pictures and trompe l’oeils are both instances of pictorial representation (and thus both examples of visual substitution) that both involve Gombrich-style illusion. Still, in section 5.2.4 I provide a tool to help distinguish ordinary pictures from trompe l’oeil and thus further distinguish the concept of “illusion” from
“delusion.” What distinguishes the one from the other is that ordinary pictures offer viewers a continued sense of what I call “medium awareness”—the ability to remain aware of the picture’s medium such that our awareness of it is not absorbed by our awareness of the content of the picture.¹⁷⁷ In remaining aware of what (or how) the picture is constructed the path to deception and false belief is barred—whatever the experienced illusion is, it is not delusion. Only in the case of trompe l’oeils is there the possibility of epistemic delusion (false belief), and even in such cases the false belief is so fleeting, so momentary, that we might question whether it is a substantial false belief at all. (I will not pursue this position in earnest, but instead simply gesture towards the idea that the epistemic delusion of the trompe l’oeil may not be as potent as has been claimed.) The notion that we remain aware of the fact that we are looking at a picture is continued in section 5.2.5 and concluded in section 5.2.6, where I talk about Gombrich’s claim that all ordinary representational pictures provide us with the “true information” that what we are looking at is a picture, not the object itself.

Two of the most important concepts involved in a Gombrichian conception of illusion are the idea that an object can serve as a substitute (previously discussed in Chapter 1) and idea that a substitute offers various affordances (discussed in Chapter 4) capable of being realized by a user; these are discussed in sections 5.3.1 and 5.3.2 respectively. Here I will expand my earlier claims about physical substitutes (such as the hobbyhorse) to include the idea that the pictures serve as visual substitutes for the things they represent. In Section 5.3.4 I then take the account of affordances and explicitly apply it to the discussion of pictures. Given the claims about the affordances a picture offers a viewer, sections 5.3.5-5.3.6 inquire how it is that a picture can be a non-delusory
visual substitute for some represented object. I maintain that a picture can be of one thing (an orange) rather than another (an apple) because the picture functions as a substitute for some subset of visual information found in the object, namely some constitutive set of relationally arranged visual elements the perception of which can trigger off a visual (i.e., recognitional) experience for the viewer that is similar, but not identical to, the visual experience triggered off by the object.

In conclusion, section 5.4 offers some final classificatory remarks on how we should understand a Gombrichian conception of illusion. If we wish to make sense of what it means for a picture to afford an illusion of what it represents (in Gombrich’s sense) we must permanently divorce the concept from delusion, false belief, and deception. Illusion exists along a spectrum of possibilities, and the Gombrichian conception of illusion apropos of ordinary pictures does not involve delusion, which will only occurs in certain instances such as the trompe l’oeil. To have a pictorial illusion (in the case of ordinary pictures) is to experience the imagined reality of what the picture represents without taking that experience to be presenting us with the actual object. The usual instances of pictorial illusion means that a viewer remains aware of the picture’s medium concurrently with the picture’s content—the artist has engaged in an act of transforming a flat marked surface into an artifact affording the non-delusional perception of the object it represents.
5.2 What illusion is not: distinguishing illusion from delusion

5.2.1 Why there remains confusion over Gombrich’s position on illusion

The first task I must undertake is to be clear about a platitude that still garners a voice in a number of critical circles, viz. that Gombrich’s position on illusion is tantamount to the assertion that illusion is simply delusion. In short, this claim states that the Gombrichian conception of illusion leads to the unsavory conclusion that seeing a picture of O generates the same experience as seeing O itself; and so, rather than believe that we are looking at picture of O, when we look at the picture we believe that we are looking at O itself.

I am reluctant to have to return to this claim, as I believe that Gombrich has been quite explicit regarding the fact that “illusion” does not amount to false belief or delusion. The terms are not synonymous, and experiencing a pictorial illusion (of the kind Gombrich has in mind) does not result in a viewer being perceptually or epistemically deceived by the picture—an illusion (in Gombrich’s sense) does not result in a viewer taking the picture to be the represented object. About this I believe Gombrich is quite clear. As it is, however, the association of illusion with delusion remains a major impasse for readers of Gombrich’s work. So if we are to grant that the kind of pictorial illusion we are concerned with (in the case of ordinary pictures) is distinct from delusion, the reader should want to know why is it that such a position continues to be advanced as the truth about Gombrich’s account. Why do theorists still take up arms against Gombrich and his “delusion” theory of representation?
There is no easy answer why this misinterpretation remains in acceptance. I will, however, wager three guesses that might shed some light on the matter. The first is that Gombrich’s primary focus in *Art and Illusion* is not illusion in art per se, nor is it the idea that illusion is or should be the aim of art. The book itself is based on a series of lectures (entitled “The Visible World and the Language of Art”); yet the book’s subtitle, *A Study in the Psychology of Pictorial Representation*, should offer no misunderstanding of what it is about. The goal of *Art and Illusion* is not to lay out a technical, detailed theory of illusion. Rather Gombrich makes use of a lecture format that allows him “the privilege of leaving stones unturned and avenues unexplored,” which in turn forces the reader to devise his or her own answers to any unresolved questions (2000, p. xxxviii).

While the accessibility of *Art and Illusion* goes a long way towards allowing it to appeal to both lay-person and academic alike, there exists a kind of open-endedness to the position Gombrich puts forward that might leave readers nonplussed upon finishing the book.

The second guess as to why “illusion” remains tethered to “delusion” in interpretations of Gombrich’s writing revolves around the concept of “illusion” itself. Readers will hold a number of prejudices, expectations, and assumptions concerning the term, and it is no wonder that its introduction elicited such a diverse range of reactions. There is no doubt as to the “loaded” nature of the word “illusion,” but it might also be wise to remember that in the English language “the connotational range of the word ‘illusion’ is greater than that of the word `delusion’” (Woodfield, 1996, p. 278).

Colloquially, we can us “illusion” to mean a great many things, while its presence in an academic setting can leave readers grasping at different definitions. For instance, many of
the optical illusion that frequent psychology textbooks remain tethered to the notion that
their perception triggers a state of delusion. What are me to make of this delusion? If this
state of delusion is understood as epistemic delusion—the generation of a false belief—
then we might be mistaken in our application. Despite the irresistible visual “pull” of
illusions such as the Fraiser Spiral, the Müller-Lyer Illusion, or the Poggendorf Illusion,
there is no chance that they trigger a false belief about what we are seeing. That we
simply cannot see them otherwise might prove to be a helpful way of understanding my
interpretation of Gombrich’s conception of pictorial illusion. In these cases it would seem
that we can maintain a persistence of illusion without the persistence of delusion (or the
fixation of belief. (More on these thoughts later on.)

The point is that the terms “delusion” and “illusion” are being run together. Again,
this might be because usage has allowed for “delusion” to be used in place of “illusion”:
we can say that an illusory experience is a delusional one, meaning that the experience
contradicts what might be generally accepted as reality or as being rational. Or perhaps
the notion of illusion is being associated with some sort of mistake, and that this, in turn,
makes us think of delusion and deception. The deeper concern is that should we fail to
qualify the details of the nature of delusion we will be left in a muddled mess. We will
simply use the term illusion to cover all the situations, leaving us with a vastly larger
number of questions and concerns: Does having an illusion automatically imply that just
our senses have been deceived? Have we formed a false belief about what we have seen?
Do we take the content of our precept to be, quite literally, something that it is not?

This concern points to a general lackadaisical attitude in applying the correct
terminology, and since the term “illusion” seems to affords a more diverse range of
possible meanings it would seem to have lead to the more immediate confusion
surrounding Gombrich’s use of the term: we are unsure how exactly Gombrich
understands the concept of “illusion”—so we postulate an answer; we are unsure how
and in what context Gombrich intends the concept of “illusion” to be applied—so we
postulate that as well. In the end, our trying to fill in these gaps with postulated answers
has resulted in a confused state of understanding surrounding Gombrich’s position. One
certainly cannot blame readers for siding with established opinions about “illusion” and
about Gombrich, especially when those opinions are from well-respected theorists. But
even if we dispense with assigning blame, we can remain troubled by the fact that
Gombrich’s account is so often rejected out of hand and that Gombrich himself has been
labeled a “heresiarch” of the illusion theory (Schier, 1986, p. 142).

Finally, it is not insignificant to remember Gombrich’s “theory” of perception
(discussed in Chapters 1 and 3) and his association of ordinary perception with the
perception of pictures. He is explicitly clear that Art and Illusion “rests indeed on the
assumption that the illusion which a picture gives us can be explained by the similarity of
our reaction to the picture with our reaction to the visible world.” (Gombrich, 1969, p. 8)
Perception involves interpretation (in the case of perceiving pictures no less than in the
case of perceiving objects in the actual world), and Gombrich warns against wholly
detaching the one from the other. In taking up this position, and while also attempting to
breath new life into the concept of illusion, it seems inevitable that one might see
Gombrich as courting the possibility of illusion being associated with delusion. The
textbook examples of optical illusions leave the lay-reader with a taste of trickery and
deception, and if Gombrich is drawing analogies between how we look at pictures and
how we look at the actual world then there is a strong possibility that the perils and problems that come with applying the term “illusion” in cases in which we look at the actual world will haunt the use of the term in describing cases in which we look at pictures.

These are but guesses as to why there remains a confusion surrounding Gombrich’s use of the concept of illusion. There may or may not be truth to any or all of them, and thus they should be used as signposts to help illuminate the rest of the discussion as it moves forward. That being said, readers should ignore them at their own risk for they serve a valuable function: like the short description of an artwork found in a museum pamphlet they can offer some perspective as to why things are the way they are.

5.2.2 An illusion of what? The uniqueness of the commonplace

The previous section laid out guesses concerning why one might think Gombrich’s conception of illusion equates illusion with delusion. While the reader should keep them in mind, I believe that we can effectively dispense with this issue and take Gombrich at his word: illusion (as he understands the notion) is distinct from delusion, false belief, deception and the like. Of course, doing this does not resolve the lingering question of what Gombrichian illusion is and why it is not delusion. The rest of this chapter will hopefully answer this rather Gordian issue, but for now let me offer some preliminary remarks to help set the stage for what is to come.
If we can put aside the complexities of the question for a moment, it might be helpful to look at the succinct, albeit facile, description of illusion offered by Catherine Wilson (1982):

one way of explaining the ability of observers to classify pictures according to subject-matter...is by proposing that when a picture is identifiable as, e.g., a picture of a woman being menaced by a lion, this is because it gives the viewer the illusion, or allows him to enter into the illusion, that he is seeing a woman being menaced by a lion.” (p. 212)

This uncomplicated account of pictorial illusion offers a simple description of our perception of a representational picture and the resulting experience of seeing that picture as of its subject-matter. We have the ability to see a flat, marked surface as of (or in terms of) a represented object or scene, and it is a general ability shared by most viewers. Wilson finds labeling this occurrence as an “illusion” to be an accurate way to describe the uniqueness of such a commonplace event; I am inclined to agree with her. One reason is that Gombrich’s concern with the most mundane and “unartistic” of images demonstrates his desire for the concept of illusion to be freed from the traditionally labeled “illusionistic art.” Images are everywhere—they are so commonplace that they have become a part of the visual landscape of our everyday life; yet they retain a kind of magic that makes their perception unique. What Wilson seems to be pointing out, and what Gombrich will agree with, is that at its simplest the idea of illusion can describe the fact that we are able to see one thing as another without courting deception. Despite its may different connotations, illusion appears to be “well suited to a philosophical explanation of the ability” to classify a picture according to subject-matter (Wilson, 1982, p. 212). I find Wilson’s position to be significant for the present discussion since it appears to exemplify a similar sort of spirit held by Gombrich regarding the term
“illusion,” namely that while there might be “some other equally general way of explaining this ability” to classify a picture according to subject-matter “it is difficult to see what it might be.” (Wilson, 1982, p. 212)

In Chapter 3 I offered an account of how a picture affords recognition of its subject-matter due, in part, to the triggering off of similar recognitional experiences. We can begin to see how that account fits with Wilson’s position that illusion is helpful for explaining our ability to classify a picture according to its subject-matter. To make this position somewhat more credible I must still dispense with two worries: the first is the tiresome issue that the illusion is tied too closely to delusion. While not wishing to beat the proverbial dead horse, the next section will offer a more detailed reply to this worry by asking what we can learn by distinguishing a trompe l’oeil from an ordinary representational picture. Here I can briefly state that a trompe l’oeil is created to visually trick the viewer; and the very best trompe l’oeils are so visually potent that a viewer may not be able to see the brushstrokes constitutive of the picture—he or she is “compelled” to see the represented object.

My goal is not to unpack the phenomenology of the visual experience of a trompe l’oeil. Suffice it to say that while the visual experience of seeing the trompe l’oeil may result in a false belief about the nature of the depicted object there remains some contention of how lasting such a false belief can be (e.g., true epistemic deception may last only for a brief moment, and the viewer may be unable to return to that state of false belief once he or she is aware of the fact that they are looking at picture). That being said, we must still ask whether we can consistently apply the term “illusion” to the trompe l’oeil and to the ordinary picture if what we mean by “illusion” is really “delusion.” If
“illusion” is truly synonymous with “delusion” then we would expect the application of the term “illusion” to be limited to cases of pictures where it is possible that deception and false belief can occur. Ordinary pictures are not trompe l’oeil; an ordinary picture is not intended to deceive or trick the viewer into taking it to be the represented object itself, nor is there any real possibility of this happening. As such, using “illusion” interchangeably with “delusion” in cases of ordinary pictures is suspect, requiring a more detailed investigation. (I will argue, in section 5.2.4, that both ordinary pictures and trompe l’oeil involve the sort of Gombrich-style illusion that I develop; the difference, however, being a matter of degree.)

The second worry I must address follows on the coat tails of the first: it holds that any illusion theory “is incompatible with our ability to attend to the purely compositional details of a picture while continuing to see it as representational.” (Wilson, 1982, p. 215) Because we cannot maintain a twofold experience of the picture (à la Wollheim), any illusion theory must fail presumably because it will fall into delusion. The worry goes like this: the illusion view states that we cannot at the same time focus on both a picture’s “compositional details” (i.e., its medium) and its representational content (i.e., what the picture is of); therefore, when we attend only to the representational content of the picture (to the exclusion of the compositional details) we lose track of the fact that we are looking at a picture. Because the illusion theory fails to allow for this twofold experience (so the worry goes) we become deceived and take the picture to be the object itself.

Chapter 4 indirectly addressed this worry apropos of Wollheim’s account of twofoldness. There I argued for a more nuanced distinction involving “being aware of” and “attending to” a picture’s medium as a way to bypass Wollheim’s criticism of Gombrich. In this
chapter I will return to this distinction to argue that by retaining an *awareness* of a picture’s medium (as opposed to attending to it) we can further distance a Gombrichian account of illusion from the possibility of falling into delusion.

5.2.3 The possibility of “epistemically benign” illusions

I want to take this section to turn briefly to the discussion of illusion offered by Richard Allen in his paper “Representation, Illusion, and the Cinema.” (1993) This investigation of Allen’s position should help to shed light on the distinction I am trying to pursue, that illusion is not synonymous with delusion. While Allen is primarily concerned with illusion as it might be applied to the cinema, his remarks concerning illusion’s application to representational pictures have some bearing on my interpretation of Gombrich. In his discussion Allen confines the use of the term “illusion” to cases of trompe l’oeil paintings (apropos of illusion’s application to pictorial representation), and intends his use of the term “illusion” to be interchangeable with “delusion.” Yet, Allen also mentions the possibility (put forward by Noël Carroll but dismissed as “unilluminating” and “trivial”) that an illusion can be “epistemically benign”—meaning that the sense in which “x is an illusion of y” does not result in false belief. I believe that these two points, which Allen does not linger over, can help to bolster my interpretation of Gombrich’s account of illusion.

The first thing to point out is that Allen’s account of illusion in representational painting is an example of a very common mistake. There, Allen holds that illusion, if it is applied to pictorial representation *at all*, must be associated with the kind of sensory and
epistemic deception exemplified by a trompe l’oeil. I have already talked at length about the nature of trompe l’oeil paintings, however Allen’s description of trompe l’oeil sufficiently captures how the notion of illusion has come to be confined to these types of pictures:

The distinction [between representational paintings and photographs] paves the way for distinguishing two different kinds of illusionism—trompe l’oeil and reproductive illusion—associated with representational painting and photography respectively… (1993, p. 22)

In a trompe l’oeil painting the viewer is deceived by the character of what is perceived. Instead of seeing an object depicted in the surface of the painting the viewer perceives the object itself [italics added]. (1993, pp. 24-25)

In a trompe l’oeil illusion, the medium itself is disguised. (1993, p. 25)

When we experience the trompe l’oeil, our customary perception of the painting as a painting is inhibited…What we take for granted in our normal experience of representational painting—that our perception of how the object is depicted enters into our perception of the object—is thrown out of kilter. (1993, p. 25)

Later on in the same article Allen tells us that illusions actually deceive in two ways: “they deceive our senses, and they lead us to make false inferences.” (1993, pp. 33-34)

Illusions, then, can lead us to “see something that does not exist and believe that it does exist [italics added].” (1993, p. 34) We can experience sensory deception and epistemic deception, writes Allen, yet the two are distinct—they can occur concurrently but the presence of one does not imply the presence of the other. Allen’s remarks concerning trompe l’oeil, however, offer an example of an illusion that involves both sensory and epistemic deception. And given his comment that the term “illusion” (when applied to representational painting) should only be associated with cases of trompe l’oeil, we can clearly see an account that equates illusion with delusion apropos of representational pictures.
I find this significant for a few reasons. First, it presents us with a clear position that equates “illusion” with “delusion.” For Allen, if we assume that there are ever any cases in which the term “illusion” can be applied then such cases must be instances of trompe l’oeil; and if we are to apply the term “illusion” only to cases of trompe l’oeil then we must be using the term “illusion” to mean “delusion.” (This, of course, returns us to the question of why there remains confusion concerning Gombrich’s account of illusion. If theorists continue to offer accounts of illusion that equate it with the notion of delusion, indirectly or not, then readers of Gombrich will continue to shackle him to such a conclusion.)

The other matter to discuss regarding Allen is his very brief discussion of the possibility of an “epistemically benign” illusion. This was first put forward by Noël Carroll, who stated that it might be the case that the phrase “‘x is an illusion of y’ simply means that ‘x looks like y’ and deception plays no role.” (1993, p. 33) It is fair to infer from Carroll’s wording that if something x is an “epistemically benign” illusion of y then x does not generate a false belief or delusion about its nature. That is, seeing x does not lead one to falsely believe that he or she is seeing y. We can re-phrase “epistemically benign” to more formally read:

\[ x \text{ is an illusion of (or generates the illusion of) } y \text{ and } x \text{ does not result in a false belief that } x \text{ is } y \]

Allen initially plays down this possibility as “unilluminating” and “trivial,” because, as he puts it: “[to] the extent that all pictorial representations might be said to look like what they depict, all pictorial representations are illusions.” (1993, p. 33) I find this to be a strange thing to say. I have not found any theorist who would equate the claim “x looks
like $y$” with the claim “$x$ is an illusion of $y$.” Most critics of an illusion theory—should they be sympathetic to some kind of resemblance theory of pictorial representation or not—will either simply not talk about a connection between “looking like” and “illusion” (e.g., Neander, 1987) or else explicitly seek to separate the idea of “looking like” from that of “illusion” (e.g., Gilman, 1992; Hyman, 2006; Peetz, 1987).

Gombrich would certainly want more from an account of illusion than the simple fact that a picture looks like what it depicts—that much is clear. I have talked (in Chapter 3) about Gombrich’s rejection of purely mimetic, resemblance, and (what he calls) objectively iconic accounts of pictorial representation. That being said, I do not think that Gombrich wholly rejects the idea of “looking like” nor even Allen’s rather rhetorical claim that “all pictorial representations are illusions.” In fact that seems to be the very point Gombrich is trying to make, and it is certainly not an unilluminating or trivial one. To say that all representational pictures are illusions highlights the kind of magical act of transformation\(^{182}\) that Gombrich talks about: we see a picture as a flat marked surface and then we see that it as a seascape, or a woman, or whatever. And to find that these perceptual experiences arise from one and the same artifact it is to be confronted with an illusion—it is to be confronted with an act of magic. Gombrich would agree with Allen’s assertion but for very different reasons.

But of course we cannot stop there. Philosophy is tasked with answering those “why” questions—adventuring out beyond the harbor of routine acceptance and into the choppier waters of uncertainty and complexity. That is why I find Allen’s remarks both helpful and confused. Apropos of representational pictures, the term “illusion” need not be limited to trompe l’oeil. This is due to the fact, mentioned previously, that while

176
illusions can afford both sensory and epistemic deception we can experience one without the other. Our senses can be deceived without our entertaining any false beliefs about what it is we are perceiving. Trompe l’oeils are intended to, at the very least, generate such a strong perceptual deception that the picture’s medium is “hidden” from us; trompe l’oeils also have the possibility of generating in us the false belief that the picture is the represented object. So while we can use the term “illusion” to describe trompe l’oeils we must be clear that what we really mean is “delusion,” which is distinct from the use of the term “illusion” to describe ordinary pictures. Thus, for clarity’s sake I am thus claiming that we should apply the correct term “delusion” to the perception of trompe l’oeils because they generate a delusory perceptual experience in a viewer.183

Ordinary pictures do not involve delusions. They have not been created to elicit beliefs about the existence of the objects represented in or by the picture (and in the next section I will talk about why ordinary pictures are different from trompe l’oeil in this regard). Allen dismisses the possibility of epistemically benign illusions because he associates the idea of “illusion” with deception and delusion; he links it with the idea of a picture merely “looking like” an object. This is too naive an understanding of illusion; it will only hold if it is the case that illusion is synonymous with looking like, something I have argued against. The discussions in Chapters 2 and 3 argued that as a relational model a picture can trigger off a recognitional response similar to that of the object, and that the idea of “likeness” can be based upon response and not necessarily verisimilitude. While the picture is like the object in the sense of triggering-off a similar recognitional response we do not believe that the picture is the object (nor do we see the object without seeing the picture). We might see the imagined reality in the picture but epistemic
deception plays no role. So in seeing the imagined reality of the picture, in recognizing the picture as of the represented object, Gombrich will say that we experience an illusion of seeing the object. We can see a picture as of a seascape but we remain aware of the fact that we are not seeing an actual seascape—we see a seascape that is not really there. If we can acknowledge that the connotational range of the term “illusion” extends beyond perceptual deception and false belief then we can begin to use the term “delusion” correctly, limiting its application to trompe l’oeils. By doing this we can begin to point out, contra Allen, a very informative and non-trivial fact about how representational pictures function.

5.2.4 Illusion, trompe l’oeil, and medium awareness

Before continuing on let me offer a brief summary of the findings of the preceding section in the form of a negative account:

1. It is possible to use “illusion” synonymously with “delusion” if a picture involves perceptual deception and false belief.
2. Trompe l’oeil pictures involve perceptual deception and false belief—they are intended to fool the eye and (possibly) to generate a false belief that a viewer is looking at the object itself.
3. Thus, in the case of trompe l’oeils “illusion” can be used synonymously with “delusion.”
4. Ordinary representational pictures are not trompe l’oeils: ordinary pictures do not involve perceptual deception and false belief.
5. Ordinary pictures do not involve “delusion.”
6. Therefore, in cases of ordinary pictures, “illusion” cannot be used synonymously with “delusion.”
The point of this systematic layout is to show that if one wanted to use the terms “illusion” and “delusion” interchangeably, the only appropriate cases would be those involving trompe l’oeil. It is only in cases of trompe l’oeils that the connotational range of the word “illusion” can be extended to include instances involving delusory experiences perceptual and epistemic deception. However, if we are talking about non-trompe l’oeil pictures—what I have repeatedly referred to as “ordinary” or “usual” pictures—then whatever illusion is, it is not the same as delusion. We can stand in front of Degas’ self portrait and remain amazed at the subtle and beautiful way in which he captured his own likeness, but never once are we fooled into believing that we are looking at Degas himself. We remain cognizant of the picture as a picture because we can remain cognizant of certain features that make the picture a picture. Belief does not play a role in the process of ordinary pictorial perception; at least not directly. But what do I mean by this?

When we see the ordinary picture we do see the picture for what it is, a flat marked surface; but we also see, and recognize, the marks as adding up to some representational content. What we do not do is form some belief about whether what we are seeing/recognizing is an actual object or not. We simply do not proceed to that step. The illusion, the magical act of transformation that the artist undertakes, lies in our seeing the picture as a picture and our seeing the picture as of the represented object. We see both of these things at the same time; and yes, belief does emerge as a part of our awareness and our switching attention between both the picture’s medium and the picture’ representational content. (I will return to discussing “awareness” and “attention” below.) The belief, then, when we see the ordinary picture, is something like the
following: “I am looking at an orange-picture;” or “I believe that I recognize an orange in
the picture;” or “I see an orange but I do not believe that I am looking at an actual
orange.” All of these are valid beliefs that can accompany the perception of an ordinary
picture that involves the sort of Gombrichian illusion I am describing. But again, what the
perception of an ordinary picture does not involve, what we are barred from doing (i.e.,
without misperceiving the picture), is forming a further belief about the nature of the
content of the picture. The belief “I am looking at a picture, and in that picture I see an
orange, but I know that I am not looking at an actual orange” accompanies the ordinary
picture, I believe, in the immediate act of perceiving the ordinary picture. It is a belief
that accompanies the continued awareness of the picture’s medium and the picture’s
content, and so long as that dual awareness persists we are in no danger of forming the
further (false) belief that the pictured-orange is a real orange. We are in no danger of
slipping from “illusion” into “delusion.” Thus, the illusion of seeing the pictured-orange
remains different from (but similar to) the perceptual experience of seeing an actual
orange, both of which remain different from (but similar to) the perceptual experience of
seeing a trompe l’oeil painting of an orange.

In the previous chapter I argued that we are able to interpret Gombrich is such a
way as to accommodate some form of twofoldness. According to my interpretation, we
can do this if we make the distinction between “being aware of” and “attending to” a
picture’s medium concurrently with our seeing its subject. We cannot attend—in a
conscious or focused way—to both the medium and the represented object at the same
time; we can, however, remain aware of—in a peripheral sense—the one when we focus
on the other. If we interpret Gombrich in this way we can see the manor house in
Wivenhoe Park or Degas in his *Self-portrait with soft hat* and at the same time be aware of the medium that makes up each picture. For Gombrich, a pictorial illusion entails a kind of slight of hand—we are afforded the ability to remain cognizant of the fact that we are looking at a picture while seeing that very same picture as if something that is not really there.

In the Introduction to *Art and Illusion* Gombrich tells us that when we look at the famous duck/rabbit picture we “do not have the illusion that we are confronted with a ‘real’ duck or rabbit. The shape of the paper resembles neither animal very closely.” (2000, p. 5) This is one of those times where Gombrich obfuscates, and so to clarify, we may wish to rephrase this statement to read that we “do not have the *delusion* that we are confronted with a ‘real’ duck or rabbit.” Further down the page Gombrich goes on to say that, “Illusion, we will find, is hard to describe or analyze” even though we know “that any given experience *must* be an illusion” (2000, pp. 5-6). The second quote is especially poignant in light of Gombrich’s assertion that despite certain “extreme cases” it remains true that “we still experience some kind of illusion when we see a picture on a wall or in a book.” (2000, p. 277) It would seem that, while illusion must be distinguished from delusion and false belief, whatever it is, illusion also remains an experience that we are clearly aware of. Pictorial illusion of the kind Gombrich is talking about is not, as I will point out below, something altogether uncommon in our day lives. It is an illusion we encounter everyday. For instance, the Liechtenstein print hanging in my living room affords me an illusion of a ball of twine, but the experienced illusion does not impinge on my actual reality; that is, it does not lead me to believe that there is an actual ball of twine hanging on my wall. Such a case of delusion only arises with the
trompe l’oeil, because it is only there that “the illusions of art [are] about our real environment” (2000, p. 277).\textsuperscript{186}

Of course even the delusory experience of the trompe l’oeil is hardly complete—the limitations of likeness constrain how much the artist can capture in the picture, leaving us with a continual awareness of the picture’s “configurational” component (to use Wollheim’s terminology). This is not to deny the captivating delusional experience that accompanies a trompe l’oeil. I simply mean to buffer the notion that such an experience will \emph{completely} engulf the viewer’s belief about what he or she is seeing. One reason for this lies is the technical issue that artists who construct trompe l’oeils must content themselves with very shallow arrangements in which the depth of field is quite small (e.g., traditional “still-life” subjects). The primary goal of the trompe l’oeil is to trick the eye of the viewer; to visually fool the beholder. Such deception can usually only be achieved by using subject matter with a shallow depth of field.\textsuperscript{187} This is due to facts about binocular disparity (i.e., the difference between what each eye perceives that permits us to plot an object in a zone of depth). Binocular disparity “only provides effective spatial information about the foreground. The disparity dwindles with the distance and [eventually] become imperceptible.” (Gombrich, 1969) So a distant mountain range will appear to flatten out because stereoscopic vision can only provide depth information about what is closer to use. (For this reason, the information that each eye receives from a far-off scene may not necessarily result in the perception of a three-dimensional scene; yet the extent to which our knowledge of the three-dimensional nature of the external world influences what we see remains unclear.)\textsuperscript{188}
The primary reason that the delusion of the trompe l’oeil is not complete is that such an experience simply cannot be sustained. I witnessed this very fact while visiting a museum with my wife. Rounding a corner, we were shocked to see a pair of shears hanging on the far wall. This seemed very out of place given where we were. Upon walking closer we realized that the shears were not actual shears at all, but merely painted ones. We tried again and again to return to the initial state of trickery, even going back to the same corner to round it and glance at the painting—to no avail. We could continue to see the picture as of a pair of shears—seeing them as very life-like shears—but we could not return to having the experience of an actual pair of shears. That path was barred once we are aware of the nature of the picture as a picture, and that happened when we became aware of certain features of the picture. We saw the edge of the painting and the texture of the canvas, a brushstroke or the light reflecting off the picture’s surface. It was at that point that the visual house of cards fell apart—it was only a picture. This initial deception will also vanish if we shift our viewing position, since we find that the position of the objects within the picture do not shift along with our movements. The delusory experience of the trompe l’oeil is potent, but a state of continual epistemic deception is not possible outside of certain experimental settings, such as viewing a picture in single-point perspective through a peephole. Again, the deception of the trompe l’oeil remains a kind of illusion, but we must classify the experience for what it is: an illusion that can involve deception and false belief. So while one can describe such examples as “illusions,” it is easier and much more unambiguous to label cases that involve such visual and epistemic trickery as “delusions.”
The reader might be taken aback that while I have made a great effort to
distinguish illusion from delusion, I accept Gombrich’s claim that the delusory
experience of the trompe l’oeil remains a “kind of illusion.” How can the delusion of the
trompe l’oeil be classified as an illusion if “illusion” and “delusion” are distinct? The
position may sound contradictory, but since Gombrich will claim that both ordinary and
trompe l’oeil pictures count as pictorial representations, the foundational elements of
Gombrich’s theory of illusion—that a picture represents by functioning as a visual
substitute—will apply equally to both. For Gombrich, trompe l’oeil pictures work as
pictorial representations in much the same way as do ordinary pictures: in both cases the
set of relationally arranged visual elements allow the picture to function as a visual
substitute for the object(s) the picture represents; and in functioning as a visual substitute
the picture creates the illusion of the seeing the represented object(s). It is functioning as
a visual substitute that is required to have a pictorial representation, and that requirement
is fulfilled by a trompe l’oeil such as Samuel van Hoogstraten’s *Trompe l’oeil* no less than
by Constable’s *Wivenhoe Park* or Degas’ *Self-portrait with soft hat*.

What distinguishes the ordinary pictures from a trompe l’oeil is that, in the
former, a viewer’s visual experience involves a continued awareness of the picture’s
medium, whereas this is absent in a viewer’s visual experience of the trompe l’oeil;
however, I do not believe that for Gombrich “medium awareness” should be considered
a feature necessary for an artifact to count as a pictorial representation. Medium
awareness will usually coincide (in the case of ordinary pictures) with the presence of
Gombrichian illusion, which I have described as the capacity of a picture to function as a
visual substitute for the object(s) it represents. The usual picture affords the viewer
awareness of the perceived object as a pictured object and not the object itself. This is not the case with trompe l’oeils: there the illusion is not countermanded by an awareness of the picture’s medium, which means that the visual experience is (very nearly) only of the picture’s content. So the primary difference between the ordinary picture and the trompe l’oeil is that for a trompe l’oeil medium awareness is very low or non-existent. That said, it should not offer any trouble for my account of medium awareness is meant to remain independent from the issue of illusion. Medium awareness is only relevant to whether a picture counts as an ordinary picture or a trompe l’oeil, and it should not be understood as the mark of whether a picture counts as a pictorial representation nor whether a picture involves Gombrich’s concept of illusion.

To make this distinction clearer, it might also help to draw parallels between the trompe l’oeil and such optical illusions as the Frazer Spiral or the Müller-Lyer illusion. These are illusions “in the strong meaning [italics added] of the word which implies an actual deception of the senses.” (Beloff, 1961, p. 62) For the trompe l’oeil and such optical illusions the illusion is irresistible—the viewer cannot help but experience it despite knowledge to the contrary. I may know that the Frazer Spiral is made of concentric circles, and I may know that the still-life of apples is only a picture—yet it is hard for me to detach myself from what I am visually experiencing. Let me be clear, though: I do not wish to suggest that the trompe l’oeil and optical illusions such as the Frazer Spiral and the Müller-Lyer exhibit the same kind of illusion. For one, the nature of the former as a pictorial representation strikes a difference between the two examples. Also, it seems that I can more easily enter into and out of the trompe l’oeil illusion than I can the illusion generated by the optical illusions. Medium awareness might be
low in the case of the trompe l’oeil, but I think that in many cases the picture will afford us some degree, however slight, of medium awareness (be it of the frame, the wall, or the glare of light). And the easiest way to break the spell is to simply move our viewing position—the deception is thus never complete, and we will remain aware of the fact that what we are looking at is a picture. On the contrary, it is difficult, perhaps impossible, to see the Fraiser Spiral or the Müller-Lyer illusion devoid of the illusion. We know the Fraiser Spiral is not a spiral, and we can narrow our focus and follow the concentric circles—yet we still see a spiral; we know the arrows of the Müller-Lyer illusion are the same length, and we can narrow our focus and concentrate on each line in turn—yet we still see two different length arrows. Knowing the way out of those illusions does not release us from the illusion in the same way as in the case of the trompe l’oeil.

The illusion of the trompe l’oeil is not identical to that of the optical illusion; nevertheless, my remarks demonstrate that it is similar in certain respects. Significantly, neither the trompe l’oeil or the optical illusions generate substantive, continual false belief. Both can generate epistemic deception: we may be tricked the first time we see the image, but once we are aware of the content of our perception there is little chance that we will experience epistemic deception again. (This does not negate the possibility that, should there be diminished or impaired viewing conditions, the image can trigger a false belief. But again, even if this occurs we will not remain in a state of epistemic deception—we will reinterpret what we see given what we already know, which will end the epistemic deception but not the perceptual deception.)

5.2.5 Perceiving a picture’s “true information”
The illusion of the trompe l’oeil is an extreme—it is the point where illusion spills over into delusion and becomes about our actual environment. The obvious question is where does the one end and the other begin? This is a valid question, but one that I must side-step. My goal is not to point out where this occurs (or if it is even possible); rather I want to offer a tool, a signpost, that can serve as a means for directing us down one path or the other. To do this we must look to the constitutive designs of the picture itself. In Art and Illusion Gombrich talks about the possibility of making an image that is a “complete portrayal” of the city of Tivoli, one that presents “no false information” and which gives “as much information about the spot as we would obtain if we looked at it form the very spot where the artist stood.” (2000, p. 90). Many trompe l’oeils (even many photo-realistic paintings) seem to come close to this ideal. But at the same time, if we assumed that such a complete portrayal were possible it “would never yield [the] illusion of looking at Tivoli through a window, precisely…[because any] such portrayal would inevitably carry with it an amount of ‘false information’, or rather, the true information that we are faced with a picture.” (Gombrich, 1973, p. 168) So on the one had Gombrich is telling us that if we want a complete portrayal of some scene, it must contain no false information about that scene; on the other hand we find him saying that such a complete picture may not be possible because the picture would always carry with it the true information that we are looking at a picture.

What does Gombrich mean by this, that ordinary pictures (i.e., non-trompe l’oeil picture) contain the “true information” that we are looking at a picture and not the object itself? I believe that it is in these sorts of pictures that the pictorial features that make the picture a picture remain clearly present to us: we can see the brushstrokes or the pencil
lines across the page; we can see the texture of the canvas or the signature of the artist—
the picture presents us with facts (the “true information”) that it is a picture, not the
object itself. All ordinary pictures possess this “true information”—specifically, pictures
that I have referred to as “central” cases of depiction. They are the sorts of common
pictures that Richard Wollheim had in mind when he put forward his notion of
twofoldness, and they are the sorts of pictures described throughout Art and Illusion.

I believe Gombrich intends the presence of this “true information” in ordinary
pictures to help countermand the sorts of startling visual deceptions (and possible
epistemic deceptions) afforded by trompe l’oeils, even though in both cases we may
recognize the picture’s subject with relative ease. Ordinary pictures afford us continual
awareness of this “true information,” whereas it is the very feature absent (or at least very
much diminished) from our encounter with a trompe l’oeil. There, the recognition of what
the picture is of (its content) overrides our recognition of the picture’s design features (the
medium), and we are unable to remain continually cognizant of the latter while we are
attending to the former. In many cases we may not even be able to see the brushstrokes in
the trompe l’oeil or the texture of the picture’s surface, so hidden is the picture’s medium.
(Though, as I have already state, compete lack of medium awareness is probably not a
possibility, which means that complete epistemic deception is not a possibility.) All we
recognize, all that we see in the trompe l’oeil, is the picture’s content. The picture’s
design features vanish, being “absorbed” by our recognition of what the picture
represents.

A picture’s “true information” can be understood as involving an awareness of a
number of different pictorial features. For instance: seeing the frame or the edge of the
canvas; being aware of the lack of movement of the internal features of the picture as we shift our viewing position; seeing that the external lighting conditions in which the picture is being viewed do not affect the lighting conditions within the picture itself; seeing the constitutive design features of the picture (e.g., the brushstrokes or pencil marks). Our perception of ordinary pictures cannot involve the belief that we are looking at the actual depicted object (in the way that the perception of trompe l’oeil pictures can involve such a belief) because of these very facts—perception of the pictures means that we are always presented with the “true information” that yields an awareness that we are looking at a picture.¹⁹⁵

Thus, I believe we have found a helpful tool for distinguishing “illusion” from “delusion.” The former (in the usual case of the ordinary picture) will entail a large degree of continual awareness of the picture’s medium which will serve to help eliminate the possibility of confusing the picture with the represented object (even in the case of highly naturalistic pictures such as *Wivenhoe Park*). Thus, in the usual case, having the illusion of seeing an ordinary picture P as of an object O will not mean that a viewer is deceived into taking P to be O itself, because in seeing P as of O the viewer retains awareness of those pictorial design features of P (i.e., the configurational component, its medium) that clearly distinguishes it as a picture of P. The similarity between the ordinary picture and the trompe l’oeil resides in the fact that both are pictorial representations (for Gombrich, both function as visual substitutes), and so both will involve Gombrich-style illusion. In both, the perception of the picture generates an illusion of the object, affording us the experience of seeing the picture as of the object. It is when the awareness of the above design features begins to disappear, and thus when
the design features can be described as absorbed by the recognition of the picture’s content, that Gombrich will say that illusion spills into delusion—it is there that the possibility of deception and false belief becomes much more real, and it there that we can see the substantive divide between “illusion” and “delusions.”

5.2.6 Concluding thoughts on the claim that “illusion is delusion”

There is certainly no shortage of prominent theorists who remain convinced that “illusion” is synonymous with “delusion”: Richard Wollheim (1963), Nelson Goodman (1976), Flint Schier (1986), and Dominic Lopes (1996) have all claimed (in one form or another) that an illusion theory of pictorial representations should really be understood as a delusion theory of pictorial representation. Given the influence that such writers carry, it is no wonder that Gombrich remains shackled to the notions of delusion, false belief, and deception.

Many of these opponents of illusion in art have rejected what Gombrich calls “the reality of illusion in front of the painting.” (1973, p. 194) This “reality” really amounts to a simple claim: “Every representational picture is the promise of depth” (Muniz, 2005, p. 17). The “promise of depth” amounts to the “reality of illusion.” Despite the fact that the depicted depth, the depicted reality, is not really there, we can recognize the picture as of its subject even in the simplest cases. The illusion need not be delusion for us to fall under the spell of the artist’s creation. Gombrich quite plainly tells us that “I fully agree that in looking at a seascape hanging on the wall of a museum we are never tricked into mistaking the painting for a window opening out on to the real sea. If this were the only
legitimate meaning of the term ‘illusion’ the matter would be regarded as closed and the
problem dismissed” (1973, p. 194) Gombrich quite plainly states that the
reality of illusion must be kept separate from that of delusion, for it is evident that we are
not fooled by the majority of pictures. In fact, inquiring about epistemic deception is
really a misguided investigation.

Those who ask about our “beliefs” in front of paintings are certainly asking the
wrong question. *Illusions are not false beliefs, though false beliefs may be caused by illusions* [italics added]. What may make a painting like a distant view through
a window is not the fact that the two can be as indistinguishable as is a facsimile
from the original: it is the similarity between the mental activities both can arouse,
the search for meaning, the testing for consistency, expressed in the movements of
the eye and, more important, in the movements of the mind. (Gombrich, 1973, p.
240)

I have talked previously about Gombrich’s position on the connection between ordinary
and pictorial perception—that the two are not diametrically opposed, and in fact that
there may be a great deal of congruity between them. It is perhaps because of the
association between them that the notion of pictorial illusion can be confused for pictorial
delusion. Yet it is also because of the association between ordinary and pictorial
perception that Gombrich sees the possibility for pictorial representations. Because
pictures are capable of capturing a similar set of relations of visual elements, and because
pictures are able to trigger off a similar recognitional response in our visual system (the
capability of arousing similar “mental activities”), pictures are able to create the illusion
of an imagined reality—it is this illusion that allows us to see pictures in terms of the
objects they represent.

There is no doubt that we use the term “illusion” to describe a variety of
situations; however, *it is a mistake* to use the term “illusion” in the present context of
ordinary pictorial representation *if we mean for it to be synonymous with delusion*. Think of watching the magician’s assistant being sawed in half. None of us thinks that the assistant is *really* being sawed in half. Yet our visual experience is of the assistant being sawed in half—we are seeing something that cannot be and we want to know how it is done. The illusion does not, however, result in a false belief about what we are looking at. If we thought that the magician’s assistant were really being sawed in half our reaction would be quite different—as it is, we are amazed, not horrified, by what we see. Similarly the depictive painting presents us with a visual experience of something that is not really there—we see a seascape or a manor house but continue to believe that we are looking at a picture of a seascape or manor house. We want to know how it is done. Gombrich is not the only one to foster the analogy between pictures and magic. In discussing Gombrich Menachem Brinker tells us that, “There is something ‘magical’ in all two-dimensional images of three-dimensional objects” (Brinker, 1996, p. 44). This act of magical transformation is, I believe, the uniqueness in the commonplace; it is the artist’s ability to transform a flat, marked surface into something more than a flat, marked surface; it is our ability to “sense three-dimensional depth in the two-dimensional painting” and all the while remain aware of the fact that “there is no real three-dimensional depth there, but only a two-dimensional image of it.” (Brinker, 1996, p. 58)

There is a wonderful passage at the end of the article “The Sky is the Limit” that should be presented in its entirety, as it serves to transition towards the impending discussion of substitution and visual affordance.

The visual information the painter can simulate [in the picture] may never actually duplicate the information we pick up from solid objects close by. But is it not possible that he can mobilize the [visual] system to produce the same phantom
sensations which come into play in those processes of search or probing for simplicity precisely in situations of inadequate information? If that is true, our perception of pictures would indeed differ from the perception of the visual world, but the right stimulation from the canvas may still engender a reaction similar to that which we experience in front of nature [italics added]. A fine landscape or seascape by one of the Dutch masters certainly does not give me the illusion that the museum wall opens into parts of Holland. But I would claim that in getting absorbed in such a painting my search for meaning between and behind the brushstrokes weaves on its surface a rich fabric of uncontradicted sensations. Following the artist’s suggestion I begin to forget the textured surface. I see the horizon curving and the sky arching over the earth, not a mediated perception so much as a mediated phantom. (1982, p. 171)

5.3 Affordances, illusion, and visual substitution

5.3.1 A mediated phantom: the emergence of a substitute

If I have been interpreting Gombrich correctly, illusion is found in the kinds of pictures we encounter everyday—pictures that clearly give away the fact that they are pictures, avoiding deception and the possibility of taking the picture to be the objects it represents. While I will talk more about how pictures generate an illusion in the forthcoming pages, for now we can begin to see a semblance of an answer in the idea of visual substitution mentioned in previous chapters. Recall that this idea holds that there is a visual connection between picture and object, a connection that entails that a picture functions as a visual substitute by being a relational model for some object (something I will talk about more in section 5.3.3). By capturing the relations between the visual elements of the represented object, the picture is able to trigger off a similar visual response to that triggered off by the represented object.

In “Meditations on a Hobbyhorse” Gombrich set out to find a word to describe the child’s hobbyhorse: “image” and “portrayal” are rejected, while the he sees as “more
accommodating [is] representation.” (1994b, p. 1) The hobbyhorse does not look at all like a horse, but it can “stand for” or “be substitute for” a horse. And so, within the context of play for the child the hobbyhorse represents in the sense of being a substitute for a horse (Gombrich, 1994b). In picking up and playing with the stick it becomes a horse. Think also of the worshiper who prostrates before an idol. The idol does not serve as a likeness or portrayal of the deity, but rather as a representation for it—that is, as a substitute, a stand-in. The idol is used as means to access the deity; it allows the worshipper to praise and worship the deity. Worshipers “are not worshipping the images [as such], but the deities which these images depict; the image only denotes the place, and gives [the worshipers] the occasion to turn to the god with offerings and prayers.” (Gomperz, 2011, p. 1) But there is no confusion or deception on the part of the worshipers, just as there is no confusion or deception on the part of the child. The idol is not the deity itself, as it might be seen face-to-face; and yet it is clearly more than a mere shaped artifact of wood or stone. “There can be no doubt that the person honoring the image feels that they are grasping the deity…in the image and through it, and yet there can be no doubt as to the difference between the two.” (Gomperz, 2011, p. 1)

In truth there is a very simple explanation for how Gombrich intends substitution to work: something counts as a substitute because of how it functions in a certain context. How an object is used by an individual is what makes it a substitute. If object A can fulfill a certain function similar to one performed by object B then A can be a substitute for B. We should not, however, be seeking identity (see Chapter 3). Identity can denote exactness, a copy—substitution denotes something less perfect but just as potent: an object that, at its simplest, makes use of the barest number of features needed to stand-in
for something else. We can exploit this idea because the human mind has an almost infinite elasticity to employ this kind of behavior; it is resourceful enough to find equivalences in the most disparate phenomena, which amounts to the ability to substitute one thing for another (Gombrich, 1994c). But while we do possess this resourcefulness it must be applied in the right way: it must fit with a possible way in which an object can be used. (I will talk about this in connection with affordances in the next section.)

Furthermore, object A might be able to fulfill a function similar to one performed by object B, but the context in which A is being used will help to determine whether it can serve as a substitute for B. A stick can function as a substitute for a horse within the context of play but it will not function as a substitute for a horse with the context of “cutting down a tree”; or within the context of “playing fetch with my dog.” There is no guarantee that the stick will be used as a horse-substitute by the child, but it is the context (of play) allows for a certain function (e.g., “ridable”) to be realized.

5.3.2 A substitute and its affordances

Let’s say that object A functions in a variety of ways [a, b…n], but not in ways [x,y…n]. We can say that the ways in which A can function are the “uses” that A affords. In Chapter 3 I discussed the fact that an affordance should be understood as something that an object “offers,” “provides,” or “furnishes” to an observer (Gibson, 1986, p. 127). Thus, the idea of how an object functions (and here I included representational pictures) can be explained in terms of the the kinds of affordances it offers an individual. Something can be used for different purposes, and thus can function in a variety of
different respects, dependent upon its different affordances. To help clarify, Gibson offers a concrete example of the affordance of support:

If a terrestrial surface is nearly horizontal (instead of slanted), nearly flat (instead of convex or concave), and sufficiently extended (relative to the size of the animal), and if its substance is rigid (relative to the weight of the animal), then the surface affords support...It is stand-on-able...walk-on-able and run-on-able.” (1986, p. 127)

So a horizontal, flat, extended, rigid surface affords support for an animal. Water does not offer such an affordance for humans, though it does for water bugs. Similarly we can take at a stick and find that it affords a number of different ways of being used by different animals: it affords nutrition for some, support for others, and protection and safety for still others.

Before continuing on it would be wise clarify my position on affordances. At the beginning of Chapter 3 I stated that my account of affordance breaks from the Gibsonian account, and therefore I am not merely offering a Gibsonian account wrapped in a Gombrich package. In the above quote we can see that for Gibson affordances were physical and manipulative in nature; they involved a kind of “doing” on the part of the object. The flat surface affords having things done to/on/with it; it affords doing something to/on/with other organisms. Affordances are very “physical” in nature. Pictures, Gibson would retort, are not capable of this. We don’t act upon the picture by picking it up or trying to manipulate it; nor does it impact me in any sort of physical sense. This is the way in which I am breaking, quite clearly, from how Gibson conceived of affordances. To make this breaker clearer, the reader will see (in following sections) that I am parsing the idea of a picture having visual affordances in terms of the picture’s being a visual substitute. By this I mean something like the following: the picture is a
substitute for some subset of visual information found in the depicted object, and with respect to certain visual systems like ours. One may take issue with my account of pictorial affordances—it should be clear, however, that my usage of the term is divorced from Gibson and the details of his ecological notion of perception.

Having done this, let us now turn back to the specific example of the hobbyhorse: how does the idea of affordance apply there? Gombrich says that there are two conditions necessary to turn a stick into the hobbyhorse: “first, [is] that [the stick’s] form made it just possible to ride on it; secondly—and perhaps decisively—that riding mattered.” (1994b, p. 7) Given the above description we can say that the stick’s form affords being ridden; the stick’s form “provides” or “offers the possibility” of riding (albeit simulated) within the context of play.\(^{196}\) (Of course, the context of play will allow for a wide variety of affordances for a stick, “ridability” being just one of them.) A wonderful example of the variety of affordances offered by the form of a simple stick can be seen in Antoinette Portis lovely children’s book *Not a Stick*. In the book, Portis’ (presumably adult) narrator utters various connotations of the warning “Be careful with that stick!” to a little piglet. After each utterance we are shown a picture of the piglet holding a stick; the piglet responds to the narrator by explaining, “It’s not a stick!” We then see a picture of how the piglet uses the stick; or, given that this is a picture book, how the piglet *sees the stick*: as a fishing pole, a band major’s baton, a horse, a paintbrush, a caveman’s spear, a knight’s sword. For the piglet a stick can be used in lots of different ways, and depending on what game the little piglet is playing will determine how the stick is to be used (e.g., “cowboy” or “painter”).\(^{197}\)
As it is, the second condition that Gombrich offers reminds us that how, or better yet, what the child is playing will limit how the stick is to be used. In this case it is riding that matters: the child makes use of the stick in a way similar to that of an individual riding an actual horse—the child mounts the stick and “rides” it around. A couple things can be said about this: first, again, the stick possess the affordance of “ridability” given the existence of actual riding activities; second, it is the immediate context of play that draws our attention to this affordance, highlighting its significance. Thus, when the stick is being played with it does not just have the capacity to be ridden (i.e., the affordance of “rideability”)—it has the affordance being realized as the child sits astride the stick. It is in realizing the affordance, in using the stick, that it becomes a horse for the child. Similarly, when the piglet “plays knight” the stick becomes a sword to slay the monstrous dragon; when he “plays artist” it becomes a paint brush. Different acts of play will change the function that matters for the context at hand. But in each case the stick serves as a sort of blank screen capable of affording a number of different uses; the stick is able to stand in for various different objects should it be used in the right way.

Essential to Gombrich’s account of the act of substitution is that a substitute’s affordances are able to fulfill “certain demands on the organism” (1994b, p. 4) In fulfilling “certain demands” a substitute’s affordances are capable of triggering-off certain responses in the organism (e.g., in the child, or, for pictures, in the viewer) that allow for the substitute to be used in place of the substituted object. The common factor, the most basic feature, that connects a substitute and whatever it is being substituted for is
“function rather than form.” (1994b, p. 4) What is of primary importance, then, is how a substitute is used (and what, as used, it triggers off in the user) rather than whether the substitute looks like or shares the external form with the substituted object. The stick does not look like a great many of the things that the child uses it for; yet for the child it works just fine as a horse, a paint brush, or a sword. Recall Gombrich’s discussion of the example “Modeled skull from Jericho” in which shells are used for eyes. Shells are not eyes, nor are they very eye-like, but when put in the right context (e.g., in the eye sockets of a skull) they fill the role of eyes by suddenly “looking” at us. Substitution depends on “use” “context” and “metaphor,” all of which are integral for one thing to be used in place of another.

This does not necessarily mean that resemblance and visual appearance are unimportant for Gombrich—at some point an individual gave the hobbyhorse a mane and eyes (for how else could it see?); as for painting, how can we say that a painting is of Lake Placid if it does not look like Lake Placid? What this shows is that what we should be concerned with in asking how a picture represents, and in asking whether a picture generates an illusion of what it represents, is whether the picture fulfills a certain function for the user. My claim is that the picture can fulfill this “minimum requirement of function” in just the same way as the hobbyhorse: by affording a certain usage, one that need not involve any sort of delusion or false belief.
5.3.3 Pictures as visual substitutes

Artifacts like the hobbyhorse and the idol serve as substitutes because they can *physically* stand-in for the substituted object (for a lack of a better term). They can be manipulated by the child or the worshiper; they can be picked-up, held, thrown, sat on—such artifacts incorporate a sort of manipulative usage, which is why I describe them as able to substitute in a *physical* sense (like an actor who stands-in for a sick member of the cast). (This understanding of affordances tracks Gibson’s original use of the term.) The important feature for a given substitute to work is that it affords a certain usage that fulfills a relevant function, which, as I have said, takes precedence over its form or appearance. Think back to my earlier example of the rock standing in for my friend Ed. We need not rely on the rock exemplifying any Ed-like characteristics for it to serve as a substitute for Ed within the context of, say, demonstrating a car accident (e.g., “This rock is Ed, and he was standing here when the truck jumped the curb.”). The rock can perform its function (being a substitute for Ed) so long as those engaged in the demonstration are aware of the rule that “This rock is standing in for Ed.”

In previous chapters I claimed that for Gombrich picture-making involves creating substitutes that are meant to *visually* stand in for an intended object or scene. In this way we can see that if a picture is to serve as a substitute for some object the relevant function that must be fulfilled is of a *visual* nature. This means that, in a sense, for a pictorial representation the function depends on the form (in the sense that an object’s visually perceptible properties are what matters). A pictorial representation’s “use” is that it visually represents something or other; so for example, a picture of a horse must have some visual connection to an actual horse for the picture to count as a pictorial...
representation of a horse. We must be able to recognize it as a picture of a horse to say that it is a representation of a horse. This is not to imply that whatever such a visual connection between object and picture is it need be all that strong. Ancient cave paintings may not be complete or might be out of proportion, and yet we can recognize the subject matter; street artists are often able to capture the likeness of a sitter—we can recognize the picture as of the sitter and not her friend—with just a few strokes of chalk or marker.

This idea that the visual substitute involves a visual connection between picture and object can be helped along by the claims raised in the previous chapters: our horse picture serves as a relational model for a horse in the sense that the picture captures certain relationships between the visual elements seen in an actual horse, which, in turn, presents a viewer with a visual equivalent to an actual horse; additionally, by presenting us with this kind of relational visual equivalent the horse-picture affords a viewer a similar visual recognitional experience to an actual horse, one in which similar visual processes are engaged—a viewer recognizes the picture as of a horse but also remains aware of the picture’s medium. The Gombrichian concept of illusion (as it applies to ordinary pictures) entails that “illusion” need not involve “delusion” so long as a viewer’s visual experience of the picture involves continued awareness of the picture’s medium. The picture can create the illusion of seeing the depicted object—a similar recognitional visual experience can be triggered off by the picture to that triggered off by the object—but the picture-as-visual-substitute will not trick a viewer into believing that he or she is looking at the actual object.
5.3.4 Pictorial affordances

As I have already discussed, as a substitute a picture (just like any other object) can offer a variety of affordances: it can be used as a door stop; an umbrella to block the rain or a blind to block the sun; a fan to cool a friend or to shoo away flying insects. There are all manner of possible uses that a picture affords, but there will be uses that are more appropriate and more relevant than others once we consider the context in question: pictorial representation. In the essay “The Philosophy of the Technology of the Gun” Evan Selinger (2012) offers a discussion that connects to this topic. There, Selinger is emphasizing that guns can be used for a wider variety of practices, but that the very design of the gun itself constrains and shapes those uses that are relevant or worth pursuing. He writes:

In principle, guns, like every technology, can be used in different ways to accomplish different goals. Guns can be tossed like Frisbees. They can be used to dig through dirt like shovels, or mounted on top of a fireplace mantel, as aesthetic objects...But while all of these options remain physical possibilities, they are not likely to occur, at least not in a widespread manner with regularity. Such options are not practically viable because gun design itself embodies behavior-shaping values; its material composition indicates the preferred ends to which it “should” be used. [Italics added]

An object’s design, its material composition, serves to shape how it is used (or at least how it should be used, as there is no guarantee that it will be used in that way). A gun affords a number of interesting uses, but its primary use (i.e., its primary affordance), given its design, is to project bullets away from the shooter at high speeds. A picture, too, affords a number of interesting uses, but as an artifact constructed of marks on a flat surface a picture also has a primary affordance: to be perceived. Pictures are hung in
galleries, used in books and magazines, and placed atop billboards to be looked at. Rather than being used in the inventive ways described, I venture the claim that a representational picture’s primary affordance is that it offers a viewer similar sorts of visual experience as does the object represented. The affordance that is relevant to the context of depiction is thus a certain visual usage, rather than a “door stop usage” or a “block the sun usage” or a “shoo the insects away usage.” The relationally arranged visual elements constitutive of the picture allow this affordance to be realized by a viewer. Representational pictures are meant to be looked at, though their primary affordance is more than that: they afford the possibility of being recognized as of something more than just a flat, marked surface. In triggering off similar visual experiences to those of the represented object, a representational picture can thus be used as a visual substitute (in conjunction with its serving as a relational model for the object(s) represented). The picture is capable of serving as a visual stand-in for the object being recognized because the picture fulfills certain visual functions for the viewer—the picture is able to function as noted, in a visual way that allows it to be seen as of the object it represents.

5.3.5 The visual substitute as a non-delusory artifact

As I understand it, the key to Gombrichian illusion lies in the idea that a picture functions as visual substitute for a represented object or scene (something I have already mentioned in previous chapters). The hobbyhorse, as a physical substitute, can easily be taken as a horse due to both its physical features and its affording “ridability.” One must
now ask whether a picture (as a visual substitute) can equally be so easily taken for the object it depicts; and if so, what makes this is possible. One difficulty in making this claim is that Gombrich describes pictures as having the potential to afford a range of visual interpretations. In *Meditations on a Hobbyhorse* Gombrich writes: “Pictures are infinitely ambiguous because they present a flat two-dimensional geometric projection of a three-dimensional reality.” (1994a, p. 157) He then continues this line of thought in *Art and Illusion* (cf. pp. 250-252). The issue, then, is this: if a picture is infinitely ambiguous, and can be said to afford an infinite range of possible interpretations, how do we arrive at a “correct” one (assuming that such an interpretation exists)?

Firstly we must remember that pictures can only offer us a limited amount of information (I discussed this in Chapters 1 and 2). As Gombrich puts it: “the artist… cannot transcribe what he sees; he can only translate it into the terms of his medium. He…is strictly tied to the range of tones which his medium will yield.” (2000, p. 36) The artist’s job, then, is to find, through a process of trial and error, those marks that can suggest the real—he or she must find a pictorial equivalent to the represented object that allows a viewer to recognize the picture as of something. A task that does not require a perfect transcription, as triggered responses sometimes need only the merest hint of similarity.

Gombrich remains firm in his belief that that pictures are both infinitely ambiguous and that pictorial perception requires viewers to project interpretations into them (see Chapter 1). At the same time, though, he is equally convicted about the belief that a picture can be “correct.” In short, the visual elements constitutive of the picture *may not afford just any* interpretation to be projected into them. If these claims are to be
believe, the kind of pictorial ambiguity Gombrich has in mind must exist alongside the
elasticity of the human mind and its proactive ability to scan a picture for meaning. We,
as viewers, are constantly searching the picture for meaning and then trying out different
interpretations to see which fits—we try out different answers to the question, “What is
that in the picture?” While we *can* try out an infinite number of possible interpretations of
a picture, there is no guarantee as to their successfulness. And yet, if I have been right in
describing representational pictures as relational models for what they depict (see
Chapter 2), Gombrich will say that not just any set of relational elements in the picture
will trigger off a visual reaction like that triggered off by the object, no matter how elastic
the interpretive mind is.

As such, for Gombrich there is no worry of the picture’s content being *wholly
determined* by the beholder’s interpretation. That is, I, as the viewer of Degas’ *Self-
portrait with soft hat*, do not *make* the picture a picture of Degas.\(^\text{199}\) I have already
addressed this worry in Chapter 1, but I find it pertinent to return to it in order to show
how it is that a picture offers us the illusion of the object(s) it represents. A viewer may
well try out any interpretation he or she likes, but given the specific constitutive set of
relationally arranged visual elements in the picture there will only be a limited range of
interpretations that coincide with the things a picture affords being recognized as (the
ways it can be seen). Perhaps there are others ways of reading Degas’ *Self portrait with
soft hat*, but interpretations such as “pink elephant” or “woman with parrot” are incorrect.
An interpretation must fit with the way(s) a picture affords being seen given its
constitutive features, and having the illusion of the represented object will depend on a
viewer having the correct interpretation(s) that coincide with that affordance. (And given

205
Gombrich’s stance that perception is direct and immediate, I think he would say that finding the “correct” interpretation will simply require looking; it will require looking at the picture, putting forward an interpretation of what we see, and then seeing whether that initial interpretation holds up to scrutiny.)

Take Gombrich’s discussion of Erwin Fabian’s *Financial Times* (cf. 2000, p. 235). The poster consists of a chimney, rendered as a caricature with top hat and two hands, reading the Financial Times paper. Gombrich says that we “accept the chimney with the top hat as an industrialist” reading a paper; so we interpret the picture as being both a chimney and caricature of an industrialist. When we look for the industrialist’s face we notice that we are “scanning the poster, looking for indications *where to anchor our projection* [italics added]. We find it [i.e., the face] somewhere along the line, and the faintest of phantom images settles on the chimney and transforms its visual character. True, *it still remains a chimney, but its is also a face* [italics added], according to the way we look at it.” (2000, pp. 234-35) There are two important points to take away from this passage: first is that when we look at a picture we do not offer an interpretation of the picture *sui generis*—rather there must be some point in the picture that serves to anchor our projected interpretation of it. By this I mean there must be some specific set of relationally arranged visual elements in the picture that serves as a place into or onto which my projection will go (e.g., I interpret the orange-colored fruit in the immediate foreground of Fantin-Latour’s *Still Life* as “an orange” because the set of relationally arranged visual elements in that part of the picture serve to anchor the interpretation, “an orange”).
So we, as viewers, are able to project an interpretation into a picture, which affords anchorage points for our interpretation(s). It is interesting that in the above quote, while we will search for a place in the picture to anchor the interpretation “industrialist” or “face,” there is no doubt in Gombrich’s mind that the other interpretation, “chimney,” is already there. The part of the picture that anchors that interpretation can trigger off the relevant visual experience, “chimney,” whether or not it actually does so. Thus, seeing the image as a chimney appears to be an objective way in which we can interpret the image. It rings of a “demand of fidelity to nature,” that there is a way to see the picture apart from trying to see the caricatural industrialist’s face (that part of the picture “still remains a chimney” even if we never find the face).

“Chimney” is an interpretation that Gombrich assumes we will have, and, furthermore, one can infer from the quote that it is an interpretation lots of viewers will have. I take this to be a significant feature of his account of illusion. Gombrich insists that “the undeniable subjectivity of vision does not preclude objective standards of representational accuracy.” (2000, p. xli)\textsuperscript{200} As I have already said, it is possible for us to interpret a picture as of a great many things, though many of the possible ways of seeing a picture are quite improbable and will never happen; this is due to the fact that the set of arrangements of visual elements that are in the picture are not infinitely large (they have been chosen over other ones). Our visual system will encounter and process the constitutive visual elements in the picture in much the same way that our visual system encounters and processes visual elements in ordinary perception. There, we do not produce interpretations randomly, but instead offer an interpretation of “that thing out there” (the kind of “thereness-thatness” principle), in part, given our expectations and
background knowledge. Gombrich will say that pictorial seeing involves many of the same sorts of activities: we offer an interpretation of “that part of the picture right there,” given our mental set, and subsequently we have an illusion of seeing O. This will be because the specific set of relationally arranged visual elements that are found in that part of the picture will anchor an O-interpretation, thus constraining both how we interpret the picture and the resulting illusion the picture generates.

Interpreting the picture as O, and thus experiencing an pictorial illusion of seeing O, is buffered from being a delusory experience due to the continued awareness of the picture’s medium (we might also refer to this awareness as a continual interpretation of the picture as a picture). As a visual substitute (which functions as a relational model in the manner I have already described), the picture affords not only an “representational” interpretation—what the picture is a picture of—but also a “configurational” interpretation—what the picture is made of. This is how Gombrich can maintain that illusion is not delusion, as well as insist that a representational-picture-as-visual-substitute is able to avoid creating a delusory experience: so long as the picture affords an awareness of the picture’s medium (the configurational interpretation) along with an awareness of the represented object (the representational interpretation) the picture can be said to afford an illusion, and not a delusion, of what it represents.

5.3.6 A substitute for what? Visual substitution and the test of consistency

I want to offer a few final comments, elaborating on the aforementioned idea that the interpretation(s) we offer of a picture are dependent upon the the picture’s functioning
as a relational model (i.e., upon the actual constitutive arrangements of relationships of visual elements in the picture serving as anchorage points for our interpretations). It might help to think of this processes in terms of Gombrich’s notion that we subject the picture to a “test of consistency” (which I first mentioned in Chapter 1). This test, simply put, is this: we read the picture as one thing or another—arriving at some interpretation of the picture—and then, if need be, we attempt to determine whether the information in the picture can so be reliably read. A picture that passes the test of consistency will afford us a consistent and reliable interpretation.

When Gombrich says that the picture can offer us a consistent reading he means that the picture affords us “the possibility of classifying the whole of an image within a possible category of experience.” (Gombrich, 2000, p. 238) *Wivenhoe Park* affords us a number of consistent interpretations of the picture as a whole (“a landscape,” “cows on a hillside,” “a manor house by a lake”), and this is due to the facts discussed in the previous section. Yet from this we can see that “classifying the whole of an image” does not mean that there is just a single interpretation or category of experience that a picture must afford. Representational pictures can remain infinitely ambiguous, and so the same set of relationally arranged visual elements in the picture can afford an infinite possible range of interpretations (at least in theory). In actuality, though, a certain picture, Degas’ *Self portrait with soft hat* for instance, cannot represent just anything; no doubt it can yield multiple consistent interpretations but this is not the same as yielding any interpretation. Testing the picture for consistency either yields us a confused, jumbled, and inconsistent interpretation or interpretations; or else it yields us a coherent, organized, and consistent interpretation or interpretations. It is in the latter case that we will say that the picture, in
yielding us an interpretation such as “a man in a hat”, has yielded us an illusion of seeing something that is not really there, but which we can nevertheless describe in the usual terms of what we see when we see a picture (e.g., “I see a man in a hat,” or something like that).

Now, this sort of test can break down, and a clear example of this is seen in the work of M.C. Escher (which I discussed in the previous chapter). Escher’s work shows how attempting to classify a picture like The Belvedere fails because different parts of the picture offer us contradictory information. The reason for this conflict, says Gombrich, is that the visual information in the picture clashes with the assumptions and expectations we bring to it.

Take an isolated part of Escher’s Belvedere, the outside terrance, for instance, with its checkerboard floor and the low walls surrounding it. This is quite a normal realistic representation of an architectural feature and does not differ much from a photograph or picture post card of such a motif…We assume that a pavement will most probably be level and a wall upright, that the slabs on the floor will be square and the seat of the bench rectangular. If the shapes representing these objects are tapering and unequal in size, this is obviously due to foreshortening and perspective. Assumptions of this kind are so ingrained in us that it needs quite a jolt to prevent our interpretation from running along these convenient groves [italics added]. (1994a, p. 157)

We assume that floors and ceilings are horizontal, that columns are upright, and that objects shirking in our visual field are farther away; in reading the parts of the picture, however, we find ourselves confronted with an inconsistent reading of the picture as a whole. The picture does not afford classification within a consistent category of experience—it teeters between interpretations depending on which part of the picture we are looking at. It looks to be of a building with stairs, but not a structure that could possibly exist in the actual world. The constitutive parts of the picture-as-visual-substitute
lead us to conflicting interpretations, and so we are continually forced to reinvestigate the picture—we are continually forced to subject it to a test of consistency. An image such as *Wivenhoe Park*, on the other hand, does afford consistent classification: it offers us the possibility of reading the picture in a coherent and consistent way—which is to say, the picture affords us consistent anchorage points for our interpretation(s). In this way our interpretation of the picture can “come to rest,” it is not “knocked out by contradiction” (Gombrich, 2000, p. 283), and it generates a reliable illusion of the object(s) represented.

### 5.4 Final thoughts

I have said that representational pictures have a certain visual connection with what they represents; put negatively, we might say that a representational picture cannot be completely visually disconnected from whatever object it is to be a representation of. A postcard of Lake Placid has some visual connection to how Lake Placid actually looks, or else we would not think that the postcard is *of* Lake Placid. If we keep this idea in mind we can now describe the connection between visual substitution and illusion in the following way: the picture-as-visual-substitute affords a viewer a certain visual responses, meaning that the picture is capable of visually stimulating us in a way that is similar to how the represented object visually stimulates us; while the picture affords a variety of responses, as a relational model the picture can be said to be relevantly similar to what it represents in the sense that the picture contains a certain constitutive set of relationally arranged visual elements (i.e., design features: marks, dabs of paint, etc.), and
it is the perception and awareness of those relationships between visual elements that constrains the possible responses that the picture affords a viewer., allowing the picture to be visual substitute for the object(s) represented. To function as a visual substitute in this manner will mean that a picture affords a pictorial illusion of what it represents; an illusion that affords the viewer a recognitional experience of the represented object while, at the same time, also affords awareness of the picture’s medium and the picture’s status as a picture, not the actual object.

The picture-as-visual-substitute, in generating an illusion, does not generate a delusory experience that leads viewers to take it (the picture) to be the object represented. Only in the case of the trompe l’oeil, when our awareness of the picture’s medium is absorbed by our recognitional interpretation of what the picture is of, can we say that “illusion” is “delusion.” Yet, there is no contradiction in claiming that illusion is present in both ordinary pictures and trompe l’oeils: both are pictorial representations, and thus for Gombrich both are visual substitutes (in the manner described) for whatever object is being represented. The distinguishing mark, afforded by ordinary pictures but not trompe l’oeils, is found in the former affording a continued awareness of that from which the illusion comes—a continued awareness of the pictorial medium.

Pictures can no doubt be seen as of a great many things, affording a number of possible interpretations many of which were never what the artist intended (this is clearly evident as the narrative of Art and Illusion moves towards the modern era, especially in Gombrich’s description of Cubism). The preceding sections have tried to demonstrate that while this fact of representational pictures remains true, a representational picture will not afford just any reading—it was created with certain visual affordances in mind; it
was created to trigger off ways of being seen due in part to the constitutive set of relationally arranged visual elements included in the picture. A picture is recognized as this or that thing, and while we may be confused as to whether it is of this thing rather than that (e.g., is that a horse or a cow? is she smiling or frowning?) our confusion is very often of a limited nature (e.g., it certainly is not a space ship; she is not sticking her tongue out).

What we want to know—what the critics want explained—is how illusion can play any role in pictorial representation: What do we mean by illusion? How can we say, in a coherent manner, that a picture involves or contributes to the illusion that the picture is of something else? Gombrich will answer the question, “Is illusion the same as delusion?” with an emphatic, “No”—there are cases (trompe l’œils) in which illusion can be considered delusion because there is no (or at least very little) awareness of the picture’s medium to countermand the experience of the picture’s content; but in the usual cases involving the kinds of ordinary “core” pictorial representations Gombrich is concerned with, illusion will not necessarily (or in general) be the same as delusion. It is true that the notion of “illusion” carries with it a large connotational range, yet I fail to see why we cannot expand its application (as Gombrich does) to explain the objects and scenes seen and recognized in pictures. Gombrich intends illusion to describe not only the act of visually substituting a picture for some object or scene, but also the visual experience we get from seeing (and recognizing) that picture as of its subject. Moreover, this experience can happen (and in the case of ordinary representational pictures does happen) without any delusion coming in.
We encounter pictures constantly throughout our daily lives. And the pictures we see represent countless objects, scenes and events. Rather obviously, looking at the picture is quite unlike looking at the wall on which it hangs, or the paper on which it is printed. We don’t see the wall as of anything other than a wall; the paper remains starkly blank. Very often we take for granted the fact that when we see a picture we see space and depth, rather than a mere flat, two-dimensional surface—we see an imagined reality, whole worlds unfolding before us, an illusion of something that is not really there. Gombrichian illusion means (in the cases of ordinary pictorial representations) that when we see a picture we remain aware of its medium concurrently with its content. A picture, as a visual substitute, affords us a non-delusional experience of an object or scene that is not really there—and we will see the picture’s content, and experience what is being represented, as an illusion.
As the tortoise said, “I can’t make progress without sticking out my neck.”
—From R. L. Gregory’s “The Confounded Eye”

6 Conclusion

6.1 Where we have come from: Gombrich’s basic theory of pictorial representation

This dissertation began as an exploration of Ernst Gombrich’s conception of illusion; specifically, how that conception might fit into an understanding of how pictures represent the objects that they do. In the course of this exploration I defended the often rejected platitude that pictorial representational involves illusion, claiming that such illusion is distinct from delusion, deception, and false belief—a viewer can experience a pictorial illusion (in the way Gombrich claims) without thinking he or she is looking at the object itself. To gain a better understanding of the Gombrichian conception of illusion I took the notion of physical substitution (as Gombrich first presented it in *Meditations on a Hobby Horse*) and expanded its application to include pictures. According to this expanded account, representational pictures function as visual substitutes for the objects they represent. Visual substitution thus became the starting point for my working interpretation of the Gombrichian conception of illusion—when a picture visually stands-in for some depicted object, the picture affords us (and we experience the picture as offering us) an illusion of that depicted object.

Elaboration was required in order to clarify this position. To provide it, I introduced the notion of relevant similarity between the picture and the depicted object. I initially described such similarity as obtaining in those cases in which we can see the picture “as of” or “in terms of” the represented object. However, this preliminary account of relevant similarly was deficient for the following reason: while we may often
talk about a picture and an object using similar sorts of language (e.g., we can describe our seeing both of an object and of the picture that represents it as a seeing that is “as of” the object), if a picture is indeed relevantly similar to some object this fact will not depend upon my saying so. Our verbally describing the picture and the object in similar terms does not make the one similar to the other. While a case of relevant similarity may sometimes be manifested in our verbal descriptions, the nature of the relevant similarity does not turn on our ability to describe the picture and the object in any particular way.

Chapters 2 and 3 took up the task of filling out the details of this conception of relevant similarity, which I claim involves two aspects: (1) the notion that a representational picture functions as a relational model for the represented object by preserving (in the picture) a set of relationally arranged visual elements that are equivalent to those found in the represented object; and (2) given the preservation of such relationships, the fact that a representational picture affords viewers a similar recognitional experience to that afforded by the represented object, such that seeing a picture of O triggers off a similar (though not identical) visual experience to seeing O itself.

A picture, as a flat, marked surface, is unable to capture all of the rich information seen in nature, and so there can be no point-for-point matching of the picture and the represented object. To pursue the claim that a picture is similar to some object requires us to point to some other ways in which relevant similarity might hold. There is, I claimed, a visual connection between picture and object—we treat the postcard of Lake Placid as being of Lake Placid—so we must ask what aspects are connecting the picture and the object. The aspects are found in our comparing not the individual visual elements in the
picture and in the object, but rather in our comparing the presence of certain *relationships* (i.e., steps or gradients) that hold between the visual elements that occur in the picture and the relationships that hold between the visual elements that occur in the object. In capturing these visual relationships the picture is able to function as a *relational model* for the represented object, and the picture can be said to be relevantly similar to the represented object even if the picture and the object are not visually identical to one another. By focusing on the idea of pictures as relational models we are able to by-pass many of the worrisome issues surrounding resemblance and mimetic theories of pictorial representation (which focused merely on the imitation or replication of nature).

Additionally, the idea of a visual connection between picture and object held that the set of relationally arranged visual elements that are in the picture has not been arbitrarily selected—for not just any set of marks can yield the intended visual image of, say, a dog if our purpose is to create a depiction of a dog. Individual pictures are created for various reasons, which need not concern us here. What matters to the existence of representational pictures as such is that there are undeniable facts about how our visual system responds to the presence of certain relationally arranged visual elements in the objects of perception. The ability of representational pictures to depict their objects depends centrally on these facts. Certain relationally arranged visual elements were included in a dog-picture because they were found to afford a certain visual response *(viz., that of “a dog”)* that is also afforded by seeing a dog.

In describing the details of relevant similarity as a relational model one can see the importance place that Gombrich’s theory gives to the visual experience(s) triggered off by the picture. But this idea requires further elaboration. A single picture can certainly
yield viewers a variety of visual experiences. Our dog-picture can be interpreted in a variety of different ways, although the range of possible interpretations is not infinite; even less so the range of likely interpretations. However, not all of these interpretations will capture the specific thing that a given representational picture depicts. We simply cannot visually interpret an ordinary representational picture any way that we like (at least, we cannot do so and be correct). How is it, then, that there comes to exist a constitutive set of relationally arranged visual elements in our dog-picture that triggers off the visual experience of a dog rather than the visual experience of an airplane or an elephant?

To address the question of how this comes about I invoked a second aspect of relevant similarity: the inclusion of a certain constitutive set of relationally arranged visual elements in a picture will afford the viewer a recognitional experience that is similar to the recognitional experience afforded by the represented object. We recognize the picture of a dog as being of a dog because of facts about how we recognize an actual dog—the two perceptual acts are not wholly different from each other. The recognitional experience triggered off when we perceive the relationships of visual elements in a picture of a dog is similar to the recognitional experience triggered off when we perceive the relationships of visual elements seen in an actual dog.

Obviously, the set of relationally arranged visual elements found in the picture cannot identical to those found in the object, and thus the two recognitional experiences will not be identical. But identity is not required for recognition to occur; this is evident from the ease with which we recognize line drawings and caricatures. All the picture need do is contain a set of relationally arranged visual elements that is equivalent to those
found in the represented object; in being so constructed, the picture-as-equivalent need only trigger off a similar recognitional response for us to see it *as of*, or *in terms of*, the represented object.

When we look at Degas’ *Self portrait with soft hat* we can see that a man is not present in front of us, and yet we see a man represented in the picture. In seeing the picture, we recognize a man, but we also recognize the picture’s medium. The experience we have thus has a twofold nature: that of the picture’s medium (the picture *as a picture*) and that of the object the picture represents (the picture’s *content*). The twofoldness of the pictorial experience has been taken as gospel since it was put forward by Richard Wollheim. Gombrich rather famously stated that we cannot see the medium and the content at the same time; we must switch interpretations if we are to see one or the other —this was used as fodder for the critic’s assault on his theory, and it has been accepted that Gombrich’s position is a failure in this respect.

However, and contrary to these criticisms, I have shown that we can interpret Gombrich in such a way as to allow for twofoldness so long as we make a distinction between “being aware of” and “attending to”: we can remain *aware* (in a peripheral sense) of both the medium and the content of a picture at once; what we *cannot* do is *attend* (in a conscious way) to both the medium and the content of a picture simultaneously. When we attend to the figure of Degas we cannot attend to the picture’s medium—we lose track of the medium in the sense that it is no longer the focus of our conscious attention. Our continual awareness of the picture’s medium, though, can be seen in the fact that we do not forget that we are looking at a picture—I remain cognizant that I am looking at a picture of Degas, not at Degas himself. Should I wish to examine
the marks that make up the painting, the subsequent attention I in turn give to the medium does not eliminate my continued awareness of the figure of Degas. The knowledge of how the medium functions—that the marks are representationally relevant—continues to influence my perception of the picture.

The investigation of twofoldness served two purposes: it was meant to vindicate Gombrich’s position on pictorial experience—a position heavily criticized by Wollheim and many a subsequent theorist; it was also a part of my larger project to demonstrate that illusion, as Gombrich understands it, is distinct from delusion, deception, and false belief and is (as Gombrich holds) an appropriate phenomenon to appeal to in seeking to understand the nature of pictorial representation. Contrary to what is claimed in other criticisms of Gombrich's views, Gombrich's theory does not commit him to the conclusion that viewers of ordinary representational pictures suffer from perceptual delusions or are deceived about the objects of their perception.

One step in the process of disentangling illusion from delusion is to look at the differences between an ordinary representational picture such as *Wivenhoe Park* and a trompe l’oeil such as Hiernault’s *Still life of the back of a painting*. One way in which the two are different is that the ordinary picture is intended to produce a very different visual experience from that of the trompe l’oeil. The latter is meant to quite literally fool the eye of the viewer (though not necessarily the viewer him or herself), who is intended to see the painting as if it were the object itself, which, in this case, is the the actual back of a painting. Here a genuine perceptual delusion or false belief is present. An ordinary picture, on the other hand, produces no such false belief, nor is it intended to. When a viewer looks at a trompe l’oeil he or she loses a continual awareness of the picture’s
medium; that is, the medium is absorbed by the recognitional experience of the picture’s content, leaving the viewer without the counterbalancing recognitional experience that he or she is looking at a picture of the object rather than the object itself. Thus, in the case of the trompe l’oeil (and not in the case of the ordinary picture), the viewer has a delusory experience. So if we are to use the term “delusion” it must be confined to those cases in which there is *actual* deception, namely trompe l’œils.

How then are we to make sense of Gombrich’s conception of illusion? If we turn back to the integrated discussions of visual substitution and relevant similarity that I have offered in this work, we can see how one can have Gombrich-style illusion of an object without having a delusory visual experience of the object itself. In order to see this point, note again that, according to my account of Gombrich’s theory, the inclusion of the set of relationally arranged visual elements in the dog-picture triggers off a certain response in a viewer; this response is similar enough to the response triggered off by an actual dog that the picture affords us a recognitional experience *as of* a dog.

In being relevantly similar to the represented object (e.g., a dog), the picture functions as a visual substitute for it. The picture is a visual stand-in, a visual equivalent, for the represented-dog; and as a visual substitute for a dog it will stand in, visually, for the actual dog, allowing us to recognize the picture as being of a dog. When this occurs—when we recognize the object represented in the picture—we find that the surface of the painting becomes transparent in our mind; in scanning the picture, the surface features of the painting will become irrelevant to our processing so long as we have turned our attention away from the medium and onto what the picture is of (the dog). Yet even in our disregarding of the picture’s medium, we maintain a continued awareness that our
visual experience is one of a picture of a dog, not of an actual dog. So we do not suffer from any perceptual delusion or false belief.

6.2 Core Necessary Conditions Revisited

In the first chapter I proposed the idea that visual substitution is necessary for, but limited to, those cases involving pictures that Gombrich and others have identified as central or core cases of depiction. Under this proposal, there are no strict necessary and sufficient conditions for depiction, or at least my interpretation of Gombrich does not propose any such conditions; instead, visual substitution counts as a core necessary condition (CNC) for something being a depiction. In concluding my discussion, I should mention briefly how the idea of a CNC fits with my expanded interpretation of Gombrichian illusion.

Recall the Chapter One argument against the above sort of similarity as being a sufficient condition for depiction. That argument turned on my example of Winston and his twin Edward: a Winston-picture’s being a visual substitute for Winston (in the sense of being relevantly similar to Winston in the manner previously described) is not a sufficient condition for its being a depiction of Winston rather than of Edward because the visual experience of seeing Winston may share more similarities with the visual experience of seeing Edward than it shares with the visual experience of seeing the Winston-picture. Thus, if being a visual substitute were a sufficient condition for depiction we would need to conclude that Edward is more of a visual substitute for Winston, and so, in turn, that Edward depicts Winston. Obviously this is not the case, and
a plausible interpretation of Gombrich’s theory should reject any claim that being a visual substitute is a sufficient condition for being a depiction.

I then drew on another example, involving a Coke bottle, to illustrate two further points. On the one hand, Gombrich is also not offering strict necessary conditions for all forms of visual representation in a broad sense of the term—I can use a Coke bottle to represent my friend Ed, but doing so does not mean that the Coke bottle functions as a “visual substitute” for Ed. On the other hand, while the Coke bottle can be used as a general representation of my friend, it is not being used as a depiction (a pictorial representation) of my friend Ed. Thus, although this example shows that being a visual substitute is not necessary for being a representation in a broad sense, the example does not show that being a visual substitute is not a necessary condition for being a depiction. The Coke bottle is simply an example of what I earlier called ‘non-depictive visual symbolism.” Given these points, so long as we are looking at the right examples there is room for visual substitution to count as a necessary condition for depiction; and those examples are the cases that I suppose that Gombrich would count as central or core cases of depiction.

The two above notions of being a relational model and of affording the viewer the proper sort of recognition provide two prominent examples of CNCs that, on my interpretation of his theory, Gombrich takes to apply to depiction and pictorial representation. But I have left open, in this dissertation, the possibility that, with further development, the present Gombrichian theory might want to recognize further CNCs. Assuming that my interpretation of Gombrich's theory is satisfactory, and taking as my starting point those pictures that are paradigm core examples of depiction (such as
Wivenhoe Park and Self-portrait with a soft hat), I believe that it is clear that such cases will involve the CNCs that I have described above. My claim is that if such core pictures function as relational models; if they afford viewers a recognitional experience as of the represented object; and if our awareness of this recognitional experience does not absorb our awareness of the picture’s medium (as occurs in the case of ordinary pictures), then such pictures will function as visual substitutes for the objects they represent.

Additionally, as visual substitutes, those pictures will exhibit Gombrichian illusion in the manner I have described: we will see the picture; we will see the picture as of the represented object, recognizing it as more than a flat, marked surface; and yet we will remain aware of the fact that our visual experience consists in looking at a picture, not in looking at the object itself.

Given this notion of a CNC, an important question to ask, as we move along the spectrum of types of pictures, will of course be what to do once we encounter those borderline or other cases where the application of the CNC becomes debatable. If pursued in detail, this question will lead into general issues about the analysis of all forms of visual representation (and not just of genuine depictions). Those issues, however, go beyond the scope of this dissertation. I can say briefly here that when we encounter such borderline cases, I think that we will simply question whether such cases count as visual substitutes at all, and thus whether Gombrich’s conception of illusion should be applied. But such cases do not refute Gombrich's theory. If my above interpretation is correct, they are simply not clear, core examples of genuine pictorial representation at all (as opposed to being visual representations in some non-depictive way). To be sure, there is indeed a spectrum of cases here. As I understand Gombrich, he supposes that it is a psychological
fact about us that, at some point along this spectrum, the experience of illusion will "click-on." Turning from a line-drawing of a hobbyhorse to a painting of a hobbyhorse we will experience a greater degree of illusion (i.e., a heightened sense of perceived depth and space); other images might even fall in between the line-drawing and the naturalistic painting, further highlighting the degrees to which illusion can be present. Moving from the hobbyhorse itself to a line drawing, however, will result in a point at which the artifact goes from just being a physical stand-in to being a visual substitute. Where and when, exactly, visual substitution emerges will, I believe, depend on how the artifact in question is being used, and thus on how it functions for the individual and, possibly, for the society or the culture.206

Here, again, I am touching on issues that go beyond the scope of this dissertation's focus on Gombrich's theory of pictorial representation. My goal in noting these issues about borderline cases is not, in any case, to determine here the exact point at which visual substitution, and with it the above two CNCs, emerges. I simply wish to note that the application of these CNCs will be limited to a range of artifacts, and that there will be fringe cases that will constrain, and possibly eliminate, their application. With that in mind, I must leave any further discussion for another time.

6.3 The end is the beginning

The principal task of this dissertation has been to address those areas where Gombrich laid out an account of what illusion consists in, as well as to fill in those spaces where he has been silent (intentionally or not). I have presented an interpretation of Gombrich that demonstrates that we can coherently understand how pictorial
representation can involve illusion without its following that viewers of picture are visually deceived. Yet, I feel that I have left the reader in much the same position as Gombrich did (albeit in much less eloquent fashion). Because my goal was not to present an unassailable account of Gombrich’s position, nor to answer the totality of critics, the reader will find that there is still much to be done. I have left a number of avenues unexplored and raised a number of unanswered questions. It will take a great deal more work in order to look down each of these avenues adequately and to address each of these questions, and it is my hope that future investigations can take up these tasks.

Gombrich’s conception of illusion is meant to illuminate the commonplace encounter we have with pictures: while we don’t believe that we see the object itself we have the illusion of the object nonetheless. It may seem as if I am stretching the term to sensationalize a mundane, everyday experience. While I acknowledge this worry, I would ask the reader to undertake one last, brief experiment—pause at the next image you encounter; look at it as just a series of constitutive flat marks, dabs of paint, arrangement of lines, or dots on a screen; then pull back and switch, seeing it as the object that is being represented. This moment of reflective pause should undoubtedly demonstrate the ease with which we interact with the multitude of images that surround us. In describing these circumstances as involving a kind of illusion, I hope to have demonstrated the exceptionally complex nature of an everyday activity that is so often overlooked.
References


Notes


2 Although, Gombrich himself was quite puzzled by the attention that the topic of illusion received after the release of *Art and Illusion*, even going so far as to say that illusion “was never... the central issue of *Art and Illusion*.” (1973, p. 195) It may seem as if I am putting words in Gombrich’s mouth by presenting illusion as the “basic thesis” of the text. I do not believe that this is so. While illusion may not have been its primary focus, *Art and Illusion* gives us both a narrative of Western representational art and loose account of how pictures depict the things they do. An integral part of both of these is the concept of illusion, understood as a specific undertaking (i.e., the style known as “illusionism”) and, as I will argue, as a general feature of image-making.

3 The most recent Millennium Edition of the book contains a new preface that is very helpful in allowing the reader to take in the original text while at the same time giving Gombrich a forum in which to comment on the reactions and reviews to *Art and Illusion* over the past forty years.

4 As I understand it depiction is a specific kind of representation, and while many different things can count as a depiction I will use the term to apply to pictures or images (since these are the kinds of things Gombrich seems to most readily be talking about). For this reason I understand “depiction” to be synonymous with “pictorial representation,” and in the dissertation the use of the term “representation” should be read as meaning “pictorial representation” (unless otherwise state). In this way I wish to distinguish depiction from other kinds of representations, specifically linguistic representations (e.g., names, sentences, musical scores, clocks, time tables, etc.). Pictorial representations (depictions) involve a uniquely visual experience, which is not found in linguistic representations. While Gombrich does think that there are similarities between pictures and language it is the former that he is primarily concerned with. With that said, I will attempt to be consistent in my use of “depiction” in describing the type of representation discussed in *Art and Illusion*.

5 Throughout the dissertation I will use “picture” and “image” interchangeably.

6 See Gombrich, 2000, p. 32

7 These paintings can be found in *Art and Illusion* on pages 65, 67, and 201 respectively

8 These paintings can be found in *Art and Illusion* on pages 55, 216, and 273 respectively.

9 I will try to further justify this point in a later chapter.

10 This does not preclude the possibility that these three seemingly distinct ways of understanding relevant similarity might actually be interconnected in some way. I briefly mention this option here, but I will have to defer further argumentation until the next chapter where I can give the issue a more thorough investigation.

11 Of course, cashing out the idea of relevant similarity in terms of a linguistic expression won’t help to describe the stickleback fish example. Something else has to be going on. So if I am going to maintain that relevant similarity is as central to Gombrich’s account of depiction as I think that it is I will have to say what, if anything, is intrinsically similar about the experience of seeing an apple (s-a) and the experience of seeing a picture of an apple (s-pa). This I will attempt to do in Chapter 2.
I will talk more about visual elements throughout Chapters 2 and 3, but see specifically §2.5.1.

See Gombrich, 2000, p. 40

See Gilman, 1992, ch 5 ff.

Gombrich says little of what he means by “the function” of a picture (or toy), and even less about what “context of action” is. Let me make an educated guess. By “function” I believe Gombrich means something like: the purpose for which an artifact is used. By “context of action” I believe he means something like: the arena or situation in which an individual interacts with, or uses an artifact. Some examples of a picture’s function might be “to portray, to illustrate, to decorate, to entice or to express emotion,” all of which are possible within a perceptual or visual context of action (Gombrich, 2000, p. 119). Similarly, a toy (e.g., a stick) might function in any number of ways within a playful context of action: as a horse, a sword, a magic wand, a lasso, a whisk.

There are two things to point out here. Firstly, it should be obvious that a painting need not be of some actual object or scene, and this points us back to the question of whether Gombrich’s theory is concerned with the depiction of particulars or kinds. The second thing to point out concerns the hobbyhorse. The hobbyhorse is not an image—it is a substitute, so one might think that in this case we are comparing apples to oranges. But this isn’t so. Representation includes the hobbyhorse just as much as the pencil sketch, the cave drawing or the oil painting because all of these exist on a continuum. The central issue is how a given artifact is to be used, what function it is meant to fulfill for the individual. (See the following endnote.)

I believe that some questions might be raised over whether or not the case of the hobbyhorse involves illusion (in the manner I develop it) in the same way as the representational picture does. Put differently, since I will focus on the connection between visual substitution and pictorial illusion, and since the hobbyhorse involves substitution but is not a pictorial representation, we must ask whether or not the hobbyhorse involves illusion at all. I must grant that this is a valid concern. The representations that I am concerned with, be they pictures or hobbyhorses, both count as substitutes; more specifically, representational pictures count as visual substitutes. That is Gombrich’s position and one that I plan to show is plausible. Hobbyhorses are not pictorial representations. So the question remains whether or not all substitutes (visual or otherwise) involve illusion. Function and the context of action (i.e., use) help to determine how or what a representation is seen as. Richard Woodfield put it in the following way: “Insofar as the child treats a hobby horse as standing for a horse, the child’s behavior simulates that of riding a horse. An independent spectator sees the stick as a horse through the child’s behaviour. Though she doesn’t thereby see the stick as a horse, like a cloud may be seen as a face or a face seen in a cloud, the spectator sees horse-type behaviour occasioned by the use of a stick.” (Woodfield, personal communication, 2012) I have placed the hobbyhorse along the long and complex spectrum of illusion. I have not offered to define the point at which an artifact moves from being a physical substitute to a visual substitute; this is partly because I do not think that finding such a point matters, but also partly because it will most likely prove impossible to do so. As I define it, illusion is about taking one thing to be something else without the resultant experience involving a false belief. This definition is primarily applied to pictorial representations, but it might also be applied, in an analogous manner, to artifacts like the hobbyhorse. The worry in doing so is that the concept of illusion will be spread too thin to have any substance. I have granted this a valid concern, but I must leave the investigation of it for another time.
To be clear: in this case “create” does not mean “copy” or “replicate.” The artist can no more create the exact visual experience in the picture than he can create, on the page, the exact visual scene. “Create” appears to mean something like “approximate” or “equivalence,” where the artist is able to produce an artifact that yields a visual experience which is like the visual experience of an actual object in some respects, but which we would not say is identical or a copy of it.

The core necessary condition, as I present it here, works alongside the idea, mentioned above in section 1.3, of relevant similarity—that a depictive picture, one that serves as a visual substitute, will be relevantly similar to the object it is depicting in that it triggers off a similar visual experience to that of the depicted object. In this section (1.5) I am attempting to show how these two ideas can work together.

If we assume that Winston and his twin are identical twins that share a number of similar visual properties (e.g., facial structure, hair, build, etc.).

See Gombrich, 2001, pp. 101-102

In discussing Gombrich’s theory I will accept this way of appealing to central or core cases where the depiction (i.e., the picture) does satisfy the necessary condition of being-a-visual-substitute of the object (or kind of object). While I will not necessarily worry about cases of non-depictive symbolism, I will mention that all pictures are capable of function as depictions and symbols. There is way to cleanly separate the two functions.

See Gombrich, 2000, p. 28.

It is important to note that Gombrich uses the term “schema” in a number of different ways. Firstly, the schema is said to act as a tool for categorizing or pigeonholing new or unfamiliar perceived information from the external world by means of established and familiar sets of categories; secondly, the schema is understood as the artist’s vocabulary, constituting of a formula or set of methods for constructing an image that can be applied, adapted, and then reapplied in different situations.

As evidence for his claim that we focus on relationships of elements rather than individual elements Gombrich points to ethnological studies. For instance, chickens expect food when presented with the brighter of two pieces of paper. “Their little brains are attuned to gradients rather than to individual stimuli.” (Gombrich, 2000, p. 50)

See Gombrich, 2000, pp. 40-44

This idea, of returning to an innocent “childish perception” of the world, is given by John Ruskin (quoted in Gombrich, 2000, p. 296). In rejecting this idea Gombrich makes the further assertion that “all paintings must be interpretations” (Gombrich, 2000, p. 384).

29 The idea of an affordance comes from Gibson (1986). While I will talk in more detail about affordances in Chapter 3, there are some initial remarks that should be made here. First, Gibson developed the idea of affordance apropos of objects in the external world, but it is unclear how (if at all) he would have applied this idea to pictures. Second, the reader will see that I have taken this general Gibsonian idea of affordance and applied it, through the lens of Gombrich’s writings, to pictures (and to the Gombrichian concept of illusion). While Gombrich accepted a number of Gibson’s key ideas no where does not explicitly state whether the idea of affordance is one of these. With this in mind, I maintain that the idea of affordance fits the alongside the rest of Gombrich’s writings on pictures and pictorial representation, as well as with Gombrich’s position on perception. As such, I mean to integrate the idea of affordance as a key component in my interpretation of Gombrich’s theory of illusion.

30 Of course, artists have come relevantly close. That is to say, there are examples of trompe l’oeil paintings that have fooled even the most knowledgeable and meticulous viewers (we can see this in the stories of Parrhasios and Zeuxis, as well as Giotto and Cimabue).

31 Gombrich’s use of “mistaken” here appears to be describing a visual mistake, not a mistake in the sense of a false belief. The artist, in most cases, is not trying to find which mark will lead a viewer to be deceived into believing that the image is a house or an apple. This sort of mistake would only arise if the artist were intending the image to be a trompe l’oeil.

32 These are two important characteristics of picturing according to Gombrich. Artists have, through much trial and error, developed the means to suggest three-dimensions through the creation of an equivalence in two-dimensions. Gombrich also uses the terms “approximate” and “abbreviation” to describe this process of suggestion. While he never offers a detailed definition of any of these words, they all appear to accompany the claim that a picture is inherently deficient at conveying certain information about its subject, which means that the artist must find a way, in two-dimensions, of presenting enough of the subject’s information for it to be recognized by a viewer.

33 Incidentally, Gombrich (in much the same way as J.J. Gibson) uses the terms “cues” and “clues” interchangeably. In the dissertation I won’t take issue with whether this is correct or not. Instead I will follow Gombrich’s approach and assume the two terms as interchangeable. The use of “cue” can be found in Meditations on a Hobbyhorse (Gombrich, 1994, pp. 155-6), and the use of “clue” can be found in Art and Illusion (Gombrich, 2000, pp. 281 and 341).

34 Walton (1990, p. 4, fn1) objects to this. He states that a child doesn’t really ride the stick (i.e., hobbyhorse) but only pretends to, so there really isn’t a univocal, single context of action shared between a rider of real horses and the child.

35 In the essay “Visual metaphors of value in art” (Gombrich, 1994) Gombrich writes that, “visual representation may have its roots in…[the] ‘transference’ of attitudes from objects of desire to suitable substitutes. The hobbyhorse is the equivalent of the real horse because it can (metaphorically) be ridden.” I think that, given Gombrich’s theory, this kind of transference can be carried over to pictures: the picture is the equivalent of the “real” object, scene, or state of affairs because it is metaphorically experienced as the object, scene, or state of affairs it depicts. The transference that occurs in the case of pictures is a visual one: the way we look at an actual object (the visual processes involved) is transferred over, as much as possible, to the way we look at a picture of the object.

36 See David Novitz’s description of the schema (1977, p. 52).
In Meditations on a Hobbyhorse Gombrich describes the significant or distinctive features as those fitting into biological or psychological locks which fulfill certain functions for the viewer: just as a baby may reject a realistic looking doll in favor of something which is “cuddly” (e.g., a blanket may be better for “cuddling” if it fulfills the function better than the realistic doll), the schematic image need only (at least at first) have those features which allow for it to “work” through the fulfillment of a function in use by the viewer. I think that in the case of pictures, the features that “work” are those that afford recognition, and certain features will do a better job at facilitating recognition than others.

I don’t want this to be confusing. It is Gombrich’s claim that the schema is essential for all depictive art, and, since all picture-making involves the creation of substitutes it follows that all substitutes will be developed through the use of a schema. Throughout art history, especially recent art history, artists have fought to rid their images of any trace of a schema (Constable being Gombrich’s prime example of this). In Gombrich’s narrative we find artists employing a noticeable schema less and less, to the point where, at the end of Art and Illusion, the Cubists and the Modernists claim to have completely eradicated a noticeable schema from their images. Does this mean that with the schema so goes the substitute? I’m not sure. There is no doubt that Gombrich’s account will find it difficult (to say the least) to accommodate these contemporary artistic styles; since there is not enough room to here address this larger concern, I will hold out concluding that this difficulty entails a coup de grace for Gombrich’s theory.

See Gombrich, 2000, pp. 138-9, and 280.

“I was looking out the window, watching for the street car, and I saw through the shrubs by the fence the brilliant red slats of the familiar truck; just patches of red, brilliant scarlet. As I looked, it occurred to me that what I was really seeing were dead leaves on a tree; instantly the scarlet changed to a dull chocolate brown. I could actually ‘see’ the change, as one sees changes in a theatre with a shift in lighting. The scarlet brown seem positively to fall off the leaves, and to leave behind it the dead brown. I tried to recover the red by imagining the truck…but I could not” (Gombrich, 2000, p. 226).

Jerome Bruner calls this “perceptual preparedness.” (1957, pp. 129-133)

Of course picture are not inkblot tests. All pictures (being two-dimensional) are incomplete and thus need to be supplemented by the beholder; but it would be a strange (and I would venture mistaken) theory that said that the beholder was wholly determinative of a picture’s content. Of course I am not wholly responsible for making Manet’s At the Races a painting of a horse race. To his credit Gombrich does not follow this line; still, he remains rather vague on how much influence a beholder does have in determining or completing a picture’s content.

For a good description of how Gombrich thinks he can bridge the gap between the claim that an image is “formed” enough to have a reliably consistent interpretation and the claim that the beholder projects an interpretation into the image, see Gombrich, 2000, p. 222.

By no means does this imply that there was a decrease in artistic skill. All that Gombrich means by this phrasing is that as artists continued to experiment with new and different pictorial devices they found that it was possible to create images with a greater visual “presence,” and for viewers to respond in similar (and possible more interesting) ways by not trying to include every stitch of detail. Less was more because making use of the “loaded brush” no longer made sense given the direction art was headed.

Here I am describing “attention” and “awareness” within the context of visual perception. In Chapter 4 I will talk more about their application to other forms of perception (specifically auditory perception).
46 See Gombrich, 1994 for a description of this change from physical substitution to more visual substitution.

47 By the phrase “individual visual elements” I mean those things like different intensities of light; variations in color and tone; and arrangements of lines responsible for size and outline shape. When looking at an object our visual system can break down what we are seeing into these different individual visual elements. However, the point of this chapter is to show that it is not these individual visual elements that are of the primary concern for our visual system, but rather the relationships between the visual elements. It is the relationship between different intensities of light; the relationship between different colors; the relationship between different arrangements of lines that are of primary concern for our visual system. Furthermore, this is a concern in the perception of pictures just as it is a concern in ordinary perception.

48 Hopkins actually offers an explicit account of the difference between having “an experience of…” and having “an experience as of…”. He writes that, “Having experience of a property is a matter of being in perceptual contact with an instance of it; having experience as of it is a rather weaker notion. One can have experience as of a property even in the property's absence; whereas one cannot then be in perceptual contact with (any instance of) it. Nor is the difference just a matter of whether the experience is veridical: experiences that are veridical by accident do not put us in perceptual contact with the properties they represent. What is required is some kind of correlation between the presence of the property and the occurrence of the experience.” (2005, p. 130)

49 See Gombrich (2000). The idea of talking about an object and a representation of it in a similar manner might be what he means when he say a representation can extend the class of an object.

50 This of course assumes that Gombrich believes pictures are in fact similar, in some way, to their depicted objects.

51 This is a paraphrase from the riddle given by the Mad Hatter to Alice in Lewis Carroll’s Alice’s Adventures in Wonderland (1998, p. 60). Carroll intended the riddle to be nonsensical and without an answer, but in the preface to the 1896 edition of the book he offered a possible solution: “Because it can produce a few notes, tho they are very flat; and it is nevar put with the wrong end in front.” Note the spelling of “never” as “nevar”, which is raven spelled backwards. So perhaps there is a way in which a raven is like a writing desk, though I would claim that if such an answer exists it is much more ambiguous and puzzling than how a picture is like an object.

52 My aim remains to relevant similarity to help clarify the notion of visual substitution, however in doing so I have perhaps inadvertently bestowed the current conception of relevant similarity (i.e., as an “as of” experience) with some special status within Gombrich’s theory. I followed this approach in Chapter 1 to help flush out the details of Gombrich’s theory of depiction. Since I do not mean for relevant similarity to retain the label of being merely a descriptive linguistic tool, and since I do not wish the phrase “as of” to evince a special status within Gombrich’s theory, I will work throughout this chapter to remedy both of these issues.


58 Much of the controversy surrounding Gombrich concerns the degree to which conventional plays role in his account of pictures. There is not enough space to begin an investigation of this immense topic. The important point to take away from these tantalizing quips is this: for Gombrich, that a picture represents something or other is not simply because it looks like it. As such, for Gombrich there are deep problems with explaining representation by appealing to a shared visual appearance or a resemblance between picture and object.

59 This is the title of Part One of *Art and Illusion*.

60 I will return to Goodman’s objections later on in the chapter.

61 Most theorists agree with Goodman’s triad of arguments against the sufficiency of resemblance: (1) its seems that two things can resemble each other without one representing the other, meaning that resemblance is symmetrical while representation is asymmetrical (Goodman, 1976, p. 3); (2) resemblance is thought to be reflexive while representation is not, meaning that while all things are said to resemble themselves they do not represent themselves (Goodman, 1976, p. 3); and finally (3) pictures are said to resemble other pictures more than they resemble the objects they depict, yet a picture represents some object and not another picture (Goodman, 1976, p. 4). See Catherine Abell (2009), David Gilman (1992), and Karen Neander (1987) for arguments against some of Goodman’s points.

62 Though Nelson Goodman does not explicitly state these as his assumptions I think it is fair to say that he holds them given his description, and criticism of resemblance in *Languages of Art* (1968). Furthermore, I believe that one could level his criticisms, *mutatis mutandis*, against my account of relevant similarity.

63 I am here using the term “absolute” to mean “non-contextual.”

64 This kind of “fruithood” comparison is something we do with children when we try to teach classification. In trying to explain to them what counts as a fruit and what counts as a vegetable we will show them different fruits and vegetables—sometimes in picture books, sometimes when we visit the grocery store. The apple and the orange both display “fruithood,” so it would seem to make sense to compare them in that fashion. The idiom “comparing apples to oranges” is used to label a comparison between two incommensurable things; my goal is not to criticize the merit of the idiom itself, rather I am simply pointing out that its application to the present case of picture and object is a bit misleading.

65 The idea that a picture and an object might share something in common is offered by Craig Files in a description of resemblance. See Files (1996), p. 403.
There could be a number of contexts in which we might want to compare the son to the mother, not all of which are visual. For instance, we might ask whether the son's mannerisms, his inflections, or his hot-headedness is similar to the mother's. This does not contradict the fact that there may very well be contexts of comparison that are visual in nature. It is with these kinds of visual contexts that I draw my analogy. When looking at a picture we often turn from picture to object and ask, “How are they similar?” by which we probably mean, “How are they visually similar?” and not “How are they similar tout court?” Just as we do not compare son and mother with regards to all and every respect, neither do we compare picture and object with regards to all and every respect. This is an important fact that is often overlooked, quite unfairly I might add, in the haste to dismiss any role the look of a picture might play when investigating the nature of depiction.

This is often the kind of learning game we play with children: “Is this a fruit or a vegetable?” we say to the child. A teacher might even ask a group of students in the midst of youthful scientific inquiry, “What is this?” Investigating what things have in common can lead to strange comparisons, for instance if we ask what a rock, a dog, and a guppy all have in common: they all contain carbon.

It is true that the context in which we compare a picture and an object is a visual one, and this is certainly similar to how a mimetic theory asks whether a picture imitates natures, or how a resemblance theory asks whether a picture resembles the appearance of an object. But where relevant similarity diverges from these other theories is precisely in what is being compared: it is not a shared point-for-point appearance, nor is it the individual visual elements of the picture and the object. All that relevant similarity asks is that certain relevant visual relationships between elements in the object be preserved in the picture, which in turn allows for the possibility (however unlikely) that a picture of O bear only the vaguest of appearances to O itself. So, in the extreme (or the abstract) cases, the relationships in a cubist work depicting O may correspond to those of O itself, even though the picture presents an unnatural depiction of O (i.e., it captures in a single picture the totality of visual relationships seen from multiple view points).

It might be the case that when “the mode of representation varies…so too will the relevant respects” (Neander, 1987, p. 216). I am not sure if changing the mode of representation is what change the relevant respects, but it certainly a very plausible possibility.

We might still ask whether there is some more abstract respect of which each of these different specific respects is an instance. In other words, is there one overall relevant respect in which pictures are relevantly similar (in the way I claim) to objects? While there does not appear to be a single, overall relevant respect, the description I give of the relational model may offer one way in which there could be such a respect: while any individual instance of a relevant respect can change from case to case (for P1 the relevant respect in which it is relevantly similar to O is A, but for P2 the relevant respect in which it is relevantly similar to O is B), the overall, abstract relevant respect common to both P and P* is whatever preserves the relations of visual elements found in the object. So in any P there is a set of relations of visual elements R, and in any O there is a set of relations of visual elements S; the overall, abstract relevant respect in which P is relevantly similar to O is just that R in P exists and corresponds (in the manner described in this chapter) to S in O. Further, different individual instances of R in P simply exhibit relational fidelity to different instance of S in O, but so long as an individual instance of R in P corresponds to an different instances of S in O we can say that there exists an overall, abstract respect in which a picture is relevantly similar to an object.


Neander makes this claim with regards to resemblance, but I think that the same sort of thing applies to my account of relevant similarity. See Neander (1987), p. 216.
I am walking a fine line with these statements, as I do not want to belittle my own position. Pictures remain flat. They are quite different from objects in the actual world, and thus they remain unlike them in very real ways. At the same time, this entire chapter is arranged around the proposal that Gombrich remains committed to the idea that in certain visual ways pictures are like the objects they represent. I must ask the reader to continue along this journey to determine whether this proposal is plausible.


It is possible that the choice of medium will also influence the degree to which the picture is similar to the object. This could be an different way to interpret the quote by Karen Neander in endnote 16: not only will the mode of representation influence which respects are relevant but it will also influence the degree to which the picture is (and which we can describe it as) similar to an object.


Bühler discusses the idea of *relationstreue* in the chapter “Symbolic Fields in Nonlinguistic Instruments of Representation” (2011, pp. 203-219).

We can think of the artifact that is more image-like, the representation that is completely realistic, as perfectly capturing the motif it is representing. It is unclear what example Bühler has in mind, though I think Gombrich’s example of the waxwork is the best we can offer.

What do I mean by this? Think of the word “horse.” Why does it correspond to the animal itself? The word “horse” is a symbol that lacks any noticeable degree of visual likeness, and the fact that the word functions representationally is not because it corresponds to the object in any direct way, but because that specific set of marks and sounds are used and have been conventionally accepted as a means for picking out that specific animal. On the other hand, the naturalistic painting of a horse is a non-linguistic symbol that is meant to correspond to a horse in a visual way. The written word may be purely symbolic, but I think that Gombrich, following Bühler’s lead, will claim that the painting remains both image-like and symbolic. It corresponds to its subject in a less arbitrary (more visual) way, but that it is a picture, and thus made of flat marks on a page makes it symbolic as well. This fits with Bühler description of one pole as more symbolic and the other as more image-like, and it allows for a range of representational forms to be both symbolic and image-like to greater or lesser degrees.

The idea that a picture could be of a particular individual (e.g., Holbein’s painting of Henry VIII or Degas’ self-portrait), or of some individual in general (e.g., any number of paintings including odalisques or cherubs) is something I touched on in Chapter 1. However, if the picture does not look like Henry VII or like an odalisque we will certainly not say that it is a picture of Henry VII or a picture of an odalisque.
Even in cases like a Cubist painting or a child’s sketch I would think that the picture was made with the intention that it should be recognized as of something. The child wants to make a picture of a “horsey” for her parent, and even though she may not be able to draw a horse all that well she still wants the parent to look at it and see that it is a horsey. While there are a great many Cubist works that require some non-visual characteristic (e.g., the title or background information) to discern their subject matter, others can still fit the example I have been describing. For instance, while Picasso’s Portrait of Abroise Vollard displays “a broken architecture of shards of flesh-or brick-coloured painting; planes that have been started and stopped, as if in a slow-motion exaggerated cartoon of the movement a painter makes between looking up, recording on canvas the detail he sees, looking back”—the painting remains “a portrait of an individual whose presence fills the painting. Vollard is more real than his surroundings, which have disintegrated into a black and grey crystalline shroud.” (Jones, 2002) In short, the subject (Vollard) is able to stand out from the background despite the broken-architectural shapes Picasso has chosen to use. Viewers may not be able to, in the words of Dominic Lopes, experience “individual-recognition” and see that the painting is of the famous art dealer (without returning to its title); but surely viewers can experience “kind-recognition” and see that the painting is of, among other things, a person (most likely male), with a head, a face, and shoulders (1996, p. 137).

One worry is that the claim there is is a set of visual elements in the object implies they the elements are decomposable. But what do we mean by decomposable? I can look at my couch and the way the the sunlight falls on it through the window and I can see that there are areas in full sunlight, areas that are in less light, and areas that are in shadow. Or I can see that at the contour of my desk can be broken down, visually, into different interconnected lines. So we can say that there is a set of visual elements that make up the objects in my visual field. And yet, the argument is not concerned with the individual intensities of light, or the individual tones or colors, or the individual lines; but rather with how they interact, how they relate to one another. So the “set” of visual elements might consist of lots of individual visual elements, and we may be able to, in an experimental or contrived setting, focus on each in turn; but Gombrich is saying that what we respond to and interact with are the relationships between the visual elements we encounter.

See Gombrich (2000), p. 94. Gombrich quotes Leonardo as describing the painter as “the Lord of all manner of people and of all things”, and of the subjects that the painter creates “he is their Lord and God.” (p. 96)

The original Italian reads: “favella, favella, che ti venga il cacasangue!”

See the chapter, “From Light into Paint,” especially pp. 40-41.


Those who have been steeped in ecological psychology may find it difficult to divorce my use of “affordance” from the details of a Gibsonian account. I do not deny that my notion of affordance draws upon some of what Gibson says about affordance (since he coined the term itself). Again, it should be clear that I am not offering a Gibsonian account of perceptual affordance; nor am I attempting to draw a strong analogy between my interpretation of Gombrich and the position given by Gibson. My account of affordance breaks from the details surrounding Gibson’s ecology of perception and offers a unique perspective of how it is that we recognize a picture’s subject. While I will point to some similarities between Gombrich and Gibson there is no doubt that the two were inexplicably at odds on a number of topics. For more on this break see the discussion of visual substitution and affordances in Chapter 5, specifically section 5.3.2.

See Gombrich (1982).
The sorts of perceptually constant features that are most often mentioned include color, shape and size. In “The mask and the face: the perception of physiognomic likeness in life and art” (1972) Gombrich mentions an additional type of constancy: physiognomic constancy. I find this significant. The reader should take note that the topics of the construction, reaction to, and recognition of physiognomic features runs throughout many of Gombrich’s works and, I would claim, bolsters my claim that the presence of perceptually constant features can play an important part in explaining how a picture is relevantly similar to an object.

My discussion of Gombrich’s understanding of perception may, I fear, have become muddled. While I take some responsibility for this confusion, a great deal of the blame must rest with Gombrich. This is because, as in other areas of his work, a systematic account remains absent (whether intentionally or not). One must search out those islands of clarity in which Gombrich expresses his position explicitly. One such island is the essay “‘The Sky is the Limit:’ The Vault of Heaven and Pictorial Vision” (1982a). There Gombrich tells us that he believes, “with [J. J.] Gibson, that normally the visual array contains all the information we need to perceive the invariant forms of edges and solids. If we did not recognize a straight line as straight and a plane as plane we would soon come to conflict.” (1982a, p. 164) Perception of the world, it would seem, follows this pattern of directness. When it comes to viewing pictures, however, Gibson and Gombrich diverge (though Gombrich might see this divergence as less extreme than Gibson does). Take for instance Gibson’s response that Gombrich “cannot accept [the] separation [between the perception of reality and the perception of pictures] for he believes the postulate of a continuity between them. In short, he believes there is no clear and logical break between the real and the illusory.” (Gombrich 1971, pp. 5-6) Gibson’s position stems from (as Gombrich puts it) Gibson’s belief that there is “no continuity between” looking at pictures and looking at the world, whereas for Gombrich, “it is precisely where the rainbow ends that art begins.” (1982a, p. 170) It is in the case of looking at pictures that we find the increasing influence of “the inside”—the active role of the beholder’s share (1982a, p. 171). The information in the picture will always be less than the information from nature itself, so the picture must be supplemented. In a way Gombrich’s position can be made somewhat clearer by looking at the fault lines that have been established between himself and J.J. Gibson; any further discussion must be left for another occasion.

While perception may not be decomposable at the level of the conscious act itself it may be decomposable after the fact.

Whether of not Gombrich holds this view is a contentious point in the literature. Richard Wollheim (1998) and Terrence Wilkerson (1978) both criticize Gombrich because they claim he does hold this position. Daniel Gilman (1992), on the other hand, is critical of Gombrich for not holding this position. In this case I believe that Gilman is wrong, and both Wollheim and Wilkerson are right, though for the wrong reasons: Gombrich does indeed affirm that the processes involved in the perception of pictures are not wholly different than those used in the perception of objects in the actual world, leading to his insistence that “the divorce between the perception of pictures and that of reality…can never be cared out completely.” (Gombrich, 1969)

Richard Wollheim (1980) explicitly denies this position. He is very wary of comparing the processes used for ordinary perception and the processes used for perceiving pictures. The phenomenological experiences are different, he says, and he chastises Gombrich for failing to make this distinction (Wollheim, 1980, pp. 214-215).
Of Schier’s theory, Jerrold Levinson writes that he “proposes that a representation is pictorial just in so far as it recruits the visual recognitional capacities that subjects already possess for familiar objects, so that a picture represents an object O if it triggers, in the subjects who view it, the same capacities for recognition that would be triggered by the sight of O in the world.” (Levinson, 2003, p. 18)

It is not insignificant to my position that these ideas, attributed to Schier, sound a great deal like the Gombrichian theory I have been defending. However, Schier’s position is perhaps a bit more extreme than Gombrich would care to embrace. For Schier there exists a “natural generativity” between looking at the world and looking at pictures of it: once we have interpreted one picture we can interpret other pictures with no further training. Thus, having the perceptual ability to recognize an apple or a goat or our spouse means that we have the ability to recognize a picture of an apple, a goat, or our spouse. It is this further claim that I am hesitant to attribute to Gombrich.

Lopes distinguishes between three things: a picture’s subject, a picture’s designs, and a picture’s content. (1) A picture’s subject is the real-world entities the picture represents; (2) a picture’s designs are the privileged set of visual properties that make up the picture, and by which the picture represents things in the world (i.e., its subject). Lopes’ list of pictorial designs includes: “marks, directions, boundaries, contours, shapes, colours, hues, relative contrasts of light and dark, and also textures, such as smoothness of surface or invisibility of brushwork.” (1996, p. 3) (3) A picture’s content are the properties a picture represents the world as having and should be separate from the properties and features of the designs that are constitutive of the picture. A picture’s lines, textures, shading, and colors should not be properties ascribed to the subject. Rather, “When we experience a picture in the right way, we have an experience which represents the world has having the properties the picture represents it as having.” A major goal in Understanding Pictures is to “distinguish between design, content, and subject, explaining content by means of design, subject, and the relation between them.” (1996, p. 4)

The bulk of his arguments for this position can be found in Chapters 2-4 of his doctoral dissertation “Lines of sight: an essay on mind, vision, and pictorial representation” (1988); the conclusion is found at the end of Chapter 5.

To be clear, while Gombrich is famously wary of turning the discussion of representation towards the “image on the retina,” which he says is as useless to the artist as the mythical image in the unconscious (Gombrich, 2000), he nevertheless draws support from ethological studies and biological facts about our visual system. See Gombrich (1982, 1996, and 2000).

The actual quote from Walton reads, “Normally the viewer of Hobbema’s millscape just looks and sees that fictionally there is a mill with a red roof near a grove of trees.” (1990, p. 311) The make-believe aspects of Walton's theory of depiction are, of course, foreign to Gombrich's view. While I do not find the inclusion of the discussion of make-believe to conflict with the immediate purpose of demonstrating a possible congruency between the perceptual activity of looking at pictures and looking at objects, it is important to note that Walton explicitly rejects Gombrich’s account of illusion. Even if Gombrich and Walton could agree on the fact that there are similarities between the visual process of looking at pictures and the visual process of looking at their objects, it should not be forgotten that they would still disagree on a great many details of how a theory of representation should be cashed out. For instance, in a footnote on page 4 of Mimesis as Make-Believe Walton lays out an explicit criticism of Gombrich’s example of the hobbyhorse.
For example, I am not sure whether Gombrich would agree with Schier’s position that the recognition of pictures only requires our ordinary perceptual abilities. For Gombrich, pictures are symbols as well as pictures, so the awareness and understanding of certain artistic conventions will need to play a role in the picturing process. Secondly, Lopes, like many writers, is hostile to much of Gombrich’s work. This can be seen throughout his book *Understanding Pictures* (1996). While I will not go into any details, readers should note the section “Depiction and Vision” (Lopes, 1996, pp. 37-51) in which Lopes explicitly discusses Gombrich and the illusion theory. Third, I am not sure how much stock Gombrich would put in Gilman’s reliance upon retinal images to prove that images and objects are similar in some manner, even if that conclusion is, as I argue, the same. To paraphrase Gombrich, reducing the discussion to the notion of the image on the retina is to rely on a will-o’-the-whisp and will not be all that fruitful of an endeavor. Finally, Gombrich would most likely call into question Walton’s analysis of representation as an act of make-believe. While Walton asks us to imagine of our perceptual experience of the picture that it is a perceptual experience of the object the picture depicts, Gombrich will assert that our perception of a picture really is, in a sense, the perception of the depicted object because of the kind of similar experiences each triggers off. (Of course I have already hinted to a similarity between Walton’s view and Gombrich’s insistence upon the active role of the beholder in “completing” the image.)

See Gombrich’s discussion of this in “A Note Further to the Drawing of Bicycles” (1999).

In employing the idea of “appearance” I am following Gombrich’s lead in accepting its use. This brings with it a host of issues, notably the disagreement between himself and J.J. Gibson (see also endnote 5). While Gombrich believes that there are appearances, like the appearance of a mountain, Gibson firmly rejects this. For Gibson we do not see aspects of objects, we simply see objects. The exchange between the two in Gombrich (1971) offers some clarity on the matter, but the issue is left unresolved. Gombrich’s parting blow, however, is significant for my purposes: “I still wonder, for instance, what Professor Gibson would say to a painter who wanted to paint the Mont Cervin out of his hotel room in Zermatt. Would he ask him to tour all the surrounding valleys in order to embody in his picture the invariant features rather than a particular aspect? And what would the finished result look like?

This account may be a bit simplistic since we do not always recognize what a picture is of. All that I am pointing out is that the picture must pass an initial test of recognition for the artist (who is a first viewer) if it is to be recognized by any other viewer at all. This does not mean that every viewer will recognize the picture every time it is seen, it just means that the act of creating the picture involves an initial act of recognition on the part of the artist.

A critic might object to my using “see…in” and “recognize…as” in the same sentence, citing a distinction between seeing-in and seeing-as. Gombrich, however, does not appear to place much weight in making this distinction (see Gombrich, 1972 and 2000). As such, I will follow suit.

I might also question whether these sorts of pictures are depictions at all. We can take a naturalistic portrait of President Barack Obama and invert the colors, flip the orientation, cut the picture into tiny squares and then rearrange the pieces, but I do not think the resulting picture would be a depiction of President Barack Obama. With a key and some training we can learn to decode the picture, and we might learn to how to classify similar pictures, but such pictures would not be depictions of their subjects; furthermore, we would not recognize the content of the picture in the same way we recognize the naturalistic portrait of Obama.
One might question how a hastily drawn line-drawing might fit into this account. The lines used to construct the image are a tangled mess, and so it is not unlike the mess of tangled lines that might be used to construct the aforementioned tangled-line picture; yet we can clearly recognize that the image is of something—we can make out the various figures even though the image is tangled and unfinished. What distinguishes the recognitional act employed in the perception of a naturalistic image from that used in the perception of a tangled-line image? And where does our line-drawing fall? We might say that such an image can demonstrate the area between the two; that while it make engage our visual system in a similar manner to the perception of objects in the actual world, there are elements of the line-drawing that require some knowledge of style and convention—some viewers may need to demonstrate some degree of learning in order to read the image. This is a difficult topic to disentangle, and any further investigation will have to be put off until another time.

In fact we never recognize every single part of a picture at once (or even always), but this is no different than what happens when we look at objects in the world. Looking out my window I cannot recognize every detail of the woman walking her dog down the street, since the limits of my visual system preclude me from seeing the stitching of her sweater, the size of the dog’s nose, and possibly even whether her pants are black or dark blue. Yet I am quite confident in describing what I am seeing (a woman waling a dog) because I recognize most parts of what I am seeing.

Let me try to flesh out this idea that a picture is a visually contradictory artifact. If we assume my claim that the relations among the visual elements in the picture are similar (i.e., equivalent) to the relations among the visual elements in the object, then we might wish to say that when we see a picture the brain actually forms two contradictory hypotheses about what is being perceived: (1) we are seeing something flat—the picture as a two-dimensional object; and (2) we are seeing something with depth and space—the picture as having perceived three-dimensions. Now, R.L. Gregory advances the idea that illusions involve contradictory hypotheses, but it is unclear whether he would accept the claim that ordinary pictures are visually contradictory since he did not believe that ordinary pictures generated illusions in the way that Gombrich believed that they did.

When we look at something like a Chuck Close painting it may seem as if the opposite is true. That is, there is not sense to be made of it by looking at the parts one-by-one, and instead we can only see what it is a picture of by taking the whole thing in at once. Yet, a Chuck Close is really no different than an ordinary representational picture: when we are too close to the ordinary picture the subject disappears and all we see are brushstrokes and blobs of paint. The same happens with the Chuck Close (albeit often on a much larger scale)—we must take the whole thing it at-a-glance, not through piecing it together. The perception of most pictures works in this way, though in most cases we can make sense of the parts of the picture in a way that we cannot with the Chuck Close (or, e.g., Impressionist or Pointillist works). We read the picture as a coherent whole, then we can look at the parts that make it up, and then we can move back-and-forth between the two. There is no break in a coherent reading. The example of the Escher picture is an example of how this process can break down.

For an brief account of how saccadic eye movements are a natural part of our interaction with everyday objects see Michael Land, Neil Mennie, and Jennifer Rusted’s paper, “The roles of vision and eye movements in the control of activities of daily living” (1999).
As you read your eyes are actually making quick movements across the page, followed by short stops; the quick eye movements are known as **saccades** while the short stops are called **fixations**. This is a common feature of our normal perception of the world: as we look around our eyes are constantly moving, (sometimes consciously and sometimes unconsciously), and while there may be large eye movements should the environment be new and unfamiliar, most often the movement of the eyes will be confined to the center of the gaze where we find those features and stimuli essential for the extraction of information.

For one of the first studies in the area of eye movements during the perception of pictures, see G.T. Buswell’s *How People Look at Pictures: A Study of the Psychology of Perception in Art* (1935). Retrieved from [http://psych.wfu.edu/art_schirillo/articles/Buswell,%201935.pdf](http://psych.wfu.edu/art_schirillo/articles/Buswell,%201935.pdf)

What is interesting about picture’s such as Escher’s, Hogarth’s *False Perspective*, and, I would argue, most impossible pictures Gombrich has in mind, is that they are, section-by-section, by-and-large naturalistic. In *The Belvedere*, for example, we can recognize people and that some are climbing up ladders, sitting on benches, and looking out into the distance; we can recognize the mountains, valley, and sky in the background; and even the building itself is constructed of many naturalistic elements (e.g., the stairs, the door, the bars on the window, and so on). It is the fact that these naturalistic elements are arranged in an unnaturalistic (i.e., impossible) way that makes the global impression we get from looking at the picture contradictory. It is only as we try to see how each of these pieces adds up to a coherent whole that we realize the coherent whole is not there.

Two things to point out: First, perhaps evidence from certain “genres” of art may weaken this claim. Art for children’s books, as well as science fiction and fantasy art all incorporate some of the intentions I have labeled as “exceptions.” But again, as was the case with *The Belvedere* described in the previous endnote, such genres will nevertheless include certain naturalistic elements that often afford the viewer effortless recognition. The subject matter may be un-natural (a nun sitting in a flying chair; space ships doing battle; a phoenix gouging the eyes of a bassalisk) but viewers can often effortlessly recognize most (if not all) of what the picture is of. The second point I wish to make is that artists have, on some level, always been about breaking with traditions and established practices. So claiming that all pictures have a single common “customary function or role” may be a bit misleading. A survey course of the history or art can easily demonstrate how, in an overly simplistic way, each artistic movement sought to challenge the presuppositions of previous movements, and thus how each sought to challenge or at least raise awareness of the different possible ways in which pictures can function for a viewing audience. And yet I wish to maintain that pictures are meant to be recognized, whether they are purely figurative, depictive pictures or whether they are more abstract or conceptual ones. In this way, then, there may yet exist a “customary function or role” that pictures are meant to serve.

More work would have to be done to confirm or disprove these claims.

Gombrich, (2000), p. 272. Gombrich rightly points out that he draws not only on Popper for these views, but also on Gestalt psychology. Although he does not accept the entirety of this school of thought, he does attempt to pull out those parts of it that are productive and “logically right”.

It should be pointed out that in the context of perception hypotheses are not to be understood as consciously or intentionally formed. I do not look at the table where my cup of coffee is resting and form the hypothesis (in my head), “I am perceiving a table with a coffee cup on it.” The hypothesis *just is* my perception of the motif, and the content of that perception (what I see) is provisional and subject to refutation. Conscious or intentional hypotheses may occur where the perceptual system is diminished or incapacitated in some way, or where the motif is indeterminate or ambiguous.

---

113 As you read your eyes are actually making quick movements across the page, followed by short stops; the quick eye movements are known as **saccades** while the short stops are called **fixations**. This is a common feature of our normal perception of the world: as we look around our eyes are constantly moving, (sometimes consciously and sometimes unconsciously), and while there may be large eye movements should the environment be new and unfamiliar, most often the movement of the eyes will be confined to the center of the gaze where we find those features and stimuli essential for the extraction of information.

114 For one of the first studies in the area of eye movements during the perception of pictures, see G.T. Buswell’s *How People Look at Pictures: A Study of the Psychology of Perception in Art* (1935). Retrieved from [http://psych.wfu.edu/art_schirillo/articles/Buswell,%201935.pdf](http://psych.wfu.edu/art_schirillo/articles/Buswell,%201935.pdf)

115 What is interesting about picture’s such as Escher’s, Hogarth's *False Perspective*, and, I would argue, most impossible pictures Gombrich has in mind, is that they are, section-by-section, by-and-large naturalistic. In *The Belvedere*, for example, we can recognize people and that some are climbing up ladders, sitting on benches, and looking out into the distance; we can recognize the mountains, valley, and sky in the background; and even the building itself is constructed of many naturalistic elements (e.g., the stairs, the door, the bars on the window, and so on). It is the fact that these naturalistic elements are arranged in an unnaturalistic (i.e., impossible) way that makes the global impression we get from looking at the picture contradictory. It is only as we try to see how each of these pieces adds up to a coherent whole that we realize the coherent whole is not there.

116 Two things to point out: First, perhaps evidence from certain “genres” of art may weaken this claim. Art for children’s books, as well as science fiction and fantasy art all incorporate some of the intentions I have labeled as “exceptions.” But again, as was the case with *The Belvedere* described in the previous endnote, such genres will nevertheless include certain naturalistic elements that often afford the viewer effortless recognition. The subject matter may be un-natural (a nun sitting in a flying chair; space ships doing battle; a phoenix gouging the eyes of a bassalisk) but viewers can often effortlessly recognize most (if not all) of what the picture is of. The second point I wish to make is that artists have, on some level, always been about breaking with traditions and established practices. So claiming that all pictures have a single common “customary function or role” may be a bit misleading. A survey course of the history or art can easily demonstrate how, in an overly simplistic way, each artistic movement sought to challenge the presuppositions of previous movements, and thus how each sought to challenge or at least raise awareness of the different possible ways in which pictures can function for a viewing audience. And yet I wish to maintain that pictures are meant to be recognized, whether they are purely figurative, depictive pictures or whether they are more abstract or conceptual ones. In this way, then, there may yet exist a “customary function or role” that pictures are meant to serve.

117 More work would have to be done to confirm or disprove these claims.

118 Gombrich, (2000), p. 272. Gombrich rightly points out that he draws not only on Popper for these views, but also on Gestalt psychology. Although he does not accept the entirety of this school of thought, he does attempt to pull out those parts of it that are productive and “logically right”.

119 It should be pointed out that in the context of perception hypotheses are not to be understood as consciously or intentionally formed. I do not look at the table where my cup of coffee is resting and form the hypothesis (in my head), “I am perceiving a table with a coffee cup on it.” The hypothesis *just is* my perception of the motif, and the content of that perception (what I see) is provisional and subject to refutation. Conscious or intentional hypotheses may occur where the perceptual system is diminished or incapacitated in some way, or where the motif is indeterminate or ambiguous.
This does not mean that we might not imagine that what we are seeing out the window is a thestral, a werewolf, or a hippogriff. Even though we tell children that there are no monsters under their bed or hiding in the woods we may still imagine that we see such fictional creatures lurking in the ambiguous shadows or crashing through the woods just out of sight. Assuming that we are not involved in playing a game in which such creatures exist, our guesses at these ambiguous shapes and noises will not be of this kind. When it comes to looking at a picture of such fictional creatures we can certainly “see” each one in a picture in the same way that we see Wivenhoe Park, an apple, or John F. Kennedy in a picture. Presumably, however, the content of such pictures will not be mere blobs or ambiguous arrangements of paint, but instead will be “finished” pictures of this or that fictional creature. Of course none of this assumes that every viewer will recognize that a picture is of a werewolf, much less that it is of a thestral or a hippogriff.

In truth there may actually a greater range of recognitional possibilities worth investigating. For instance: (1) I recognize the picture as a simple physical object, without yet noting that it is a picture or representation (this would be the first part of the “recognitional duality); (2) I recognize that it is a picture or representation of some sort, but I don't yet recognize what it represents; (3) I recognize, in some inferential or indirect way, that it is a picture that represents, say, an apple; but I don't actually see the apple 'in' the picture (this would be like the tangled-lines example); (4) I actually see the apple 'in' the picture (this would be the second part of the “recognitional duality). More discussion is needed to determine how, if at all, these possibilities would fit into Gombrich account. I am indebted to Robert Howell for bringing these possibilities to my attention. The idea of a “recognitional duality” fits somewhat with Dominic Lopes’ position in Looking at Pictures (1996).

We do not confuse the two types of recognizing (and thus we are not deceived) in the case of an impossible or inconsistent picture (e.g., The Belvedere) because the representationally-relevant characteristics of the picture conflict or contradict with each other. In most cases, whether we are looking at an impossible picture or a usual picture, we retain an awareness of those characteristics of the picture that make the picture a picture. It is hard to imagine a case in which we look at an impossible picture and confuse it for the depicted object, given the inconsistency between the picture’s constitutive parts. The parts of Escher’s The Belvedere may conflict or contradict each other, but we remain aware of the fact that we are looking at a picture and not the actual depicted objects. This is also made apparent by the fact that any of the impossible or inconsistent elements of a picture will continually conflict with those characteristics of a picture that make it a flat-artifact and not the depicted object (e.g., the flatness of the canvas, the frame, the glare of the lights, and so on).

At first pass this statement may sound too hyperbolic. And yet, it not only fits Gombrich’s discussion of extending concepts to cover a representation of X as well as X itself (see his discussion of the Renaissance bronze crab in 2000, pp. 114-115; see also his claim that an image extends a class of objects: 1996, p. 2, and 2000, pp. 100-102), it also does the work of capturing the nature of visual equivalence, the idea that a picture can trigger off similar (though not identical) visual responses and can thus be similar to an object in certain relevant ways. As such, I do not see why the statement “The picture is the subject though not the subject itself” cannot make perfect sense within the context of the interpretation I have given.

Though perhaps this is exactly the type of thing that happens in play. The child, picking up a stick, recognizes it as a horse and begins to ride it; the little girl, holding a blanket, recognizes it as a baby as she cradles and coos to it. And all of these examples, it should be remembered, are discussed in connection with the beginnings of substitution in “Meditations on a hobby horse” (Gombrich, 1996).
Of course, taking one three-dimensional object to be another three-dimensional object is the kind of thing that happens in the case of viewing a representational sculpture: we recognize the sculpture as something else, namely the motif it represents. So perhaps the language of “recognizing as…” describes representational sculpture as well if not better than representational pictures.

There will undoubtably arise cases in which an “ordinary” picture includes some or even a great many visual elements that border on those capable of deception. I do not propose there to exist any formula to determine at what point such “ordinary” pictures become deceptive, and thus leave a viewer without the first kind of recognizing I mentioned. It is quite obviously a gray area that would have to be addressed on a case-by-case basis. All that I am proposing here is that pictures that fall within the core or central cases I believe Gombrich has in mind (i.e., the ordinary ones) will afford two kinds of recognition which I take to be obvious and uncontroversial.


See my discussion of constancy, especially color constancy, in Chapter 2

To make a definitive claim is beyond the scope of my expertise. I am here making an empirical proposal that there is a certain set of relations that underlies a picture’s representing an object as well as the fact that viewer’s can see the picture and have an experience “as of” an object or see the picture “in terms of” an object. But to conclusively claim that these relations do underlie my assertions is an empirical claim that requires empirical studies or experiments to prove or disprove it. All I am claiming here is that the picture preserves an equivalent set of relationally arranged visual elements that are capable of maintaining perceptually constancy. Such relationships will not be exactly the same as those found in the object, but rather, I have been claiming, the best analogue set of relationally arranged visual elements within the specific medium.

In what is to follow the notion of perceptual constancy should be seen as a sufficient condition for recognition—perceiving perceptually constant features will afford us recognition of an object, but not every act of recognition will involve the perception of features that are perceptually constant.

In The Image and the Eye Gombrich writes that the idea of constancy “covers the totality of those stabilizing tendencies that prevent us from getting giddy in a world of fluctuating appearances. As a man comes to greet us in the street, his image will double in size if he approaches from twenty yards to ten. If he stretches out his hand to greet us, it becomes enormous. We do not register the degree of these changes; his image remains relatively constant and so does the colour [sic] of his hair, despite the changes of light and reflection.” (1982, p. 18).

By “color value” I mean the lightness or darkness of a color. Contrasts in color value can separate objects in space (e.g., a dark color value surrounding a light color value could suggest that the light color value area is being illuminated by a light source and is thus closer to us than the dark color value area). Gradations of color value demonstrate mass or the contour of a contiguous space. In addition to these possibilities, Gombrich also mentions that certain physiognomic features might afford constancy as well. See also endnote 89.

I want to explicitly point out, here, that the relation R in P should be understood as “relation R in the depicted object O in P.” This is important because R in P is meant to serve as an equivalent to S in O, and I do not want the reader to be confused by the different designators “R” and “S.”

Gombrich (2000), pp. 33-90. See especially section 3 (pp. 49-52).
To be clear, I am not claiming that there is one single relationally arranged set of visual elements that is constitutive of the totality of an object’s characteristics. Rather, when we perceive, e.g., the wheat stack, we will actually perceive a number of different relations between a number of different visual elements. For example: an object’s contrast of color value might be constitutive of the relation $S_{f_1...f_n}$; it’s size might be constitutive of relation $X_{e_1...e_n}$; it’s shape might be constitutive of relation $Y_{d_1...d_n}$, and so on. In this way, when we look at the actual wheat stack we perceive a host of different relationally arranged sets of visual elements that present us with a number of “wheat stack” features: size, shape, and the contrasts and gradients of color value (i.e., the lightness and darkness of the wheat stack).

I think that, at its simplest, the physiognomic features Gombrich has in mind are those associated with the features of a human face (e.g., ears, nose, mouth, eyes, etc.). What is quite amazing, writes Gombrich, is that our visual system seems to be “hyper-sensitized” to “single out the expressive features of a face from the chaos of sensations that surrounds it,” which Gombrich describes as our ability to “extend the group of shapes that can be read as a physiognomy” (Gombrich, 1996, p. 6). While it may prove impossible to determine which variables or features of a given face are the ones that afford us physiognomic constancy of that face, it remains a rather obvious fact that we are able to pick out the face of a friend, a child, or a spouse from a crowd of others. For more on physiognomy see Gombrich (1969); Gombrich (1992); and Gombrich (2000), especially section ten, “The experiment of caricature.”

It would seem that the range of possible relations of elements is vast, and indeed this is the case: the artist has a daunting task of which ones to choose from, and there are a large number that can trigger recognition. And yet, the artist still works through trial and error to find ones that serve her purpose, ones that work for her and for the viewer. Perhaps she could have achieved the same or similar results with a different set of relations of elements. Even if that is true, it still remains that in choosing a subject the artist constrains the possible relations of elements she will be choosing from. The same subject rendered in pencil lines and in oil paints will probably require different relations of elements; two different subjects rendered in oil paints might also require choosing from different relations of elements. While recognition can certainly be triggered off by a number of different relations, even still the possible ones are narrowed down as the artist sets about the task of actually making the picture.

Dan Gilman offers a more detailed discussion of these points in the section “Perspective” in the first chapter of his dissertation, Lines of sight: an essay on mind, vision, and pictorial representation. He also cites Max Black (1972) and Nelson Goodman (1976).

This claim is a bit misleading. Even in the most optimal of viewing conditions we never perceive all of an object’s visual features. This is simply a fact about how humans perceive the world. We cannot see around corners or behind objects without shifting our viewing position. It might be wise to point out that some might claim that Cubist works have sought to overcome this perceptual limitation by presenting the viewer with all of an object’s visual features at the same time. If this is the case then they are certainly not “naturalistic” in the usual sense of the word, and thus they depict in an a rather un-naturalistic manner.

I believe this is simply another way of saying that pictures are translations—they cannot copy all of the information found in an object itself, and so each pictorial medium (e.g., pencil, oil, charcoal, film, and so on) must work to accommodate the relationships within its own limitations.
There are a couple things to point out here. (1) There could be a way in which rectangles and triangles are used to “symbolize” a circular rim, though in such a case I am unsure whether we would say that they “depict” the circular rim. (2) There is also the possibility that very small rectangles and triangles could be arranged to depict a circular rim (e.g., an entire picture might be created consisting of nothing but very small rectangles and triangles as the the sole “depictive elements” in the same way that very small square pixels are arranged to depict objects on a computer screen. I am inclined to think that a case such as (2) would be different than the sorts of depictive pictures I have in mind, but further discussion would be needed to clarify this position. (I would like to thank Robert Howell for bringing this possibility to my attention.)

Gombrich gives this example in *Art and Illusion* (2000, p. 344).

I have taken the concept of “characteristic” viewpoint from Gombrich, though he uses “shape” instead of “viewpoint” to describe the sort of thing I am talking about. In either case, I think that the idea of a characteristic way of looking at an object is synonymous with the psychological concept of the “canonical view.” We can see examples of the canonical view when we look at a pedestrian crossing sign, or a “no dog’s allow” sign. In these cases the point of the sign needs to be understood with the least amount of work on the part of the reader (viewer), so creating a picture of a pedestrian or a dog the can be read quickly and easily means using some canonical or characteristic shape or viewpoint.

Gombrich’s example is that of a penny. The characteristic view of a penny, the one which gives us the most information and by which we can most easily characterize it, is straight on with the image of Abraham Lincoln facing towards the viewer.


Lacking any formal training in psychology or experimental philosophy, I have no way of verifying whether *all* features of perceived objects are really perceptually constant features in the manner I have described. Certainly there are a range of features that we might call the “usual” perceptually constant features, but there may also be a range of “unusual” perceptually constancy features. If so, does their being unusual mean that they are a different kind of feature, or might we simply want to say that they allow perceptual constancy to occur albeit in a different way? For instance, I can recognize my wife when looking straight on at her (in the usual way), but I might also recognize her when I look down at her from a building (an unusual way): is perceptual constancy working in the same way in both these cases, or are there wholly different features allowing recognition to occur? I do not know.

I should make it clear straightaway that while my primary concern is the switching that occurs between a meaningful and a meaningless interpretation, a viewer can switch between different meaningful interpretations, and a viewer can also switch between different meaningless interpretations. I turn to the issue of switching between different meaningful interpretations in sections 4.4.1 and 4.4.2. I do not, however, discuss the switching between different meaningless interpretations of a picture. That is, I do not discuss the possibility of switching between seeing different non-representationally relevant ways of seeing the picture. Such switching might involve seeing (i.e., focusing on) the flatness of the canvas; switching and seeing the texture of the canvas; switching again and seeing the glare of the light off paint, and so on. I have, in the text, subsumed all such “meaningless” interpretations under a single heading. The reader must excuse me for perhaps doing an injustice to the many diverse ways in which we see the non-representational elements of a picture; nevertheless, this approach has proven to be the easiest means for discussing the primary issue of Gombrich’s notion of meaningful/meaningless switching.
In ordinary perception the notion of “being aware” is meant to capture a kind of peripheral-perception; that is, an act of perception that is not quite at the conscious level. Two such examples come to mind. (1) When I am in conversation with a friend and I notice that, behind him, someone has walked in, I am able to retain my focused attention on my friend while, at the same time, being aware that someone has entered the room. (2) When I reach out and grasp the pencil on my desk I am attending to the pencil, but I am nevertheless aware of my hand. In sections 4.5-4.5.2 I will attempt to extend this idea of peripheral awareness to the perception of pictures, claiming that while we can remain aware of the picture’s medium at the same time as attending to the picture’s content, we cannot attend to both medium and content at the same time.

See Wollheim, 1980.

The term “disjunction” is used by Wollheim (1980, p. 214). The quote in Art and Illusion from which Wollheim draws his conclusions says that rather than copy what they saw artists had to “manipulate those ambiguous cues on which we have to rely in stationary vision till their image was indistinguishable from reality. In other words, instead of playing ‘rabbit or duck’ they had to play ‘canvas or nature,’ played with a configuration of colored earth which—at a distance at least—might result in illusion.” (Gombrich, 2000, p. 29)

The first reason Wollheim gives, “concerns the range of things that we may see in something as opposed to those which we may see something as.” (1980, p. 210) We can only see particulars (“something seen is a particular”), and so we can only see things as particulars; “but, as for seeing-in, we may see not only particulars in it, but also states of affairs.” So we can see a state of affairs in P but we cannot see P as some state of affairs. Wollheim gives an example: “If I am looking at x, and x is a particular, I can see a woman in x, and I can also see in x that a woman is reading a love-letter: but, whereas I can see x as a woman, I cannot see x as that a woman is reading a love-letter.” The second reason for the change from seeing-as to seeing-in is that the latter cannot be constrained by what Wollheim calls “the requirement of localization”. That is, when I see x as y I can specify which bits of x (up to the whole) that I see as being y. I cannot do the same sort of thing for seeing-in: I cannot specify which parts of x allow me to see y in x.

I will not take up the issue of whether Wollheim is correct in his view that seeing-as is deficient in this regard, though it should be noted that I am not all that convinced by his arguments. In fact, it might be the case that a seeing-as theory can be salvaged by the same kind of distinction (between “being aware” and “attending to”) that I make in this chapter. If this is the case, and the “being aware”/”attending to” distinction can apply to seeing-as and seeing-in, then so much the worse for seeing-in. Unfortunately, there is not ample space to address this issue so I will have to postpone such a discussion until another time.

I am skeptical of this jump. Wollheim offers no evidence or support for his claim that we cannot simultaneously see x as y and see how x sustains our perception of it as y. As I understand it, twofoldness is the unique distinguishing feature of seeing-in, which means that if seeing-as afforded the same kind of dual simultaneous experience there would be nothing special about seeing-in as opposed to seeing-as. Perhaps Wollheim is rushing his conclusion, but since I do not have any evidence to counter his move I will leave my remarks as they stand.

It is unclear why Wollheim uses “unlimited” to refer to the breadth of the simultaneous experience of seeing the object and the medium. Is it that we can enter this dual experience at any time during our perception of a representational picture? Or, put more accurately: can we see the represented object and the picture’s constitutive medium whenever we look at a representational picture, and thus we have limitless access to this dual experience so long as we continue engage in perceiving the picture? It's unclear what benefit using this term has.
It might be somewhat unfair to offer a defense of Gombrich while following Wollheim’s later conception of twofoldness, since it was the early conception that precipitated Wollheim’s critique of Gombrich. Wollheim’s response to Gombrich initially involved laying out a conception of twofoldness that involved two separate experiences. The shift to a conception of twofoldness involving a single experience with two aspects certainly did not, however, dismiss Wollheim’s complaints of Gombrich’s position, nor did it serve to reconcile any of the wider differences between them. It is for that reason that continuing along the path I have laid out should not be seen as problematic. I simply feel that it is pertinent to point out this subtle point, even if it is a case of splitting hairs.


Bence Nanay (2005) claims that Wollheim equivocates between two senses of twofoldness, each of which takes up a different position on the nature of the configurational aspect: the first sense involves a viewer being aware of two aspects (the represented object and the way it is represented), while the second sense involves a viewer being aware of two entities (the surface and the represented object). This may seem a semantic quibble, but it may make for an important distinction. Let me put it this way: there are certain features of the picture that are constitutive of "how" it represents an object and are thus said to be representationally relevant features (e.g., the arrangement of colors, shapes, and lines)—these relate to the first sense above. Other features, while part of the experience of looking at a picture, are simply features of the picture as a picture and are not considered to be representationally relevant (e.g., cracks in the dried paint, the frame)—these relate to the second sense. So, when we are aware of the configurational aspect in the second sense we are aware of features of the picture that do not contribute to it as a representation (i.e., features, like the cracks in the dried paint, that are not relevant to the picture’s being a representation); when we are aware of the configurational aspect in the first sense, we see features of the picture that add up to some representational content (e.g., a manor house). All of this is to say that our understanding of the configurational aspect of twofoldness is not settled; and, in fact, it makes a significant difference how we understand it.

One of the major stumbling blocks for Wollheim’s account, at least according to critics, is that he does not offer any details regarding how the two aspects are related to each other.

See the previous quote from page 5 of Art and Illusion.

I think that the impossibility of simultaneously holding multiple meaningful interpretations is further explained by the difficulty in dislodging an interpretation once it has taken hold. While this is especially true of our perception of random arrangements of lines and shapes, I believe Gombrich intends it to also be true of both simple line drawings and naturalistic paintings. See his discussion in 1962, ff.; and 2000, pp. 226-228, and 236.

It is not impossible that we could see a landscape picture as some sort of Ames-demonstration-like arrangement of lines and surfaces that, when seen from a certain angle, looks like a landscape. But while this is not an impossibility, it remains highly unlikely to be realized in the ordinary cases of picture perception for the very simple reason that such an interpretation will almost never enter our minds.
By “conflicting” I mean the holding of representationally relevant interpretations of the same image—seeing P as of O and as of R at the same time. For example, seeing Self portrait with soft hat as a picture of a Degas and as a picture of a flowering poinsettia is impossible. We simply cannot hold both visual interpretations at the same time. We may certainly attempt to see the painting as being of a flowering poinsettia (though doing so would be an incorrect interpretation of the picture), but in so attempting we must set aside seeing the painting as of Degas. Significantly, this does not exclude the possibility of interpreting the picture as representing something and having some symbolic value. I can see a painting as of a man hanging on a cross with a dove flying about him; I can also see the scene as symbolizing the Holy Spirit descending down from heaven. Both interpretations can be held, I believe, at the same time. This is, in fact, the very point Gombrich makes in the introduction to Symbolic Images: Studies in the Art of the Renaissance. Images, while being representations of objects, scenes and events, are also symbols, which is where the role of conventions comes into play. Thanks to Richard Woodfield for pointing this out to me.

While we may not visually recognize the cues that indicate flatness nor the marks constitutive of the medium, we may nevertheless visually experience both of them, if only at an unconscious perceptual processing level.

I have already stated that Gombrich is quite clear in divorcing his claims concerning illusion from that of deception and false belief. That said, I will have to hold off a more detailed discussion of illusion until the next chapter.

For my purposes I think that this is a sufficient enough explanation of what happens when we perceive a trompe l’oeil. I do want to point out that there may be a few other possibilities of what occurs when we look at a picture: (1) it might be that we forget the medium (i.e., forget the medium intellectually while registering it perceptually); (2) we may visually recognize, in a deliberate conscious way, the medium as such (i.e., “these are dabs of paint on a flat page” or something like this); (3) we might be visually conscious of the marks on the page while not giving them our attention (i.e., we would know if they were changed, or at least we would know something in our perception was different); (4) the marks on the page are registered by our perceptual system even though we are not or may not be conscious of them, even minimally, at all (i.e., the marks may register at an unconscious perceptual processing level). These are all possibilities given my distinction (in the previous chapters) of “being aware of” and “attending to,” but I will not pursue them in any detail.

“Simple seeing-in” is distinguished from “pictorial seeing.” The former kind of seeing seems to be just the type of seeing involved in recognizing pictorial content without being aware of or connecting with the picture as a picture: for instance, the awareness “directed to or had in connection with postcards, passport photos, magazine illustrations, comic strips, television shows, or movies.” (Levinson, 1998, p. 228) Levinson goes on to use a number of idioms to describe the latter kind of seeing. For example, pictorial seeing is: seeing pictures as pictures; seeing pictures as pictorial contrivances; seeing or appreciating pictures aesthetically. Pictorial seeing-in thus seems to involve some more conscious or intentional (i.e., Nanay’s description “higher-order”) activity than simply seeing an object in a picture.
The way in which I understand and apply “attending to” and “awareness of” appears to correspond, roughly, to the kind of distinction between thetic and non-thetic consciousness, as well as the principle of figure/ground proposed by Gestalt theory. As I understand it, the notion of thetic and non-thetic consciousness holds that when I pick up, say, a piece of wood I am thetically aware of the roughness of the wood, but non-thetically aware of my hand as such—the roughness of the wood becomes the primary object of my consciousness while my hand persists as a conduit towards that primary object of consciousness. As a conduit leading to thetic consciousness, the non-thetic consciousness acts “as a kind of background hum or drone to our conscious thetic activity.” (bat020, 2010) This idea is also echoed in the Gestalt principle of figure/ground. As the individual perceives he or she takes in the environment as a whole, and in responding to the environment as a whole the individual will respond to the stimuli that are constitutive of that whole. Some of these stimuli the individual is aware of (figure), and others the individual is not aware of (ground). In choosing to focus on, say, the duck it becomes the figure, while at the same time I relegate the rest of the image (i.e., the rabbit) to the ground. Looking at more complex scenes, whether in the external world or in a picture, requires one to choose (sometimes automatically) what is to emerge as the figure and what is to recede as the ground. The choice will result, Gombrich will say, from a complex amalgam of physiology, expectations, and background knowledge; thus, we return to the interplay between the artist’s constructed image (what is given) and the proactive role of the beholder (what is projected). The notion of thetic/non-thetic consciousness and figure/ground do not perfectly track the one I have proposed; yet, in gesturing to them I can only serve to bolster my claim that there are important differences between our “being aware” and “attending to” objects of perception.

The reader should know that Nanay develops two modified definitions of twofoldness based on the claim that Wollheim actually uses of the concept “twofoldness” in different ways. While the distinction between “being visually aware” and “attending to” is important for these two definitions, I do not think that the details of his account will affect my claim that Gombrich allows for some conception of twofoldness. While there is room to debate these details, and hopefully the interpretation I propose will precipitate such a debate, this is not the place to do so.

I equate “peripheral attention” with the idea of “being visually aware,” which I describe in Chapter 3. See also endnote 165 for a further list of possible ways we might experience pictures.


Nanay refers to this type of conscious activity as being a “higher-order” conscious activity. I have opted to drop this descriptor, instead opting to stick with the phrase “attending to.”

Actually, I think that the farther away multiple objects are from a viewer the more easily it is to (consciously) attend to them at the same time within the visual field. The farmhouse and the herd of cows in the distance; the gas station and the city bus; the goal keeper and the striker scoring a goal: in all these cases, depending on the point of view and the distance from which I am viewing the scene, I may be able to take in (i.e., attend in a conscious way to) both objects simultaneously. It is much more difficult, nay impossible, to consciously attend to both the farmhouse and the herd of cows when the herd of cows is direct in front of me and the farmhouse is in the distance. In the example from the text, perhaps if my friend and the customer were standing side by side at a certain distance from me I could be visually conscious (i.e., aware) of both of them at the same time without having to give both of them my full visual attention. But again, if they are standing at a relatively close proximity to me it will still be difficult to devote dual visual attention to both of them at the same time. At such a close distance I may simply not be able to devote simultaneous visual attention to multiple objects within my visual field, though I may still be able to maintain some level of conscious visual awareness (in the sense of the distinction made in Chapter 3).
This also seems to be a conclusion shared by Andrew Harrison (2001). He writes that: “Switching between two exclusive ‘picture interpretations,’ or ‘recognitions’ is quite a different matter from switching between some ‘picture interpretation’ and another and no pictorial interpretation at all. It is the latter which surely has to constitute simply attending to the marked surface in its own right. However, such a corrective, when right enough, might well be used to strengthen Gombrich’s point, which was that once we do attend to a marked surface as a picture this excludes attending to it as ‘merely’ a marked surface.”(p. 41)

See Chapter 3.

Wollheim refers to this kind of seeing as “straightforward seeing” (1980, p. 217).

I will talk more about the details of Gombrich’s theory of illusion, as well as the various criticisms of it, in the next chapter.

Recall that the goal of Chapter 4 was to demonstrate that, the critics claims to the contrary, we can interpret Gombrich as allowing for a simultaneous awareness of a picture’s medium and its content.

My thoughts on why confusion remains around the neck of Gombrich’s theory of illusion are echoed by Richard Woodfield (albeit in a much more general tone), and his entry in Stuart Brown’s *Thoemmes Continuum Dictionary of 20th Century Philosophers* seems to offer a suggestion as to why this is. There he writes: “Originally delivered as a series of public lectures in Washington’s National Gallery, [Art and Illusion] was aimed at a cultivated public rather than a specialist scientific audience. As a consequence he used memorable formulations, such as ‘making comes before matching’, vivid illustrations, such as his use of Jastrow’s duck-rabbit to illustrate perceptual ambiguity, and striking metaphors. By contrast his later collection of essays *The Image and the Eye* largely resulted from presentations to scientific audiences with a consequently more rigorous use of examples and language. Unsurprisingly he has been accused of a ‘change of heart’, shifting from conventionalist humanist to positivist scientist. In truth his work has been completely consistent from his earliest publications in the Viennese journal *Kritische Berichte* to his posthumous publication *The Preference for the Primitive*.” (Retrieved online; pp. 1-2)

Gombrich points this out in a few different places. In the article “Illusion and Art” (1973) he is both “gratified and puzzled by the attention which my discussion of illusion has been accorded, for it had never been the central issue of Art and Illusion.” (p. 195) The book was drawn from a series of lectures entitled “The Visible World and the Language of Art,” and the subsequent title was chosen just as much for brevity’s sake as it was for the fact that it retained the word “Art.” “We never dreamed that this title [“Art and Illusion”] would convey to some that I considered illusion, or even deception, the main aim of art.” (1973, p. 195) More recent evidence comes from the Preface to the 2000 Edition of *Art and Illusion*. There Gombrich again regards with surprise the fact that the title, *Art and Illusion*, “has misled some people to think that I value, or even advocate, illusionistic painting.” (p. xxv) For though “the study of art and the study of illusion cannot always be kept apart” the book, *Art and Illusion*, “is not intended as a plea, disguised or otherwise, for the exercise of illusionistic tricks in painting today.” (2000, p. 7)

For more on the specific style of “illusionism” and “illusionistic art,” as well as how it differs from trompe l’oeils, see the first chapter of M.L. d’Orange Mastai’s book *Illusionism in Art: Trompe l’oeil a History of Pictorial Illusionism* (pp. 7-25).
On a side note, Catherine Wilson (1982) makes a rather telling remark regarding the possible confusion surrounding the idea that a picture creates the illusion of what it represents. At one point Wilson is discussing what it means to see a picture of a duck as a duck. To say, “I see it as a duck,” just happens to mean the same thing as, “I see a picture of a duck” when what is being seen is in fact a picture of a duck. She then goes on to say that, “If it is misleading to say that a picture of a lion gives you the illusion of seeing a lion it is equally misleading to say that you see it as a lion.” (1982, 217) I take this statement to be imply that if the opposite should be the case and it is not misleading to say that I see the picture as a lion, then equally it must not be the misleading to say that the picture of a lion gives me the illusion of a lion. The former is a common enough phrase, and to dispense with it that would require such an acceptance I do not believe to be possible since it would be at odds with how we normally talk about our experience of seeing pictures. Wilson’s remarks aim to show us that there need not be some mystical voodoo surrounding the notion of a picture generating an illusion—so long as we clarify what we mean by “illusion” (and it could mean a great many things) we can use the concept in a coherent, and meaningful way to describe the everyday occurrence of seeing a picture.

See Gombrich, 2000, pp. 104, and 195

The discussion thus far warrants a few further comments. Allen is wrong to equate “illusion” and “delusion,” and thus he is wrong to confine “illusion” only to cases of trompe l’oeils. Ordinary pictures do not create delusory experiences while trompe l’oeils do, therefore the term “delusion” should be reserved only for cases of the latter. However, this does not negate the fact that Gombrichian illusion (as I interpret it) is involved in both ordinary pictures and trompe l’oeils. Because both types of pictures count as pictorial representations, and because both function as visual substitutes in the manner I have described, each can be said to involve Gombrich-style illusion. The difference is that the illusion of the trompe l’oeil becomes delusion because we lose an awareness the picture’s medium—our experience of what makes the picture a picture is absorbed by our recognition of the picture’s content. Throughout the dissertation I maintain the terminological distinction of confining the term “delusion” to trompe l’oeils and “illusion” to ordinary pictures because that is the easiest way to understand what a pictorial illusion is. While I explicitly maintain this distinction throughout the dissertation it is important to keep in mind the fact that “illusion” involves a spectrum of degrees, and is thus found in both the ordinary picture and the trompe l’oeil.


I say that the connotational range of illusion could be extended to included delusion, though it need not be; furthermore, to do so would require one to explicitly state how the term “illusion” is being used. I am quite happy to allow “illusion” to be used synonymously with “delusion” if in so doing: (1) the use of “illusion” is confined to the appropriate set of examples, viz. instances of trompe l’oeils and not ordinary pictures; and (2) that what one really means when one uses the term “illusion” to describe a trompe l’oeil is that it involves “delusion.” In this way, so long as one uses the appropriate set of examples and clarifies the meaning of the terms being used, illusion can be used to describe a delusory state. That being said, this type of clarity is rarely (if ever) done, which simply perpetuates the confusion over what is meant by an illusion.

It is unclear whether trompe l’oeil are the only pictorial representations that lead to false belief about the objects being represented. Certain photorealistic paintings, such as Ralph Goings’ Ralph’s Diner or Richard Estes’ Paris Street Scene, might be just as visually jarring as a trompe l’oeil—possibly even more so, since the subject matter of a photorealist painting is not limited to a shallow depth of field like that of the trompe l’oeil. I cannot take the time to investigate such paintings, so I must leave the question of their significance unanswered.
This does not negate the possibility that an artist might employ far-off landscapes in the service of visual deception: the artist could surround the painted landscape with a painted or actual window, and then require the viewer to stand at a certain distance, thus leading the viewer to think that he or she was looking out onto an actual landscape. Such a scenario would be even more contrived than viewing conditions for a usual trompe l’oeil, and thus would be even rarer.

When I stand on the porch of my brother and sister-in-law’s house I know that Mount Mansfield is farther away than everything else in my field of vision even if the information received by my visual system does not present a sense of depth that allows me to perceive this. But whether we can measure the extent to which my knowledge influences what I see, or affects how the perceived visual information is processed, may prove to be impossible.

I believe this is an important point and I do not wish the viewer to pass over it too quickly. Most discussions of trompe l’oeils persist as though the epistemic deception (i.e., false belief) were commensurate to the perceptual deception in all respects. I believe this is false; and further thought can reveal this to be so. Yes, a trompe l’oeil generates a perceptual delusion; yes, it can also generate an epistemic delusion. The former will persist indefinitely so long as our awareness of the medium is absorbed by our awareness of the content; the latter can only occur for a finite time as the facts about the picture qua picture will impinge upon our perceptual experience and destroy any substantive false belief we hold about what we are seeing. A trompe l’oeil does not necessarily generate an unending epistemic lacunae that automatically trips us up. While it can sometimes lead to false belief that lasts for a limited length of time, a perpetual state of false belief is simply not a possibility outside certain contrived experimental settings.

Again, the trompe l’oeil’s deception is never complete. Three comments on this point: (1) It is interesting that many people encounter trompe l’oeils, not hanging in a gallery, but in a book or on a computer screen. They "know" that this picture is a trompe l’oeil; and they know that it is supposed to create an experience that deceives the viewer into taking the picture to be the object. They can see how life-like or realistic the picture is, but are they really deceived? The picture, even in color, is flattened out and lacks the true deceptive nature that it carries when seen in person. (2) Perhaps a viewer is deceived the first time he or she sees a trompe l’oeil, but subsequently the proverbial curtain is pulled back and you know how the trick is done. One might fail to discern the medium, but there is still the edges of the frame, the wall on which the pictures hangs, the sheen of light on the picture surface, etc. All those things countermand the experience that this is the object itself; thus, the deception falls short. (3) One might question whether the experience of seeing a trompe l’oeil can ever involve a false belief. That is, is it ever the artist’s goal to create a trompe l’oeil picture that the viewer truly takes to be the actual three-dimensional object it represents? The critic might say that an audience that is knowledgeable of the trompe l’oeil-style will never truly be epistemically deceived; the artist knows this. The artist might succeed in creating a visual deception, but she is aware that she has no chance (and no intention) of inducing a false belief about the status of the picture because the audience is not sufficiently naive. While I think this is a plausible position to take there is not enough space to pursue it in earnest. It would require deeper investigation into what we mean by “false belief” and whether there is a role to play for certain sub-doaxastic belief states such as “alief.”
It might be better to say that once I am aware of the trompe l’oeil illusion it is easier for me to enter into and out of it—for until I am aware that my visual experience is as of a picture I still find myself in the hold of epistemic deception, unable to control what my visual experience is of. My wife could no more control the fact that her visual experience was one of actual scissors than she can control the fact that her visual experience of an apple is one of “red” or “roundness.” Once one becomes aware that the visual experience of “actual scissors” is actually one of “painted scissors” one can enter into and out of the illusion more easily. Though, such an awareness does not destroy the illusion altogether—I still see the picture as of scissors, though now I know that they are not actual scissors no matter how realistic they look.

The Müller-Lyer illusion might not be universally experience. Some studies have shown that it can vary across cultures (Segall, Campbell and Herskovitz, 1963) and across age groups within a single culture (Ahluwalia, 1978). The hypothesis is that whether an individual experiences the illusion can depend upon whether he or she is exposed to a Western carpentered world (consisting of rectilinear objects that signify the corners of objects in perspective).

See the discussion of this in section 1.5 in Chapter 1.

The idea that we are unable to see the medium because it is absorbed by our recognition of the picture’s content comes from Katerina Bantinaki (2007).

Let me clarify: our perception of ordinary pictures can certainly involve the belief that we are looking at the depicted object—we would just be wrong. We would be mistaken to have such a belief. There are no doubt times when, for any number of reasons, we see an ordinary picture as an actual object, but in such cases we are just wrong about our perception. Perhaps the lighting is off; perhaps we are tired, or perhaps our senses are overstimulated from an exceptionally good meal. Whatever the case maybe be it is more than possible that we can make a mistake about the nature of an ordinary picture. But that is just what it is, a mistake, and we should acknowledge that in most cases perception of the picture will go off as it should: we recognize the depicted object but do not believe that it is the actual object.

Clearly the kind of “riding” the child performs with the hobbyhorse is different than how one actually rides a horse, but there do remain similarities (e.g., sitting astride the object and then moving the stick as you would move with a horse). To the extent that the goal is to do something “riding-like,” the fact that I can do these things makes the stick function for me as a horse.

Portis seems to imply that we, as outsiders (and probably more specifically, as adults) simply cannot “see” all of the different ways in which a stick can be used by the child. This is true when we watch children playing. It may look to me as if my nephew is simply climbing the snow bank, but within the context of play (i.e., in his mind) he is scaling the tallest mountain, being careful not to fall off the side of the climb. We cannot tell what is going on from “the outside.” Should we enter into the context of play, however, the snowbank (or the stick) can become something more than the physical object that it is.

I use the term “perceive” to mean “being looked at” or “being seen.” There remains the possibility that those without sight can nonetheless encounter and interact with a pictorial representation (as evidenced by the work of John Kennedy, 1974 and 1993). Perhaps a blind individual can “read” a picture that has raised bumps and lines, touching it as she would braille; and perhaps there is more to be said over the nature of pictorial representations. I cannot devote the appropriate space to such topics here, so I will leave the discussion as it is. For the purposes of this investigation I will focus on the traditional understanding of a pictorial representation as a marked artifact that depicts a subject-matter and is seen by some viewer.
This is a worry echoed by Katerina Bantinaki: “Gombrich argues that it is the viewer’s mental set that determines what she will see in a given picture [italics added]” (2007, p. 274), rather than the design itself, “which makes the object perceived really a product of the viewer’s mental set.” (2007, p. 276)

I believe that we can safely assume from what Gombrich says here that coupled with his account of projection he believes there to be interpretations common to or shared by viewers. The beholder’s share to project interpretations into an image notwithstanding, there will be interpretations of a picture that will be common to different viewers—one might call these interpretations “objective.” Gombrich can draw this conclusion because (1) picture’s work as relational models for what they depict, capturing the relationships between visual elements found in nature; (2) our visual system’s response to seeing arrangements of visual elements triggers certain responses, and assuming that our visual system is working correctly (and thus that we are no mistaken or hallucinating) the relationships of visual elements constitutive of the picture will trigger off a limited range of visual responses—and with it a limited range of interpretations of what is being seen. So it is in this way that Gombrich might conclude that (all? most?) viewers will arrive at a chimney interpretation; furthermore, it is is this way that Gombrich might label such common or shared interpretations as “objective.”

Gombrich mentions the “thereness-thatness” principle in 2000, p. 259.

Now, it is also true that (in a sense) all pictures offer us contradictory information as well: a representational picture presents us with both two-dimensional visual information and three-dimensional information (or, at least, two-dimensional information that can be interpreted as three-dimensional information). So at the same time a picture contains (a) features that are flat, and two-dimensional; and it contains (b) features that can be visually interpreted as three-dimensional depth or space. Again, the sense of three-dimensional visual information is not tied to an actual three-dimensional object but to a two-dimensional object—the picture. The sense of depth in Wivenhoe Park is not the result of the picture opening up into an actual landscape, and so it is in this sense that a picture presents us with contradictory information, affording a dual interpretation—flat artifact and an object with depth and space.

There are two points to make here: first, the reader should remember that I am focused on clear (core) cases of pictures that depict objects or scenes. We can debate the fringe cases where the kind of thing being represented is not all that concrete (e.g., representing emotions and feelings, or where the picture is said to represent some abstract thing or concept); but in bringing in the discussion of affordances I am explicitly tying it to those clear (core) cases of representational pictures. Second, even if a picture were intentionally created in a haphazard manner (e.g., “I am just going to see what happens when I do this…”, which, incidentally, is something Picasso would often do), again, I do not believe that the specific arrangement of design features in the picture can be seen as just anything. The design features may afford multiple interpretations, but they will not be unlimited. So even a randomly created image will afford a limited number of interpretations given the constitutive design features it has.

Gombrich talks about the “transparency” of the pictorial medium in “The Evidence of Images” (Gombrich, 1969).
Recall that pictures, at least the sorts of representational ones Gombrich is concerned with, can function as both representation and symbol. (See also endnote 162). This is rather obviously evident from the sorts of allegorical images that emerged during the Medieval time; but it also possible that just about any image can carry with it an underlying symbolism. The intentions of the artist; the mores and moral codes of a particular society; certain historical markers; the variety of cultural stereotypes; certain religious imagery—all of these have the potential to impregnate an image with symbolic value. Such symbolic value can thus exist alongside the representational content of an image, possibly going unnoticed by viewers who are unaware of those clues that hint at the symbol’s existence. Most viewers (myself included) are ignorant of the historical, cultural, sociological, and religious circumstances out of which images were created. Ignorance of such things may not hinder the recognitional experience of the picture—I recognize the picture as of a man nailed to cross—but being unaware of the picture’s symbolic value (or elements in the picture that function symbolically, e.g., a dove, symbolizing the Christian idea of the Holy Spirit, descending towards the man on the cross) may hinder a fuller awareness of how the image was intended to function. All of these points make the perception of representational images a much more complex activity, but any investigation of this complexity must be put aside for another occasion.

As Gombrich seems to understand it, visual substitution tracks the changes in the function of an artifact. For instance, societies that worship an idol take it to be, quite literally, the deity—they worship the idol as the deity; the child turns the hobbyhorse into a horse, but she does this within the context of play—the worshipers, however, are not playing at worshiping the deity, they worship it. “Play” is important for distinguishing the two. While the same horse-object can afford a variety of uses (being worshiped and being played with) the context of action shows us how use, by an individual, can constrain the possible ways in which the object functions. The representational image shows us yet another change in function: it no longer serves as a physical stand-in, in the way that the hobbyhorse does, but instead affords a viewer a certain visual usage. The hobbyhorse is the substitute, whereas the image serves to suggest a reality outside itself that a viewer endows with three-dimensional space and which affords the viewer a certain visual recognitional experience. The “use” of the representational image, how it is intended to function, distinguishes it from the idol and the hobbyhorse (even though all three can, generally speaking, be labeled as “substitutes”). See also