

4

Fact and Fiction in the Neuropsychology of Art

*Roman Frigg and Catherine Howard*¹

The time-honoured philosophical issue of how to resolve the mind/body problem has taken a more scientific turn of late. Instead of discussing issues of the soul and emotion and person and their reduction to a physical form, we now ask ourselves how well-understood cognitive and social concepts fit into the growing and changing field of neuropsychology. One of the many projects that have come out of this new scientific endeavour is Zaidel's (2005) enquiry into the neuropsychological bases of art.²

Zaidel's book is widely considered to be a landmark piece in the field, and for this reason it deserves careful consideration. Her approach to the neuropsychology of art combines neurological, evolutionary, and cognitive perspectives. One aspect of this programme is the study of brain damage in established artists: the relationship between art and certain brain regions is investigated through an examination of the difference between pre- and post-damage works of a painter suffering brain damage. This aspect of Zaidel's programme is the focus of this essay. Our conclusion will be critical: the methods used are inappropriate to the subject matter, and progress in the study of neuropsychology of art is more likely to be made if the topic is approached in a different way. We should emphasize, however, that we do not reach this conclusion because we generally have a sceptical perspective on reductionist programmes such as Zaidel's. In our view there is no good reason to believe that we could not, at least in principle, one day have a thorough understanding of the physiology behind our ability to create, imagine, enjoy, and recognize art. The aim of Zaidel's project is a good one.

In her investigation Zaidel uses a range of established artists (i.e., artists who are recognized as such by art history and society broadly construed) from a range of time

¹ The authors are listed alphabetically; the paper is collaborative. To contact the authors write to r.p.frigg@lse.ac.uk and z3192298@student.unsw.edu.au

² Related research programmes are those of Vilayanur S. Ramachandran and Semir Zeki. For a critical discussion of these see Hyman (2010).

periods and artistic persuasions as objects of study. Our main criticism is that this focus is mistaken for two reasons. First, to address the questions that neuroscience can reasonably be expected address, there is no need to focus on established artists. In fact, art school undergraduates would be no less valuable subjects. Second, the choice of historical figures as subjects forces Zaidel to extract her data from sketchy and incomplete historical records (incomplete in that they do not provide detailed information about the exact nature of brain damage suffered), which rarely, if ever, provide data that are fit for purpose. Studying 'ordinary' art students rather than established historical figures also makes this problem go away. And this is what we recommend should happen: rather than relying on gappy historical records, data should be gathered on living painters using state of the art technology, and these painters can be chosen arbitrarily since nothing depends on their standing in the art world.

1. Learning from brain damage in established artists

Zaidel investigates the neuropsychology of art by looking at subjects who have suffered brain damage and then exploring the artistic implications of this damage. To do so she has collected anecdotal and historical data about artists throughout history. Zaidel's research is extensive, and covers artists from all areas of art: the visual arts, literature, music composition and performance, and even film. Zaidel's goal is to find what common elements span these varied concepts of art, and ultimately to find out how our brains process, create, and recognize works from any field *as art*. However, Zaidel is not just interested in any old piece; her focus is works of art that have been produced by established artists, that is ones whose status as artists is confirmed both by professional art history as well as society at large and whose works remain recognized and hallowed as art throughout time and within recognized and prestigious establishments. In other words, in her study of the neurology of art, Zaidel is interested in studying artists and works that belong within the upper echelon of the artistic world: 'Art', with a capital A, as we shall call it.

In this chapter we aim to give an account of the method Zaidel uses to achieve her goal. For the sake of clarity, we will limit ourselves to those studies relating to visual artists, specifically painters. We will also limit these examples to painters who have suffered specific neurological damage. Despite the fact that Zaidel argues from a range of conditions, both physical (such as defective colour vision) and neurological (such as brain damage caused by strokes), we limit our attention to neurological conditions since these are more pertinent to the issue at stake and help us bring into focus her claims regarding the role of neurology in the study of art.

We will use two of Zaidel's examples to examine her method (this is for ease of presentation only; our points could just as well be made using other cases). The first is

the case of the Bulgarian painter Zlatio Boyadjiev (pp. 30–1).³ In 1951 Boyadjiev, then 48 years old, suffered a stroke within the left hemisphere of his brain. This left Boyadjiev with permanent right hand paralysis (his dominant hand) and mixed aphasia. After his stroke, Boyadjiev learned how to use his left hand to paint. After gaining proficient skill at this new task, he continued to produce work that was acclaimed by critics and exhibited in museums. His ability to realistically depict figures was unchanged. However, compared to the pre-stroke period his use of colours became less exuberant and less varied; the perspective was less convergent (with some pieces lacking depth altogether); a left–right mirror symmetry became typical of many of his compositions; and there were fewer figures present in each piece (creating a feeling of less overall complexity in the piece).

The second example is that of portrait painter William Utermohlen (pp. 43–4). From the age of 57 onwards, Utermohlen experienced a slow deterioration of global cognitive ability due to the onset of Alzheimer’s disease. An MRI scan revealed generalized cortical atrophy with no asymmetry. As the disease progressed, Zaidel notes that Utermohlen’s trademark use of realistic depictions tapered off, and a use of more abstract techniques ensued. Facial distortions were introduced into his portraits; perspective and depth were slowly lost over time; and although colour and form were still used, colour was left unblended and paint was applied with broad brush strokes which gave the paintings an air of patchiness. Eventually he gave up colour altogether and resorted to just sketching with a pencil.

The thrust of these examples (and many others in the book) is to correlate alterations in artistic style with alterations in the brain, and then draw a conclusion about which parts of the brain are involved in the production of artworks. In more detail, the method consists in the following steps:

1. Examine an artist’s work before damage.
2. Examine the same artist’s work after having suffered brain damage.
3. Identify aesthetically relevant aspects in the artist’s work that have changed.
4. Assume that these changes appear exclusively due to the brain damage.
5. Identify the nature of the brain damage suffered.
6. Conclude that the change in the brain is responsible for the change in artistic style and that therefore the part of the brain affected by the disease is responsible for the production of the pre-stroke style.

In Zaidel’s case studies the last stage is often left implicit, but the overall discussion clearly indicates that such an ‘attributive’ step is the aim of the exercise. In the first case, for instance, the conclusion clearly is that damage to the left hemisphere was responsible for the changes in Boyadjiev’s style. This, then, provides us with a possible clue as to

³ Unless indicated otherwise, page numbers in this chapter refer to Zaidel’s (2005).

where in the brain certain artistic capabilities are located. For want of a better term, we refer to this method as the *method of diachronic difference*, MDD for short.⁴

In this chapter, we aim to closely scrutinize Zaidel's method and argue that it suffers from serious problems. Our criticisms fall into two categories. The first is Zaidel's application of MDD. The second criticism focuses on the use of MDD as an adequate cornerstone for research of this nature. Here we will examine similar projects that have tried to use MDD in this fashion, and the problems that subsequently arose affecting their study. We will argue that Zaidel's study of brain damage in artists is not immune to these same problems. Finally, we propose an approach that avoids these pitfalls and may lead to a great wealth of interesting information regarding neurology, aesthetics, and art.

2. Aiming too high

The first problem we will deal with is Zaidel's use of MDD. Our problem becomes palpable as soon as we are pressed to lay out what the above examples really show. Upon closer inspection this seems to be rather little. The first example suggests that there is a connection between brain damage to the left hemisphere and the absence of perspective as well as an unimaginative use of colour; the second case suggests that there is a connection between abstraction and the malfunctioning of the cortex as a whole. Even if one grants that these findings generalize to other cases—which is by no means clear; in fact, Zaidel (p. 31) herself points out that the loss of the ability to produce perspectival representations is often associated with lesions of the right rather than the left hemisphere—they are too unspecific to give us serious hints about the connections between the workings of the brain and the production of art. Unless the brain damage is clearly located, we cannot draw useful conclusions about which parts or functional units of the brain are involved in the production of art, which undermines the very project of a neuropsychology of art. Locating a capability somewhere in the left hemisphere, or even the entire cortex, is not specific enough to provide serious clues about the neurological underpinnings of art. And this problem is by no means restricted to the two examples we have chosen. The book gathers together a large collection of cases, but only few, if any, receive detailed treatment. For the most part the descriptions are short and coarse.

By and large the brevity of the discussion is owed to the fact that historical records provide only limited information. Zaidel places great emphasis on investigating established artists—Otto Dix, Louis Corinth, and Willem de Kooning are but a few of the eminent artists discussed—and where she considers lesser(-known) artists she is careful to choose only those who are of some art-historic interest. The basic problem is that our knowledge about the conditions of these artists is, as a matter of fact, limited. For

⁴ This name echoes Mill's 'method of difference' (1843, Book 3, Ch. VIII, Sec. 2). We add 'diachronic' to make explicit that we are comparing properties at different times.

one, diagnostic techniques of the kind we have today were not available at the time and hence we often only have the crudest of descriptions of their neurological conditions; for another, medical records often contain only limited information about brain damage suffered by an artist. Although Zaidel is not to blame for the sketchiness of available historical records, there is a serious question about whether such records are able to form the basis of a research programme in neuroscience. They are not. Far more detailed knowledge of the neurological conditions of patients would have to be available in order to form a sound basis of an investigation of the functioning of the brain in the production of art. Current neuroscience has evolved to the point where we understand in great detail the structures of various regions of the brain and the interconnections between them. In fact, we now have a very detailed map of the brain telling us even how individual columns (a small bundle of neurons operating as a unit) work and in which functions of the brain they are involved. When studying specific brain damage to attribute artistic skills and practices to brain regions, we can only make useful claims through locating damage on this finely drawn chart. It is no longer reasonable to make significant claims with a loose description of the location of the relevant damage; locating damage in the right or the left hemisphere, or even in the visual cortex as a whole, no longer furthers our understanding of the functioning of the brain in general, or the neurology of art in particular.⁵

The focus on eminent historical figures should also be given up for another, and independent, reason. In order to see what this reason is, let us first give a more abstract characterization of MDD:

1. Examine a person's capability of doing X before brain damage.
2. Examine the same person's capability of doing X after the person has suffered brain damage.
3. Identify relevant differences in the person's capability of doing X before and after having suffered brain damage. Call this difference ΔX .
4. Assume that the brain damage suffered is the only cause of ΔX .
5. Identify the nature of the brain damage suffered. Call the difference between the brain's condition before and after brain damage ΔC .
6. Conclude that ΔC is the cause of ΔX .

This method suffers from all the well-known difficulties of inductive reasoning,⁶ but these do not concern us here; we assume that they can be circumvented successfully in

⁵ Section 3 of Zaidel's contribution to this book may be read as suggesting that the aim of the investigation is not so much correlating certain artistic skills (or loss thereof) with certain precisely circumscribed parts of the brain, but rather to show that artistic talent, skill, and creativity are diffusely located in the brain and do not have a particular centre. Even if this is the goal, the problems remain the same: our failure to precisely locate certain conditions does not imply that they have no precise location.

⁶ Let us mention but some: the inductive step of generalizing from the sample investigated to the entire reference class is notoriously problematic; the assumption that there are no confounding factors (premise 4) is highly problematic in the current context and has to be accepted as an article of faith; and neurological data

the relevant cases because our criticism of Zaidel's approach is orthogonal to concerns about induction. The point we want to emphasize is that there are at least two conditions on *X* for a legitimate application of MDD. The first condition we call the *requirement of specificity*. It must be clear what it means to do *X*—there cannot be any ambiguities about the concept itself—and we need operational criteria to unequivocally decide on any given occasion whether a person is doing *X* or not. If there are ambiguities about *X*, or if in an experiment we cannot decide whether the person really does *X*, then looking for the neurological foundations of *X* is an ill-defined problem.

The second condition we call the *requirement of individuality*. This condition requires that whether the person under investigation does *X* (or fails to do so) only depends on that person's mental state and not on other persons' mental states.⁷ Let us illustrate this with an example. Whether John feels pain depends only on the state of his brain (and possibly his nervous system); other people's mental activities play no part in John's being in pain. By contrast, whether John is behaving politely is not only a function of his brain state. Behaving politely depends on other people's preferences and on social norms which are beyond the control of the individual. For instance, while in traditional Islamic cultures it is impolite to eat with your left hand, it is impolite in most Western societies not to use both hands to eat. If John eats with both hands, he is behaving politely or impolitely depending on the cultural beliefs of people around him. But John's action is the exactly the same in either case—he is eating with both hands—and hence the mental actions that guide his behaviour must be the same too.⁸ For this reason, the concept of behaving politely cannot be studied from a neuroscience point of view, because the subject matter of such an investigation is an individual's brain state. It is a necessary condition for activity *X* to be open to neurological investigation that the execution of *X* be a function only of the individual's state of mind, and not of what happens in the person's social environment.

The *X* at stake in Zaidel's investigation is 'producing a work of Art'. Unfortunately producing a work of Art clearly fails to satisfy the first condition, and, at least given current views concerning the nature of art, it does not satisfy the second condition either. As regards the first condition, it is a matter of fact that there is no agreement

are nearly always messy, even when experiments are conducted under strict control, and hence the neurological condition itself is in part speculative.

⁷ Doing *X* may also depend on facts about the world, as long as they do not depend on other persons' beliefs and desires. For instance, if we assume externalism to be true, then it must be the case that there are objects in the world for a person to have object-involving thoughts. This dependency on facts is harmless in the context of the current discussion. What is at issue here is dependence on other persons' mental states.

⁸ One can, of course, investigate a person's disposition or willingness to respect a given set of social rules or the execution of a particular action that is in line with the rules (e.g., keeping at a certain distance when talking to someone). But this is not the same as investigating behaving politely *per se*: someone can be willing to respect rules and yet fail to do so, or keeping a certain distance to your interlocutor may be polite in one context but not in another one.

about what defines a work of Art.⁹ Classical definitions which construe art as imitation or representation (Plato), a medium of transmission of feelings (Tolstoy), intuitive expression (Croce), or significant form (Bell) are believed by many to be seriously defective. More recent approaches include functionalism (Beardsley), proceduralism (Dickie), or approaches emphasizing historical reflexivity (Danto). For our current concerns the relative merits of these approaches are immaterial. The salient point is that there is a plethora of different schools of thought and that producing a piece of art means something different to each of them. Hence there is no unanimous view of how a piece of art is to be identified, which undermines any attempt to identify the neurological basis of the production of art.

Assume now, for the sake of argument, that this dispute could be resolved. Of course we can only speculate about how the conflict will be resolved, but at least given the state of play in the current discussion, this resolution is unlikely to be in line with MDD. Most contemporary definitions of art in one way or another explicitly appeal to social practices and the role of institutions (this is explicit, for instance, in Dickie's institutional theory) in order to define art: what turns an artefact into a work of art is neither a particular property of the object itself nor a specific characteristic of its process of production, but rather the role it plays in certain social practices. This suggests that art is more like behaving politely than like feeling pain, and hence does not lend itself to a neurological investigation at all since it fails on account of the requirement of individuality.

Although Zaidel does not discuss this point, there is implicit acknowledgement of it in her discussion, since any reference to artistic value (or any of its cognates) are conspicuously absent from her discussion of *actual cases* (as opposed to programmatic statements). The discussion of specific bodies of work focuses on the use of perspective, the choice of colour, the level of abstraction, the curviness of lines, the smoothness of boundaries, the thickness of layers of paint, the characteristics of brush strokes, the presence of symmetries, the distortion of objects, the choice of motives, and the like. What Zaidel is really examining is how neurological damage affects the specific representations that are being processed by the artists, and how in turn the damage affects their ability to create their brushstrokes and accurately represent what they desire on their canvas. These (and related) aspects no doubt play a role in the appreciation of an artwork, but they are not specific to Art with capital A and it seems unnecessary to restrict attention to these 'upper echelon artists' if all that is at stake are specific technical aspects of their work. To discuss, say, the use of perspective we don't have to look at geniuses of the calibre of van Gogh; the canvases produced by Sunday painters, primary school teachers, distraught managers seeking relaxation in painting, and commercial painters producing pieces that are sold on tourist markets are not less valuable as 'data points'. Once we limit our attention to a specific aspect of pictorial

⁹ For a survey of the various positions held in this debate see Davies (2005).

representation (like the use of perspective), it just doesn't matter any more whether the pieces we look at count as Art, or whether they are merely canvases produced for any number of other reasons. In other words, the fact that the subjects studied were established artists seems to play no role at all in the conclusions that we are supposed to draw from the cases!

This is not merely an academic point. In fact, its practical implications can hardly be overstated. It is the focus on eminent figures that forces Zaidel to use uncontrolled and sketchy data extracted from artists' histories as a basis for claims concerning a low level neurological story for a change in aesthetic perception of an artist's work, and we have argued above that this is detrimental to her research programme. Once we recognize that there is no necessity to focus on established artists and that the neurological studies could just as well be carried out on the participants in the painting class at the local community college, we free the investigation from the straitjacket of historical records and open up the possibility of gathering detailed and reliable data using cutting edge technology. And this is exactly what we think should happen. Rather than keep relying on notoriously gappy historical records, data should be gathered on living painters, irrespective of their standing (or even participation) in the art world. In fact, this is the only way forward if we really want to understand the neurological basis of how visual representation works.

In sum, the *X* in the investigation should not be 'producing a work of Art', but a particular skill like 'producing a perspectival ink drawing' or 'producing a representation that is truthful with respect to colour'. This shift not only makes the problems with specificity and individuality go away (and hence allows us to use MDD); it also gives investigators the possibility to choose contemporary subjects who have no recognition in the art world (and who may not even aspire to so being recognized) on which data can be generated using cutting edge technology. It is such data rather than sketchy historical records that should be used in a study of the neurology of art.

3. Qualms about top-down approaches

Before outlining in more detail (in Section 5) what we regard as a more promising approach to neuroaesthetics, we would like to have closer look at a parallel case: the debate over the neurological underpinnings of the cognitive notion of a visual representation, which also suggests that the focus on Art is a red herring.

Let us start by setting the stage. There are two possible approaches to cognitive neuroscience: the traditional top-down approach, and the more controversial bottom-up approach, where, in this context, by 'top' we mean the level of mental phenomena as we experience them (seeing a house, feeling sad, wanting to sleep, etc.) and 'down' refers to the level of brain states (patterns of neuron firings, etc.). The difference between the two is best illustrated through a slightly revised version of Dennett's bridge analogy, in which he describes top-down cognitive neuroscience as 'reverse engineering' (Dennett 1998: 255). Consider someone who has no knowledge of engineering, but

who takes an interest in the workings of a particular bridge. This person could start by looking at the bridge and identifying certain functional units like the deck, the pillars, the pillar foundations, the anchorage blocks, the suspenders, and the suspension cables. There are certain ways in which these fit together and understanding how the bridge works involves understanding how these parts fit together: the pillar foundations carry the pillars, the suspensions are fixed at the top of the pillars, the suspension cables are connected to the suspensions and they carry the platform. These macro concepts provide a good description of the bridge and the person can now tell what each of the bridge pieces do, which parts are integral to structural integrity, etc. But this does not yet satisfy the person; she also wants to know why and how the parts can perform their function. So she starts looking at the materials used and the way they are connected: the foundation blocks are made from concrete, the pillars are metal bars riveted together in particular way, the deck is a combination of stone and steel, etc. The more she knows about the constitution of each part, the better she understands how these parts work and why the entire bridge holds together. In sum, the person *first* develops a well worked out macro-theory of the bridge, and *then* asks what kind of micro-constitution allows the various parts identified in the macro-theory to work in the way they do.

This approach contrasts with a bottom-up method. This method does not require a well worked out macro-theory in order to start investigating the micro level; in fact, it tells us not to have one—and if we happen to have one to put it to the side for the time being. The way forward, on this view, is not to look for the micro underpinnings of a finished macro-theory, but to start tampering with the micro structure directly and observe how changes at the micro level bear on broad issues we are interested in. In the case of the bridge, for example, the bottom-up method denies that we *first* have to describe the bridge in macro terms like ‘pillar’ and ‘foundation’ and *then* ask what the constitution of a pillar and a foundation is; instead it invites us to ponder directly what effect it would have on the stability of the bridge if we replaced, say, rivets by screws, used wood instead of steel in certain places, etc. This would lead to an understanding of the effect of micro changes on the functioning of the bridge, but without presupposing a particular macro conceptualization of it.

Starting the investigation with the notion of Art, and a view about what art is and how it has to be understood, amounts to adopting a top-down approach: it presupposes a clear understanding of the phenomenon under investigation and then asks how the different elements of our understanding of Art are realized at the neurological level.¹⁰ It is easy to see why a top-down approach seems appealing for Zaidel’s project. For one, the top-down method seems generally tidier and more systematic than the somewhat messy and anarchic bottom-up approach. For another, especially in a field

¹⁰ In fact, she is specifically evoking common cognitive science methodology and not the straight neurological methodology the title of her work implies.

in which there already exists a body of knowledge it would seem to be natural to take this knowledge into account.

Is this the right choice? This is not the first time that a problem of this nature has been presented to the philosophical community. A similar issue has been well discussed throughout the philosophical and cognitive neuroscience literature about the long-standing debate between Kosslyn and Pylyshyn concerning the cognitive notion of visual representation (VR). We will now discuss this case in some detail, which brings us to the conclusion that, first impressions notwithstanding, a bottom-up methodology seems to be more appropriate for neuroaesthetics than a top-down approach.

VR is defined in this discussion as what it involves for someone to imagine something (Pylyshyn 2003); that is, what it involves for someone to have an image of something in her head. As imagining something is a cognitive concept, evoking such things as images and imagining and thought (not traditionally language used in strict neuropsychology), the cognitivists, most recently encouraged by Pylyshyn, were the first to try and create a 'cognitive map' or theory that would help explain this common experience. This theory of VR would outline precisely what is happening when we imagine something visually; it would, as it were, provide us with the anatomy of visual representation, detailing which kinds of cognitive capacities are involved in imagery, how they fit together, and how they relate to the capacities involved in other tasks such as verbal expression. To come back to Dennett's analogy, this theory of VR would be like the description of the bridge in terms of concepts like 'pillar' and 'platform'. Such a theory is regarded by cognitivists as the indispensable starting point of every investigation into how VR functions at the level of the brain—without such a macro-theory, so the cognitivists think, one cannot even begin addressing this issue.

However, Pylyshyn quickly ran into a problem. Due to cognitive theory being so dependent on semantic content, there seemed to be no real way to create one static theory that would account for all forms of VR. For instance, if a cognitivist decided to examine the cognitive aspects that were evoked when you imagined a man walking along the beach at sunset, you could be said to evoking cognitive semantic concepts such as water, sand, outdoors, humans (man), sunset, etc. But if you were to imagine a fresh plate of sheet metal being produced by a machine in a factory, you would evoke semantic concepts that involve industrial centres, machinery, fire, ore, coal, etc. A cognitive theory of VR would consist of a general schema of which both concrete cases would be an instance (in pretty much the same way in which the trajectories of planets orbiting around the sun and of heavy objects falling from towers are instances of the Newtonian laws of motion). Unfortunately it turned out that there seems to be no way to collate these two very distinct and different visual representations into any one unified cognitive theory. In fact, it turns out that it is not even clear what format such a theory should have—an issue that was the source of a rousing debate: in even considering the problem of visual representation, Pylyshyn states that you are automatically evoking the argument that 'thought' is pictorial (as opposed to linguistic) in format. However, understanding the 'format' of 'thought' does not seem possible until

you have a better idea of what you are looking at, and as the cognitivists seem unable to agree what VR really is on a cognitive level, a unified theory of VR seems unattainable.

This has severe consequences. Due to this inability to assume any format, no upper ground was solidified as a basis for a top-down cognitive to neurological explanation of the cognitive/phenomenal aspects of VR. For Pylyshyn, if there is no solid 'upper ground' theory of VR, then a top-down neurological explanation of VR simply does not exist and not only has the project of a neurological reduction failed, but there seems no viable starting point for researching and understanding the complexities of VR at all.

Art seems to be much like VR in that it is a multilayered, highly complex and context-sensitive concept, and hence it is unlikely that there will ever be a unifying theory about Art that would provide the starting point for a neurological top-down reduction.¹¹ As a consequence, Pylyshyn's conclusion concerning the prospects of a neurological reduction of VR seems to carry over to Art unscathed: it can't be done!

This leaves two options: biting the bullet and regarding reduction as impossible, or choosing an alternative methodology. Neurologist Kosslyn opted for the latter. He decided to forgo all cognitive definitions, explanations, and assumptions of VR and worked to develop a theory of VR from the bottom-up perspective. In essence this means that research is guided not by a high-level theory but by practical issues; more specifically, a subject is confronted with a particular task and then it is observed which parts of the brain are involved in tackling that task. This can be done either by observing the subject's brain in an MRI or CT scanning device while she deals with the task, or by observing subjects suffering from a particular brain condition and then comparing them with healthy subjects. Consider, for example, Maguire *et al.*'s (2006) study of navigation. In order to find out what role the hippocampus plays in our ability to navigate in spaces that we have become acquainted with a long time ago, a taxi driver with bilateral hippocampal lesion as well as several control taxi drivers were asked to navigate in an interactive virtual reality simulation of central London. The investigation found that the hippocampus is not required for general orientation relying on main roads, but that it plays a role when it comes to navigating in areas off the main roads. The salient point here is that no cognitive theory of navigation—let alone a complete and accurate one—is presupposed; the investigation bypasses this step and simply asks what is and what is not important when a subject is actively participating in the activity you wish to study.

Through use of this method, Kosslyn discovered that the reason why VR intuitively seems to be so similar to vision is that the areas of the brain used for VR share

¹¹ In her review of Zaidel's book, Franz (2006) seems sympathetic to the idea that Zaidel's research question does not seem well defined, which could be one of the reasons why it is so hard for Zaidel to develop an easy to state research programme. Brown (2006) also feels that the cognitive aspects of Zaidel's approach have not been given enough justice, which could be another reason why a top-down methodology does not seem to fit her current project.

two-thirds of the cortical space of normal vision. In fact, visual representations are actually represented within the visual cortex in a way very similar to how the visual cortex processes normal sight: the difference between VR and normal sight seems simply to be the origin of the information processed. In the case of normal vision, the information comes directly through our eyes from the real world. In the case of VR, the information is taken from our memory systems and then collated throughout the visual cortex (Kosslyn *et al.* 2003). Despite not having a theory of visual representation, Kosslyn was, through bottom-up exploration of the issue, able to find out how it worked and develop a working theory that helps us understand more about this phenomenon.

Since Art is much like VR, this case shows the dangers involved in subscribing to a top-down method: we may be led to nihilism where interesting insights could be gained by using other methods because a misplaced focus on Art (and the high level theory attached to it) comes to stand in the way of progress that could be made in our understanding of neurological processes that are involved in aesthetic judgements.

4. The neuropsychology of art revised

The arguments in Sections 3 and 4 converge towards the same point: the focus on Art is a dead end for neuroaesthetics. Progress is more likely to be made if the investigation, first, focuses on 'local' themes rather than Art; second, uses controlled laboratory environments to generate data rather than trying to extract information from historical records; and third also pays attention to aspects of the appreciation of art. Let us address these points one at a time, beginning with the second.

As we have pointed out in Section 3, the focus on established artists is both unnecessary and detrimental to the project of neuroaesthetics. The project of gathering information about eminent historical figures from historical records is doomed to failure from the start because these records do not provide information which meets the needs of modern neuroscience as regards precision, detail, and specificity. Instead neuroscientists interested in aesthetics should investigate the workings of the brain of persons engaging in art-related practises no matter what their standing (or even participation) in the art world and observe their brains while they do so using all means available in a modern laboratory (for instance MRI, fMRI, PET, EEG, and MEG). This is because relevant issues can be investigated using any subjects with some artistic talent or interest: one does not need the great artists; simple art school students would do. The data collected would be far stronger and more controlled than those gleaned from partial and possibly misleading medical records, and they would provide a more solid and reliable basis for an understanding of the brain processes involved in art-related activities.

Studying living subjects would have two further advantages. First, although attributing changes in the production of art to a specific change in the brain is a first step, it is no more than that. A complete investigation on the neuropsychology of art would

need to include a way of obtaining information regarding brain activity throughout all stages of the artistic process. It would be helpful to understand how each region is involved in first conception, then first sketches, then filling out the colour and the idea all the way to the final draft, and then work completion. It may even be useful to examine the difference in response to the artist looking at the final work. This would give us data representing a holistic approach to the neuropsychology of art, and a rough sketch as to how the system as a whole is structured. If a rough sketch of the whole system can be obtained, then any information we can gather regarding specific system damage can be discussed in its context.

Second, studying broken systems is a difficult job that neurologists excel at: one could claim that the majority of modern neuroscience is based on breaking and examining the brain.¹² However, other methods have become available in recent years: we now have the technology and knowledge to bypass studying broken systems and instead focus on watching the brain function correctly. That is, we no longer need to break aspects of the brain to find out what they do, we can watch the brain work as a whole system uninterrupted while real artists work using the above-mentioned technologies.

Let us now return to the first of the above issues: what questions should be addressed by neuroaesthetics. There seem to be at least three kinds of questions that neuroscience could meaningfully address (needless to say, this list is in no way exhaustive). The first family of questions concerns technical aspects in the execution of a piece of visual art. Neuroscience is well equipped to study the effects of the use of perspective, colour, shading, symmetries, distortions, thickness of lines, and many more, for instance by observing what happens in a person's brain when they produce or see, say, a drawing using one-point perspective and comparing it to the brain of people who do (or are not able to) see or produce such images. Indeed, there already is a wealth of information available about some of these issues. For instance, the functioning of the visual cortex and the eye itself have been studied extensively, which can lead to studies that help us understand the physical (and emotional) response we feel when looking at certain colours, or certain shades, or certain oblique lines.

Some studies help us make leeway on these difficult topics. For instance, extensive studies have already been carried out on how our brain uses ganglion cells to see and interpret colour, light, and shading (Livingstone and Hubel 1988), how our brain can see and distinguish between different types of colour (Martin 1998), and how we physiologically pick up movement and lines within our visual field (Merrigan and Maunsell 1993). Understanding vision in this way could well be the first step to understanding how these specific images or specific colours/tones/shades affect other neural centres engaged in memory or emotion, ultimately helping us understand the neurology of how art affects us.

¹² For a closer examination on the problems presented by studying broken systems, see Glymour (1994).

The second group of issues centres on the phenomenon of creative impulse. One person looks at a sunset and simply wants to enjoy its beauty while another person sees the same sunset and is compelled to recreate it in drawing or painting, or capture the emotion that it evokes in some way. Is there a difference in the brain function of these two persons? Studying the difference between what someone with creative impulse does with their perception in comparison with someone who does not have any creative drive at all may give us a starting point to understanding the creation of art.¹³

The third group of questions focuses on the appreciation of art. Although the emphasis in studying art usually is on the artist, the audience plays an important role too—pieces of art are produced with the intention that they be appreciated *as such*. This can lead to the study of those who call themselves ‘art appreciators’. What goes on in the brain of someone who appreciates, say, a drawing *as a piece of art* as opposed to regarding it simply as a provider of information about the spatial arrangement of certain objects? One could study this difference, for instance, by investigating the difference between a subject observing a picture just as a picture (for instance a photograph in a newspaper article) and compare this to the neurological effect of them observing that same photograph as a work of art in an art exhibition. This in itself would give us some clue as to how the brain processes and recognizes art in different forms, and allow us to examine different perspectives of someone experiencing art.

None of these three areas of study amounts, in itself, to the study of art, but it is plausible that studies such as these will give us vital information that could, at some later point, also contribute to the neurological study of the more cognitive, phenomenological, and aesthetic aspects of art. Finding the necessary neurological conditions for artistic skill (technique), creative impulse, recognition, and appreciation may very well be the key to truly understanding the neuroscience of art, and potentially also help us understand how our capability to produce art relates to (and depends on) our general cognitive capacities for everyday survival activities such as general motor skills as well as auditory, spatial, and visual mechanisms.¹⁴

References

- Brown, Steven (2006), ‘Bringing Science to Art’. *Trends in Cognitive Sciences*, 10/8: 356–7.
 Davies, Stephen (2005). ‘Definitions of Art’, in Berys Gaut and Dominic Mciver Lopez (eds.), *The Routledge Companion to Aesthetics*. 2nd edn. London: Routledge, pp. 227–40.

¹³ Zaidel seems to agree with this at least. In her introductory discussion of what she considers to be artistic she includes anything that can be seen as using a creative impulse as being artistic (even to some degree science). Studying this creative urge in its own right could be seen as a worthwhile starting point.

¹⁴ We would like to thank Chris Bertram, Martin Frigg, Andrew Goldfinch, Peter Goldie, Julian Reiss, and Elisabeth Schellekens for helpful comments on earlier drafts. Thanks also go to Dahlia Zaidel for her feedback and her forbearance.

- Dennett, Daniel (1998). 'Cognitive Science as Reverse Engineering: Several Meanings of "Top-Down" and "Bottom-Up"', in *Brainchildren: Essays on Design and Mind*. Boston: MIT Press, pp. 249–60.
- Franz, Liz (2006). 'Arts and Minds'. *Lancet* (April): 302.
- Glymour, Clark (1994). 'On the Methods of Cognitive Neuropsychology'. *British Journal for the Philosophy of Science*, 45: 815–35.
- Hyman, John (2010). 'Art and Neuroscience', in Roman Frigg and Matthew Hunter (eds.), *Beyond Mimesis and Convention: Representation in Art and Science*. New York: Springer, pp. 245–62.
- Kosslyn, Stephan M., Giorgio Ganis, and William L. Thompson (2003). 'Mental Imagery: Against the Nihilistic Hypothesis'. *Trends in Cognitive Sciences*, 7/3: 109–11.
- Livingstone, Margaret and David Hubel (1988). 'Segregation of Form, Colour, Movement, and Depth: Anatomy, Physiology, and Perception'. *Science*, 240: 740–9.
- Maguire, Eleanor, Rory Nannery, and Hugo Spiers (2006) 'Navigation around London by a Taxi Driver with Bilateral Hippocampal Lesions'. *Brain*, 129: 2894–2907.
- Martin, Paul R. (1998). 'Colour Processing in the Primate Retina: Recent Progress'. *Journal of Physiology*, 513: 631–8.
- Merrigan, William H. and John H. R. Maunsell (1993). 'How Parallel are the Primate Visual Pathways?' *Annual Review of Neuroscience*, 16: 369–402.
- Mill, John Stuart ([1843] 1973). *A System of Logic: Ratiocinative and Inductive*, ed. John M. Robson. Toronto: Toronto University Press.
- Pylyshyn, Zenon (2003). 'Return of the Mental Image: Are There Really Pictures in the Brain?' *Trends in Cognitive Sciences*, 7/3: 113–18.
- Zaidel, Dahlia W. (2005). *Neuropsychology of Art: Neurological, Cognitive, and Evolutionary Perspectives*. New York and Hove, UK: Psychology Press.