LOOKING FOR BEAUTY IN THE BRAIN

ETHAN WEED

The emerging research area of neuroaesthetics has provoked a good deal of discussion. Although it seems reasonable to describe the experience of aesthetic enjoyment as a mental event, and it also seems reasonable to claim that mental states must be related to brain states, the search for specific brain states that correlate with aesthetic enjoyment is tricky, despite the many recent advances in brain-imaging technology. Correlating the aesthetic experience with specific brain states involves defining the aesthetic experience. By applying a model from the world of empirical consciousness research to three neuroaesthetic experiments, I show that each of these studies approaches the object of study, the aesthetic experience, from a different perspective. By employing a framework to make explicit the sometimes implicit assumptions involved in neuroaesthetic research, I hope to open a new avenue for the continuation of an already fascinating discussion.

Auf der Suche nach der Schönheit im Gehirn
Das neu aufkommende Forschungsgebiet der Neuroästhetik hat zahlreiche Diskussionen ausgelöst. Wenngleich es einleuchtet, dass man die Erfahrung ästhetischen Genusses als mentalen Vorgang beschreiben kann, und die Forderung, dass mentale Zustände zu neuronalen Zuständen im Gehirn in Beziehung gesetzt werden müssen, ebenso selbstverständlich erscheint, ist die Suche nach mit ästhetischem Genuss verbundenen neuronalen Zuständen trotz der zuletzt großen Fortschritte auf dem Gebiet des sog. Brain-Imaging schwierig. Das In-Beziehung-Setzen von ästhetischer Erfahrung mit bestimmten neuronalen Zuständen verlangt auch eine Definition dessen, was ästhetische Erfahrung ist. Indem ich ein Modell aus dem Bereich der empirischen Bewusstseinsforschung auf drei neuroästhetische Experimente anwende, zeige ich, dass sich jede dieser Studien ihrem Objekt, d.i. der ästhetischen Erfahrung, von einem anderen Standpunkt aus nähert. Indem ich einen begrifflichen Rahmen nutze, der geeignet ist, die oft impliziten Annahmen neuroästhetischer Forschung explizit zu machen, hoffe ich, einen neuen Weg zu weisen, wie eine faszinierende, laufende Diskussion fortgesetzt werden kann.

I. INTRODUCTION: LOOKING FOR BEAUTY IN THE BRAIN

In an excellent introduction to the new field of neuroaesthetics (unfortunately only currently available in Danish), Martin Skov argues that a full understanding of the aesthetic experience must include a neurological component.[1] The argument is straightforward: an aesthetic experience cannot be explained by a description of the aesthetic object alone. The aesthetic *experience* of an object involves an interaction between the object and a human mind, and since (depending on one's philosophical leanings) the mind must be either identical to or at least intimately involved with brain-states, this implies a description of

[1] Skov, Martin (2005). 'Hvad er neuroæstetik?' *Kritik*, 38 (174), p. 1.

brain processes. Skov sketches the process of the aesthetic experience as a flow-chart in which form (specific qualitative aspects of the object) is perceived and processed by the machinery of human cognition, leading to some sort of phenomenal experience, which he calls 'effect':

$$\text{FORM} \rightarrow \text{COGNITION} \rightarrow \text{EFFECT}$$

Thus, understanding what makes a painting or a piece of music beautiful entails an understanding of what makes us *think* the painting or the music is beautiful, and that ultimately implies an understanding of underlying neural processes. What is it that happens in the brain that makes aesthetic experiences different from normal experiences? Aesthetic experience is based on perception; we see a painting or a landscape, hear music or the wind, and so forth. But it is a perceptual experience with some sort of value marker attached to it – it is a *beautiful* painting, piece of music, and so on. What in the brain adds this extra element? This is the question neuroaestheticians would like to answer.

Of course, looking for beauty in the brain isn't easy. As Skov points out, the philosophical question of what we mean by aesthetics at all may seem insurmountable in itself. Add to this the general paucity of our current understanding of the functioning of the brain, and the task of nailing such an elusive target as aesthetic experience to specific brain structures seems nearly hopeless. Then again, as neuroaesthetic theorists like Zeki[2] and Ramachandran and Hirstein[3] hasten to point out, the field is still in its infancy. In addition, non-invasive brain-imaging techniques such as PET (Positron Emission Tomography) and fMRI (functional Magnetic Resonance Imaging) now make it possible to measure aspects of the neural activity of healthy subjects performing cognitive tasks. The remarkable spatial resolution that these techniques provide allows the researcher to make at least informed guesses about the brain structures involved in cognitive work, although the results from brain-imaging studies are rarely as clear-cut as one would like. In the following pages, I will consider some of the methodological and theoretical questions involved in designing an experiment that aims at finding a neural correlate of the aesthetic experience. The majority of studies in the field have to do (perhaps for practical reasons) with aesthetic *reception*, and not with artistic *creation*. It is therefore worth mentioning, in passing, that a pioneering study by Solso[4] demonstrates the possibility of using

---

2 Zeki, Semir (2001). 'Closet Reductionists.' *Trends in Cognitive Science*, 5 (2), pp. 45–46.

3 Ramachandran, Vilayanur S., & Hirstein, William (1999). 'The Science of Art: A Neurological Theory of Aesthetic Experience.' *Journal of Consciousness Studies*, 6 (6–7), pp. 15–51.

4 Solso, Robert (2000). 'The Cognitive Neuroscience of Art: A Preliminary fMRI Observation.' *Journal of Consciousness Studies*, 7 (8–9), pp. 75–85.

fMRI to study the brain activity of an artist at work. In the present paper, however, I will only be looking at studies that investigate aesthetic reception. I will examine what I will call the 'linking problem', and consider in particular the ways in which the researcher measures the aesthetic experience.

II. NEURAL CORRELATES AND THE LINKING PROBLEM

In the ideal neuroaesthetic experiment, we would put the subject in a scanner, elicit an aesthetic experience in the subject by exposing her to selected stimuli, then elicit a non-aesthetic experience by exposing her to other, similar stimuli, and compare the resulting brain-images in order to identify the specific brain regions involved in the aesthetic experience. Although this is the general template for a neuroaesthetic experiment, it is, of course, a fantasy. There are a host of practical and theoretical problems that confound this seemingly simple design. I will not be able to deal with the majority of these issues here,[5] but will limit myself to what Overgaard[6] has called the 'linking problem'. The link that Overgaard refers to is the link between what the researcher would like to measure, and what she actually measures.

Overgaard's treatment of the linking problem originates in the context of empirical consciousness research. Briefly, the majority of research on the neural correlates of consciousness (NCCs) aims at finding the specific brain-states that correlate with specific contents of consciousness. For instance, if I look at a butterfly, the state of activity that my brain is in at the moment of seeing the butterfly must in some concrete way be different from the state that my brain is in when perceiving a cat. These brain states are the correlates of *content consciousness.*[7] Overgaard succinctly summarizes the definition of a content-NCC: 'A neural system is a neural correlate of content of consciousness [*sic*] if it directly correlates with a state of having one particular conscious experience and if it does not correlate with every other conscious or any unconscious state.'[8] Although Overgaard accepts Chalmers' general definition of a content-

---

5 See Savoy, however, for a brief but pointed overview of some of the practical problems involved in functional brain scanning studies, Savoy, Robert L. (2005). 'Experimental Design in Brain Activation MRI: Cautionary Tales.' *Brain Research Bulletin*, 67, pp. 361–67, and Uttal for an in-depth and critical consideration of a wide range of problems in the use of fMRI techniques in cognitive neuroscience, Uttal, William R. (2001). *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain.* Cambridge, MA, and London: Massachusetts Institute of Technology.

6 Overgaard, Morten (2003). 'On the Theoretical and Methodological Foundations for a Science of Consciousness.' *Bulletin fra Forum for Antropologisk Psykologi*, 13, p. 28.

7 See Chalmers, David J. (2005). 'What Is a Neural Correlate of Consciousness?' Retrieved August 22, 2005, from http://consc.net/papers/ncc2.html.

8 Overgaard (2003), 'On the Theoretical and Methodological Foundations', pp. 26–27.

NCC, he has a specific criticism of the assumption of 'direct correlation'. What the researcher would like to do, he points out, is to correlate a specific conscious state, such as the perception of a butterfly, with a specific brain state. However, the experimenter can only access the brain state indirectly, through some sort of index. In the case of fMRI, which is likely to become the most common tool in neuroaesthetic research, the brain images which are taken to be a measure of the brain state are a representation of relative changes in oxygen-levels associated with increased blood-flow in the brain, which is assumed to be a response to an increase in the metabolism of neurons involved in a specific cognitive task. But of course it is not only the parts of the brain that are specifically related to any one cognitive task or mental state that are active at the moment the task is performed. The brain is constantly active in all sorts of ways, many of them not specifically cognitive. So in order to isolate the aspects of brain activity that are relevant to the experimental task, researchers are forced to compare two data-sets: one in which the subject performs the required activity, and another in which she either performs a different task, or does nothing at all. The control data-set is subtracted from the task data-set, so that the resulting image is generated based on the difference between the two. The fMRI technique is thus 'currently limited to seeing the *differences* between brain states'.[9]

To make matters worse, the researcher has no direct access to the conscious state that she is investigating either. Although we all have introspective access to our own consciousness, a third-person researcher has only indirect measurements of the conscious states of her subjects, that is, some form of report. The following model illustrates this point:[10]

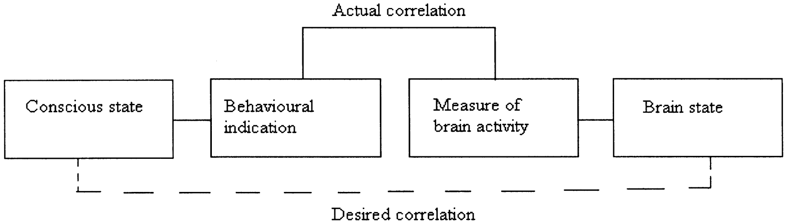

Figure 1

Although Overgaard's model is meant as a description of the current state of affairs within empirical consciousness research, I believe it also provides an

9 Savoy (2005), 'Experimental Design in Brain Activation MRI', p. 363.

10 Adapted from Overgaard (2003), 'On the Theoretical and Methodological Foundations', p. 28.

interesting framework for examining and evaluating certain aspects of some current research in neuroaesthetics. This is the task to which I will now turn.

III. KAWABATA AND ZEKI: NEURAL CORRELATES OF BEAUTY

In an article from 2004, Kawabata and Zeki[11] tackle the question of beauty in the brain head-on. Subjects were first asked to rate a large quantity of paintings from four different categories (landscape, portrait, abstract, and still life) as either beautiful, ugly, or neutral. 300 paintings from each category were viewed, giving a total of 1200 paintings. Of these, paintings classified as beautiful, ugly, and neutral from each genre were selected to be viewed in the scanner, with a total of 192 paintings viewed during scanning. The selection of paintings that subjects viewed in the scanner was based on each individual subject's prior categorization of the paintings. As the subjects viewed the paintings in the scanner, they rated them again as beautiful, ugly, or neutral. The scans were then classified according to the button-press rating that subjects made while in the scanner, allowing Kawabata and Zeki to contrast scans of subjects viewing ugly and beautiful stimuli, using scans of subjects viewing neutral stimuli as a baseline.

In terms of our linking model, Kawabata and Zeki's study looks like this:

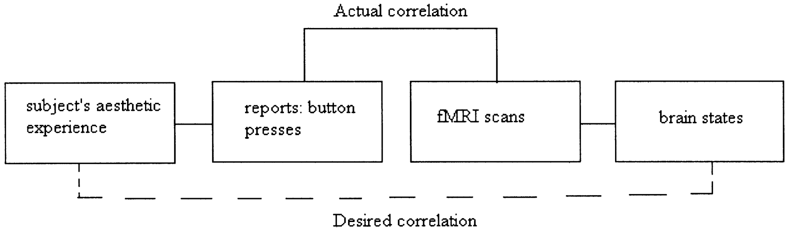

Figure 2

What is revealed in this diagram of the experiment? Although all the usual caveats about the relation between *actual* brain states and the images generated by the processing of the MR data apply, this will of course be the case in all such studies. What is more interesting for our purposes is the link between the reports that the subject makes by means of button presses and the actual experience that the subject has had while in the scanner. The authors would like to investigate the effects of beautiful objects on the brain and, recognizing that different people have different tastes, they have avoided the problem of choosing an object to use as the 'beautiful' stimulus by allowing the

[11] Kawabata, Hideaki, & Zeki, Semir (2004). 'Neural Correlates of Beauty.' *Journal of Neurophysiology*, 91, pp. 1699–705.

subjects to choose their own stimuli from a large selection. Nonetheless, the method of reporting, button presses, begs the question of the nature of the experience of the subject while in the scanner. All that we know about the subjects' experience is reflected in their selection of button one, two, or three to indicate the degree of their aesthetic experience or lack thereof – and notice that it is precisely the subjects' *experiences* that we are interested in. We would like to know what, in *neural terms,* their experience of beauty is. And the measure of beauty that is correlated with the scan results in Kawabata and Zeki's study is a button-press report for each of 192 paintings, where the average response time per painting was just under a second.

IV. RAMACHANDRAN AND HIRSTEIN: A PROPOSED EXPERIMENT

In their controversial article 'The Science of Art',[12] Ramachandran and Hirstein propose a psychophysical experiment that could be carried out in order to test some of their specific hypotheses about the relation between neural activity and the aesthetic experience. To my knowledge, they have not actually carried out the test, and indeed they mention that it would be 'laborious'. Yet, since it is more design than results that interest us here, I believe it will be illuminating to apply our linking model to Ramachandran and Hirstein's proposed experiment.

The authors posit eight ways in which artists can exploit aspects of normal perception in order to elicit certain effects in the viewers of their works. For instance, they write:

> The third important principle […] is the need to *isolate* a single visual modality before you amplify the signal in that modality. For instance, this is why an outline drawing or a sketch is more effective as 'art' than a full colour photograph. This seems initially counterintuitive since one would expect that the richer the cues available in the object the stronger the recognition signal and associated limbic activation. This apparent objection can be overcome, however, once one realizes that there are obvious constraints on the *allocation of attentional resources* to different visual modules. Isolating a single area […] allows one to direct attention more effectively to this one source of information, thereby allowing you to notice the 'enhancements' introduced by the artist.[13]

The experiment they propose involves exposing subjects to caricatures or outline drawings of famous faces and to photographs of the same faces, while

12 Ramachandran & Hirstein (1999), 'The Science of Art'.
13 Ibid., p. 24.

measuring skin conductive response (SCR) as an index of limbic activation. This is slightly different than the sort of experiment we have considered thus far, and will require a slight modification of our model:

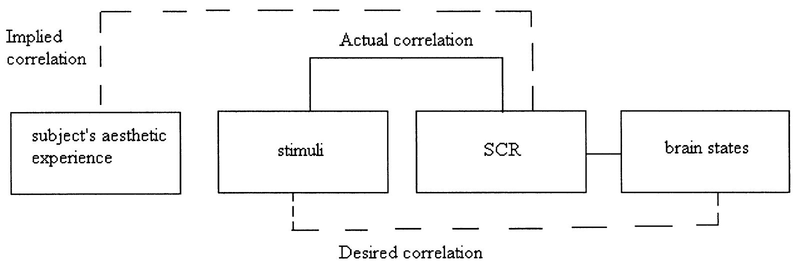

Figure 3

In the Ramachandran and Hirstein experiment, the link between the subject's aesthetic experience and her report has been eliminated. This is because the authors would like to eliminate the reliance on introspective reports as a measure of subjective experiential states. They write: the size of the SCR is a direct measure of the amount of limbic (emotional) activation produced by an image. It is a better measure, as it turns out, than simply asking someone how much emotion he feels about what he is looking at because the verbal response is filtered, edited, and sometimes censored by the conscious mind – so that your answer is a 'contaminated signal'.[14]

The Ramachandran and Hirstein experiment, then, uses a lie-detector to evaluate subjects' aesthetic experience. Not only does this indicate a different attitude towards the validity of introspective reports from that seen in the Kawabata and Zeki experiment, it also reveals a different conception of the object of study. The subject's aesthetic experience in this case is not her *conscious* experience of an object as beautiful, but rather her *subconscious* limbic response to the object.

Notice also the dual role of the SCR measurement. SCR in this experiment is the measurement of brain activity. It is hypothesized that SCR will also be an indication of the subject's aesthetic experience of the object, a sort of involuntary report. This link, which I have labelled 'Implied correlation' in Figure 3, is justified by the theoretical connection between limbic activation during low-level visual processing and a later, higher-level aesthetic experience. The use of SCR instead of fMRI thus has important implications. fMRI, for better or worse, is a technique that looks at the whole brain (although in so-called "regions of

14 Ibid., p. 32.

interest" or ROI studies certain areas of the brain are isolated for investigation). SCR, on the other hand, which is a measure of changes in electrical conductivity produced by sweating, is assumed to be a measure of activity in the limbic system, the area of the brain generally and loosely associated with emotion. By using an experimental paradigm that pinpoints one particular brain structure, the authors have *a priori* determined that that structure is the seat of the aesthetic experience.

In contrast to the Kawabata and Zeki study, the stimuli suggested are not chosen for their beauty (or ugliness), but rather as examples of a specific technique that can be exploited by an artist. This is also an approach employed by Zeki and Marini[15] in a fascinating fMRI study in which they explore the use of colour in Fauvist art by investigating the differences in brain activity in subjects viewing images of objects in their natural colours and in unnatural colours. Ramachandran and Hirstein are interested in correlating an aspect of the stimulus with brain activity in a particular region; thus, their desired correlation is *not* between the subject's aesthetic experience and a particular brain state, but rather between the stimulus itself and a brain state. The actual correlation in the Ramachandran and Hirstein experiment is between the stimulus and the SCR. While this gives a more direct link between the objects of desired correlations, it has the curious effect of pushing the aesthetic experience of the subject, which we might otherwise assume to be the object of a neuroaesthetic study, to the side, connected to the experiment only by a theoretical extension from the SCR results.

V. BLOOD AND ZATORRE: SHIVERS DOWN THE SPINE

One of the most well-known and also most unusual studies conducted on brain activity in response to art is what has come to be known as the 'chills study' by Blood and Zatorre.[16] Conducted with a PET scanner, the chills study attempts to correlate an aesthetic experience of music as defined by the experience of 'chills' or 'shivers-down-the-spine' with specific brain states. Subjects were trained musicians (at least eight years of training), all of whom reported experiencing the well-known phenomenon of 'shivers-down-the-spine' when listening to certain pieces of music. Subjects selected their own stimulus, providing 'one piece of music that consistently elicited intensely pleasant emotional responses, including chills'.[17] Prior to scanning, subjects also rated

---

[15] Zeki, Semir, & Marini, Ludovica (1998). 'Three Cortical Stages of Colour Processing in the Human Brain.' *Brain*, 121, pp. 1669–85.

[16] Blood, Anne J., & Zatorre, Robert J. (2001). 'Intensely Pleasurable Responses to Music Correlate with Activity in Brain Regions Implicated in Reward and Emotion.' *Proceedings of the National Academy of Sciences*, 98 (20), pp. 11818–23.

[17] Blood & Zatorre (2001), 'Intensely Pleasurable Responses to Music', p. 11818.

the intensity of their emotional response to the other subjects' chosen stimuli; one of these, which the subject had rated as emotionally neutral, was used as a control. Two baseline tasks, silence and noise, were also scanned. During the scan, measurements were also taken of subjects' heart rate, respiration, EMG (muscular electrical activity), electrodermal response, and skin temperature. After the scan, subjects were asked to rate their own response to the four stimuli (their own chills-music, neutral music, noise, and silence). Figure 4 gives a linking diagram for the chills study:

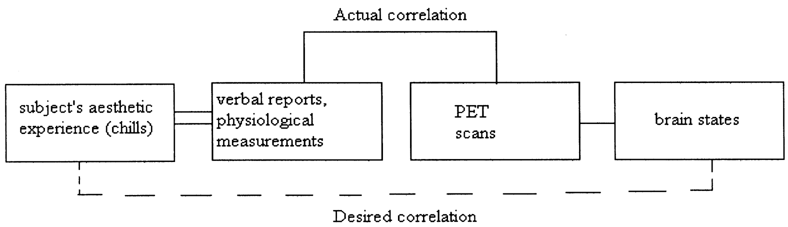

Figure 4

Although the study was carried out with a PET scanner and not an MR scanner, these two techniques (which both measure changes in blood-flow in the brain associated with neural activity) share many characteristics, and for our purposes the differences do not play a large role. None the less, it may be worth mentioning a few things in passing. First, the PET scanner has the advantage (for a study on the subtle effects of music) of being quieter than the MR scanner, which can be quite noisy. Second, the PET scanner has a poorer temporal resolution than the MR scanner, so the images generated may reflect more extended processes and miss some of the quicker changes. Then again, both PET and fMRI scans have very poor temporal resolution (seconds) compared to the speed at which neural changes actually occur (milliseconds). The general comments on the subtraction method of data extraction mentioned in the introduction, above, hold true for both PET and fMRI studies.

What is of more interest in the present context is the way in which Blood and Zatorre get at the aesthetic experience of the subject. Perhaps the most important aspect of the study, from a methodological point of view, is the choice of a very specific indicator of the aesthetic experience, the experience of shivers-down-the-spine, that can be measured not only physiologically (changes in heart rate, EMG, and respiration correlated significantly with reported shivers) but also *introspectively*. The use of the chills paradigm introduces a *phenomenological* measurement that can be directly correlated with a third-person, objective measurement.

In Figure 4, I have indicated this with a double link between the aesthetic experience and the measurement of the experience. Blood and Zatorre have *two* measurements of their subjects' mental experience, allowing them to interpret the brain images generated by the PET scan with greater confidence. One of the greatest difficulties in conducting brain-imaging studies of cognitive tasks is that one can never really know what the subject is doing in the scanner.[18] The chills paradigm tackles this problem by means of a flanking manoeuvre, in which the object of study is pinned down from two sides. Jack and Roepstorff[19] refer to this combination of physiological, phenomenological, and brain-scan measurements as *triangulation*.

Whereas Ramachandran and Hirstein's proposed experiment *defines* the aesthetic experience as the SCR measurement of limbic activity,[20] regardless of what the subject may introspectively *claim* about her aesthetic response to the stimuli, Blood and Zatorre's subjects were actually chosen on the basis of their 'reports of frequent, reproducible experiences of chills in response to certain pieces of music'.[21] This is another aspect of the chills study that sets it apart not only from both Ramachandran and Hirstein's proposed study and the study conducted by Kawabata and Zeki, but also from the vast majority of studies conducted in the field of cognitive neuroscience.[22]

VI. INTROSPECTION: THE EXPERT WITNESS

Whereas Ramachandran and Hirstein suggest using 'naive experimental subjects'[23] and Kawabata and Zeki's subjects had 'no special experience in

18 See Lutz, Antoine, Lachaux, Jean-Phillipe, Martinerie, Jacques, & Varela, Francisco J. (2002). 'Guiding the Study of Brain Dynamics by Using First-person Data: Synchrony Patterns Correlate with Ongoing Conscious States during a Simple Visual Task.' *Proceedings of National Academy of Sciences*, 99 (3), pp. 1586–91.

19 Jack, Anthony I., & Roepstorff, Andreas (2002). 'Introspection and Cognitive Brain-mapping: From Stimulus-Response to Script-Report.' *Trends in Cognitive Sciences*, 6 (8), pp. 333–39, and Jack, Anthony I., & Roepstorff, Andreas (2003). 'Why Trust the Subject?' *Journal of Consciousness Studies*, 10 (9–10), pp. v–xx.

20 To be fair, I should note that Ramachandran and Hirstein are of course aware that there is more to the aesthetic experience than can be measured by galvanic skin response. Ramachandran writes that SCR 'provides a convenient and reliable index of arousal. As such, it provides *one* measure of response to art – but certainly not a complete measure'. Ramachandran, Vilayanur S. (1999). 'Author's Response.' *Journal of Consciousness Studies*, 6 (6–7), p. 74. My point is merely that within the context of the proposed experiment, the aesthetic experience is *experimentally* defined as limbic activity as measured by SCR. Interestingly, in this context, Blood and Zatorre found no correlation between SCR and the chills phenomenon.

21 Blood & Zatorre (2001), 'Intensely Pleasurable Responses to Music', p. 11818.

22 Jack & Roepstorff (2002), 'Introspection and Cognitive Brain-mapping'; Varela Francisco J. (1996). 'Neurophenomenology: A Methodological Remedy for the Hard Problem.' *Journal of Consciousness Studies*, 3 (4), pp. 330–49.

painting or art theory',[24] Blood and Zatorre's subjects were musicians, with eight to ten years of musical training. Their justification for this unusual choice is the following:

> Musicians were used in this experiment based on the premise that this population is more likely to experience strong emotional responses to music; however, music training is not necessary to experience these responses.[25]

Not only were the subjects trained musicians, but '[i]ndividual subjects were selected on the basis of their reports of frequent, reproducible experiences of chills in response to certain pieces of music'.[26] In order to get at the aesthetic experience, Blood and Zatorre thus called on a panel of 'expert witnesses'. This move is a response to a classic problem in experimental psychology: how to make objective measurements when the object of study (the mind) is by definition a subjective, first-person affair.[27] That this problem is relevant to the field of neuroaesthetics is made abundantly clear by some of the passionate responses to Ramachandran and Hirstein's article, as well the insistence of cooler heads on the experienced ineffability of the aesthetic experience.[28] It is simply difficult (or impossible, depending on whom you ask) to describe in words the experience of aesthetic enjoyment. Nonetheless, that intangible, ineffable aesthetic experience must *also* be an absolutely concrete brain state (or at least *correlate* with one,[29] or correlate with relative changes or modulations of a more basic background brain state[30]). And, as Overgaard's linking model illustrates, if we want to correlate brain states with subjective mental states, then we must make use of some sort of measurement of those subjective states.

The three neuroaesthetic studies presented in my article illustrate three different methods of dealing with this problem. It would be tempting to conclude from the above that Blood and Zatorre are more interested in the conscious experience of their subjects than the other researchers, that Ramachandran and Hirstein are only interested in subconscious limbic response and not interested in conscious experience, and that Kawabata and Zeki lie somewhere in-between. However, this would do a great disservice to all

---

23 Ramachandran & Hirstein (1999), 'The Science of Art', p. 32.
24 Kawabata & Zeki (2004), 'Neural Correlates of Beauty', p. 1699.
25 Blood & Zatorre (2001), 'Intensely Pleasurable Responses to Music', p. 11818.
26 Ibid.
27 Marcel, Anthony J. (2003). 'Introspective Report: Trust, Self-knowledge, and Science.' *Journal of Consciousness Studies*, 10 (9–10), pp. 167–86.
28 Ex. De Clerq, Rafael (2000). 'Aesthetic Ineffability.' *Journal of Consciousness Studies*, 7 (8–9), pp. 87–97.
29 Chalmers (2005), 'What Is a Neural Correlate of Consciousness?'.
30 Searle, John R. (2000). 'Consciousness.' *Annual Review of Neuroscience*, 23, pp. 557–78.

of these researchers. In fact, all three studies are essentially interested in *both* the conscious, experiential aspects of aesthetic experience *and* in the underlying brain mechanisms that can at least be said to *contribute* to the conscious experience, if not to *constitute* it. Although Ramachandran and Hirstein (as well as Zeki) have been accused of taking a reductionist approach to art, the only reason for investigating such 'low-level' brain processing as limbic activity[31] or colour processing in V4[32] in the appreciation of art is precisely to get at that 'higher-level', ineffable quality that we experience as aesthetic. These approaches break down the problem into more manageable pieces. And while Blood and Zatorre make an explicit and thorough use of introspective reports in their experimental design, they are not conducting a study on the neural correlates of consciousness, but rather on the effects of a stimulus (music) on the brain.

Having said this, I must add that I believe the application of Overgaard's linking model to the neuroaesthetic studies described in this paper shows Blood and Zatorre's study to be the most methodologically convincing. Their approach to experimental design is not only an application of Jack and Roepstorff's 'triangulation',[33] but also an application of what Varela has called 'neurophenomenology'.[34] Varela's reference to phenomenology is not merely a reference to 'introspection' or 'retrospection' of the sort advocated by Jack and Roepstorff. It is an explicit reference to the philosophical tradition of Phenomenology, which relies on trained observers making careful, precise observations of their own inner states. When studying the correlation of brain activity with such subtle mental states as emotions, Varela argues, we need subjects who are able to accurately distinguish between different emotions: it cannot be assumed that just anyone can be relied on to provide the sort of precise reports of mental states that are needed to make convincing correlations with observed brain activity. By using trained musicians who recognize and consistently experience the phenomenon in question in response to specific passages of music, Blood and Zatorre can be more certain of eliciting the effect that they wish to study. The difficulty, of course, in trusting any subject, even the 'expert witness', is in being certain that she is able to discern and accurately report on her mental states. By use of triangulation, that is, by conducting retrospective interviews with subjects after the scanning

31 Ramachandran & Hirstein (1999), 'The Science of Art'.

32 Zeki, Semir (1999). *Inner Vision: An Exploration of Art and the Brain.* Oxford: Oxford UP; Zeki & Marini (1998), 'Three Cortical Stages of Colour Processing in the Human Brain'.

33 Jack & Roepstorff (2002), 'Introspection and Cognitive Brain-mapping'; Jack & Roepstorff (2003), 'Why Trust the Subject?'.

34 Varela (1996), 'Neurophenomenology'.

experience, and by taking physiological measurements and then correlating those with the retrospectively reported experience of chills or absence of chills, Blood and Zatorre are able to correlate their PET scans with a measurement of subjective experience that is at least as precise a measurement of that experience as the scans are a measurement of brain states (if not more so).

In addition, unlike the Kawabata and Zeki study, in which subjects reported their aesthetic response by means of button-press during scanning, the introspective reports used by Blood and Zatorre are retrospective. Although this distances the introspective report temporally from the actual experience (but not much), it has the advantage of freeing the brain-scanning data from the interference of the need to make a report. Blood and Zatorre's subjects simply lay in the scanner and passively listened to the stimuli, unlike Kawabata and Zeki's subjects, who actively engaged in a process of rating many pictures at high speed (less than a second per picture). Choosing a retrospective reflection as a means of quantifying the aesthetic experience instead of an immediate report via button-press thus adds a degree of ecological validity to the experiment. Listening passively (albeit in a scanner) to music is more like the sort of everyday aesthetic experience we would like to describe in neurological terms than quickly rating pictures is.

VII. REFINING THE NEUROAESTHETIC EXPERIMENT: LOOKING FOR THE EXPERIENCE OF LOOKING AT ART

In the following pages, I would like to examine ways in which the field of neuroaesthetics might progress, building on the general framework of the chills study. The three experimental designs we have considered so far are based on a stimulus-response paradigm. The subject is exposed to a stimulus, and her response is recorded. But, argues Ellis,[35] this does not capture the way we actually 'use' art. Approaching both perception and art from a Gibsonian, 'ecological' perspective, Ellis writes: 'Because neither perceptions nor emotional responses are really passive "responses" at all, art does not cause us to feel a certain way. Instead, we "use" art for the purpose of symbolizing our emotions.'[36] Thus, '[p]aintings, rather than *causing* us to see and feel certain ways, only *provide us with an opportunity* to do so'.[37] As Skov points out,[38] the aesthetic *experience* is not contained in the aesthetic object itself, but involves an interaction between

35 Ellis, Ralph. D. (1999). 'The Dance Form of the Eyes: What Cognitive Science Can Learn from Art.' *Journal of Consciousness Studies*, 6 (6–7), pp. 161–75.

36 Ibid., p. 161.

37 Ibid., p. 164.

38 Skov (2005), 'Hvad er neuroæstetik?', p. 1.

the object and a human mind. Reception of a piece of art takes place against an experiential background or context. We do not respond to a painting in the same way every time we look at it. Not all of Blood and Zatorre's subjects experienced chills every time their own selections were played in the scanner, although they chose the music because it often elicits chills. A poem that on one day seems to express the ultimate truth about life, the universe, and everything can seem flat and uninteresting the next day. This experiential background or context has, of course, its neural correlates, and this must surely be an important factor in designing an experiment that seeks to correlate brain activity with aesthetic response. We all have our various moods, worries, preoccupations, hopes, and so forth, and we bring these along with us into the brain scanner.

This is far from a trivial problem. Lutz et al write: 'Even during well-calibrated cognitive tasks, successive brain responses to repeated identical stimulations are highly variable.'[39] That is, if we put a person in a scanner, and expose her to the same stimulus ten times, we are likely to scan ten different neural patterns, all of which would have to be categorized as a 'response' to the same stimulus.[40] Clearly, this has important implications for our confidence in the scanning images as somehow characterizing the neural substrate of the same experience. Which of the ten scans should we choose as the 'right' one?

One likely reason for the great variability in response to the same stimulus, suggest Lutz et al, is variability in the neural background into which the stimulus is introduced. Lutz et al describe an experiment in which they employ the techniques of Varela's 'neurophenomenology' in order to characterize this background state, and to demonstrate how this phenomenological data can be used as a co-variant in data analysis.

The experimental task involved fusing an autostereogram (the well-known 3D illusion). During the experiment, EEG recordings were taken from 62 electrodes. As soon as the subjects had fused the stereogram and experienced the illusion of depth, they pressed a button, ending the trial. However, prior to the experiment, the subjects underwent an intensive training programme, in which they learned to observe their own conscious states and report on them. By responding to open questions, they worked with the authors to characterize their experiences prior to the resolution of the task, and on the moment of experiencing the illusion. The subjects practised observing

39 Lutz, Lachaux, Martinerie & Varela (2002), 'Guiding The Study of Brain Dynamics', p. 1586.

40 For more on intersession variability within a single subject see McGonigle, David J., Howseman, Alistair M., Athwal, Balwindar S., Friston, Karl J., Frackowiak, Richard S., & Holmes, Andrew P. (2000). 'Variability in fMRI: An Examination of Intersession Differences.' *NeuroImage*, 11, pp. 708–34.

and reporting on their experiences while fusing the stereograms until 'they found their own stable experiential categories to describe the main elements of the cognitive context in which they perceived the 3D shapes'.[41] These stable experiential categories were called Phenomenological Clusters (PhCs). The PhCs described the subjects' feelings of their own preparedness prior to fusing the image, and the quality of their perception of the illusion. The PhCs fell into (generally) three categories: a feeling of 'steady readiness', in which subjects felt 'present' and 'prepared'; a feeling of 'fragmented readiness', in which subjects made an effort to be prepared, but were 'distracted' or 'tired'; and a feeling of 'unreadiness', in which subjects felt totally unprepared, and only saw the illusion because their eyes were in the right position. These three phenomenological categories corresponded with qualities of the perception of the illusion, such that when subjects felt *steady readiness*, they experienced the illusion as 'continuous' or 'satisfying'; when they felt *fragmented readiness*, their experience of the illusion was also more 'discontinuous'; and feelings of *unreadiness* were accompanied by a feeling of 'surprise' on seeing the illusion.

The authors found a statistically significant correlation between trials when grouped by either the phenomenological data or by the behavioural data (reaction time), thus anchoring the introspective data to a third-person, objective measurement as seen in the chills study above. Furthermore, they found a striking correlation between the introspective reports of the feelings of readiness or unreadiness prior to seeing the illusion and the results of the EEG measurements. When subjects reported feeling ready to fuse the stereogram, EEG recordings revealed widespread frontal activity *prior* to the introduction of the stimulus, and the behavioural data showed a quicker response time. In contrast, when subjects reported feeling unprepared, little or no frontal activity was registered prior to the presentation of the stimulus, activity was scattered across the entire cortex upon presentation of the stimulus, and reaction times were slower.

It would seem, then, that Lutz et al succeeded in showing that the neural context into which the stimulus is inserted can, to some degree, be characterized by recourse to introspective data, and that these data can be correlated with both behavioural and with neural measurements. While Blood and Zatorre[42] make no reference to the neurophenomenological program, the Lutz et al study[43] (one of whose four co-authors is Francisco Varela) is an explicit attempt to implement the program in an empirical context. Where Blood and Zatorre's

[41] Lutz, Lachaux, Martinerie & Varela (2002), 'Guiding the Study of Brain Dynamics', p. 1587.

[42] Blood & Zatorre (2001), 'Intensely Pleasurable Responses to Music'.

[43] Lutz, Lachaux, Martinerie & Varela (2002), 'Guiding the Study of Brain Dynamics'.

expert witnesses were musicians who claimed to experience often shivers-down-the-spine while listening to certain pieces of music, Lutz and co-authors have carefully trained their panel of expert witnesses on a very specific task.

What are the implications of these differences? The elegance of the chills study is in its simplicity and its high degree of ecological validity. In contrast, the Lutz et al study required painstaking preparation, in which experimenters and subjects worked together to generate the 'Phenomenological Clusters' that were used as a variable in data analysis. This of course implies a distancing of the experience under study from a 'natural' context, that is, a lower ecological validity. Although the intent of the study was to characterize the phenomenal and neurological context into which the response to the stimulus was introduced, the very methods used to achieve that characterization must surely have an effect on the state itself. It is one thing to look at an illusion, and something else to observe oneself looking at an illusion. The very act of self-observation is a psychological difference that is likely to have an effect on the neural activity we measure. Then again, it could be argued that although introducing this element of rigorous self-observation alters the mental context of the subject, the additional information that can be gained about that context is worth this sacrifice. Lutz et al subjects were able to identify three different mental states that served as a background for their experience of the illusion. Blood and Zatorre's subjects, in contrast, only characterized the experience of the stimulus itself, and said nothing (because they were not asked) about their own state prior to the stimulus.

Another interesting aspect of the Lutz et al experiment is the degree of attention that was paid to differences between individual subjects. The experimental cohort was small (four subjects); this may be a necessity in a study that looks so intensively at individual experience. Even with only four subjects, the authors were obliged to add a fourth category to the list of PhCs because one of the subjects insisted that he experienced a feeling of 'open attention without active preparation' that was clearly distinct from both *steady readiness* and *fragmented readiness*. One can easily imagine that with a larger number of subjects the number of PhCs might quickly become unwieldy. Then again, variation between subjects is a general problem in brain-imaging studies; different people have different brains that respond to the same stimulus in different ways on different days. In order to be able to say something general about a group of subjects, researchers are often obliged to average together results from all their subjects, producing a brain-image which is in fact a map of nobody in particular's neural activity. The other option is equally disagreeable, however: saying something concrete about only one subject on one day (or even one particular trial). Although Lutz et al's study cannot be said to in any way resolve this problem, it can at least be said to acknowledge and take seriously the neural context into which experimental stimuli are introduced. This must be an essential step toward getting at a neural characterization of Ellis's idea of the active *use* of art, as opposed to a passive *response* to art.

VIII. CONCLUSION: WHERE TO GO FROM HERE?

I have used Overgaard's linking model to examine the methodologies of three studies investigating the effects of aesthetic stimuli on the brain. We have seen that different methodologies rest (sometimes implicitly) on different theoretical presuppositions or attitudes, and that the choice of experimental design has implications for the sorts of questions one can ask and the sorts of answers one can get. In particular, we have seen that imposing one or the other methodological framework always implies defining the aesthetic experience of the subject in one way or another, in order to be able to measure it as a variable. I have argued that Blood and Zatorre's chills study is a particularly good example of how one can define the aesthetic experience in a way that is both precise within the context of the experiment *and* meaningful in general terms outside the context of the experiment. Most of us know the experience of shivers as a response to beautiful music. The chills study thus manages to be both precise and ecologically valid.

In addition, we have looked at an experimental paradigm that (while not specifically neuroaesthetic) addresses questions of background neural context that would seem to be of importance in designing a study of the way a subject, in a certain mental (and therefore brain) state, interacts with an aesthetic object. Ideally, this would allow us to design experiments with a higher degree of ecological validity: if part of the aesthetic experience has to do with the general mental state that we are in at the moment of perceiving the object, then it seems reasonable to try to characterize that general mental state in phenomenological and behavioural/physiological terms so that it can provide a context for the data analysis. Paradoxically, we have seen that the neurophenomenological method, employed as a means to do just that, in fact removes the experience further from the 'natural' sort of experience we would ideally like to measure, and introduces a new background element, the conscious monitoring of one's own experience, into the experience itself. To my knowledge, whether or not this introduction of what we might call 'meta-introspective' activity into the subject's task has any measurable effect in terms of neural measurements like fMRI remains untested.

To date, the simplest and most elegant study of aesthetic response that I have seen is Blood and Zatorre's chills study. Although the chills study also depended on introspective reports, no special training of the subjects was involved. The *choice* of subjects, however, was unusual: where the norm is to use naïve subjects, Blood and Zatorre chose trained musicians as their subjects. This might be one route towards future refinements of neuroaesthetic experiments: rather than train one's expert witnesses, the researcher could take advantage of subjects' prior training. Professionals trained in aesthetic appreciation could be used as scanning subjects, perhaps in combination with the Phenomenal Clusters technique. Although subjects like art critics and museum curators spring to mind, the possibilities are wide. Castriota-Scanderbeg et al,[44] for instance, conducted an fMRI study comparing brain activity in experienced wine-tasters and novices. Vuust et al[45] conducted an MEG study on the responses of trained jazz musicians and people with no musical training to rhythmic cues. Although my examination of the various studies treated in this paper has focused on the difficulties involved in designing an experiment that can convincingly correlate subjective, aesthetic experiences with concrete brain activity, I would like to end on an optimistic note. fMRI and PET techniques, in conjunction with other measures of neural activity such as EEG and MEG (which have a better temporal resolution), offer an opportunity to observe brain activity in healthy subjects under reasonably comfortable conditions. The challenge will be in designing methodological paradigms that can take advantage of the strengths of these techniques, minimize their weaknesses, and take seriously the dual nature of the object of study: the brain *and* the experience.

Ethan Weed, Aarhus University Hospital, Denmark
ethan@pet.auh.dk

REFERENCES

Blood, Anne J., & Zatorre, Robert J. (2001). 'Intensely Pleasurable Responses to Music Correlate with Activity in Brain Regions Implicated in Reward and Emotion.' *Proceedings of the National Academy of Sciences*, 98 (20), pp. 11818–23.

[44] Castriota-Scanderbeg, Alessandro, Hagberg, Gisela E., Cerasa, Antonio, Committeri, Giorgia, Galati, Gaspare, Patria, Fabiana, Pitzalis, Sabrina, Caltagirone, Carlo, & Frackowiak, Richard (2005). 'The Appreciation of Wine by Sommeliers: A Functional Magnetic Resonance Study of Sensory Integration.' *NeuroImage*, 25, pp. 570–78.

[45] Vuust, Peter, Pallesen, Karen J., Bailey, Chris, van Zuijen, Titia L., Gjedde, Albert, Roepstorff, Andreas, & Østergaard, Leif (2005). 'To Musicians, the Message Is in the Meter: Pre-attentive Neuronal Responses to Incongruent Rhythm Are Left-lateralized in Musicians.' *NeuroImage*, 25, pp. 560–64.

Castriota-Scanderbeg, Alessandro, Hagberg, Gisela E., Cerasa, Antonio, Committeri, Giorgia, Galati, Gaspare, Patria, Fabiana, Pitzalis, Sabrina, Caltagirone, Carlo, & Frackowiak, Richard (2005). 'The Appreciation of Wine by Sommeliers: A Functional Magnetic Resonance Study of Sensory Integration.' *NeuroImage*, 25, pp. 570–78.

Chalmers, David J. (2005). 'What Is a Neural Correlate of Consciousness?' Retrieved August 22, 2005, from http://consc.net/papers/ncc2.html.

De Clerq, Rafael (2000). 'Aesthetic Ineffability.' *Journal of Consciousness Studies,* 7 (8–9), pp. 87–97.

Ellis, Ralph. D. (1999). 'The Dance Form of the Eyes: What Cognitive Science Can Learn from Art.' *Journal of Consciousness Studies*, 6 (6–7), pp. 161–75.

Jack, Anthony I., & Roepstorff, Andreas (2002). 'Introspection and Cognitive Brain-mapping: From Stimulus-Response to Script-Report.' *Trends in Cognitive Sciences*, 6 (8), pp. 333–39.

Jack, Anthony I., & Roepstorff, Andreas (2003). 'Why Trust the Subject?' *Journal of Consciousness Studies*, 10 (9–10), pp. v–xx.

Kawabata, Hideaki, & Zeki, Semir (2004). 'Neural Correlates of Beauty.' *Journal of Neurophysiology*, 91, pp. 1699–1705.

Lutz, Antoine, Lachaux, Jean-Phillipe, Martinerie, Jacques, & Varela, Francisco J. (2002). 'Guiding the Study of Brain Dynamics by Using First-person Data: Synchrony Patterns Correlate with Ongoing Conscious States during a Simple Visual Task.' *Proceedings of National Academy of Sciences*, 99 (3), pp. 1586–91.

Marcel, Anthony J. (2003). 'Introspective Report: Trust, Self-knowledge, and Science.' *Journal of Consciousness Studies*, 10 (9–10), pp. 167–86.

McGonigle, David J., Howseman, Alistair M., Athwal, Balwindar S., Friston, Karl J., Frackowiak, Richard S., & Holmes, Andrew P. (2000). 'Variability in fMRI: An Examination of Intersession Differences.' *NeuroImage*, 11, pp. 708–34.

Overgaard, Morten (2003). 'On the Theoretical and Methodological Foundations for a Science of Consciousness.' *Bulletin fra Forum for Antropologisk Psykologi*, 13, pp. 6–31.

Ramachandran, Vilayanur S. (1999). 'Author's Response.' *Journal of Consciousness Studies*, 6 (6–7), pp. 72–75.

Ramachandran, Vilayanur S., & Hirstein, William (1999). 'The Science of Art: A Neurological Theory of Aesthetic Experience.' *Journal of Consciousness Studies*, 6 (6–7), pp. 15–51.

Savoy, Robert L. (2005). 'Experimental Design in Brain Activation MRI: Cautionary Tales.' *Brain Research Bulletin*, 67, pp. 361–67.

Searle, John R. (2000). 'Consciousness.' *Annual Review of Neuroscience*, 23, pp. 557–78.

Skov, Martin (2005). 'Hvad er neuroæstetik?' *Kritik*, 38 (174), pp. 1–10.

Solso, Robert (2000). 'The Cognitive Neuroscience of Art: A Preliminary fMRI Observation.' *Journal of Consciousness Studies*, 7 (8–9), pp. 75–85.

Uttal, William R. (2001). *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain*. Cambridge, MA, and London: Massachusetts Institute of Technology.

Varela, Francisco J. (1996). 'Neurophenomenology: A Methodological Remedy for the Hard Problem.' *Journal of Consciousness Studies*, 3 (4), pp. 330–49.

Vuust, Peter, Pallesen, Karen J., Bailey, Chris, van Zuijen, Titia L., Gjedde, Albert, Roepstorff, Andreas, & Østergaard, Leif (2005). 'To Musicians, the Message Is in the Meter: Pre-attentive Neuronal Responses to Incongruent Rhythm Are Left-lateralized in Musicians.' *NeuroImage*, 25, pp. 560–64.

Zeki, Semir (1999). *Inner Vision: An Exploration of Art and the Brain*. Oxford: Oxford UP.

Zeki, Semir (2001). 'Closet Reductionists.' *Trends in Cognitive Science*, 5 (2), pp. 45–46.

Zeki, Semir, & Marini, Ludovica (1998). 'Three Cortical Stages of Colour Processing in the Human Brain.' *Brain*, 121, pp. 1669–85.