

A Neurocognitive Approach to Music Reading

LAUREN STEWART

*Wellcome Department of Imaging Neuroscience,
University College, London, United Kingdom, and
University of Newcastle upon Tyne, NE1 7RU, United Kingdom*

ABSTRACT: Music reading offers a unique perspective on the acquisition of a notational system. Many people cannot read music, but a large proportion are motivated to learn. Musical literacy is therefore amenable to studies of acquisition in a way that language literacy is not. The studies reviewed here investigate how musical symbols on the page are decoded into a musical response. The studies address the nature of the mental representations used in music reading, as well as their instantiation within the brain. The results of a musical Stroop paradigm are described, in which musical notation was present but irrelevant for task performance. The presence of musical notation produced systematic effects on reaction time, demonstrating that reading of the written note, like the written word, is obligatory for those who are musically literate. Spatial interference tasks are also described that suggest that music reading, at least for the pianist, can be characterized as a set of vertical to horizontal mappings. These behavioral findings are mirrored by the results of an fMRI training study in which musically untrained adults were taught to read music and play piano keyboard over a period of three months. Specific learning-related changes were seen in the superior parietal cortex and fusiform gyrus, for melody reading and rhythm reading, respectively. These changes are suggested to correspond to the acquisition of processes that deal with the extraction of spatial and featural properties of notation, respectively.

KEYWORDS: music reading; learning; fMRI; automaticity; melody; rhythm

INTRODUCTION

While language notation primarily functions to carry semantic meaning, music notation carries instructions for the production of a musical performance. Music reading thus lies at the interface between perception (encoding of visual symbols on the page) and action (production of the musical response). Although it is undeniably the case that music reading is “more than a visuo-motor task,”¹ it is also the case that without a theory of music reading at this level of transcription, studies of musical expertise and the role of higher cognitive processes in music reading will necessarily be limited in their scope.

Address for correspondence: Lauren Stewart, Institute of Cognitive Neuroscience, Alexandra House, 17 Queen Square, London WC1N 3AR. Voice: +44-20-7679-1160; fax: +44-20-7813-2835.

l.stewart@ucl.ac.uk

Ann. N.Y. Acad. Sci. 1060: 377–386 (2005). © 2005 New York Academy of Sciences.
doi: 10.1196/annals.1360.032

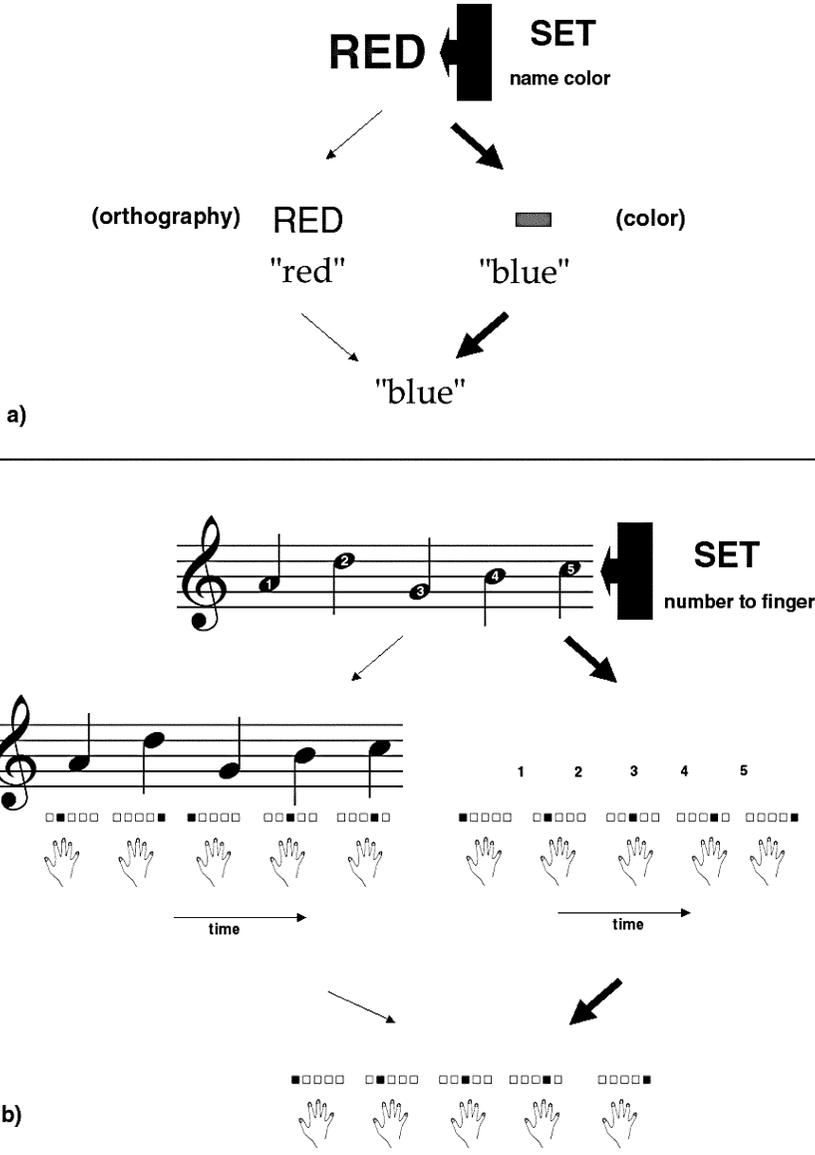


FIGURE 1. Schematic of the language Stroop task and the musical stroop task. In the language Stroop task (a), the participant is required to name the ink color of a color word. In the incongruent situation, the color name and the color word do not correspond (as above). In order to make the appropriate response (naming the color), the participant must inhibit a tendency to read the color word. In the musical Stroop task (b), the participant is required to make a sequence of keypresses, on the basis of the numbers, superimposed on musical notes. In the incongruent situation, the numbers and the notes do not correspond (as above), and the participant must inhibit a tendency to play the notes. Participants who cannot read music have only one mapping, from the numbers to the fingers, and are therefore unaffected by the note/number congruence of the pairing.

AUTOMATICITY IN MUSIC READING

In investigating how musical symbols are decoded into a musical response, one question that can be asked is whether this process occurs automatically. Does musical literacy result in obligatory processing of musical notation, in much the same way that standard literacy results in the automatic processing of the written word?² Reports from accomplished musicians suggest that this may be the case:

I have come to the conviction that sight-reading has something to do with the speed at which the visual image is converted into a muscular act; and with some people this is a very rapid transformation. As a matter of fact, it is so rapid, that the transformation of a visual image into its pianistic result is so rapid that a great deal of what happens escapes awareness—it just happens.³

The classic Stroop task² demonstrates the obligatory nature of word reading. Naming the ink color of a color word is faster when the ink color and the color word correspond, for example, the word *red*, written in red ink, compared to the word *red*, written in blue ink, because reading words is more automatic than naming colors. In order to test whether this automaticity extends to music reading, a musical version of the classic Stroop task⁴ was used, as illustrated in FIGURE 1. Pianists and nonmusicians placed their right hand over the keys G to D of a piano keyboard and made keypresses according to the appearance of numbers, between 1 and 5, on the screen. If the number 1 was seen, they pressed the key beneath their thumb as quickly as possible. If the number 2 was seen, they pressed the key beneath their index finger as quickly as possible, and so on. The numbers were superimposed onto musical notes, which were either congruent or incongruent with respect to the numbers. In a congruent trial, the number 1 would be superimposed onto the note G, the number 2 superimposed onto the note A, and so on. In an incongruent trial, the number 1 may be superimposed onto the note A, the number 2 on the note D, and so on, as shown in FIGURE 1. Baseline trials were also included in which numbers were presented in a straight line and in the absence of musical notes.

Both musicians and nonmusicians were easily able to perform the number-to-finger mapping. However, the pianists took several hundreds of milliseconds longer to make a sequence of keypresses when the notes and numbers were incongruent compared with trials in which only numbers were present and were significantly facilitated when notes and numbers were congruent. The nonmusicians showed no



FIGURE 2. Musical Stroop task in a nonmusical context. Example of a motorically matched pair of stimuli: congruent (a) and incongruent (b).

effect of congruence, performing at the same speed regardless of the note/number correspondence. A similar effect was also observed when the musical notes were removed so that the numbers were presented in vertical locations that were either musically congruent or incongruent (FIG. 2).

The experiments previously described used a fixed association between particular fingers and particular piano keys (e.g., thumb above G, index finger above A). However, pianists do not form absolute associations between individual fingers and individual notes; rather they can play any note with any finger, and the choice of fingering will depend on the musical context. In order to show that the effect we observed was independent of the previous association between fingers and notes, the musical Stroop experiment was repeated with the small difference that the numbers now referred to particular keys rather than particular fingers, and participants responded using a single finger. The systematic effects on response time in pianists were still seen. Equally, when the participants responded by pressing keys on a computer keyboard, rather than a piano keyboard, the effect was still observed. Thus the effect does not depend upon which effectors are used; nor does it depend on a musical context for the response. Instead, the critical locus of interference seems to be the spatial location at which the response is made. Overall, these experiments provide evidence that musical notation is automatically read by those who are musically literate, even when the process of reading music is detrimental to task performance.

SPATIAL MAPPINGS IN PIANISTS

On the basis of these findings, it was hypothesized that music reading involves a set of spatial mappings from a note's position on the musical staff to the specified location of the instrumental response. Notes on the musical staff are organized along the vertical dimension so that, for pianists, notes that are vertically lower on the staff correspond to keys that are further to the left on the keyboard, while notes that are vertically higher correspond to keys that are further to the right on the keyboard. Musicians and nonmusicians were compared on a nonmusical version of the musical Stroop task (FIG. 3), which again required them to place their right hand over five adjacent keys and to think about their fingers in terms of the numbers 1 to 5.

Numbers appeared in different vertical locations, but the spatial location at which they appeared was always irrelevant. The subjects were required to perform the same number-to-finger mapping as before (1 = thumb, 2 = index finger, and so on) using a computer keyboard. A congruent trial was one in which the number 1 (specifying a keypress beneath the left-most effector) would appear in the vertically lowest box, the number 2 in the second lowest box, and so on (a piano-like vertical to horizontal relationship between visual stimulus and its associated response).

There were two kinds of incongruent trials: random and systematic. An incongruent (random) trial was one in which the number 1 would appear in a box other than the vertically lowest one, the number 2 would appear in a box other than the second to lowest one, and so on. An incongruent (systematic) trial was one in which the number 1 would appear in the uppermost box, the number two in the box second from top, and so on (opposite to the mapping used for congruent trials). Such stimulus-response mappings would be incongruent with respect to those used by pianists when reading from a musical score. Pianists took longer to make the same sequence

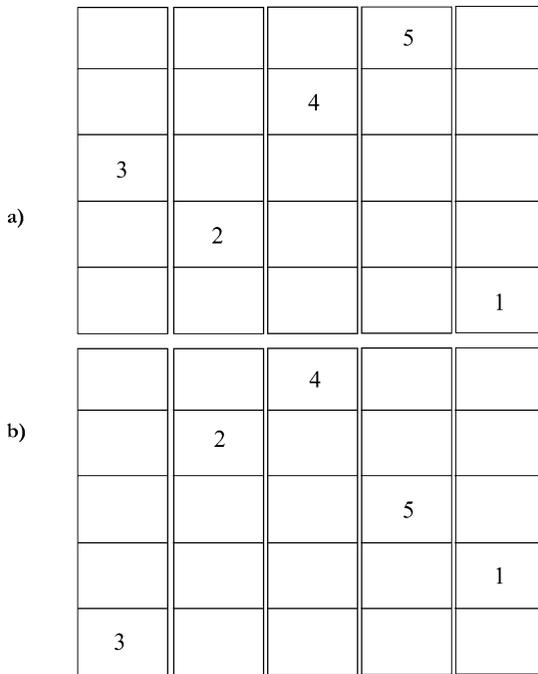


FIGURE 3. Nonmusical Stroop task. In congruent trials (a), the number-to-finger mapping corresponds to the spatial stimulus–response mapping used by pianists when playing from a musical score. In incongruent trials (b), numbers were presented in spatial locations that did not correspond to these stimulus–response mappings. Each trial consisted of 5 stimuli, specifying a sequence of 5 different keypresses. Stimuli appeared consecutively, triggered by a keypress response to the previous stimulus.

of keypresses when the numbers appeared in incongruent locations (whether random or systematic), compared to congruent locations. Even when neither the stimulus nor the response elements of the task were in any way musical, the spatial mappings inherent to the reading and playing keyboard music could still be seen. The demands of constantly mapping from vertically organized visual stimuli (musical notes on the staff) to a set of horizontally organized response elements (piano keys) thus appear to forge a set of spatial mappings that are in evidence even outside of a musical context.

NEURAL CORRELATES OF MUSICAL LITERACY ACQUISITION

These cognitive studies of music reading have shed light on the nature of the representations that are involved in reading music. However, it is also pertinent to ask which brain areas are involved in this transcriptional process. Music reading has the advantage of being a real-world skill that can, to some extent, be fractionated into behavioral components that are already well understood. By studying the neural

basis of music reading at a basic transcriptional level, we can ask how these components are integrated into a complex skill. A previously conducted study scanned professional pianists as they read, played, and listened to music within the confines of a positron emission tomography (PET) scanner.⁵ When playing music from a score, with auditory feedback, was contrasted with reading from a score and hearing it played, the superior parietal cortex was activated, leading the authors to suggest that this area is involved in mediating the “sensori-motor transformations for visually guided skilled actions and finger positioning,” although this study lacked a nonmusical sensorimotor transformation for comparison, so the musical specificity of these results was open to debate.

However, a subsequent study⁶ contrasted the pattern of brain activity involved in playing a note, based on a musical note on the staff, with playing a note based on a verbal label corresponding to that note or based on a number corresponding to the required finger. When musical transcoding was compared to either verbal or numerical transcoding, activation of right superior parietal cortex was found. The major difference between the musical transcoding and the verbal and numerical transcoding is the reliance of the transformation on the spatial information contained within the musical stimulus. Thus we can infer that the activation of the superior parietal cortex appears specific to the kind of sensorimotor transformations required for reading music.

Converging evidence for the role of the superior parietal cortex in music reading comes from an fMRI study that measured changes in the brains of individuals as they learned to read music and play keyboard over a period of three months.⁷ Adult par-

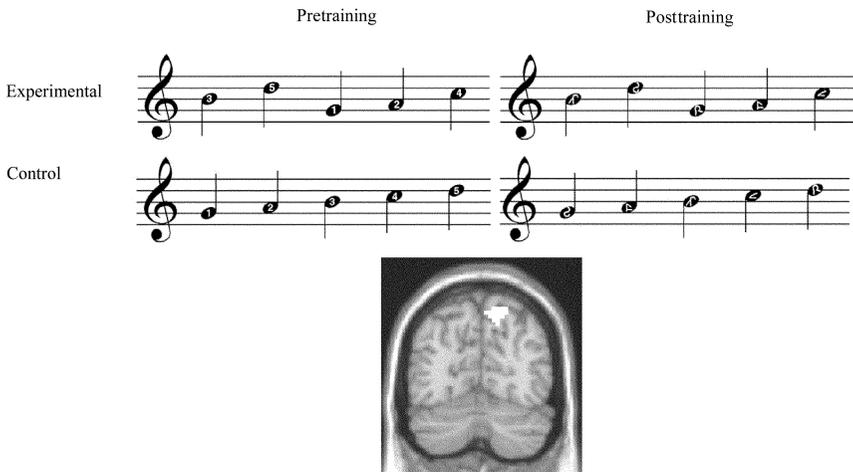


FIGURE 4. fMRI study: reading music for melody. *Upper panel:* Example of stimuli. Pretraining: participants performed a number-to-finger mapping. Experimental trials were unpredictable in their specified motor response compared to control trials, which specified an ascending or descending sequence of responses. Posttraining: the same stimuli were used, but nonsense symbols replaced the numbers, so that participants now performed a note-to-finger mapping. *Lower panel:* Statistical parametric map, rendered onto a normalized structural image, showing a learning-related change in superior parietal cortex (coordinates: 18, -72, 57).

ticipants were scanned twice, once before training and once after training. During both scanning sessions, the participants saw musical notes that varied in pitch (between G and D) and responded by making the appropriate keypresses on an MRI-compatible keyboard. Before training, when musical notes held no meaning for the participants, they simply made their responses based on numbers that were superimposed onto the notes (notes and numbers were always congruent). After training, these numbers were replaced by visually similar but meaningless symbols, so that stimuli were visually similar across the pretraining and posttraining scanning sessions (FIG. 4). Now participants made their responses based on their newly acquired musical literacy skills. A training effect was seen in the right superior parietal cortex, less than half a centimetre from the area identified in the musical transcoding study.

The neuroimaging studies that have so far been mentioned have concentrated on delineating the processes involved in reading musical notation for pitch. However, apart from knowing what to play, it is equally important to know when to play it. The rhythmic instructions in musical notation are conveyed not by spatial information contained within the notation but rather by its visual features: whether a notehead is filled or open, or whether there is a stem or a dot. In the same training study men-

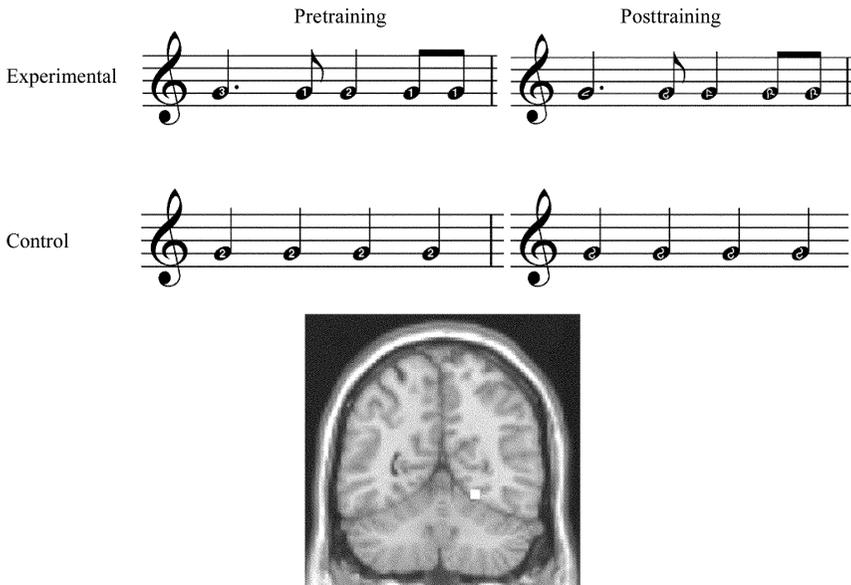


FIGURE 5. fMRI study: reading music for rhythm. *Upper panel:* Example of stimuli. Participants produced rhythms, using only a single key. *Pretraining:* the numbers indicated the duration of each keypress. An auditory beep was used to give the duration of the single beat. Experimental trials always comprised three different durations, whereas control trials specified a sequence of keypresses of the same duration. In order to keep the number of keypresses constant across control and experimental trials, control trials were either a bar of eight short duration keypresses, a bar of four medium duration keypresses, or a bar of two long duration keypresses. *Lower panel:* Statistical parametric map, rendered onto a normalized structural image, showing a learning-related change in the fusiform gyrus (coordinates: 21, -57, -12).

tioned above, one condition required participants to tap out an equitonal rhythm from notation. Before training, numbers that were superimposed onto the noteheads conveyed note durations, but after training, these symbols were once again replaced with visually similar meaningless symbols so that rhythmic literacy was required to successfully decode the notation (Fig. 5).

A pre-/posttraining comparison of the brain's response to rhythmic notation revealed learning-related changes in the fusiform gyrus.⁸ This region and adjacent regions in the ventral occipital cortex have been shown to be important for visually discriminating between exemplars of a particular category, for instance, faces,⁹ buildings/scenes,¹⁰ letters,¹¹ and now, rhythmic notation. It is striking that reading music for pitch and for rhythm, respectively, involves the dorsal and ventral visual processing streams.^{12,13} Of course, it is usual for musical notation to contain both pitch and rhythmic information, and future studies will be necessary to investigate how these two aspects of music reading draw upon processing in both of these streams to produce a response that is integrated in space and time.

The melody and rhythm reading conditions previously described required participants to attend to these different dimensions of musical notation and respond accordingly, based on their explicit knowledge concerning the meaning of musical notation. In order to investigate whether the brain automatically processes musical

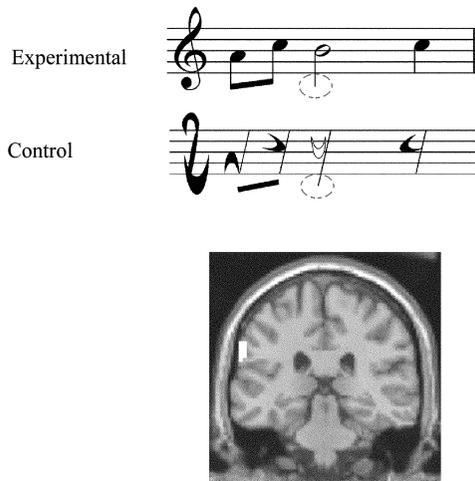


FIGURE 6. fMRI study: implicit music reading. *Upper panel:* Example of stimuli. Participants performed a visual feature detection task in which the target (a vertical stem that protruded above or below the horizontal staff) could form part of a musical stimulus (musical trial, above) or a nonmusical stimulus (nonmusical trial, below), constituting experimental and control trials, respectively. Participants indicated whether the target was ascending or descending, using an arbitrary up/down mapping to the index and middle fingers. Musical trials were visually and motorically matched with nonmusical trials, but only the musical trials were musically interpretable. *Lower panel:* Statistical parametric map, rendered onto a normalized structural image, showing a learning-related change in the left supramarginal gyrus (coordinates: $-63, -30, 27$).

notation, participants engaged in a third condition, in which notation was present but irrelevant. The task was a visual one that required participants to indicate whether a visual ascender or descender (vertical stem extending above or below the five horizontal lines) was present (see FIG. 6). Control trials were similar, but the ascenders and descenders were embedded within a kind of notation that was visually similar to musical notation but could not be musically deciphered. The prediction was that musical notation and the visually matched nonmusical notation would both be meaningless before training but that after training the musical notation alone would automatically activate elements of a newly acquired musical lexicon.

A comparison of the brain's activity before and after training revealed a learning-related effect in the left supramarginal gyrus. As predicted, this region did not distinguish between the musical notation and the nonmusical notation before training but was more active for musical notation than for nonmusical notation after training. Neuroimaging studies, transcranial magnetic stimulation (TMS) studies, and patient studies have highlighted the importance of this area in the processing of "motor intention."¹⁴⁻¹⁸ A plausible interpretation of the activation change seen in the supramarginal gyrus during the implicit music reading condition, therefore, is that during the course of training, participants learned to make specific keypresses in response to particular musical notes. After training, the visual appearance of musical notes was automatically and unconsciously interpreted as an instruction to act, though processing was terminated before the response was executed. Support for this interpretation comes from the demonstration that the musical novices showed a musical Stroop effect after training,⁷ suggesting that the mere presence of musical notation is sufficient to result in musical response preparation.

CONCLUSIONS

The studies described above reveal that music reading, at least for keyboard performance, requires coordinated sensorimotor translation between a set of vertically organized stimuli and a horizontally organized set of responses. These spatial mappings, which are likely to relate to the decoding of pitch from notation, develop in the early stages of skill acquisition and are associated with functional changes in the superior parietal cortex. The ability to decode the rhythmic elements of notation, by contrast, appears to depend upon a visual discrimination process subserved by the fusiform gyrus. These "what" and "when" aspects of music reading seem to map onto the what and when occipitoparietal and occipitotemporal streams, respectively.

ACKNOWLEDGMENTS

The work referred to in this article was supported by the Medical Research Council and the Wellcome Trust. The behavioral experiments also involved Uta Frith and Vincent Walsh. The neuroimaging work was conducted in collaboration with Rik Henson, Knut Kampe, Robert Turner, Uta Frith, and Vincent Walsh.

[Competing interests: The author declares that she has no competing financial interests.]

REFERENCES

1. SLOBODA, J.A. 1984. Experimental studies of musical reading: a review. *Mus. Percept.* **2**: 222–236.
2. STROOP, J.R. 1935. Studies of interference in serial verbal reactions. *J. Exp. Psychol.* **6**: 643–662.
3. WOLF, T. 1976. A cognitive model of sight-reading. *J. Psycholing. Res.* **5**: 143–171.
4. STEWART, L., V. WALSH & U. FRITH. 2004. Reading music modifies spatial mapping in pianists. *Percept. Psychophys.* **66**: 183–195.
5. SERGENT, J., E. ZUCK, S. TERRIAH & B. MACDONALD. 1992. Distributed neural network underlying musical sight-reading and keyboard performance. *Science* **257**: 106–109.
6. SCHÖN, D., J.L. ANTON, M. ROTH & M. BESSON. 2002. An fMRI study of music sight-reading. *Neuroreport* **13**: 2285–2289.
7. STEWART, L., R. HENSON, K. KAMPE, *et al.* 2003. Becoming a pianist. An fMRI study of musical literacy acquisition. *Ann. N. Y. Acad. Sci.* **999**: 204–208.
8. STEWART, L. 2003. *Neurocognitive Studies of Music Reading*. PhD Thesis. University College. London.
9. KANWISHER, N., J. MCDERMOTT & M.M. CHUN. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* **17**: 4302–4311.
10. EPSTEIN, R. & N. KANWISHER. 1998. A cortical representation of the local visual environment. *Nature* **392**: 598–601.
11. COHEN, L., S. DEHAENE, L. NACCACHE, *et al.* 2000. The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* **123**: 291–307.
12. MISHKIN, M. & L.G. UNGERLEIDER. 1982. Contribution of striate inputs to the visuospatial functions of parieto-occipital cortex in monkeys. *Behav. Brain Res.* **6**: 57–77.
13. GOODALE, M.A. & A.D. MILNER. 1992. Separate visual pathways for perception and action. *Trends Neurosci.* **15**: 20–25.
14. RUSHWORTH, M.F., T. PAUS & P.K. SIPILA. 2001. Attention systems and the organization of the human parietal cortex. *J. Neurosci.* **21**: 5262–5271.
15. RUSHWORTH, M.F., A. ELLISON & V. WALSH. 2001. Complementary localization and lateralization of orienting and motor attention. *Nat. Neurosci.* **4**: 656–661.
16. HARRINGTON, D.L. & K.Y. HAALAND. 1992. Motor sequencing with left hemisphere damage. Are some cognitive deficits specific to limb apraxia? *Brain* **115**: 857–874.
17. KIMURA, D. 1993. *Neuromotor Mechanisms in Human Communication*. Oxford University Press. New York.
18. SIRIGU, A., E. DAPRATI, P. PRADAT-DIEHL, *et al.* 1999. Perception of self-generated movement following left parietal lesion. *Brain* **122**: 1867–1874.