

Composing in the scanner: A functional magnetic resonance imaging single case study on visual and auditory imagery

Morteza Izadifar¹  | Taoxi Yang¹ | Marco Paolini² | Arusu Formuli¹ | Yan Bao^{3,4} 

¹Institute of Medical Psychology, Ludwig Maximilian University Munich, Munich, Germany

²Department of Radiology, University Hospital, LMU Munich

³School of Psychological and Cognitive Sciences, Peking University, Beijing, China

⁴Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China

Correspondence

Arusu Formuli, Institute of Medical Psychology, Ludwig Maximilian University, Goethestr. 31, 8033 Munich, Germany.

Email: arusu.formuli@med.uni-muenchen.de

Professor Yan Bao, School of Psychological and Cognitive Sciences, Peking University, 5 Yiheyuan Road, Beijing 100871, China.

Email: baoyan@pku.edu.cn

Abstract

“Seeing with the mind’s eye” and “hearing with the mind’s ear” are two common indicators of musical imagery, and they can be referred to as “visual” and “auditory” musical imagery. However, a question remains open, that is, whether visual and auditory imagery of the same musical composition share the same neural mechanisms. Moreover, how can neural mechanisms guarantee the temporal flow of “musical imagery”? To answer these questions, we report here a preliminary single case study using functional magnetic resonance imaging with an eminent composer who imagined one of his compositions in two states of mind as compared to his resting-state activity. In the visual imagery condition, he imagined visually the score of his composition in a continuous way. In the auditory imagery condition, he imagined auditorily the same musical composition with pauses. In spite of the modality and temporal differences, the two types of mental imagery showed similar temporal durations for the same musical composition. However, different patterns of neural activation were observed for visual and auditory imagery with one important exception, that is, a common activation pattern was observed in the left medial temporal gyrus in both visual and auditory imagery. We speculate that the left medial temporal gyrus may play an important role in the creation of apparent temporal continuity in musical imagery and perhaps even in conscious information processing in general.

KEYWORDS

auditory imagery, functional magnetic resonance imaging, music, temporal continuity, visual imagery

INTRODUCTION

For a better understanding of how cognitive contents like thoughts, feelings, memories, or images are neurally represented, different behavioral and imaging paradigms have been used, including mental imagery procedures (e.g., von Trott Zu Solz et al., 2017). One open question is whether “seeing with the mind’s eye” or “hearing with the mind’s ear” as indicator of mental imagery may be equivalent (or at least similar) to the perceptual process itself. Independent of controversies about potential similarities in neural activation patterns of imagery and visual or auditory perception, the critical point has been raised that imagery suffers severe limitations, and should not be used as an experimental paradigm due to its subjectivity (e.g., Pylyshyn, 1973; Watson, 1913). However, in spite of such methodological criticism important insight has been gained about the neuro-cognitive machinery using imagery as a

tool (e.g., Griffiths, 2000; von Trott Zu Solz et al., 2017; Zatorre & Halpern, 1993). Although the term “imagery” might suggest that it refers only to the visual modality, auditory (musical) and tactile imagery are also recognized (e.g., Halpern et al., 1995; Kraemer et al., 2005; McGuire et al., 1996; von Trott Zu Solz et al., 2017;). In fact, musical imagery (Bailes, 2019; Cotter, 2019; Halpern, 2001; Weber & Brown, 1986; Zatorre & Halpern, 1993) can be quite dominant like echoing a musical motif in our mind for hours. An impressive individual example for the power of auditory imagery is Ludwig van Beethoven who became deaf later in his life and in spite of his deafness still was a creative composer. This process of composing presumably happened on the basis of auditory imagery initiated and then followed by the visual equivalent in the musical scores. This duality of the two involved sensory modalities is another example of complementarity as a generative principle which can be observed in several

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. PsyCh Journal published by Institute of Psychology, Chinese Academy of Sciences and John Wiley & Sons Australia, Ltd.

domains of human cognition like in top-down *and* bottom-up visual processing, in content of conscious representation *and* logistic functions (like musical motifs *and* their temporal dynamics), or aesthetic evaluations when linking aesthetic appreciation *and* cognitive appraisal (Bao & Pöppel, 2012; Bao, von Stosch, et al., 2017b; Bao, Yang, Lin, Fang, et al., 2016a; Zhang et al., 2017; Zhou et al., 2016).

Considering that both auditory and visual imageries for the same musical composition share identical musical content, one question is whether the underlying neural mechanisms of auditory and visual imageries are similar or different. Furthermore, another question is, how can neural mechanisms guarantee the temporal flow of “musical imagery”? With respect to the latter point we would like to cite a statement by the eminent conductor Bruno Walter (1957, p. 36): “Kontinuität des Tempos gehört zu den wichtigsten Grundsätzen des Musizierens” (“The continuity of tempo is one of the most important principles to make music”). Thus, temporal control is essential in music, and the open question is whether such temporal control is equivalent (or at least similar) in imagery compared to the perceptual process itself. It has been suggested that specific “time windows” of finite durations are essential for temporal control. In the low-frequency domain, a time window of 2–3 s has been experimentally identified for different kinds of cognitive processes like speech segmentation, intentional movement control, sensorimotor synchronization, bistable perception, or aesthetic evaluation (Bao et al., 2015; Wang et al., 2015; Yu & Bao, 2020; Zhao et al., 2018). Furthermore, this time window may also be essential for the control of tempo in music as suggested previously (Pöppel, 1989).

Although auditory imagery has been studied experimentally using behavioral paradigms or imaging techniques like magnetoencephalography (Schürmann et al., 2002), positron emission tomography (Halpern & Zatorre, 1999; Zatorre et al., 1996), functional magnetic resonance imaging (fMRI; Halpern et al., 2004; Kraemer et al., 2005; Yoo et al., 2001) with healthy subjects or brain-injured patients (Zatorre & Halpern, 1993), there has been a neglect in analyzing the temporal continuity of imagery. Thus, temporal processing is an issue with respect to the question how information processing in the time domain is implemented on the neural level. Does the brain process information in a continuous form or in a sequence of discrete instants? Introspectively, subjective time has been considered as a continuous flow. William James (1890) has been one of the pioneers assuming consciousness as “sensibly continuous”. Indeed, our intuition suggests that time is continuous rather than characterized by discrete temporal steps. However, experimental evidence shows that the intuition of a continuous time in perception and cognitive activities in general is misleading, and that this intuition can perhaps be considered as an illusion (Bao et al., 2015; Bao, Yang, Lin, & Pöppel, 2016b; Eriksen & Collins, 1968; Pöppel, 1970, 1971, 1997; Zhou et al., 2014). For instance, if one takes a look at the receptor surface of information processing in the visual and auditory modality, then discrete time sampling with distinct time windows is actually a necessity. The different transduction time in these sensory

modalities, the physical characteristics of stimuli with different intensities or distances to the observer, or the divergent neuro-anatomical projection to central regions of the brain have as a consequence that physical time cannot be mapped directly onto central areas of the brain like the primary visual or auditory cortices. The resultant distortion can, however, be overcome by time windows of finite duration within which information being processed at different physical times is treated as co-temporal. Such a mechanism of overcoming ill-defined temporal information can technically be implemented by neural oscillations with different frequencies (Bao et al., 2015; Pöppel, 1970, 1971, 1997). It is then a general question how information represented in such sequential time windows with different frequencies are connected with each other to result in apparent continuity. What are the neural mechanisms behind these constructions of continuity, in particular with respect to the “functional glue” of the time window of 2–3 s that has been considered as the “subjective present” (Bao et al., 2015; Pöppel, 1997; Yu & Bao, 2020)? Perhaps by focusing on mental imagery with and without interruptions to the same musical composition one could scratch the surface of an answer.

Thus, in the experiment described, we aimed to obtain a better understanding of the underlying neural mechanisms of auditory and visual imagery with an identical musical composition, and furthermore, to get a hint how apparent continuity might be created in musical imagery. Since there exists only a small number of exploratory results implemented on musicians who have received long-term musical training to do both auditory and visual musical imagery, and different musical compositions might result in diverse mental processes in different individuals due to the different degrees of familiarity and understanding of the compositions etc., we decided to conduct an exploratory single case fMRI experiment trusting in the power of such studies (Flyvbjerg, 2006). Single case studies have been proven to be of heuristic value in many domains of cognitive processes like in better understanding speech, referential or procedural memory as well as visual, musical, or tactile imagery (Bao, Pöppel, et al., 2017a, Bao, Yang, et al., 2017c; Pöppel & Bao, 2018; von Trott Zu Solz et al., 2017).

MATERIALS AND METHODS

Participant(s)

The 71-year-old German composer P.M.H. took part in the experiment after giving informed written consent. He had corrected-to-normal visual acuity, and he had no history of neurological diseases. He was enthusiastic about being involved in the study, and he understood the demands upon his time, effort, and experimental conditions. The study was approved by the ethical committee of Ludwig Maximilian University Munich, Germany, in agreement with the Declaration of Helsinki. Prior to the experiment the composer had composed a piece of some 6 min duration without prescribed limitations to genre or instrumentation.

Material

The composition was used as stimulus material for musical imagery in the fMRI experiment with a block design. The composer was firstly measured the resting-state activity for 7.5 min; during this condition, the composer was instructed to keep his eyes closed and was told not to think about anything particular. After this initial phase, two experimental conditions with auditory and visual imagery were tested. The reason why we implemented the experiment into two auditory and visual conditions lies in two facts: (a) we did not see in the literature the different temporal aspects of auditory and visual imagery among musicians; and (b) we wanted to know whether different brain patterns are observed during two diverse kinds of imagery.

Procedure

A pause of 5 min was inserted between adjacent conditions. In the auditory condition, the composer was asked to imagine auditorily his composition; in this experimental phase, the musical imagery was separated into four parts, each lasting 1.5 min. The reason behind separation of auditory imagery into four sections was determined according to the composer's idea. Between sequential parts, he was required to pause for some 30 s. No explicit feedback was given with respect to the amount of time that had passed. The composer indicated the start and the end of each part by pressing a button with his right index finger. In the visual condition, the composer used visual imagery going through the score of his composition. No pause was inserted into the continuous imagery. The composer simply indicated the beginning and the end of his visual imagery. Both visual and auditory imagery are routine tasks of professional musicians. In a report 2 months later after the experiment, the composer confirmed that he imagined his composition in the interrupted auditory phase as music

“hearing it with the mind's ear”, and in the uninterrupted visual phase he went through the score “seeing it with the mind's eye”.

Tools/apparatus

Brain imaging data was obtained with a 3T standard clinical MRI scanner (Ingenia; Philips Medical Systems Nederland) with a 32ch receiving head coil. For BOLD signals, T_2^* -weighted EPI sequences were used (TR = 2500 ms; TE = 30 ms; flip angle = 90° ; acquisition matrix = 80×80 ; number of slices = 43; slice thickness = 3 mm, no gap between slices). In total, one run of 372 functional volumes was acquired for the subject. Structural data was acquired with a T_1 -weighted scan of each participant's brain anatomy (1 mm \times 1 mm \times 1 mm; 240×240 matrix, field-of view = 220 mm).

All neuroimaging data were preprocessed and analyzed using SPM12 (Statistical Parametric Mapping V12, <http://www.fil.ion.ucl.ac.uk/spm>). The functional scan volumes were subjected to spatial realignment to correct for head motion. In further preprocessing analysis, the mean functional image was co-registered to the anatomical image, normalized to the Montreal Neurological Institute (MNI) template provided in SPM12, and spatially smoothed to reduce noise using a Gaussian kernel of 8 mm full width at half maximum. The task was modeled as a block design. We applied a general linear model (GLM) using predictors convoluted with a typical hemodynamic response function. Two experimental conditions (auditory imagery and visual imagery) were compared to the resting-state condition respectively to generate statistical parametric maps of the t-tests (auditory condition vs. resting-state condition, and visual condition vs. resting-state condition). For these statistical tests, significant voxels initially passed a voxel-wise statistical threshold of $p \leq .001$, and a cluster-level threshold was obtained at the statistical significance level of $p < .01$.

TABLE 1 Brain regions that respond to the comparisons of auditory imagery versus resting-state, and the visual imagery versus resting-state

Brain regions	Montreal Neurological Institute coordinates			z scores	Number of voxels
	x	y	z		
Auditory imagery versus baseline					
LMTG	-57	-10	-25	5.27	71
LITG	-48	-13	-37	4.37	26
LTP	-54	5	-31	3.79	70
Visual imagery versus baseline					
LMTG	-54	-10	-25	10.94	172
LSTG	-60	-31	11	8.98	276
LMTG	-60	-19	-7	7.87	48
LCN	-24	-1	23	7.78	37
LSG	-45	-40	35	7.78	258
LPG	-45	-22	35	7.64	26

Note. The resting-state activity serves as a baseline in the experiment. Abbreviations: LMTG = left middle temporal gyrus; LTP = left temporal pole; LITG = left inferior temporal gyrus; LCN = left caudate nucleus; LPG = left postcentral gyrus; LSG = left supramarginal gyrus; LSTG = left superior temporal gyrus.

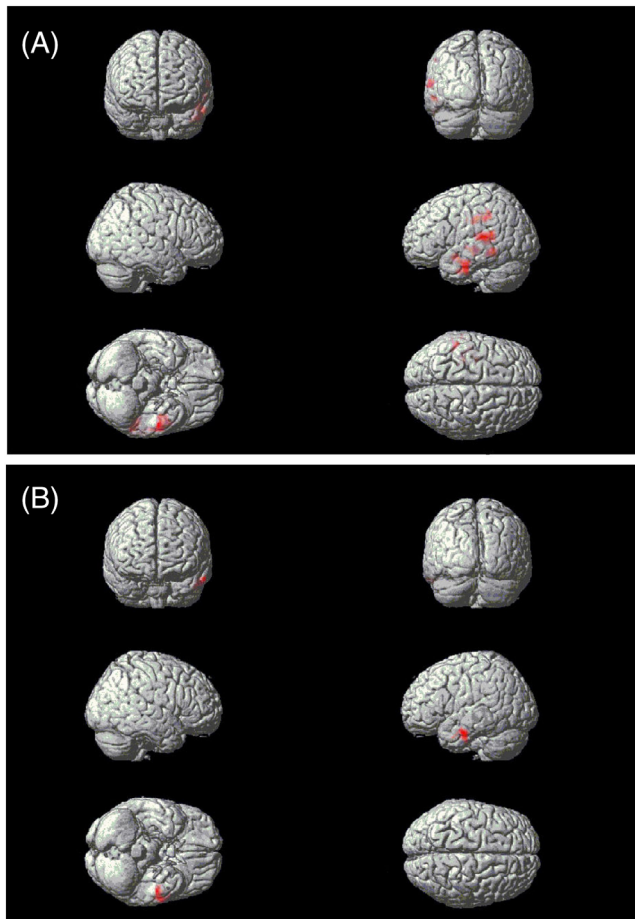


FIGURE 1 Differences in brain activation for visual imagery without a pause versus resting state (A) and for auditory imagery with pauses versus resting state (B). Activations are cluster-level corrected with p (familywise error; FWE) $< .05$. The x , y , and z coordinates are normalized using a standard brain from the Montreal Neurological Institute and displayed in radiological convention

RESULTS

Before we focused on the underlying neural mechanisms of auditory and visual imagery of musical composition, we firstly checked the behavioral data with respect to the time of imagery. Compared to the consciously perceived time of the musical composition, that is, 6 min when the composition was made by the composer, does visually imagining the same musical piece without interruption share the same temporal duration? Moreover, when the imagery is interrupted with its four segments as in the auditory condition, does the mental imagery still hold the same temporal constant?

When the composer made his continuous visual imagery with his “inner eye” on the score of the composition, he showed a rather accurate temporal control; the visual imagery phase lasted 6 min and 9 s. The durations of the four sequential parts of the auditory imagery were 80.8, 80.9, 85.3, and 129.3 s, respectively; thus, the three first parts were shorter than the target duration of 90 s, and the fourth part was considerably longer compared to the preceding three parts and the

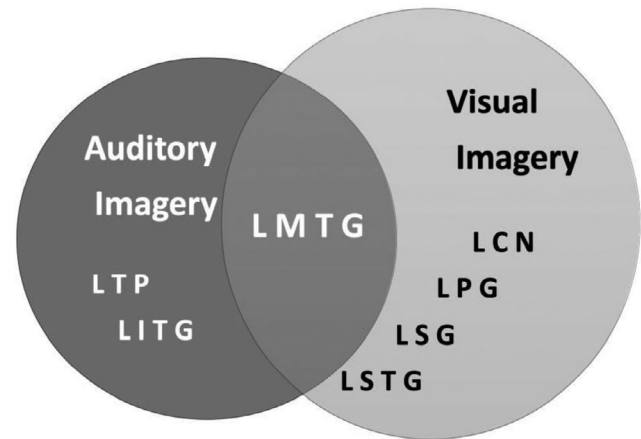


FIGURE 2 Venn diagram illustrating the overlap between auditory imagery and visual imagery. Circle size is proportional to the size of the networks identified. Abbreviations indicate the brain regions where activities were observed: LMTG = left middle temporal gyrus; LTP = left temporal pole; LITG = left inferior temporal gyrus; LCN = left caudate nucleus; LPG = left postcentral gyrus; LSG = left supramarginal gyrus; LSTG = left superior temporal gyrus

target duration of 90 s. To put the four parts together, the whole auditory imagery lasted 6 min and 27 s.

With respect to the imaging data, the neural network for visual imagery versus resting-state included the supramarginal gyrus and the postcentral gyrus, both in the left hemisphere. In the temporal cortex, the middle temporal gyrus and the superior temporal gyrus in the left hemisphere were also activated. In addition, an activation of the left caudate nucleus was observed, which is considered part of the basal ganglia (Table 1, Figure 1A). In the comparison of auditory imagery versus the resting-state, an activation pattern was observed in the middle and the inferior temporal gyrus of the left hemisphere. In addition, there was a significant activation of the temporal pole in the left hemisphere (Table 1, Figure 1B). The Venn diagram illustrates the separation and the overlap of the two experimental conditions (Figure 2). Although the patterns of neural activation are different, there is one important exception, that is, the left middle temporal gyrus (LMTG) is shared both by visual and auditory imagery.

DISCUSSION

Our behavioral results of the total imagery timing showed rather similar temporal durations for both visual and auditory imagery, which were quite close to the consciously perceived time of the musical composition. However, the experimental segmentation of the composition may have had an influence on the temporal control of mental imagery. If one takes only the first three segments of auditory imagery and compares this duration with visual imagery, one may conclude that the visual imagery for the corresponding part had a longer duration; with this conclusion, one assumes that following the

score of the composition visually did not change in “mental speed” throughout the entire composition of 6 min. However, as the sum of the four segments in auditory imagery is not too much different from the target duration, one may also speculate that the different parts of the score had actually comparable durations during visual and auditory imagery. Another possibility could be that the composer reported that he got tired during visual imagery at the end and therefore he may also have slowed down during the last part of visual imagery, similar to the auditory imagery. An answer to this question on the behavioral level has to remain open. Whether an answer can be obtained with imaging technology is a major challenge as it requires identification of distinct patterns of neural activation both for visual and auditory imagery at corresponding phases of the composition. After having analyzed the fMRI data in this experiment it has to be concluded that it is not possible to get an answer as the patterns of neural activation are very different for visual and auditory imagery.

The imaging data in our experiment clearly showed very different neural patterns for auditory imagery and visual imagery. This was consistent with mental imagery as a unique mental phenomenon being separated from the auditory and visual sensory channels in general. Thus, mental imagery seems to be modality-specific, even though common musical composition is used for different modality imagery. On the other hand, since both auditory and visual imagery in our experiment share the same semantic content with similar temporal duration, a common neural mechanism, if it exists, should be related to such commonality. Indeed, the temporal gyrus in the left hemisphere was activated for both auditory and visual imagery conditions, as shown in the Venn diagram (Figure 2). This new observation may lead to a potential answer to the question how apparent continuity is created at neural level. The fact that discontinuous mental imagery in another modality (auditory modality) share a common neural underpinning with continuous mental imagery in visual modality has actually provided better evidence for understanding the temporal continuity of musical composition, since otherwise it will be difficult to exclude the modality contribution if discontinuous and continuous mental imagery are compared in the same modality. Thus, we speculate that the LMTG with its common neural pattern in the two experimental conditions could be a signature of temporal continuity in musical imagery, which serves the function of filling the temporal gap of sequential musical representations.

Is there any other evidence available for supporting such a speculation mentioned above? It has for instance been shown that the LMTG is an important region for the semantic component in language (Acheson & Hagoort, 2013). As semantics is not limited in typical situations to separate time windows, it can be assumed that apparent continuity (or the “illusion of continuity”) may be a consequence of the semantic component in language with its longer time constant. What an utterance means within one time window is carried over to the meaning in the next time window; thus, a long-term narrative can provide in a top-

manner a temporal glue. Supporting evidence comes from some lesion studies; injuries in this region may cause alexia and agraphia for *kanji* characters used in Japanese writing, thus destroying the capacity to extract the meaning (Sakurai et al., 2008). The LMTG is also activated during poem composition (Liu et al., 2012), and it has been shown to be part of a conceptual/semantic network (Wei et al., 2012). Another study shows that theta-band oscillations in the LMTG imitate novel word consolidation (Bakker-Marshall et al., 2018). In some schizophrenic patients it has been reported that auditory verbal hallucinations are related to a cortical thinning in the LMTG (Cui et al., 2018). Furthermore, an increased LMTG was found in both low- and high-function autism in another study (Jia et al., 2018). Although these different results can be considered only as circumstantial evidence in favor of our speculation and not as a proof, they at least do not contradict it.

In conclusion, auditory and visual imagery of the same musical composition are characterized by different neural patterns, but meanwhile share a common activation of the LMTG. On the basis of previous studies (Cui et al., 2018; Jia et al., 2018; Liu et al., 2012) together with the present experiment, the LMTG might be involved in creating temporal continuity of mental imagery. If one considers some general features shared by language and music from a cultural and evolutionary perspective (Jackendoff, 2009), our experiment also suggests similarities in composing music and linguistic artifacts. However, it has to be further investigated whether the LMTG and its network is indeed causally linked to neural mechanisms that create continuity of conscious information processing.

ACKNOWLEDGMENTS

Financial support for this study has come from the Hanns Seidel Foundation Munich, Germany, to M.I., the Chinese Scholarship Council to T.Y., the Zonta-Club II, Munich, Germany, to A.F., and the National Natural Science Foundation of China (Projects 31771213 and 31371018) to Y.B.

DISCLOSURE OF CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORCID

Morteza Izadifar  <https://orcid.org/0000-0002-7207-7336>
Yan Bao  <https://orcid.org/0000-0002-5907-3955>

REFERENCES

- Acheson, D. J., & Hagoort, P. (2013). Stimulating the brain's language network: Syntactic ambiguity resolution after TMS to the inferior frontal gyrus and middle temporal gyrus. *Journal of Cognitive Neuroscience*, 25(10), 1664–1677. https://doi.org/10.1162/jocn_a_00430
- Bakker-Marshall, I., Takashima, A., Schoffelen, J. M., Van Hell, J. G., Janzen, G., & McQueen, J. M. (2018). Theta-band oscillations in the middle temporal gyrus reflect novel word consolidation. *Journal of*

- Cognitive Neuroscience*, 30(5), 621–633. https://doi.org/10.1162/jocn_a_01240
- Bailes, F. (2019). Empirical musical imagery beyond the “Mind’s Ear”. In M. Grimshaw-Aagaard M., Walther-Hansen M., Knakkegaard M. (Eds.) *The Oxford handbook of sound and imagination*, (pp. 445–466). Oxford University Press.
- Bao, Y., & Pöppel, E. (2012). Anthropological universals and cultural specifics: Conceptual and methodological challenges in cultural neuroscience. *Neuroscience & Biobehavioral Reviews*, 36, 2143–2146. <https://doi.org/10.1016/j.neubiorev.2012.06.008>
- Bao, Y., Pöppel, E., Wang, L., Lin, X., Yang, T., Avram, M., Blautzik, J., Paolini, M., Silveira, S., Vedder, A., Zaytseva, Y., & Zhou, B. (2015). Synchronization as a biological, psychological and social mechanism to create common time: A theoretical frame and a single case study. *Psych Journal*, 4, 243–254. <https://doi.org/10.1002/pchj.119>
- Bao, Y., Pöppel, E., & Zaytseva, Y. (2017a). Editorial: Single case studies as a prime example for exploratory research. *PsyCh Journal*, 6, 107–109. <https://doi.org/10.1002/pchj.176>
- Bao, Y., von Stosch, A., Park, M., & Pöppel, E. (2017b). Complementarity as generative principle: A thought pattern for aesthetic appreciations and cognitive appraisals in general. *Frontiers in Psychology*, 8, 727. <https://doi.org/10.3389/fpsyg.2017.00727>
- Bao, Y., Yang, T., Lin, X., Fang, Y., Wang, Y., Pöppel, E., & Lei, Q. (2016a). Aesthetic preferences for eastern and western traditional visual art: Identify matters. *Frontiers in Psychology*, 7, 1596. <https://doi.org/10.3389/fpsyg.2016.01596>
- Bao, Y., Yang, T., Lin, X., & Pöppel, E. (2016b). Donders revisited: Discrete or continuous temporal processing underlying reaction time distributions? *Psych Journal*, 5, 177–179. <https://doi.org/10.1002/pchj.142>
- Bao, Y., Yang, T., Zhang, J., Zhang, J., Lin, X., Paolini, M., Pöppel, E., & Silveira, S. (2017c). The “third abstraction” of the Chinese artist LaoZhu: Neural and behavioral indicators of aesthetic appreciation. *Psych Journal*, 6, 110–119. <https://doi.org/10.1002/pchj.167>
- Cotter, K. N. (2019). Mental control in musical imagery: A dual component model. *Frontiers in Psychology*, 10, 1904. <https://doi.org/10.3389/fpsyg.2019.01904>
- Cui, Y., Liu, B., Song, M., Lipnicki, D. M., Li, J., Xie, S., ... Wang, H. (2018). Auditory verbal hallucinations are related to cortical thinning in the left middle temporal gyrus of patients with schizophrenia. *Psychological Medicine*, 48(1), 115–122. <https://doi.org/10.1017/S0033291717001520>
- Eriksen, C. W., & Collins, J. F. (1968). Sensory traces versus the psychological moment in the temporal organization of form. *Journal of Experimental Psychology*, 77(3), 376–382. <https://doi.org/10.1037/h0025931>
- Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, 12(2), 219–245.
- Griffiths, M. (2000). Internet addiction-time to be taken seriously? *Addiction Research*, 8(5), 413–418. <https://doi.org/10.3109/16066350009005587>
- Halpern, A. (2001). *Cerebral substrates of musical imagery*. Bucknell University Press.
- Halpern, A. R., Bartlett, J. C., & Dowling, W. J. (1995). Aging and experience in the recognition of musical transpositions. *Psychology and Aging*, 10(3), 325–342. <https://doi.org/10.1037/0882-7974.10.3.325>
- Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, 9(7), 697–704. <https://doi.org/10.1093/cercor/9.7.697>
- Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, 42(9), 1281–1292. <https://doi.org/10.1016/j.neuropsychologia.2003.12.017>
- Jackendoff, R. (2009). Parallels and nonparallels between language and music. *Music Perception: An Interdisciplinary Journal*, 26(3), 195–204. <https://doi.org/10.1525/mp.2009.26.3.195>
- James, W. (1890/2013). *The principles of psychology* (Vol. 1). Read Books Ltd.
- Jia, C., Hu, X., Guo, K., Yang, P., Situ, M., & Huang, Y. (2018). Increased left inferior temporal gyrus was found in both low function autism and high function autism. *Frontiers in Psychiatry*, 9, 542. <https://doi.org/10.3389/fpsyg.2018.00542>
- Kraemer, D. J., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Musical imagery: Sound of silence activates auditory cortex. *Nature*, 434(7030), 158. <https://doi.org/10.1038/434158a>
- Liu, S., Chow, H. M., Xu, Y., Erkkinen, M. G., Swett, K. E., Eagle, M. W., Rizik-Baer, D. A., & Braun, A. R. (2012). Neural correlates of lyrical improvisation: An fMRI study of freestyle rap. *Scientific Reports*, 2, 834. <https://doi.org/10.1038/srep00834>
- McGuire, P. K., Silbersweig, D. A., Wright, I., Murray, R. M., Frackowiak, R. S. J., & Frith, C. D. (1996). The neural correlates of inner speech and auditory verbal imagery in schizophrenia: Relationship to auditory verbal hallucinations. *British Journal of Psychiatry*, 169(2), 148–159. <https://doi.org/10.1192/bjp.169.2.148>
- Pöppel, E. (1970). Excitability cycles in central intermittency. *Psychologie Forschung*, 34(1), 1–9. <https://doi.org/10.1007/BF00422860>
- Pöppel, E. (1971). Oscillations as possible basis for time perception. *Studium Generale*, 24, 85–107. J.T. Fraser et al.: “The study of time”, Springer, Berlin (1972) 219–241.
- Pöppel, E. (1989). The measurement of music and the cerebral clock: A new theory. *Leonardo*, 22, 83–89. <https://doi.org/10.2307/1575145>
- Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in Cognitive Sciences*, 1, 56–61. [https://doi.org/10.1016/S1364-6613\(97\)01008-5](https://doi.org/10.1016/S1364-6613(97)01008-5)
- Pöppel, E., & Bao, Y. (2018). The power of single cases: Examples from the early visual pathway and from visual art. In B. Gulyas & J. Vassbinder (Eds.), *Cultural patterns and neurocognitive circuits II – East-west connections* (pp. 33–62). World Scientific Press.
- Pylshyn, Z. W. (1973). What the mind’s eye tells the mind’s brain: A critique of mental imagery. *Psychological Bulletin*, 80(1), 1–24. <https://doi.org/10.1037/h0034650>
- Sakurai, Y., Mimura, I., & Mannen, T. (2008). Agraphia for kanji resulting from a left posterior middle temporal gyrus lesion. *Behavioural Neurology*, 19(3), 93–106. <https://doi.org/10.1155/2008/393912>
- Schürmann, M., Raij, T., Fujiki, N., & Hari, R. (2002). Mind’s ear in a musician: Where and when in the brain. *NeuroImage*, 16(2), 434–440. <https://doi.org/10.1006/nimg.2002.1098>
- von Trott Zu Solz, J., Paolini, M., & Silveira, S. (2017). Does touch inhibit visual imagery? A case study on acquired blindness. *PsyCh Journal*, 6, 139–140. <https://doi.org/10.1002/pchj.162>
- Walter, B. (1957). *Von der Musik und vom Musizieren (about music and to make music)*. S.Fischer Verlag.
- Wang, L., Lin, X., Zhou, B., Pöppel, E., & Bao, Y. (2015). Subjective present: A window of temporal integration indexed by mismatch negativity. *Cognitive Processing*, 16, 131–135. <https://doi.org/10.1007/s10339-015-0687-8>
- Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, 20(2), 158–177.
- Weber, R. J., & Brown, S. (1986). Musical imagery. *Music Perception*, 3(4), 411–426.
- Wei, T., Liang, X., He, Y., Zang, Y., Han, Z., Caramazza, A., & Bi, Y. (2012). Predicting conceptual processing capacity from spontaneous neuronal activity of the left middle temporal gyrus. *Journal of Neuroscience*, 32(2), 481–489. <https://doi.org/10.1097/00001756-200110080-00013>
- Yoo, S. S., Lee, C. U., & Choi, B. G. (2001). Human brain mapping of auditory imagery: Event-related functional MRI study. *Neuroreport*, 12(14), 3045–3049. <https://doi.org/10.1097/00001756-200110080-00013>
- Yu, X., & Bao, Y. (2020). The three second time window in poems and language processing in general: Complementarity of discrete timing and temporal continuity. *PsyCh Journal*, 9, 429–443. <https://doi.org/10.1002/pchj.390>
- Zatorre, R. J., & Halpern, A. R. (1993). Effect of unilateral temporal-lobe excision on perception and imagery of songs. *Neuropsychologia*, 31(3), 221–232. [https://doi.org/10.1016/0028-3932\(93\)90086-f](https://doi.org/10.1016/0028-3932(93)90086-f)
- Zatorre, R. J., Meyer, E., Gjedde, A., & Evans, A. C. (1996). PET studies of phonetic processing of speech: Review, replication, and reanalysis. *Cerebral Cortex*, 6(1), 21–30. <https://doi.org/10.1093/cercor/6.1.21>
- Zhang, J., Yang, T., Bao, Y., Li, H., Pöppel, E., & Silveira, S. (2017). Sadness and happiness are amplified in solitary listening to music. *Cognitive Processing*, 19(1):133–139. <https://doi.org/10.1007/s10339-017-0832-7>

- Zhao, C., Zhang, D., & Bao, Y. (2018). A time window of 3 seconds in the aesthetic appreciation of poems. *PsyCh Journal*, *7*, 51–52.
- Zhou, B., Pöppel, E., & Bao, Y. (2014). In the jungle of time: The concept of identity as an escape. *Frontiers in Psychology*, *5*, 844. <https://doi.org/10.3389/fpsyg.2014.00844>
- Zhou, B., Pöppel, E., Wang, L., Yang, T., Zaytseva, Y., & Bao, Y. (2016). Seeing without knowing: Operational principles along the early visual pathway. *PsyCh Journal*, *5*(3), 145–160. <https://doi.org/10.1002/pchj.141>

How to cite this article: Izadifar, M., Yang, T., Paolini, M., Formuli, A., & Bao, Y. (2022). Composing in the scanner: A functional magnetic resonance imaging single case study on visual and auditory imagery. *PsyCh Journal*, *11*(5), 684–690. <https://doi.org/10.1002/pchj.522>