EXPERIMENTAL STUDIES IN MUSICAL IMAGERY: IMPLICIT AND EXPLICIT COGNITION

Freya Bailes
LEAD–CNRS, Université de Bourgogne, France

ABSTRACT

'Musical imagery' is a term used to refer to the imagined auditory experience of music. A phenomenon of both short- and long-term memory, imagery is a conscious inner hearing, but one that draws on an implicit knowledge of musical structure. Imagery experiments are described in which both implicit and explicit forms of musical knowledge were tested. A primary aim was to adapt implicit priming methods to compare musical perception and musical imagery. Another objective was to explore the explicit cognition of familiar music in terms of the conscious mental representation of pitch, timing and timbre.

Each imagery trial consisted of hearing a brief extract of music, before mentally continuing the piece as if actually hearing it. At a certain moment in the mental continuation (image), either the original or modified music was re-introduced, with the task to determine whether this target seemed to be 'in tune/out of tune', 'in time/out of time', or the 'same/different' (depending on the experimental focus). Accuracy and response times were recorded and compared with data for an equivalent perception task in which the music was actually heard up to the target moment (with no necessity to image through a perceptual 'gap').

Five experiments based on this model are reported. Results allow an inference of the veridicality of imagery for different musical dimensions. The measurement of response time enabled the sensitive detection of behaviour that was different in kind (and not simply magnitude), when imaging and perceiving music. The finding is explained in relation to the balance of implicit and explicit processing called upon by the different task demands. It is suggested that this balance affects the veridicality of imagery for music.

1. BACKGROUND

In listening to music, a recently heard passage affects the hearing of succeeding events, both in terms of an abstract mental representation, and, in the case of familiar music, as a conscious image of events to come. Implicit musical knowledge, such as the harmonic 'rules' of Western classical music, could affect the perception of this music, as could an explicit knowledge of the particular piece. What is not known is the balance of implicit and explicit cognition involved when imagining rather than perceiving.

'Musical imagery' is a term used to refer to the imagined auditory experience of music. A phenomenon of both short- and long-term memory, imagery is experienced as a

conscious inner hearing, but one that draws on an implicit knowledge of musical structure. The growing interest in musical imagery (indicated by Godøy & Jørgensen's 2001 edited volume on the subject) is an indication of its importance for musical activity. For example, a musical image is evoked in sight-reading, in composition, to anticipate sound in performance, to recall sound, and even as an aspect of music listening.

It is difficult to introspect or even to describe mental representation without some reference to perception. Halpern (1992) reports a simple experimental technique with which it is possible to test whether musical imagery can be experienced in real time as a perceptual analogue. In one study (1988), participants saw a song lyric, and then had to judge whether a second lyric belonged to the same song as the first. Reaction time was measured and found to increase the greater the time distance between the lyrics, suggesting real-time imaging. Halpern's work sets a precedent for experimental techniques using 'real' music and imagery. Empirical research on musical imagery has been otherwise lacking (notable exceptions include work by Levitin & Cook, 1996; Repp, 2001; and Weber & Brown, 1986), and it is not known to what extent different dimensions of musical experience, such as timing, pitch and timbre, are experienced veridically in their conscious mental image.

A particular theoretical concern that remains unresolved is the extent to which the veridicality of imagery is dependent on experiment task type. Intons-Peterson (1980) suggests that certain auditory properties, including timbre, may be optionally important to perception and imagery. Imagery experiments inevitably call for a balance of explicit inner hearing and implicit musical knowledge. Differences in the strength of musical components in the mental image might be framed in terms of the balance of implicit and explicit cognition evoked by the imaging context, or the experimental task.

This paper reports a series of imagery experiments in which both implicit and explicit forms of musical knowledge are tested.

2. AIMS

Aims common to all five experiments were:

- To adapt implicit priming methods to compare perceived and imaged mental representations of music.
- To explore the conscious experience of musical imagery for various sound dimensions in familiar music.
- To describe the veridicality of musical imagery in terms of the balance of implicit and explicit cognition called upon by the experimental task.

ISBN 3-931852-67-9
ISSN 1617-6847

It was hypothesised for each experiment that perceived and imagined music would produce quantitatively different behaviour patterns, with mental imagery for music a weak version of perception for the corresponding music. Response times in the adapted 'priming task' would be faster and accuracy higher under perception than imagery conditions. It was also anticipated that the implicit cognition of elements such as tonal and metrical hierarchies, which impact on music perception, would similarly affect imagery.

3. METHOD

Participants either learned or were tested for prior familiarity with the musical material. Each imagery trial consisted of hearing a brief extract of the music, before mentally continuing the piece as if actually hearing it. At a certain moment in the mental continuation (image), either the original or modified music was re-introduced, with the task to determine whether this target seemed to be 'in tune/out of tune', 'in time/out of time', 'same/different', or sung by the phoneme 'di' or 'dou' (depending on the experimental focus). Accuracy and response times were recorded and compared with data for an equivalent perception task in which the music was actually heard up to the target moment (with no necessity to image through a perceptual 'gap').

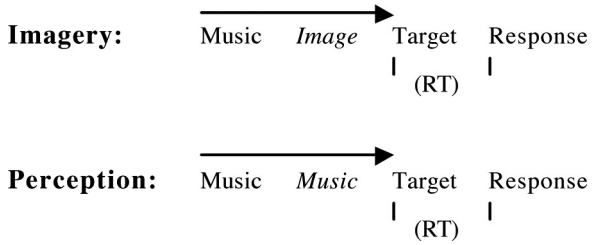

Figure 1: In the imagery condition, participants continued the music as a mental image up to the target moment, while perception tasks presented music continuously. Response time (RT) was measured between the onset of the target and computer key press.

Five experiments were based on this model. Participants in the first four were music students, while the final experiment concerned 'non-musicians' (psychology undergraduates with no musical training).

In order to control for the possibility that participants did not in fact comply with instructions to consciously represent the music through the 'gap', an additional inference condition was introduced in experiments two and four: Before participants had learned the experiment material, they performed the equivalent of the imagery task, but without instructions (or the possibility) to image between opening and target music. It was expected that imagery results would be superior to inference results.

The method was based on priming techniques, in which participant attention is explicitly drawn to a discrimination task, while a further variable is manipulated without the participant's knowledge. For example, experiment five (see table 1 below) drew attention to the phoneme on which the target note was sung, while the hidden variable of interest was whether the target melodically fitted with the context (perceived or imaged), or belonged to a different piece of music. When a target fits, facilitation in discriminating the phoneme should occur, but where the target does not belong, response should be slower and less accurate. The principal advantage of employing priming methods as a research tool in memory is their implicit nature: participants are not required to explicitly retrieve information about the prime, yet their performance may be affected by their exposure to it. This is arguably an advantage when attempting to avoid an over-reliance on conscious introspection about imagery phenomena. However, in the studies reported in this paper, some introspection was encouraged after the experiment, as participants in the first four studies were additionally interviewed about their imagery experience.

3.1 Material

Experiments differed as to the musical material employed, prior familiarity with the material, and the parameter in focus for the discrimination task (such as pitch, timing, timbre, and song lyric). The following table summarises information for each of the five experiments.

Crucially to an exploration of the role of implicit and explicit cognition, the relationship of the discrimination parameter to the perceived, imagined, or 'inference' context was different for each experiment. For example, in order to judge the intonation of a target (experiment one and two), it is necessary to relate directly to the heard or imaged context. Recalling the instrumentation of a specific target (experiment three) is a more absolute task, and the judgement of a sung phoneme (experiment five) has no explicit link to the perceived or imaged context.

Experiment	Material	Familiarity	Discrimination task	Priming variable
One	Nursery rhymes	Assumed	'in tune'/'out of tune' 'on time'/'out of time'	Tonal hierarchy Metrical hierarchy
Two	Song	Introduced	'in tune'/'out of tune'	Tonal hierarchy
Three	Timbre sequence	Introduced	'same'/'different' instrumentation	None
Four	Pop music	Introduced	'filtered'/'unfiltered'	None
Five	Nursery rhymes	Assumed	phoneme 'di' or 'dou'?	Target fit

Table 1: Table listing the material used for each experiment, with the mode of familiarisation (assumed or introduced), the discrimination task, and the 'priming' variable of interest (experiments one, two, and five only).

ISBN 3-931852-67-9
ISSN 1617-6847

4. RESULTS

Data were collected and analysed according to a repeated measures design. Response timings (RT) and accuracy rates were gathered. Accuracy rates were analysed using a generalised chi-square, while RT were analysed using a multifactorial analysis of variance with one between-subjects factor (order) and within-subjects factors depending on the particular experiment, but always including 'condition' (perception, imagery).

4.1 Experiments One and Two

Results for both experiments one and two show that an imagined musical context produces longer and less accurate response than its perceptual counterpart ($F_{1, 12} = 18.77$; $p < 0.001$, and $\chi^2 = 75.6$, $p < 0.001$ respectively). Condition as a factor in reaction time ($F_{2, 24} = 17.9$; $p = 0.0001$) was more significant in the second than the first experiment because of the addition of an inference task in which response was at its slowest. Overall, these results constitute empirical support for the supposition that imagery is a less veridical experience than the perception of music, but they indicate that a mental representation of the music aids task discrimination in comparison with the image-free inference condition.

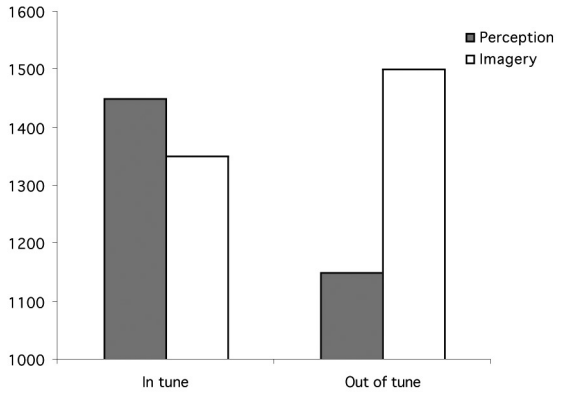

Figure 2: Mean RT (ms) for the 'strong' target of the tonal hierarchy in the pitch test of experiment one. Data show a tendency for 'in tune' perceived targets to elicit slow response compared with equivalent imagery targets.

For the pitch tasks of both experiments, it was found that the predicted effects of tonal hierarchy on perception also applied in an imagery context, with tonally 'strong' targets eliciting faster and more accurate response than weaker targets.

In both the timing task and pitch task, RT (but not accuracy rates) tends towards an inverted pattern between imagery and perception conditions for strong hierarchical targets. For example, 'out of tune' targets elicit slightly longer RT than 'in tune' targets in imagery, and shorter RT than 'in tune' targets in perception (figure 2). This pattern probably reflects the different demands of each condition: because the perception task requires a comparison between the prime and target, an unaltered target fitting well into the context might paradoxically cause a delayed reaction due to a need to 'confirm' the perceptual judgement. In the imagery task, participants hearing the target in the context of their image might be more immediate in their detection of the correct response. Task demands seem to elicit two different processes: in the perception task confirmation of a note's suitability is demanded, while imagery and inference tasks ask for a more active comparison of the target to an image (or opening extract in the inference condition).

4.2 Experiments Three and Four

As predicted, participants were better able to perform the timbral discrimination task under perceptual rather than imagery conditions, though in experiment three this result is reflected only in accuracy rates ($\chi^2 = 7.31$; $p < 0.005$), and not in response time. Timbral knowledge appears to be largely tacit, given few post-experiment reports of any ability to inwardly hear, or 'image' it. Taken as a whole, these results represent a more implicit awareness of timbre than has been supposed.

The question arises as to whether a conscious image of timbre facilitated task performance. In experiment four, the inference condition allowed a comparison between behaviour with and without a mental representation of the music. Differences between imagery and inference data show that performance with knowledge of the music was superior to that without (inference tasks took the longest time, with imagery tasks taking significantly less time to perform, $p < 0.005$). Thus the imagery task evoked a more developed mental representation than the implicit musical knowledge adequate to the inference task.

It has been seen that in experiments one and two, marginally longer RT were measured for 'different' ('out of tune' or 'out of time') targets than 'same' targets when imaging music, with the opposite pattern for perception RT. This pattern was not significant in experiment four, and reversed in experiment three. It seems plausible that this interaction of condition by target type reflects the relative or absolute connection of the discrimination task parameter to its context. Experiment three did not produce such a pattern, and also dealt with a parameter of sound that is perceived more in absolute than relative terms (namely instrumentation). Experiment four manipulated timbre through filters, entailing a more relative than absolute discrimination: In this experiment, the tendency for the interaction discussed was present for only one piece of musical material. This tendency is never significant, and is associated with the relationship of the discrimination parameter to the context.

4.3 Experiment Five

The final experiment was conducted with nonmusicians, whose task was to discriminate between the phonemes being sung at the target moment (method developed by Bigand, Tillmann, Poulin, D'Adamo & Madurell, 2001). Results from this experiment repeat the quantitative differences between perception and imagery conditions found previously, with perception contexts eliciting faster response than imagery contexts ($F_{1, 12} = 20.6$, $p < 0.001$). However, the qualitative distinction between the two tasks that had been suggested in experiments one and two is not apparent. Simply, RT is faster under both perception and imagery contexts

ISBN 3-931852-67-9
ISSN 1617-6847

when the target belongs to the piece rather than being a misfit ($F_{1, 12}$ = 82.7, p = 0.000001). This highly significant 'priming' effect is remarkable given that participant attention was focused on the phoneme discrimination rather than the purely implicit relationship (fit) of the target to the heard or imagined context.

5. CONCLUSIONS

As the experiments adapted priming techniques, response to the variable of interest was not consciously mediated, allowing a comparison of the implicit processing proper to perceived and imaged musical representations. These experiments also examined the explicit mental representation of music: the comparison of results from 'inference' and imagery conditions in experiments two and four suggests that a conscious musical image of the experimental material facilitated task response. Participant feedback in the form of post-experiment interviews (experiments one to four) indicates that musical imagery was experienced, but that the melodic elements of pitch and timing seem to be a stronger dimension of the conscious mental representation of music than timbre.

In the first two experiments, the measurement of response time enabled the sensitive detection of behaviour that was different in kind (and not simply magnitude), when imaging and perceiving music. However, this result was not significant, and seemed to diminish in inverse relation with the 'absolute' character of the discrimination task. In other words, the greater the independence of the discrimination parameter from the musical context, the more similar the behaviour observed when perceiving and imaging. This suggestion has an important bearing on theories of equivalence between perception and imagery. Intons-Peterson (1992) spoke of the inadequacy of such theories to account for the ancillary features of auditory imagery such as loudness, timbre, and kinaesthetic dimensions. She suggests that the more an experiment requires participants to draw on 'real-world knowledge', the more the results from perception and imagery tasks will be dissimilar. The current research agrees with this insofar as more sophisticated musical tasks elicited qualitatively different perception and imagery behaviour, while 'absolute' phoneme judgements did not.

The method outlined in this paper is an effective means of comparing musical perception and imagery. However, this comparison is clearly dependent on the relationship of the discrimination task to the context. Future imagery research should attend carefully to the balance of implicit and explicit cognition required by the experimental task.

6. REFERENCES

1. Bigand, E., Tillmann, B., Poulin, B., D'Adamo, D. A., & Madurell, F. (2001). The effect of harmonic context on phoneme monitoring in vocal music. *Cognition*, 81, 11-20.

2. Godøy, R. I. & Jørgensen, H. (2001). *Musical Imagery.* Lisse: Swets & Zeitlinger Publishers.

3. Halpern, A. R. (1988). Mental scanning in Auditory Imagery for Songs. *Journal of Experimental Psychology*, 14 (3), 343-443.

4. ___. (1992). Musical Aspects of Auditory Imagery. In D. Reisberg (ed.) *Auditory Imagery.* Hillsdale, NJ: Lawrence Erlbaum Associates. pp. 1-27.

5. Intons-Peterson, M. J. (1980). The role of loudness in auditory imagery. *Memory & Cognition,* 8 (5), 385-393.

6. ___. (1992). Components of Auditory Imagery. In D. Reisberg (ed.) *Auditory Imagery.* Hillsdale, NJ: Lawrence Erlbaum Associates. pp. 45-71.

7. Levitin, D. J. & Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception and Psychophysics*, 58 (6), 927-935.

8. Repp, B. (2001). Expressive Timing in the Mind's Ear. In R. I. Godøy, & Jørgensen, H. (eds.). *Musical Imagery*. Lisse: Swets & Zeitlinger Publishers. pp. 185-200.

9. Weber, W. J. & Brown, S. (1986). Musical Imagery. *Music Perception*, **3** (4), 411-426.

ISBN 3-931852-67-9
ISSN 1617-6847