Perception, 2014, volume 43, pages 825–828

doi:10.1068/p7787

SHORT AND SWEET

Is he playing solo or within an ensemble? How the context, visual information, and expertise may impact upon the perception of musical expressivity

Donald Glowinski[1], Arianna Riolfo[2], Kanika Shirole[1], Kim Torres-Eliard[1], Carlo Chiorri[2], Didier Grandjean[1]

[1] Swiss Center for Affective Sciences (CISA), University of Geneva, Campus Biotech, Geneva, CH-1211, Switzerland; [2] DISFOR, Università degli Studi di Genova, Department of Educational Sciences, Genoa, Italy; e-mail: donald.glowinski@unige.ch

Received 6 May 2014, in revised form 18 August 2014

Abstract. Visual information is imperative when developing a concrete and context-sensitive understanding of how music performance is perceived. Recent studies highlight natural, automatic, and nonconscious dependence on visual cues that ultimately refer to body expressions observed in the musician. The current study investigated how the social context of a performing musician (eg playing alone or within an ensemble) and the musical expertise of the perceivers influence the strategies used to understand and decode the visual features of music performance. Results revealed that both perceiver groups, nonmusicians and musicians, have a higher sensitivity towards gaze information; therefore, an impoverished stimulus such as a point-light display is insufficient to understand the social context in which the musician is performing. Implications for these findings are discussed.

Keywords: music, visual perception, musicians and nonmusicians

1 Introduction

Paradoxically as it might sound, people depend primarily on visual information when making judgments about music performance in concerts while consistently reporting that sound is the most important source of information in evaluating performance in music (Tsay, 2013). Platz and Kopiez (2012) conducted a meta-analysis of 15 aggregated studies on audiovisual music perception and concluded that the visual component is not merely a marginal phenomenon in music perception and it provides important information in the communication of expressivity. Visual information is also an integral part of inferring communicative processes within an ensemble—for example, a nod of the head would indicate a synchronous start (Palmer & Deutsch, 2012) or gaze interaction would be used to capture coperformers' attention (Thompson, Graham, & Russo, 2005). Thus, the evaluation of the expressive and communicative aspects in the music context may depend upon the available bodily information during the performance.

Drawing upon the methodological approach developed by Juslin (2013), we formalize the process by which one understands how available body expression and the context of the musical performance are used to capture crucial information (see figure 1). By investigating how the cues related to the performance map onto the perceptual cues of the perceivers, our study offers some insight into the efficacy of the communication process starting from available sensory information.

The present study makes three strategic decisions to evaluate the perception of coordination and emotional expressivity interactions within the context of music ensemble. First, it adopts a comparative approach: comparing a performing musician's behavior while playing in a string quartet and while playing solo. Second, it considers two types of visual

stimuli, a video and a motion capture (mocap) display that allow identifying the relevant perceptual information. Third, it evaluates expert musician and nonmusician perceivers and aims at characterizing their respective perceptual strategies to identify the social context of the performance (solo vs ensemble) and to evaluate the musician's expressivity.

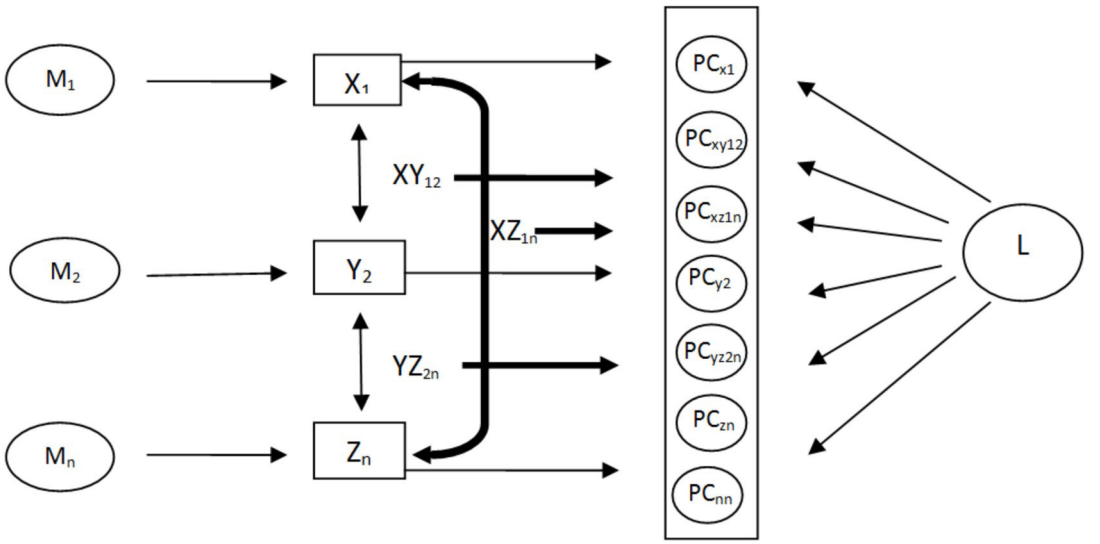

Figure 1. The revisited framework affords us the opportunity to systematically examine the impact of music performance contexts in the production and perception of the expressive behavior of musicians (adapted from Juslin, 2013). Notes: M: musician(s); X, Y, Z: expressive cues produced by musicians' behavior; XY, YZ, XZ: interactions of these cues due to the music performance contexts; PC: perceived cues by listener (L) on which are based the attributions of expressivity.

2 Experimental design

Forty-nine participants (F = 39%, mean age 29±13 years, musicians = 61%) were randomly assigned to watch either audiovideo recordings or point-light displays based on the collected motion capture of a violinist (see figure 2). In a random half of the trials the musician was playing solo, and in the other half was playing with other musicians of a string quartet; but this information was not available to participants, who had to report whether they thought the performance was solo or ensemble and how expressive they judged the performances to be. Grounded in signal detection theory, a number of accuracy indices [area under the curve (AUC), sensitivity, specificity, predictive power] were computed. Factorial analysis of covariance models were also specified including age, gender, and self-reported empathy level as covariates and expertise (musician vs nonmusician) and display (video vs mocap) as factors.

Both nonmusicians [mean AUC = 0.59, 95% confidence interval (CI): 0.54–0.64] and musicians (mean AUC = 0.66, 95% CI: 0.61–0.70) performed statistically better than chance (AUC = 0.50). Results also showed that sensitivity to solo performance (p = 0.011, effect size r = 0.35) and ensemble predictive power (p = 0.012, r = 0.36) were statistically higher in musicians compared with nonmusicians. Sensitivity to ensemble (p = 0.024, r = 0.32), overall accuracy (p = 0.016, r = 0.33), and AUC (p = 0.019, r = 0.33) were higher in the video condition compared with the mocap condition for either group (see figure 3). The inspection of a posteriori tests in the display by an expertise interaction effect revealed two main outcomes: (i) musicians performed better than nonmusicians in the mocap condition but not in the video condition, as revealed by the higher sensitivity to solo (p = 0.010, r = 0.38)

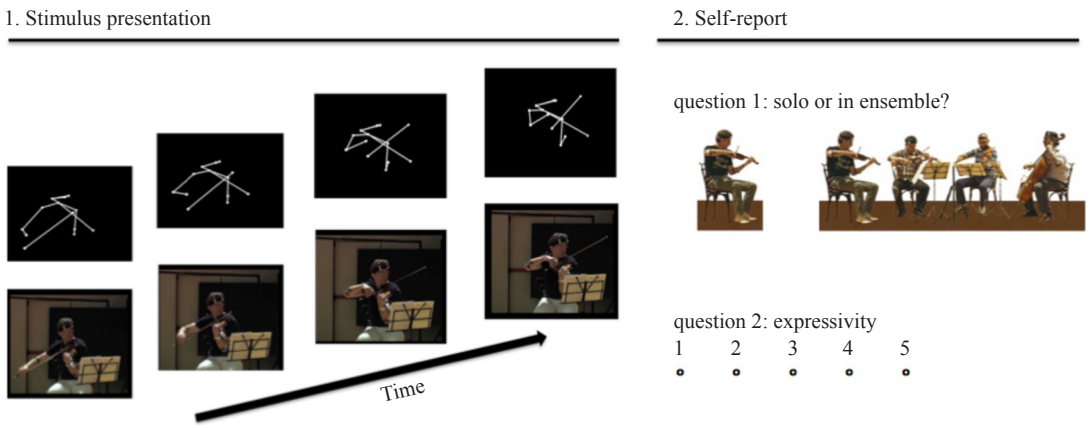

Figure 2. [In color online, see http://dx.doi.org/10.1068/p7787] Illustration of the experimental protocol used for the present study. Each stimulus duration was about 30 s. For each level of expertise (musician vs nonmusician) a first group was watching a video and a second group of perceivers was watching a point-light display based on the motion capture data of the musician performance. Each experimental condition consisted of 20 different trials showing the violinist performance achieved in isolation ($n = 10$) and with the group ($n = 10$). A total of 980 rated clips was considered in calculating the perceivers' indices.

and ensemble predictive power ($p = 0.005$, $r = 0.41$), and (ii) nonmusicians performed better in the video condition than in the mocap condition, as indicated by higher overall accuracy ($p = 0.039$, $r = 0.31$) and AUC ($p = 0.037$, $r = 0.32$). Summarizing, results showed that the accuracy of perceivers' judgment in distinguishing the two conditions (solo vs ensemble) appears to be substantially affected by the kind of display and, to a lesser extent, by their expertise (see figure 3). Usual audiovideo recordings may contain a number of visual features (eg gaze) which might be helpful in the perceptual task, whereas such features are missing in impoverished stimuli such as point-light stimuli.

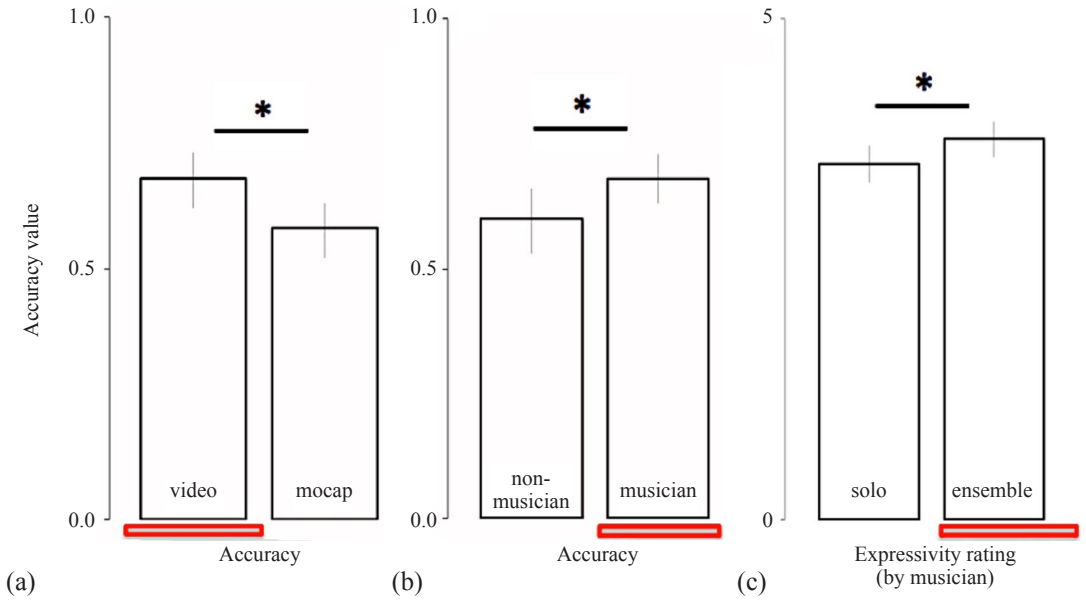

Figure 3. Accuracy of distinction between solo and ensemble performance and expressivity ratings. The histogram represents the value of accuracy indices when (a) considering video stimuli versus a point-light display derived from motion capture data of the violinist (mocap) and (b) when comparing nonmusicians and musicians. (c) The interaction effects of expressive ratings × social context (solo vs ensemble) for musicians. Asterisks denote significant differences.

With regard to the perception of expressivity, linear mixed modeling revealed no significant main effects of the difference between the solo and the ensemble conditions. However, we found that experts rated the performance as more expressive if they believed (rightly or wrongly) that the performance was in ensemble (see figure 3), whereas nonexpert perceivers showed the opposite pattern. These results seem to reflect different a priori beliefs of perceivers according to their level of expertise, and support the hypothesis that evaluation of emotional expressivity in music may depend upon the expertise and the perceived social context.

3 Conclusions

The results support the hypothesis that both musician and nonmusician perceivers evaluate the expressive behavior of a violinist differently. Using the visual stimuli presented, musicians can correctly identify social cues and expressive behaviors with greater ease than their nonexpert counterparts. For example, when the violinist is playing in the group, they utilize their skills to recognize some of the signs of communication between musicians that are not clear to nonmusicians, and this seems to help them to correctly identify solo conditions with a higher frequency of success than nonmusicians. Besides this, they also have less difficulty in the point-light display conditions. With respect to the existing literature, our results provide a further insight into the importance of rich visual information when evaluating expressivity and the social context of music and how expertise influences such perception.

Acknowledgment. Thanks are given to the School of Music Conservatorio Paganini for granting access to the school to record the participants' performances.

References

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, **10**, 235–266.

Palmer, C., & Deutsch, D. (2012). Music performance: Movement and coordination. In D. Deutsch (Ed.), *The Perception of Music* (3rd ed., pp. 405–422). Amsterdam: Elsevier.

Platz, F., & Kopiez, R. (2012). When the eye listens: A meta-analysis of how audio-visual presentation enhances the appreciation of music performance. *Music Perception: An Interdisciplinary Journal*, **30**, 71–83.

Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, issue 156, 203–227.

Tsay, C. J. (2013). Sight over sound in the judgment of music performance. *Proceedings of the National Academy of Sciences of the USA*, **110**, 14580–14585.