Title	Solo Vs Ensemble performance	
Question of interest	Are there specific non-verbal behavioral variables that may be automatically measured and that enable to distinguish between performing an action alone or jointly in a group?	
Leaders	UNIGE	
Other SIEMPRE groups involved	QUB, UNIGE-CH	
Referent scenario	Scenario 1: String Quartet	
Research objectives	1. Develop techniques for automated analysis of multimodal recordings of a musician's performance in two conditions: solo Vs ensemble performance. 2. Design a perceptual experiment to evaluate the difference between Solo Vs Ensemble performance conditions, using audiovisual recordings. 3. Identify a set of non-verbal cues that characterize the social behaviour of the musician: communicative gestures to regulate the ensemble performance, and continuous movement features enabling to distinguish between the two modalities. 4. Correlate the results of the perceptual experiments (participants' ratings) with the results from the automated behavioral analysis of musicians.	
Theoretical hypotheses	Playing music jointly with others may affect individual behavior. Joint performance requires strategies to cope with others' intentions and to adapt one's behavior accordingly. The success of the interaction may depend upon one's ability to anticipate and manage others' actions and ensure efficient group coordination. Techniques for automated analysis can be developed and assessed with perceptual ratings: external observers may be able to identify through a set of non-verbal cues the social behavior of the performer.	
Operational hypotheses	There are non-verbal visible behavioural cues in music performance that may help an external observer to distinguish between a performance interpreted alone (solo) or within an ensemble. Two types of non-verbal cues can been distinguished: key gestures using upper-body parts (e.g., head gestures) to capture others' attention and to coordinate the ensemble (Davidson et al. 2006); non-verbal behavioral variations, which are continuous perturbation of movement. These behavioral cues may refer to implicit adaptation and co-ordination process of musicians during the performance (Glowinski et al. 2011).	
Relationship with the objectives of the project	Investigate social behavior in music performance and identify the set of non-verbal cues explaining the	

	phenomenon.	
Time schedule	• Multimodal recordings at UNIGE in Spring 2011 (student quartet Music Conservatory, See D2.1); • Multimodal Recordings in July and September 2011 with Quartetto di Cremona (UNIGE); • Data analysis and perceptual experiment • Perceptual ratings of the videos of Quartetto di Cremona (Solo Vs Ensemble conditions, blind rating), Spring – Summer 2012; subjects ratings performed at UNIGE-CH and UNIGE. • Comparison of subjects' ratings with results from automated analysis.	
Methods	Automated analysis techniques described in D1.3.	Link to D1.3
Participants	• Data recordings: String Quartet of Music Conservatory; Quartetto di Cremona. • Subjects ratings: Students from UNIGE-CH (spring 2012) Students from UNIGE (summer-fall 2012)	
Materials	• Music score .a 2 min music piece by Schubert, No.14 in D minor, "Death and the Maiden", Allegro; the music score has been segmented into five musical phrases. • Musicians' post-hoc ratings of their performance After each performance, musicians rated the level of performance satisfaction, expressivity and group cohesion on a discrete 7-point Likert scale. • Ratings of participants to the perceptual experiment _____ • Video (High Definition JVC cameras YHD251 camcorder) Videos recorded from a tripod placed in front of the scene of solo performance of violin 1 and 2 and of the string quartet • Medium Shot Video (High Definition JVC cameras GYHD251 camcorder) Videos of solo performance of violin 1 from waist up.	Link to data

_____

- Audio Ambient mic

Audio recorded from stereo pair microphones of solo performance of violin 1 and 2 and of the string quartet.

- Audio Single mic

Audio recorded using piezoelectric-microphones attached to the instrument.

_____

- MoCap data

behavioural data of solo performance of violin 1 and 2 and of the string quartet. Reflective markers of the Qualisys system (www.qualisys.com) were placed in upper-part body joints, Figure xx show the 3D points corresponding to the 16 markers placed on the musicians' joints and the 3 markers placed on the instrument.
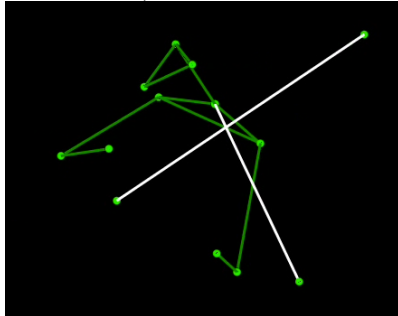

Marker label (for violin 1)
1- LFHD left head
2- RFHD right head
3- OZ   back head
4- C7   neck
5- LSHO left shoulder
6- RSHO right shoulder
7- LELB left elbow
8- RELB right elbow
9- LWRB left wristband
10- RWRB right wristband
11- SACR back
12- LASI left hip
13- RASI right hip
14- IVIBowUp

	15- IVIBowDown 16- IVISc Original real-time applications based on the EyesWeb XMI software platform have been developed to synchronize the Qualisys MoCap data together with video and audio data. Samples can be seen in the media section of the EU ICT FET SIEMPRE web pages (www.siempre. infomus.org). • MoCap rendering MoCap rendering contains the mocap of solo performance of violin 1. Mocap data has been synchronized with single piezzo mic audio using the smtpe time code, visualized and exported as .avi format. The Medium shot view (from waist up) has been selected to compare the rendered mocap with the video file (Medium Shot Video). • SMPTE The smpte timecodes provide a time reference for synchronization of the multimodal data. SMPTE contain binary coded decimal *hour:minute:second:frame* identification, using 25 frames/sec. • SMPTE list The list of SMPTE timecodes corresponding to the onset/offset of the performance start and end point of each trial and of each of the five music segment within each trial. (<mark>see also D2.1 First series of experiment</mark>).	
Data format	• Music score (.pdf) • Musicians' post-hoc ratings (excel) • Ratings of participants (excel) • Video (.avi, 720 x 576, 25 fps pro scan) • Medium Shot Video (.avi, 720 x 576, 25 fps pro scan)	

	• Audio Ambient mic (.wav,48000 Hz,16Bit) • Audio Single mic (.wav, 48000 Hz,16Bit) ——————— • MoCap data (.qtm, c3d, .tsv, 120 fps) • MoCap rendering (.avi) ——————— • SMPTE (.wav) • SMPTE list (.doc)	
Experimental protocol/procedure	<u>Musicians' recordings</u> The four musicians of the Quartetto di Cremona was invited in **two** sessions to play a fragment of classical music of duration of about 2min (Allegro of the String Quartet No 14 in D minor, known as Death and the Maiden, by F. Schubert). Performances were of three types: 1) solo of first violin 2) solo of second violin and 3) ensemble performance. <u>Subjects ratings:</u> Each participant was presented with a set of 60 samples selected from the full set of audio-video recordings of the first violin's performance. A procedure based on random permutation of pre-established lists of samples ensured that the Solo and Ensemble conditions as well as the five musical segments be presented with the same frequency. The whole procedure consisted in three main phases: 1) After each audio-video sequence, the participants had to report whether they reckoned the performance being a solo or an ensemble one and then to rate their level of confidence in the correctness of their answer using a visual analogic continuous scale (from 1 to 100). 2) The second part of the questionnaire investigated the participants' perception of the musician's expressivity and expressed emotions. They were asked to assess the level of expressivity and the level of expressed emotions of the performance by rating the 9 GEMS dimensions. 3) At the end of the session, the participant was asked to report which musician's body features (e.g., head, arm, instrument movement) she/he most focused on to assess the performance.	
Measures	• Automated multimodal analysis:	

	- Distance of the musician's head with respect to the *ear*, i.e. the subjective center of the string quartet - Regularity of the distance variations between head and quartet's *ear* measured with *Sampen* method • Musicians and Participants' ratings	
Results	Automated analysis Empirical evidence shows that SampEn values of musician head distance with respect to the string quartet's *ear* can account for the difference between Solo Vs. Ensemble conditions. Playing with others decreases the entropy of human movement related to a point situated in space, which has a social value (the ear stands as common spatial landmark to facilitate joint action). It is thoroughly logical that someone who is part of a joint action tends to make her behavior more regular: it facilitates a global alignment of the ensemble. This result is independent from the musician and from the music segment. This result confirms recent findings by Vesper et al. 2011. The authors observed that participants, who were instructed to coordinate key presses in a two-choice reaction time task, decrease the variability of their actions in a joint context compared with the same task performed individually. A hypothesis suggested by the authors is that reducing variability, hence increasing behavioural regularity, enables achieving better predictability Perceptual Experiment The experimental data collected so far using audio-video recordings have indicated that non-expert participants may have difficulties in distinguishing two modalities of interpretation of a first violinist: when playing alone (solo) and when playing with the other musicians of a string quartet (ensemble). However, the analysis of the participants' ratings, including their evaluation of musician's expressivity and emotions, seemed to suggest original strategies for decoding social behavior: when perceiving the Ensemble condition, participants tended to be sensitive to the music segment where the first violinist has clear leadership and they tended to assess correctly identified solo and ensemble performances with higher ratings of Nostalgia and Sadness.	
Descriptive results		
Inference statistics		
Additional results	-A journal paper submission on automated analysis is in preparation	

	- On the perceptual experiment, a conference paper has been published: **Can naïve observers distinguish a violinist's solo from an ensemble performance? A pilot study** Glowinski, D., Torres-Eliard, K., Chiorri, C., Camurri, A., Grandjean, D. Third International Workshop on Social Behaviour in Music at ACM ICMI12, October 22-26, 2012, Santa Monica, USA.	
Discussion	<u>Automated Analysis</u> Additional evaluation could be envisaged to assess explicitly how behaviour regularity facilitates temporal coordination in String Quartet. Recent work focusing on entrainment in small music ensemble (e.g., duet, quartet) use quantitative methods such as recurrence plot analysis to evaluate the degree of synchronization between musicians. Correlation analyses between the synchronization indexes and entropy through SampEn could help in assessing whether such relationship between reducing variability and increasing coordination exists in the string quartet. Another question of interest is the following: even if the observed coordination between musicians is intentional, it is still not clear whether musicians rely on explicit knowledge of the relation between variability and coordination performance or whether they were using this strategic relation consciously. Actually, people may not plan to change their own behaviour in this specific way to enable their co-actor predicting better their upcoming actions. As pointed by Vesper et al. 2011, they may rather form a general intention to be as coordinated as possible, "triggering a particular modus operandi of the action system that rendered the timing of actions less variable" (p.529). Another result of interest from our study is the difference observed between the two musicians. Both tend to increase their behaviour regularity when playing in the ensemble. However, this difference is higher in the case of musician n°2. These differences in behavior regularity may be related to the role adopted by musicians within the quartet. <u>Perceptual Experiment</u> Future work is needed and may include the use of point-light displays of the first violinist based on the collected motion capture data during the recordings. This new material, which captures in more detail the kinematic features of the performance, should enable to	

achieve a better understanding of the behavioral cues used by the participants. Other possible tracks for future research may include some changes in the procedure used to collect participants' data, such as: (i) addressing one modality at a time to have a better control on behavioral cues that have effect on participants' ratings; (ii) addressing experts (creation of focus group) and (iii) correlating the results of the perceptual experiments (participants' ratings) with the results from the automated behavioral analysis of musicians.