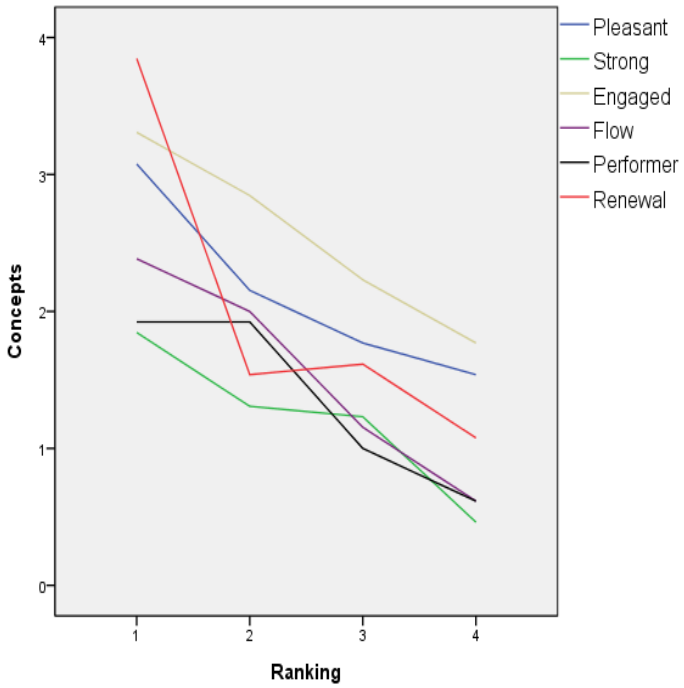
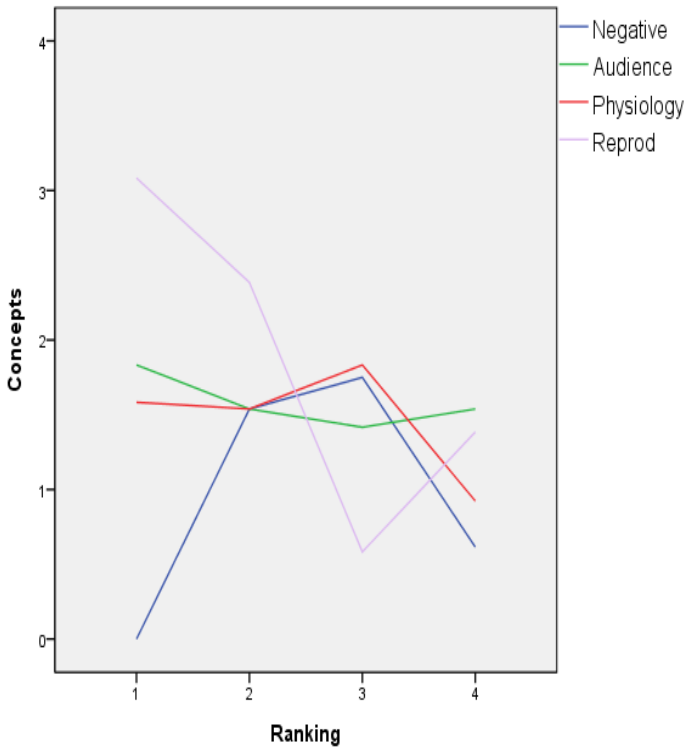
Title	Multimodal investigation of audience responses to live musical performance
Question of interest	Using different techniques to assess the dimensions of audience engagement in live performance and the eMAP features relevant to them.
Leaders	QUB
Other SIEMPRE groups involved	
Referent scenario	Scenario 3: Audience
Research objectives	This series of experiments implemented a full multimodal experiment schedule to investigate audience responses to different live music scenarios. The aim was to give us an indication of which measures are most informative and influential in determining audience enjoyment of live music performance, and if there are are inter-relationships between measures at different levels (psychological, kinetic, physiological, etc.). Because some of the measures are time-varying, relationships may include synchronies between measures and their relationship to the performance. The measures were subjective response (Quality of Experience questionnaire and continuous mechanism), physiology (GSR and pulse), motion capture and post recording video rating.
Theoretical hypotheses	Measures will be able to discriminate between different performances within and between concerts, and synchronies between different measures will be visible at certain points throughout the performances. Hence ideally, measures will be able to discriminate between differing levels of an audience's engagement, and will show congruence whilst doing so.
Operational hypotheses	There will be a significant effect of liking/engagement on all measures in the experiment. There will be correlations and synchronies visible between the continuous measures employed in the experiments (physiology, subjective response, motion capture and post recording video rating)
Relationship with the objectives of the project	The series of experiments aimed to establish the framework for large multimodal experimentation in a live music performance environment, a key aim for SIEMPRE.
Time schedule	Experiment 1: May 2011 Experiment 2: Dec 2011 (the focus of this report) Experiment 3: Jan 2012 Experiment 4: Mar 2012 (at *Sonorities*) Experiment 5: Mar 2012 (at *Sonorities*) Experiment 6: Mar 2012 (at *Sonorities*) Experiment 7: Nov 2012

Methods	
Participants	The pilot studies had small numbers of participants (15-20) 55 participated in the main experiments, 18 with sensors and all with questionnaires. The pilot experiments featured a largely student population; the others were genuine concert-goers with a range of ages and backgrounds.
Materials (music)	The pilot and January experiments presented contrasting musical genres (Irish traditional and experimental electronic music), chosen to ensure that audiences gave contrasting responses. The Sonorities experiments presented three concerts, giving a wide variety of styles within the electro-acoustic genre.
Data format	Data is in a variety of formats. Video: avi Audio: wav Physiological & continuous self report: text files Motion capture: Qualysis Questionnaire: SPSS
Experimental protocol/procedure	The second experiment followed the same format as the first but adapted the design of the faders and used the shorter questionnaire derived from analysis of the first pilot. It is shown below.  In the experiment 14 audience members attended a concert with 4 separate and contrasting performances of 10 minutes each. Each audience member answer a shortened version of the QoE questionnaire used in the previous experiment, and manipulated a fader measuring their engagement with the performances throughout the concert.The concert was entirely experimentally controlled (participants and performers recruited by researchers), and set in a room in the SARC building at QUB. Two members of the audience were also attached to physiological measures of GSR and ECG predominantly as a technical test for future experiments in this series.

Measures	*Continuous Qualitative Response:* The interface for this is a slider device with a spring mechanism which requires increased force to move to higher values (negatively scaled). The participants were asked to rate their engagement. Following the first pilot experiment, the slider was concealed so that responses were not visible to onlookers. *Retrospective Questionnaire:* We employed two versions of the questionnaire, a long version in the first pilot and the first main experiment and a shortened version (based on analysis of data from the longer version) in other concerts. *Physiological Measures:* For the second pilot experiment two participants were fitted with Galvanic Skin Response and ECG sensors to test the correlation between continuous qualitative response and physiological data. For the subsequent three experiments we increased the number of participants with physiological sensors to twelve on the January concert and 18 in the Sonorities concerts. *Motion Capture:* In the January concert participants were fitted with a silver ball on a hairclip to track their head movements via a motion captures system (Qualysis). This was done to assess group synchronization. *Post-Recording Rating:* After the experiment external judges will study the video and audio of the experiment and rate the participants on levels of engagement using the continuous qualitative response mechanism. This remains to be done.
Results	
Descriptive results	<u>Questionnaire</u> The questionnaire data have been analysed and show that a modest number of dimensions capture most of the variability in the data. Logistic regression indicated that over 90% participants can be categorised on the basis of the responses. The motion capture data suggest that there was very little movement during any part of the concert, and we do not expect to find differences in that respect. Analysis of the physiological and slider data is under way. Results from the December experiment illustrate the issues that are revealed by the questionnaires. Most of the factors measured correlate with participants' overall ranking of enjoyment, as shown below.

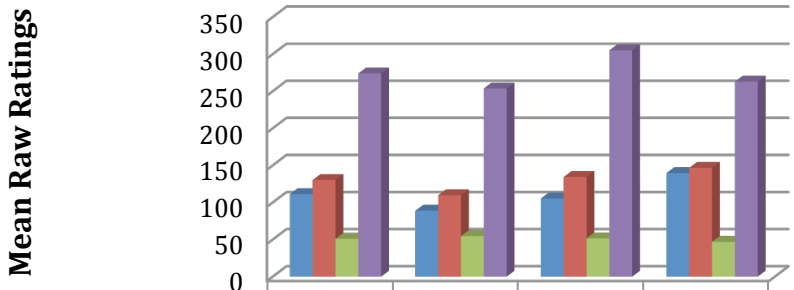


However, not all of the factors behave in this way. For example, as the graph below shows, there is an inverted U relationship between negative emption and overall enjoyment. When audience members were really negative about a performance, they did not feel negative emotion about it: they lacked emotion of any kind. Analysis of the other measures needs to be guided by this kind of information.



Faders

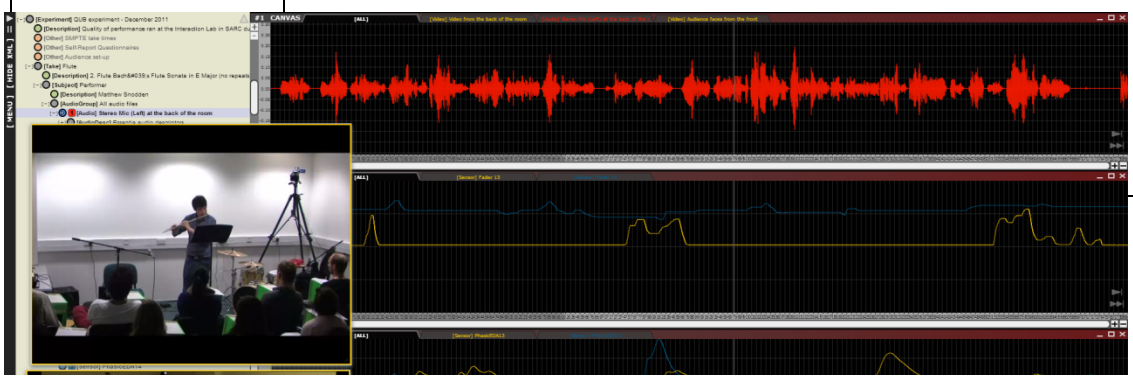
The basic requirement for the faders is that they can distinguish different levels of engagement with the performance, and given that they are continuous these can be both between and within performances. A simple way to show that they were successful in doing so is by dividing each performance into thirds and comparing the mean rating of each third. Doing so reveals that not only were certain performances preferred to others (and that this agreed with questionnaire data) but also that participants tended to grant a "grace period" of a few minutes before fully judging a performance (see figure below).



	Overall	First Third	Middle Third	Final Third
Classical	111.21	88.79	105.33	139.51
Singer-Songwriter	130.21	109.72	134.3	146.6
Traditional Irish	50.96	54.56	51.53	46.81
Experimental	273.84	253.66	305.07	262.8

Physiology & Synchronization

Once physiolical data had been extracted and pre-processed with the corect algorithms (as developed by SARC) preliminary ananlysis focused on finding moments of synchrony with the faders. This was done visually with the use of the RepoVizz tool as shown:



From this many instances could be found in which subjective responses matched physiological arousal, and we are currently working to categorize these and examine the relationship between the two measures.

Inference statistics

We show here the contrasts between responses to the two parts of the concerts in December and January respectively. They show that the questionnaire effectively captures the differences in response.

December (shortened questionnaire, experimental setting)

Concept	df	F	sig (p)
Emotion (pleasant)	4,51	4.355	.004*
Emotion (negative)	4,51	1.139	0.35
Strength	4,51	3.554	.013*
Engagement	4,51	3.529	.019*
Attention (audience)	4,51	2.542	0.052
Physiology	4,51	0.444	0.776
Presence	4,51	4.271	.005*
Reproduction	4,51	5.812	.001**
Performer	4,51	3.831	.009*
Renewal	4,51	2.308	.037*

January (full questionnaire, concert setting)

Concept	t	df	Sig.
Emotion	11.18	45	<.001
Social	4.239	46	<.001
Performer	9.086	45	<.001

	Attention	5.687	45	<.001
	Renewal	5.194	46	<.001
	Physiology	-2.096	45	0.042
	Presence	3.137	45	0.003
	Reproduction	8.026	42	<.001
	Aesthetics	6.665	44	<.001
Additional results				
Discussion	There are a number of important findings to come out of this experiment. Firstly it confirmed the viability of the contrasting performances with multimodal measures experimental scenario developed previously, and this design will be extended throughout all audience experiments in QUB. Performances during the same concert were shown to offer contrasting experiences and measures were successful in capturing this. With regard to specific measures the shortened QoE questionnaire was effective at discriminating between performances, though obviously not to the same degree as the longer version. Given that the short questionnaire is implementable in actual paid concert scenarios however it may prove even more useful for the goals of SIEMPRE. Faders were used far more than in the previous experiment, a phenomenon attributed to the obscuring boxes used here for the first time and a small training time with the participants just before the experiment began. They now show varied differences between and within performances, and will be used in the follow-up experiments. Although only 2 participants had physiological measures attached there were encouraging results of their interest to the project, as well as a confirmation of the technical ability to record the measures during the performance alongside everything else.			