KING'S College LONDON

CMPCP

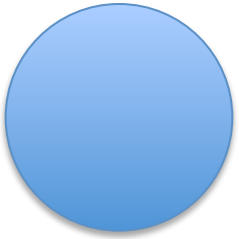
AHRC Research Centre for Musical Performance as Creative Practice

How musicians' and non-musicians' approaches to gestural representations of sound differ: findings from a motion-capture experiment

Mats Küssner

Department of Music, King's College London

AHRC Research Centre for Musical Performance as Creative Practice

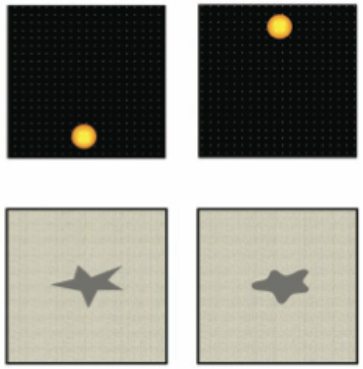


Cross-modal correspondences

- louder sounds associated with
  - larger objects
  - greater brightness
  - higher contrast
- higher pitch associated with
  - greater brightness
  - higher elevation in space
  - smaller objects
  - spikier shapes

Where do they come from?

- Innate

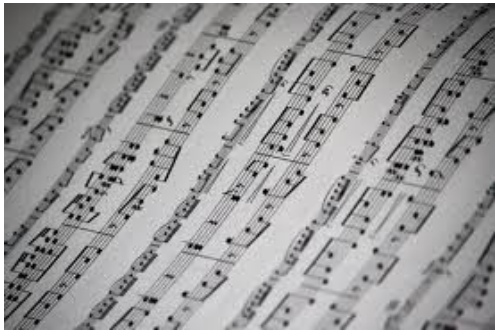


taken from Walker et al., 2010

taken from Ludwig et al., 2011

- Learned
  - (passive) statistical learning

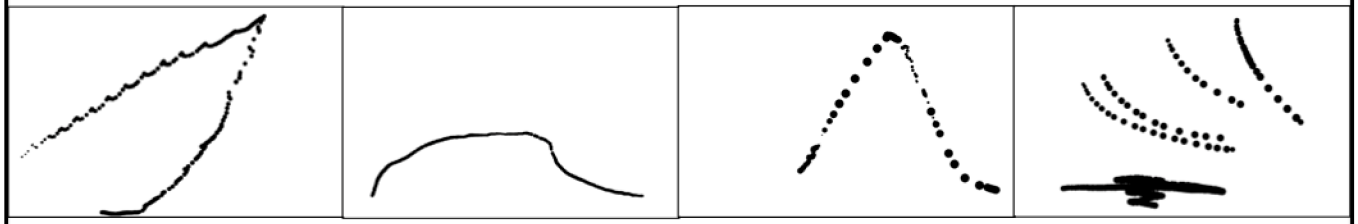
  - metaphor
  - specific training



Küssner (2013)

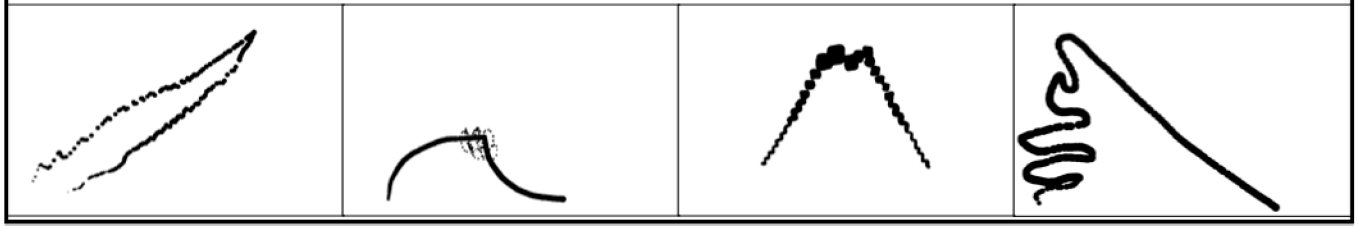
Performance part: sound stimulus 14 (pitch: up-down, amplitude: constant, tempo: decelerando-decelerando)

ppt_55 (non-mus)	ppt_02 (composer)	ppt_54 (visual artist)	ppt_24 (dancer)
glo_rho (pitch): 0.5216*	glo_rho (pitch): 0.9324*	glo_rho (pitch): 0.8773*	glo_rho (pitch): 0.4096*
glo_rho (loud): 0.3753*	glo_rho (loud): 0.5175*	glo_rho (loud): 0.4169*	glo_rho (loud): [0.1574*]
rho (pitch): 0.9601*	rho (pitch): 0.9762*	rho (pitch): 0.9875*	rho (pitch): 0.2433*
rho (loud): 0.3384*	rho (loud): -0.4141*	rho (loud): 0.0573	rho (loud): [0.1104]

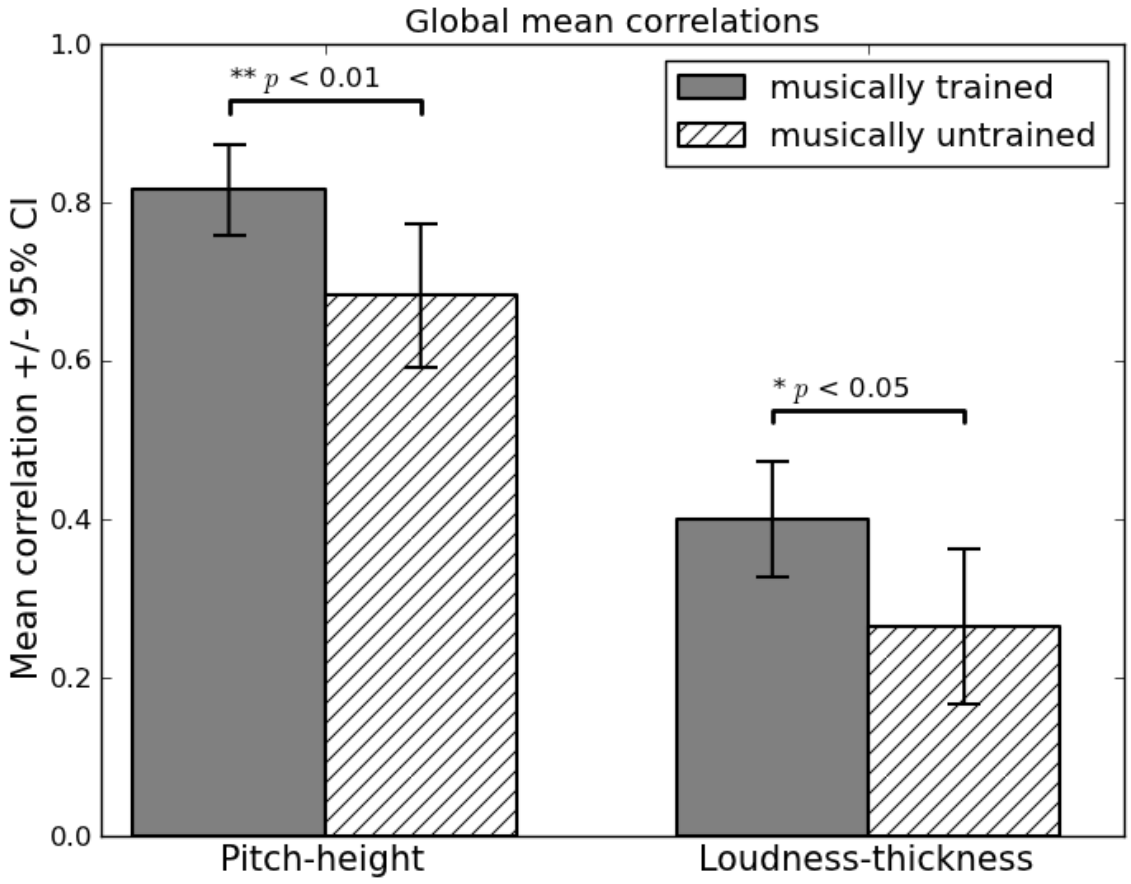


Contemplation part: sound stimulus 14 (pitch: up-down, amplitude: constant, tempo: decelerando-decelerando)

ppt_55 (non-mus)	ppt_02 (composer)	ppt_54 (visual artist)	ppt_24 (dancer)



Küssner & Leech-Wilkinson (in press)



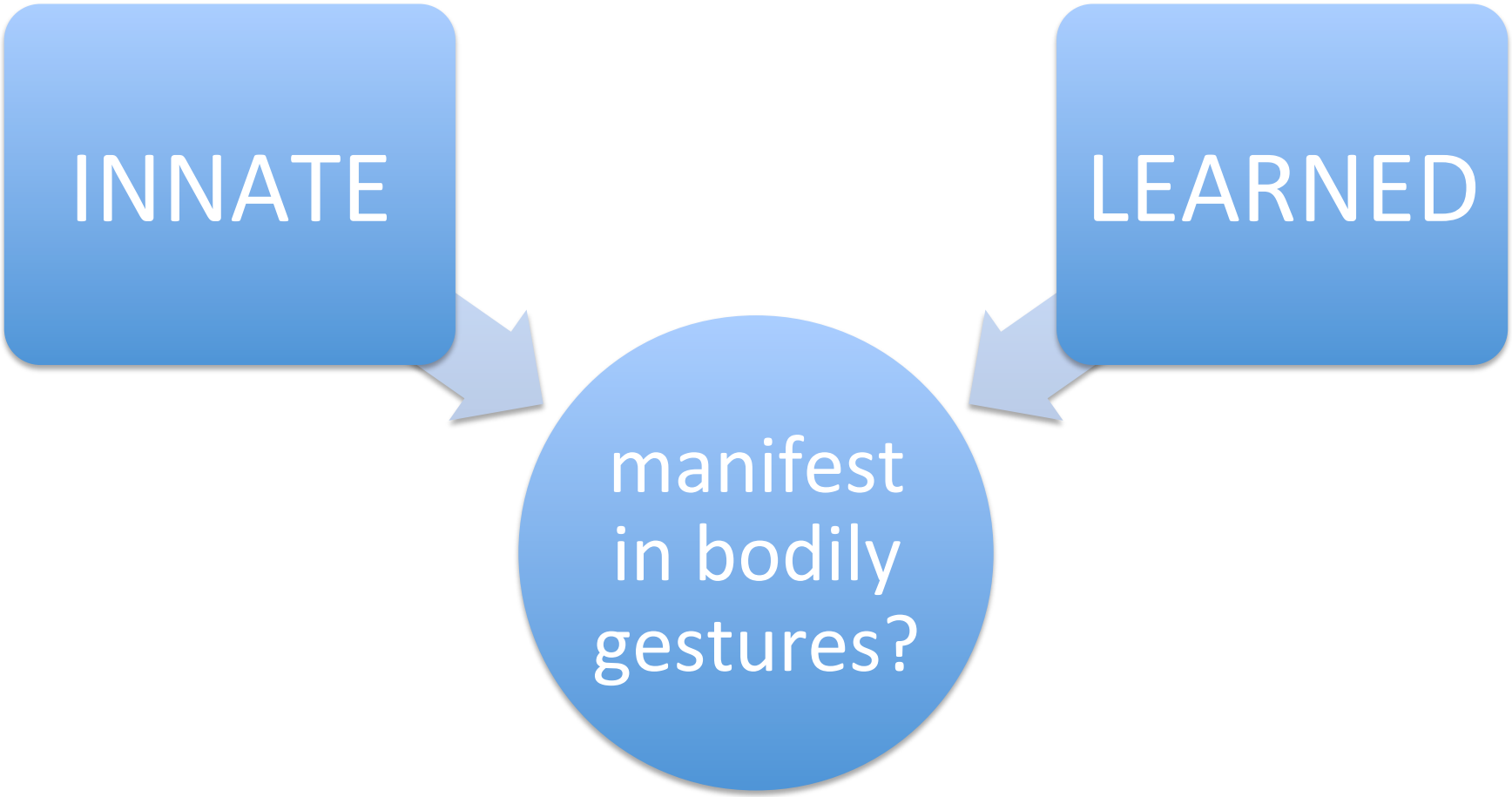
Interactions and Asymmetries

- Eitan & Granot (2006)
  - pitch is mapped onto all three spatial axes
  - pitch fall strongly associated with verticality, but pitch rise only weakly so
  - increasing loudness associated with approaching and accelerating motion (but not an ascent), while decreasing loudness associated with moving away and with descending motion.

Eitan & Granot, 2006

- "This article presents an empirical investigation of the ways listeners associate changes in musical parameters with physical space and **bodily motion**."

Cross-modal correspondences



Research Questions

How does musical training influence gestural cross-modal mappings of musical characteristics?

To what extent does a real-time visualization of the gestures influence these mappings?

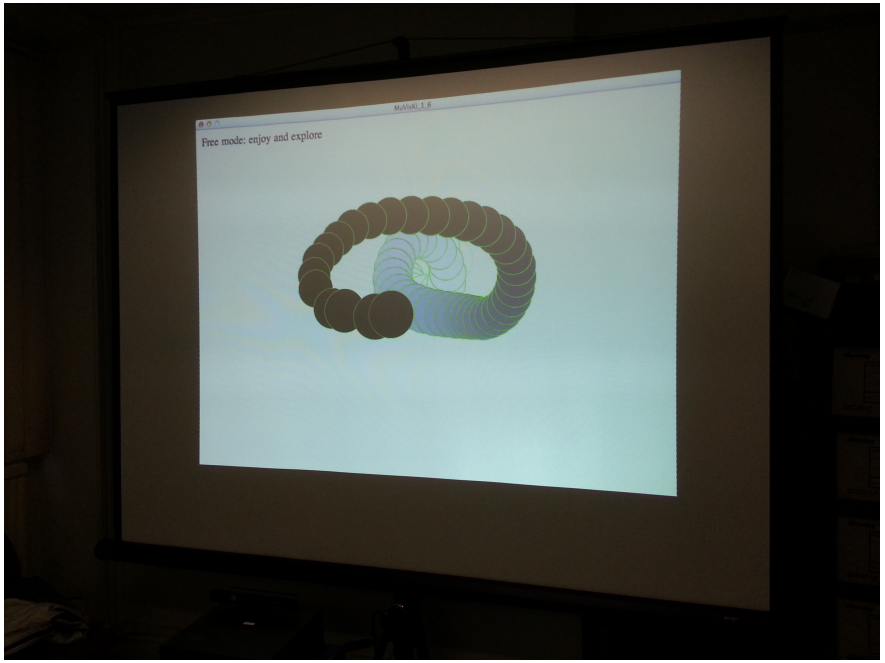
Participants

- 64 (32 female, mean age: 29.63 years [SD=12.49])
- 32 musicians (16 female, mean age: 30.09 [SD=13.66])
  - 8 keyboard / wind / string / composer
  - ≥ Grade 8 ABRSM, ≥ 4 hours per week
- 32 non-musicians (16 female, mean age: 29.16 [SD=11.39])
  - ≤ Grade 1 ABRSM, stopped playing more than 6 years ago and never played longer than 2 years

Stimuli

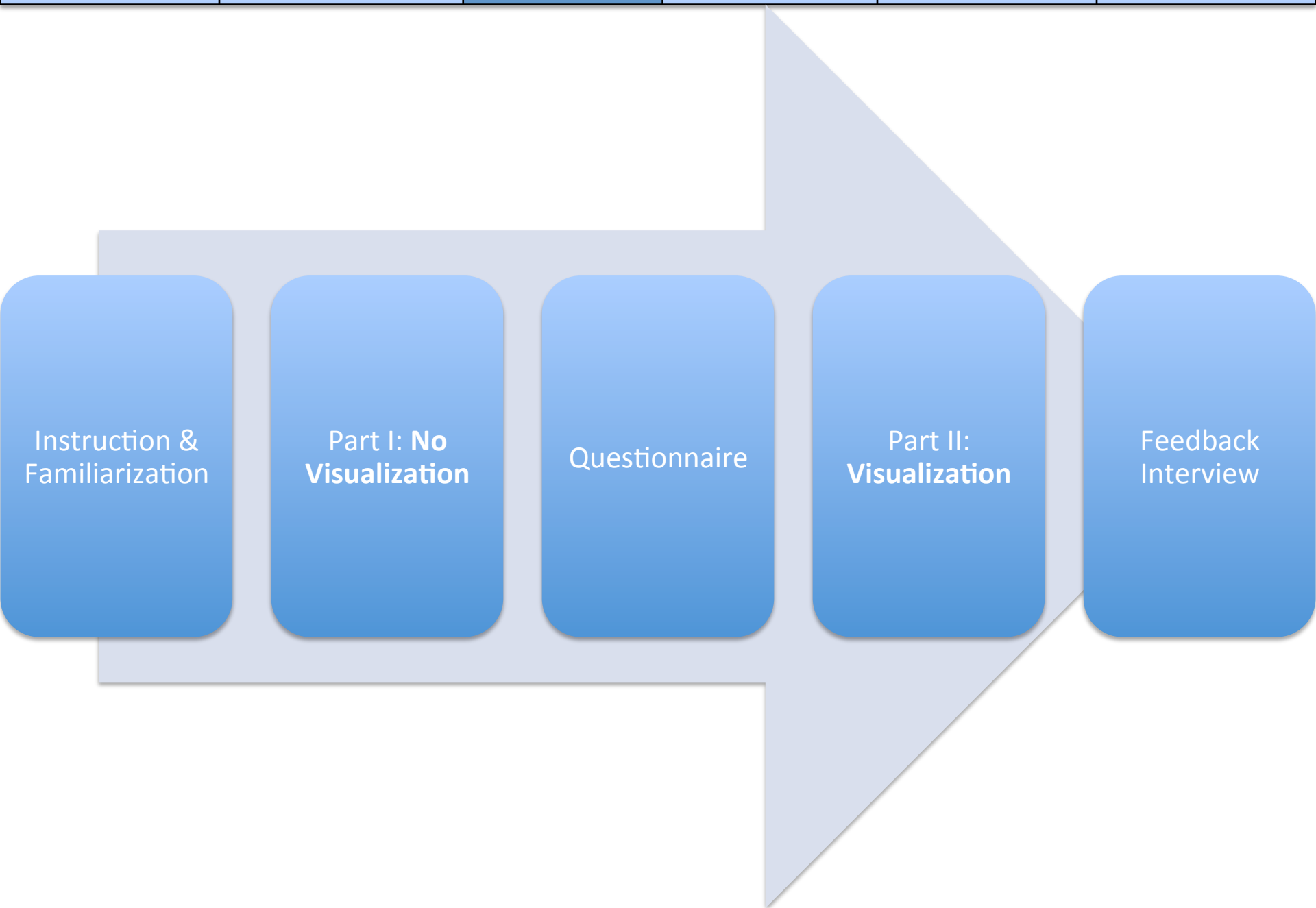
Table 1. Overview of experimental sound stimuli

No.	Length	Frequency (Note name)	Amplitude	Rate of frequency change
1	8 sec	constant (D4)	constant	N/A
2	8 sec	constant (D4)	decreasing - increasing	N/A
3	8 sec	constant (D4)	increasing - decreasing	N/A
4	8 sec	up - down (B2–D4–B2)	constant	equal
5	8 sec	up - down (B2–D4–B2)	constant	decelerando - decelerando
6	8 sec	up - down (B2–D4–B2)	constant	accelerando - accelerando
7	8 sec	up - down (B2–D4–B2)	decreasing - increasing	equal
8	8 sec	up - down (B2–D4–B2)	decreasing - increasing	decelerando - decelerando
9	8 sec	up - down (B2–D4–B2)	decreasing - increasing	accelerando - accelerando
10	8 sec	up - down (B2–D4–B2)	increasing - decreasing	equal
11	8 sec	up - down (B2–D4–B2)	increasing - decreasing	decelerando - decelerando
12	8 sec	up - down (B2–D4–B2)	increasing - decreasing	accelerando - accelerando
13	8 sec	down - up (D4–B2–D4)	constant	equal
14	8 sec	down - up (D4–B2–D4)	constant	decelerando - decelerando
15	8 sec	down - up (D4–B2–D4)	constant	accelerando - accelerando
16	8 sec	down - up (D4–B2–D4)	decreasing - increasing	equal
17	8 sec	down - up (D4–B2–D4)	decreasing - increasing	decelerando - decelerando
18	8 sec	down - up (D4–B2–D4)	decreasing - increasing	accelerando - accelerando
19	8 sec	down - up (D4–B2–D4)	increasing - decreasing	equal
20	8 sec	down - up (D4–B2–D4)	increasing - decreasing	decelerando - decelerando
21	8 sec	down - up (D4–B2–D4)	increasing - decreasing	accelerando - accelerando


Free mode: enjoy and explore

Gesturing conditions

- Instruction: represent sound gesturally while it is played
- "No Visualization": participants saw a white screen in front of them
- "Visualization": participants saw real-time visualization on screen in front of them
- Participants were presented with the same set of 18 stimuli in both conditions

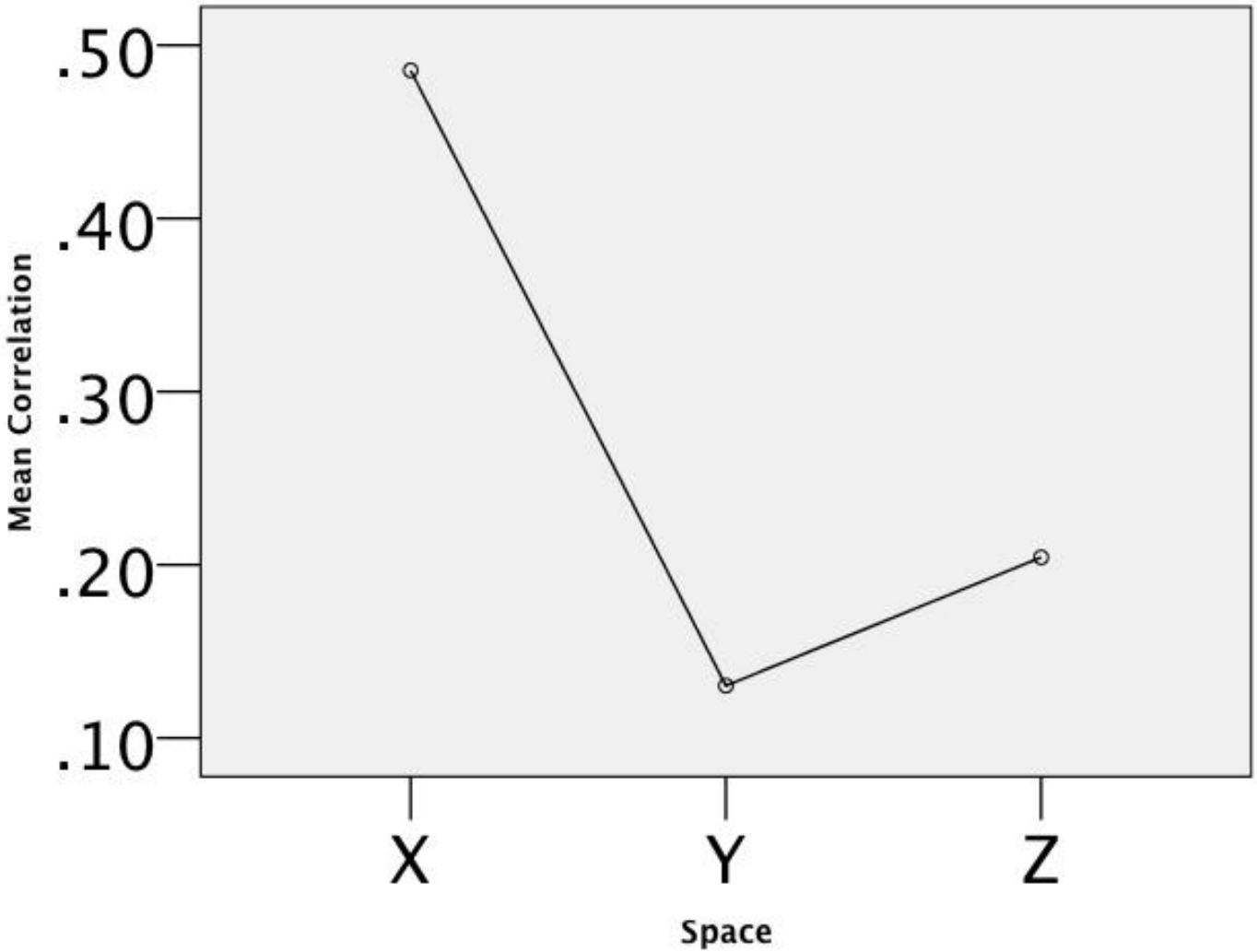

Instruction & Familiarization
Part I: **No Visualization**
Questionnaire
Part II: **Visualization**
Feedback Interview

Analysis

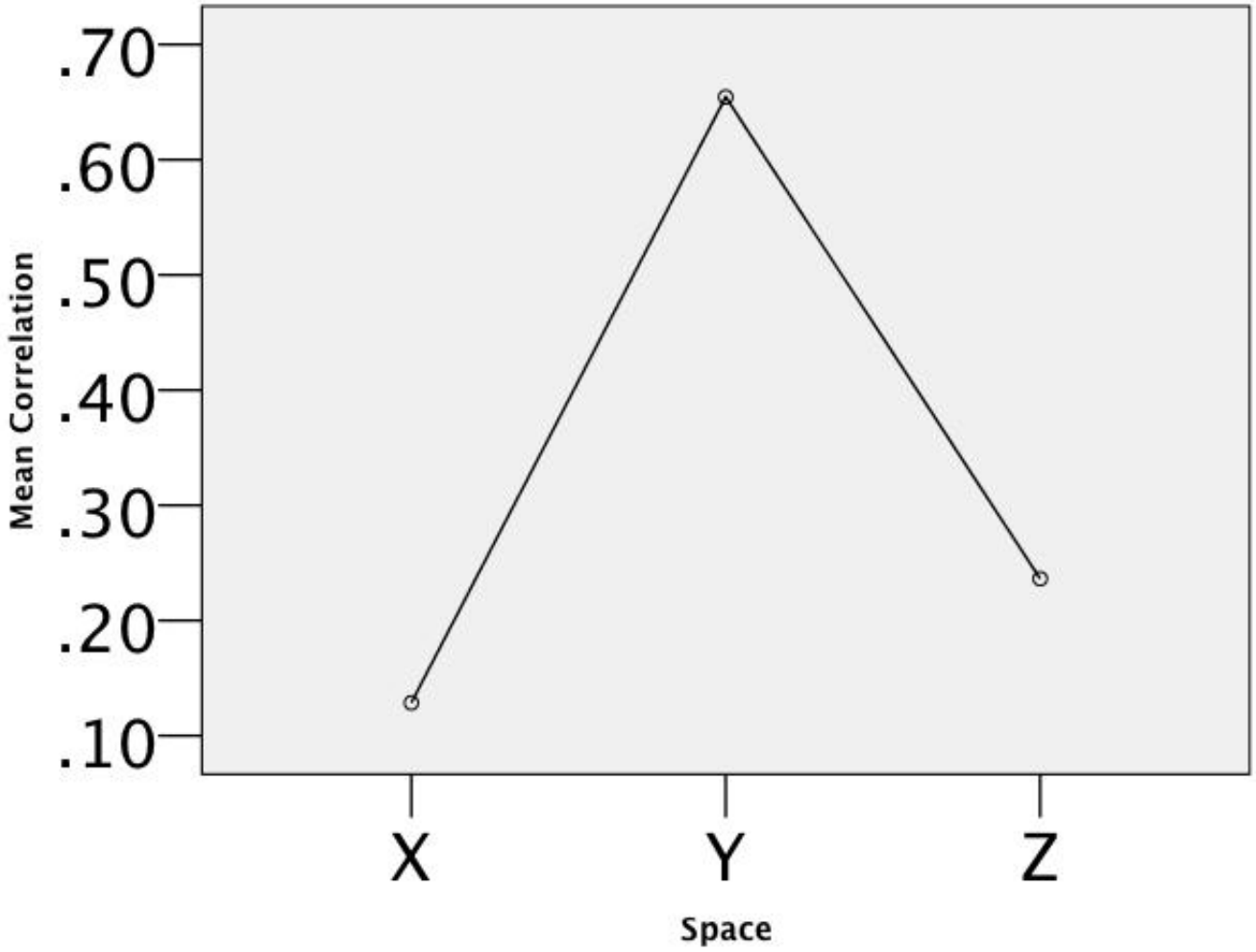
- Global correlations between sound characteristics (pitch, loudness, [time]) and movement (X, Y and Z)
  - resulting in 9 correlation coefficients per participant
- mixed ANOVAs for each sound characteristic with within-subjects factors ‘space’ and ‘vision’ and between-subjects factor ‘musician’
  - dependent variable: correlation

Which spatial axes did participants use to represent (elapsed) time, pitch and loudness?

TIME: main effect of 'space'



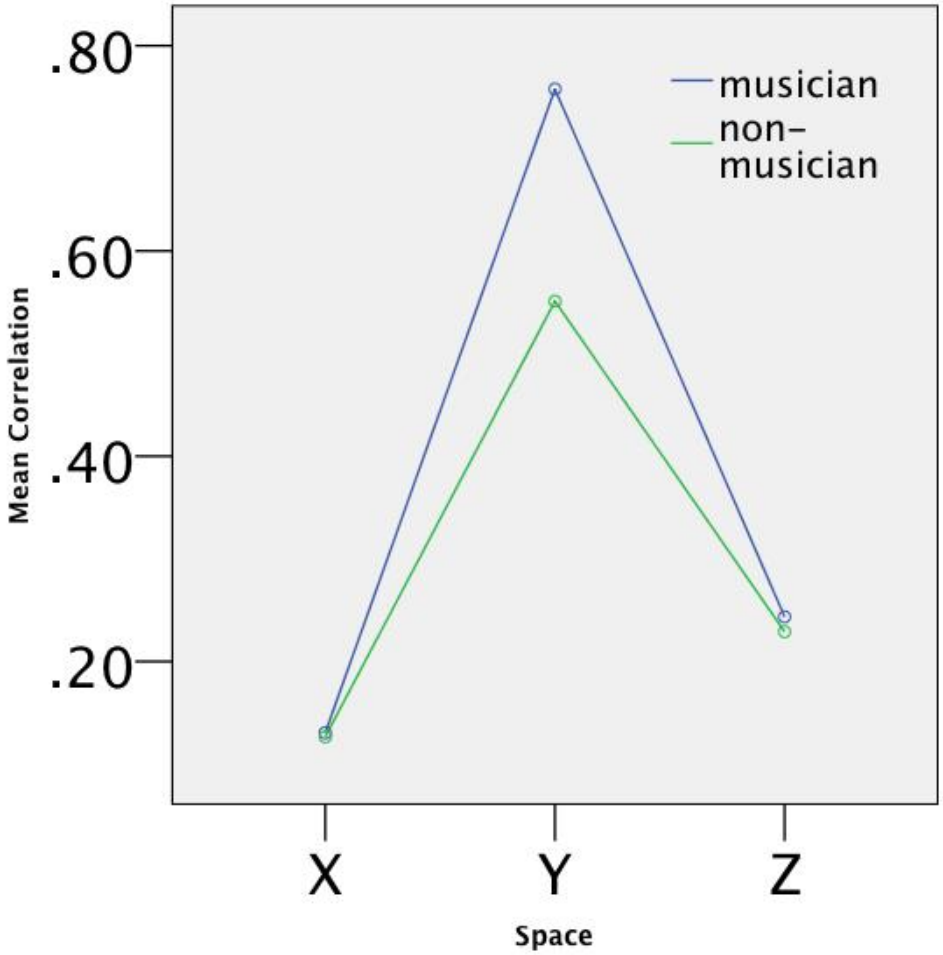
PITCH: main effect of ‘space’



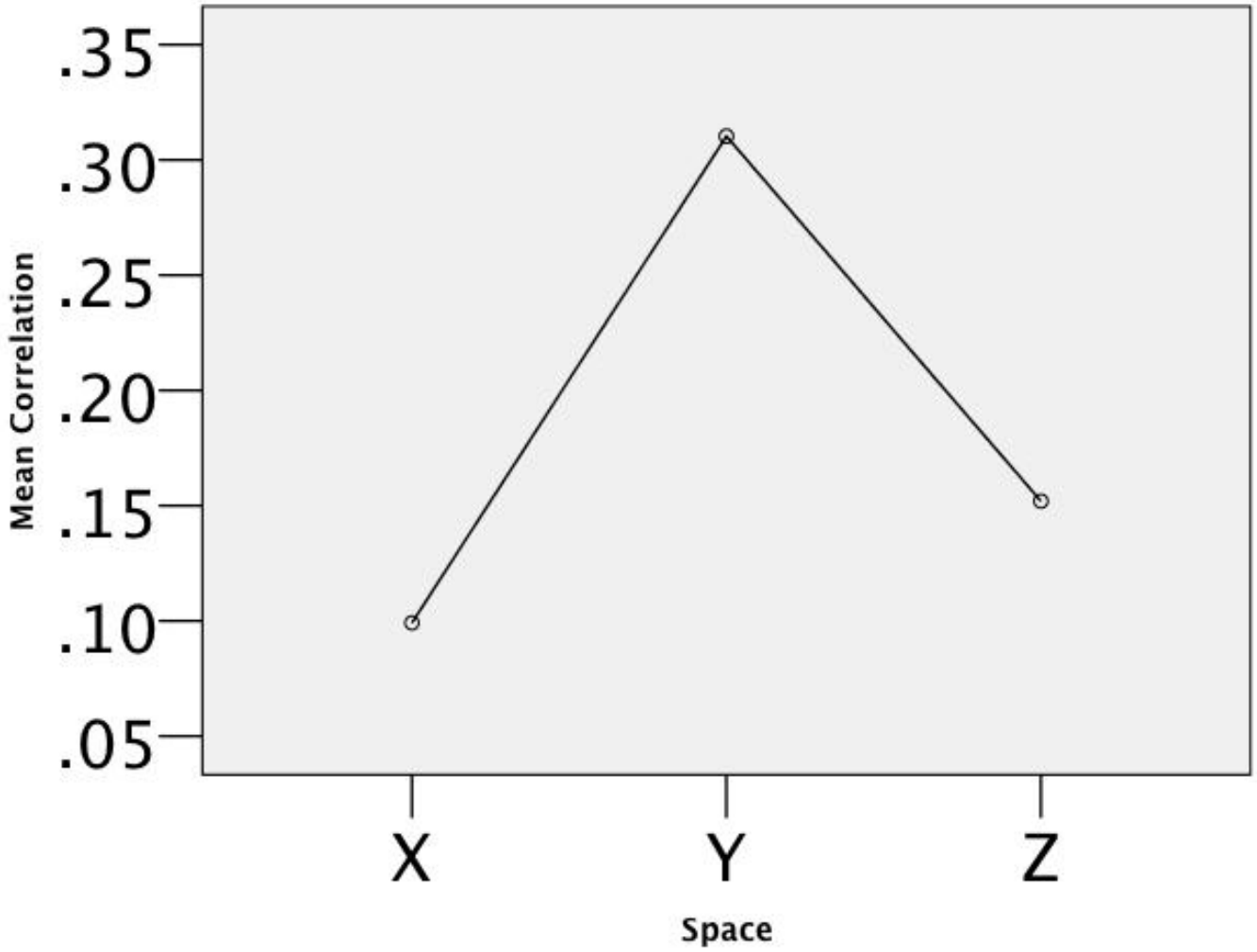
PITCH: **main effect of 'vision'**:
non-visual: M = .354 (SEM = .016)
visual: M = .325 (SEM = .015)

PITCH: **main effect of 'musician'**:
musician: M = .377 (SEM = .020)
non-mus: M = .302 (SEM = .020)

PITCH: interaction between 'space' and 'musician'

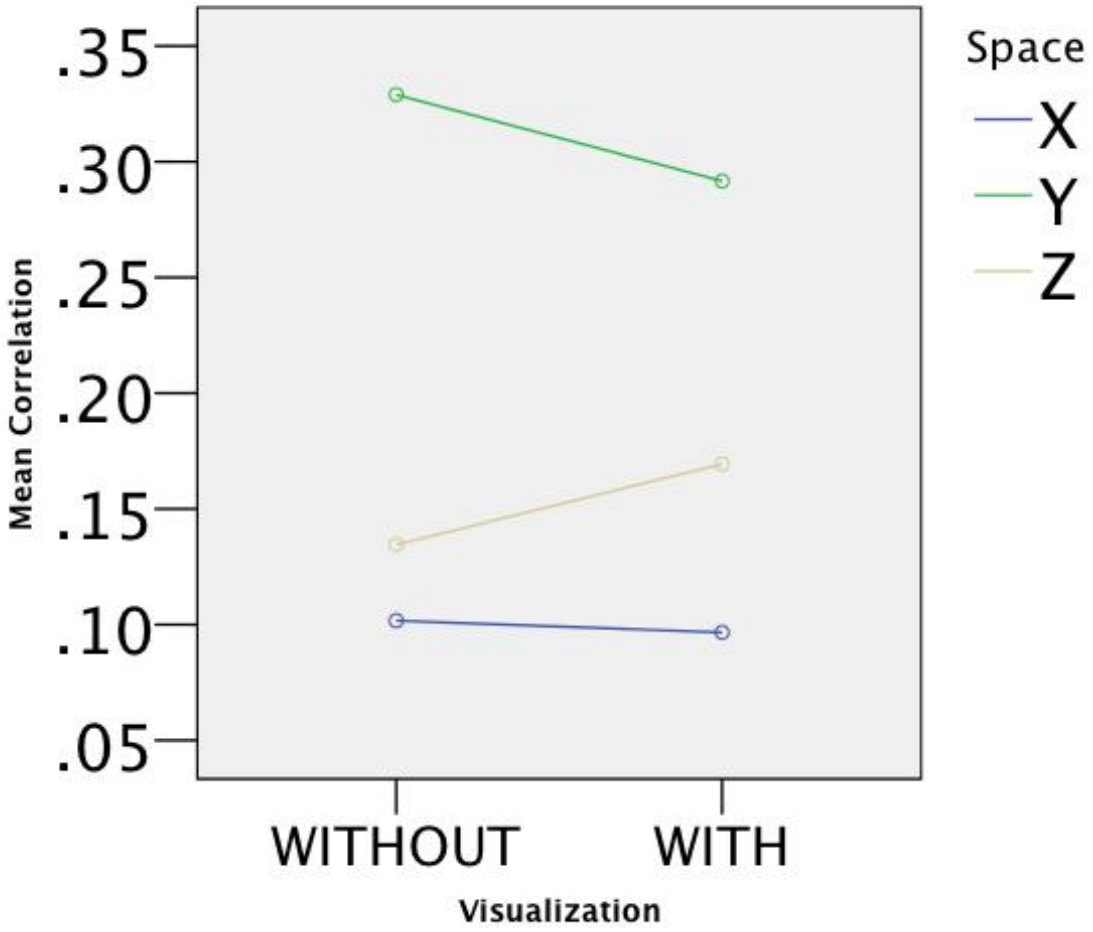


LOUDNESS: main effect of ‘space’



LOUDNESS: **main effect of 'musician'**:
musicians: M = .203 (SEM = .011)
non-musicians: M = .171 (SEM = .011)

LOUDNESS: interaction between 'vision' and 'space'



What was the **direction of movement** that participants used to represent (elapsed) time, pitch and loudness?

TIME: X-AXIS

Mean correlation between time and movement along the x-axis

			non-visual: negative (right to left)	non-visual: positive (left to right)
Musicians	visual	negative (right to left)	**3** nv: –.072 (.036) v: –.039 (.030)	**3** nv: .100 (.053) v: –.091 (.059)
		positive (left to right)	**3** nv: –.297 (.090) v: .783 (.105)	**23** nv: .628 (.367) v: .593 (.386)
Non-musicians	visual	negative (right to left)	**8** nv: –.310 (.275) v: –.292 (.206)	**4** nv: .370 (.356) v: –.082 (.065)
		positive (left to right)	**1** *nv: –.166 *v: .010	**19** nv: .586 (.336) v: .630 (.296)

PITCH: Y-AXIS

- ALL 64 PARTICIPANTS ACHIEVED POSITIVE CORRELATION COEFFICIENTS!
- Main effect of **'vision'** (M_{nv} = .679, M_{v} = .630)
- Main effect of **'musician'** (M_{mus} = .758, $M_{non\text{-}mus}$ = .551)

LOUDNESS: Y-AXIS

- All but 2 non-musicians (nv: –.001, v: –.061 // nv: .223, v: –.018) showed **positive correlations in both conditions**
- ANOVA with 32 musicians and 30 non-musicians revealed **significant main effect of 'vision'** (M_{nv} = .335, M_{v} = .299) and **'musician'** (M_{mus} = .341, $M_{non\text{-}mus}$ = .293)

LOUDNESS: Z-AXIS

Mean correlation between loudness and movement along the z-axis

			non-visual	
			negative (backwards)	positive (forwards)
Musicians	visual	negative (backwards)	**11** nv: –.158 (.149) v: –.143 (.119)	**3** nv: .042 (.026) v: –.102 (.037)
		positive (forwards)	**7** nv: –.128 (.057) v: .190 (.147)	**11** nv: .165 (.154) v: .218 (.140)
Non-musicians	visual	negative (backwards)	**16** nv: –.161 (.110) v: –.168 (.102)	**5** nv: .071 (.042) v: –.158 (.103)
		positive (forwards)	**7** nv: –.122 (.064) v: .138 (.140)	**4** nv: .062 (.052) v: .196 (.171)

Main findings

- Gestural representation of pitch is more consistent and stable (in terms of direction) than gestural representation of time or loudness
- Visual feedback of gestures triggers loudness representation on the z-axis
- Musicians represent pitch and loudness, but not elapsed time, more consistently with arm gestures compared to non-musicians

Limitations

- Participants were forced to use different dimensions for different musical characteristics (i.e. they couldn't use verticality for both pitch and loudness)
  - test musical parameters separately; however, in music they don't occur in isolation
- Forward movement resulted in larger disk
  - future studies should have two conditions (forward movement associated with larger **and** smaller disk size)

Conclusion

- Audio–visuo-spatial mappings with real (as opposed to imagined) bodily movements reveal a very strong association between pitch and height—one that is stronger for the group of musicians and stronger than e.g. between time and left-to-right movement—suggesting that even if pitch-height mappings prove innate, audio-visual correspondences are readily enhanced by cultural factors such as training.

Special thanks to…

…Dan Tidhar for developing the motion capture software.

CMPCP

AHRC Research Centre for Musical Performance as Creative Practice

…CMPCP and the 'Shapes' team (D. Leech-Wilkinson & H. Prior) for their continuous support.

King's College London

…KCL for a Graduate School Conference Fund Grant.

- Eitan, Z., & Granot, R. Y. (2006). How music moves: musical parameters and listeners' images of motion. *Music Perception, 23*(3), 221-248.
- Küssner, M. B. (2013). Music and shape. *Literary and Linguistic Computing*. Advance Access published January 15, 2013: 10.1093/llc/fqs071.
- Küssner, M. B., & Leech-Wilkinson, D. (in press). Investigating the Influence of Musical Training on Cross-Modal Correspondences and Sensorimotor Skills in a Real-Time Drawing Paradigm. *Psychology of Music*.
- Ludwig, V. U., Adachi, I., & Matsuzawa, T. (2011). Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (Pan troglodytes) and humans. *Proceedings of the National Academy of Sciences, 108*(51), 20661-20665.
- Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science, 21*(1), 21-25.