How audio and visual cues combine to discriminate the tempo of swing groove drumming

Scott Love, J.M. Hillis, Carl Haakon Waadeland, Davide Rocchesso, Federico Avanzini, Sofia Dahl, Frank Pollick

How Audio and Visual Cues Combine to Discriminate Tempo of Swing Groove Drumming

Scott Love[1], James M. Hillis[1], Carl Waadeland[2], Davide Rocchesso[3], Federico Avanzini[4], Sofia Dahl[5], Frank E. Pollick[1]

[1]Department of Psychology, University of Glasgow; [2]Department of Music, Norwegian University of Science and Technology; [3]Department of Computer Science, University of Verona; [4]Department of Information Engineering, University of Padova; [5]Cognitive and Systematic Musicology Lab , School of Music , Ohio State University

s.love@psy.gla.ac.uk

http://paco.psy.gla.ac.uk/

Introduction

• Studies show that when multiple sources of sensory information about a single environmental property are available, more precise estimates of that property can be formed by combining the different sources.

• For physical dimensions such as object size (Ernst & Banks, 2002), surface slant (Knill & Saunders, 2003) and object location (Alais & Burr, 2004), studies show that humans integrate different sensory sources in a statistically optimal fashion.

• According to these models, to maximize the precision of the combined estimate, each cue must be weighted in proportion to its reliability.

• **We investigated the integration of auditory and visual cues for a more complex physical property: beat tempo.**

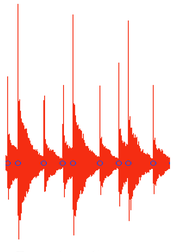
Stimulus Production

• Stimuli were created from 3D motion capture data (240Hz) of a drummer who was asked to perform swing groove drumming at 90BPM.

• This movement data was converted into a visual point-light display (60Hz) with points at the shoulder, elbow, wrist, hand and two drumstick points.

• Sounds were obtained by simulation of the first 25 modes of a circular membrane (Rocchesso & Avanzini, 2004). Parameters for the sound model were the physical parameters of the membrane and the time and impact velocity of a strike taken from the motion capture data.

• To create a standard stimulus a segment containing 9 drum impacts was chosen and found to be 97BPM . Comparison stimuli were created by multiplying the time of impact and total duration to produce 6 tempi above and 6 below the standard in 2BPM increments.



Example audio-wave

References:

Alais, D & Burr, D (2004). The ventriloquist effect results in near-optimal bimodal integration. *Current Biology 14*(3) p257-262.

Ernst, M.O. & Banks, M.S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429-433

Knill, D.C. & Saunders (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research, 43,* 2539-2558.

Rocchesso, D. & Avanzini, F. (2004) Computation of Nonlinear Filter Networks Containing Delay-Free Paths. *Proceedings of the 7th Int. Conference on Digital Audio Effects (DAFx '04)* 113-118

Methods

• 3 conditions: Audio-alone(A), Visual-alone(V), Audio-Visual(AV).

• 2IFC design with a random trial-by-trial allocation of which interval contained the standard stimulus.

• 7 novices and 1 expert drummer completed 2 one hour sessions on different days. Each session contained 3 blocks, one for each condition with 10 repetitions of each standard/comparison pair (e.g. 97 vs 85BPM).

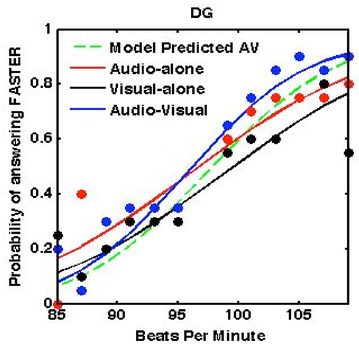
• 240 trials per person, 20 trials per data point.

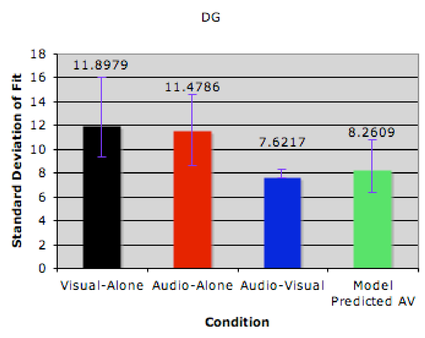
• Participants were asked to indicate which of the two intervals contained the faster drumming tempo.

• The model predicted fits are based on Ernst & Banks (2002).

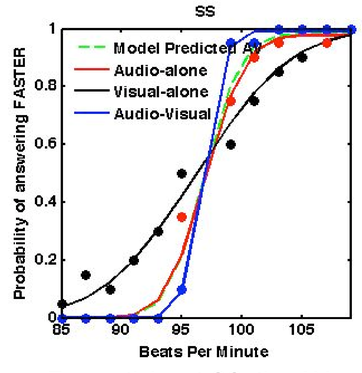
Examples of Novice Results

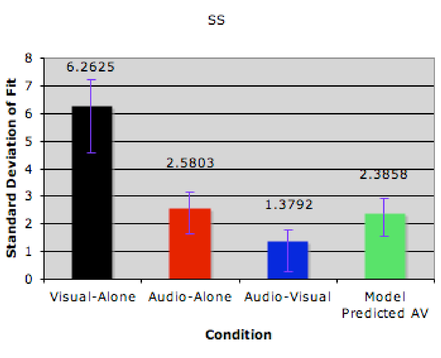
• Standard deviations are taken from the best fitting cumulative gaussian functions, with lower standard deviations indicating better performance.





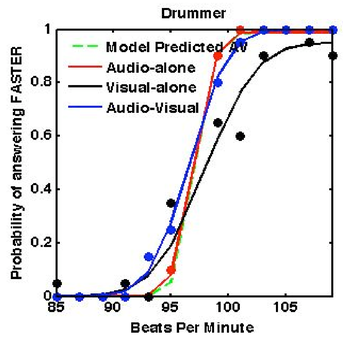
• For participant DG, the AV confidence intervals only overlap with the model predicted confidence intervals.

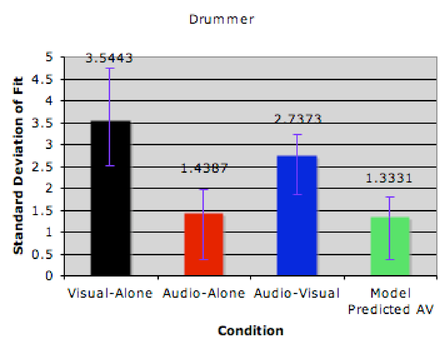




• For participant SS, the AV confidence intervals overlap with both the A and model predicted confidence intervals.

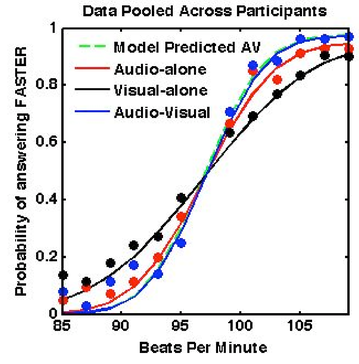
Drummer Results

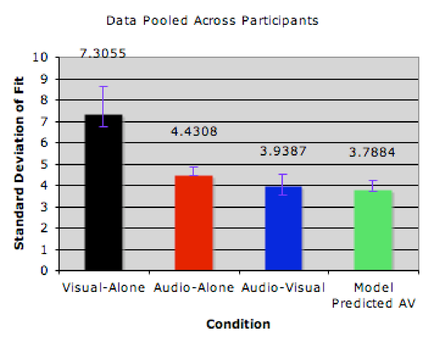




• For the drummer, the AV confidence intervals overlap with both the V and A confidence intervals but not the model predicted confidence intervals.

Average Novice Results





• On average, the AV confidence intervals overlap with both the A and model predicted confidence intervals.

Conclusions

• There are considerable individual differences in performance on each condition of the task and on the relationship between observed and model predicted results on the AV condition.

• On average combined AV standard deviations are within the model predicted standard deviations.

• However, it is not possible to draw strong conclusions as to whether the cues are being combined in a statistically optimal manner because it is still possible that only the most reliable single cue is being utilised.