WHO COULD ASK FOR ANYTHING MORE?
MIXED MODALITY RHYTHMIC DISCRIMINATION APTITUDE IN HUMAN SUBJECTS

Perception And Psychophysics

BACKGROUND

Is rhythm perception and its internal representation connected uniquely to the auditory system, or is it perhaps mo dality independent? This paper begins to explore the ability of human subjects to discriminate change in rhythms across sensory modalities to examine if a bias towards a given modality exists.

AIMS

An experiment was carried out in order to determine the ability of human subjects to perceive rhythms across different modalities. Both audio and visual rhythms were used.

METHODS

The subjects underwent binary forced-choice trials in which they were asked to compare two rhythms and determine whether or not they were perfectly identical. Rhythms could be audio or visually based. A rhythm was defined as a set of stimuli separated in time by periods of silence. The periods of silence were 2^n divisions of a set "measure length" of 4000ms, simulating quarter, eighth, and sixteenth *rests*.

The subject could be asked to compare two rhythms of any of the possible modality pairs: audio-audio, audio-visual, visual-audio, or visual-visual. The second rhythm could be exactly the same timing as the first (though potentially in a different modality) or the rhythm could be altered. Rhythms were altered by extending the length of a rest selected at random (uniformly) forward in time by one of five possible perturbation lengths: 200ms, 300ms, 400ms, 500ms, or 600ms.

A rhythm consisted of a series of purely audio or purely visual beats. The audio beats were computer generated sinusoidal waves at 440Hz, which were played from stereo speakers on either side of the subject. The visual beats were represented by a white square (~5.5 degrees of the visual field) projected on to a screen with a digital projector (60Hz refresh rate) from behind the subject. Both audio and visual beats had a duration of 200ms. The set of rest lengths in the rhythms were generated algorithmically to confom to western music rules for a 4/4 time signature.

The set of rest lengths in the rhythms were generated algorithmically. This was done by probabilistically dividing a whole note (4000ms beat duration) into two half notes (2000ms beat duration). Each of these half notes had could be divided into quarter notes, and so on. The division was allowed to progress to the point of sixteenth notes (250ms beat duration). This way all regular rhythms had the total rest duration. The final assortment of rest lengths was then shuffled randomly. Duplicate rhythms were removed and new rhythms were added until the desired number of unique trials was obtained. The numbers of rests ranged from 5 to 9 within a given rhythm.

There were a total of 200 trials. 100 of the comparisons had two identical rhythms? 25 for each modality pair. The other 100 comparisons had altered second rhythms. These were composed of one rhythm for each of the 5 different delay times (200ms, 300ms, 400ms, 500ms, 600ms). For each delay time, a rhythm was created for each of the 5 different lengths of rhythms. This set of 25 rhythms was repeated identically for each of the four modality pairs (audio-audio, audio-visual, visual-audio, visual-visual). The same random sequence of these 200 rhythms was presented to each subject.

During a trial, the subject was shown a fixation cross for one second before the first rhythm was presented. Then there was a 1.5 second delay before the fixation cross appeared again and the second rhythm was presented. This was followed by a prompt for a response by mouse-click. No interference tones or visuals were used between rhythms.

There were 10 subjects, with ages ranging from 21-29 years of age (mean 24.4, std 3.86). Subjects were of varying musical ability.

RESULTS

Modality

All the trials were separated into the four possible modality pairs and analyzed. While some subjects' performance seemed to vary a given modality pair, there was no visible correlation between modality performance with overall score or musical ability. Results across modality pairs for each subject were, on the whole, similar to each other with the mean standard deviation being 6.83%.

Using one-way ANOVA analysis with Kruskal-Wallis and multiple comparison tests, it is impossible to conclude any major statistical difference between the scores in the various modalities (p=0.4911). This is a potential indication that the processing of rhythm could indeed be modality independent.

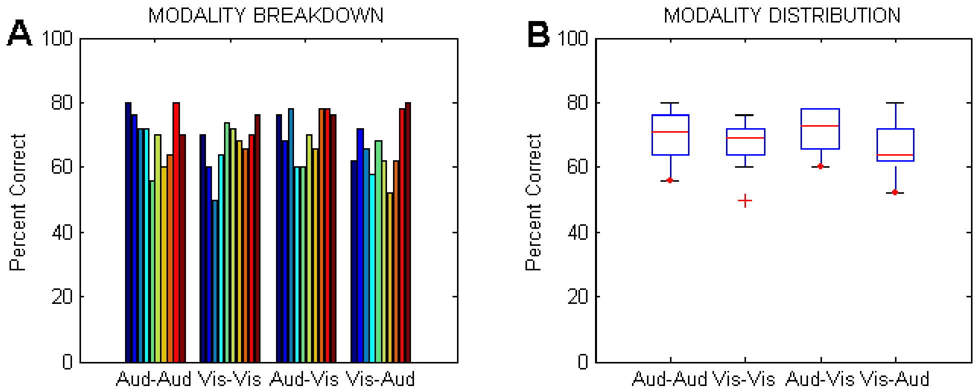

Figure 1. The overall modality breakdown of trials. The bar graph (A) shows the scores of the individual subjects (bars) for each modality, with the box plot (B) showing the variation. In the box-plots the boxes represent the upper and lower quartiles, the lines represent the median, and the whiskers represent maximum and minimum values, with outliers shown as a +.

It is, however, tempting to note the slightly better performance audio-audio and audio-visual comparisons (Fig. 1b). Subjects who scored significantly better on the "audio first" comparisons indicated that those comparisons were easier for them due their familiarity with audio stimuli, which allowed them to get a better sense of what the first rhythm was, thus rendering comparisons easier. These results are consistent with the findings of a previous study that showed a proclivity towards audio representations of rhythm over visual [4]. The experiment presented in this paper, however, uses significantly more trials as well as more accurate timing measures, thus potentially decreasing an audio representation bias due to familiarity. Many subjects also noted that audio stimuli appeared to have a more concrete onset time than the visual stimuli, even though the visual stimuli were identical in duration and accurate to ±2ms in onset time. These comments lend themselves to the idea that humans perform rhythmic comparison tasks by pattern matching or forming a rhythmic template [9], which they store in short-term memory. As there was a non-statistically significant difference in the "audio first" comparisons (with a large difference for several individuals), this suggests that slightly better performance may be due simply to an increased familiarity with forming audio rhythmic templates, but not because of an innate ability or preference to do so.

Delay

In an analysis of the comparison trials with different rhythms, the trials were separated based on perturbation length. It stands to reason that the greater the length of the delay in the second rhythm, the more pronounced the difference between the two rhythms would become. Hence, as the delay increases, so should the identification accuracy. This was clearly demonstrated by one subject, WW (Figure 2a). This subject, however, was the only subject whose results displayed such an "ideal" correlation. The rest of the subjects had different and surprising result. Six out of the ten subjects had peak performance at a 400ms delay (Figure 2b). The remaining three subjects had a peak at a 500ms delay (Figure 2c). Such peaked results seem contrary to logical expectation and seek explanation. It should also be noted that when these three subsets were broken down into modality pairs or broken down based on rhythm length, the same curves by in large persisted.

It should be clearly noted that information from the "ideal" curve cannot be used for generalized conclusions, as there was only one subject showing this behavior. The curve is called "ideal" only insofar as it displays the behavior that should be intuitively expected from an increasing perturbation length. This may in fact be anomalous behavior, and further numbers of subjects must be tested to see if this type of curve occurs in others. WW was an accomplished drummer and also tap danced, so it does not seem out of the realm of possibility that this subject actually could perceive rhythm more accurately than most of the population.

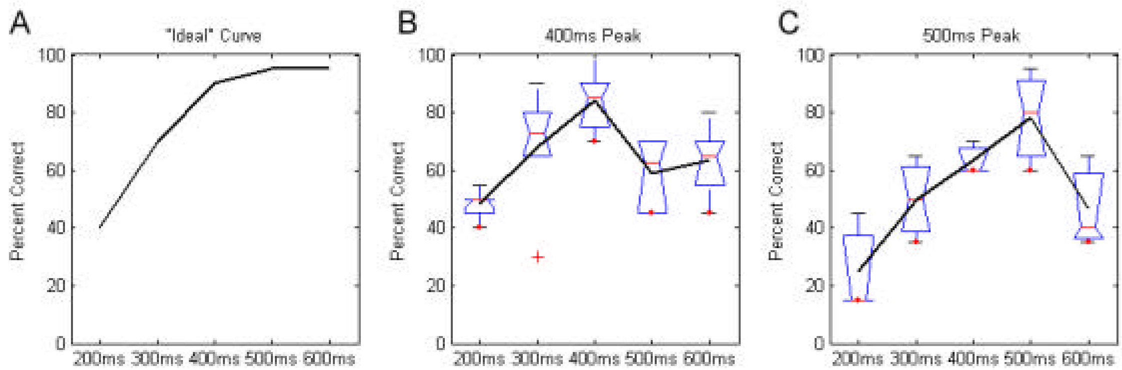

Figure 2. The "ideal" curve from subject WW (A) is shown in comparison to mean curves from the six subjects who peaked at 400ms (B) and the three subjects who peaked at 500ms (C). In B and C the mean result is shown as a black line. In the box-plots the boxes represent the upper and lower quartiles, the lines represent the median, and the whiskers represent maximum and minimum values, with outliers shown as a +.

A possible explanation for the peak at 400ms may be what could be called *beat skipping*. In the context of the rhythm paradigm, 500ms was the length of an eighth rest. This was the only one of the possible delay times that was equivalent to an even 2^n division of the 4000ms "measure" length. Therefore, there is the possibility that a subject could lose track of the rhythm, and because the 500ms delay is a standard metric division, the rest of the altered rhythm would still maintain metrical regularity. In essence, 500ms is a length of perturbation that would still keep all following beats falling on a perfect subdivision of a measure. Hence, the performance kept rising until the maximum time that was not a perfect subdivision of a measure (400ms) and then dropped at 500ms. This explanation makes the assumption that the stimuli are perceived as "instantaneous" to the subjects.

It is difficult to say anything conclusive about the subjects who peaked at 500ms, as there were only three subjects who exhibited this behavior. One possible explanation of this behavior is to look at the results as an "ideal" curve with a drop at 600ms.

For the 400ms peak group, the score for a 600ms perturbation was fairly similar to the score at 500ms. A reason for this could be that the subjects could not perceive the difference between the 500ms and 600ms delays. It stands to reason that a perceptual mechanism similar to Weber's Law of Just Noticeable Differences is coming into play. As the length of a perturbation increases, the smaller percentage difference a 100ms variation becomes—300ms is 50% longer than 200ms, whereas 600ms is only 20% longer than 500ms, yet both have a difference of 100ms. Thus, at longer delays, the scores should become more similar. This is also an aspect of the "ideal" curve, which has a decreasing rate of change as the delay length increases.

CONCLUSIONS

The results of this experiment point toward the possibility of rhythm perception being independent of modality. Far more work must be done to prove this conclusively. Further psychophysical testing of this nature on a larger scale would help solidify the statistical results of this study. Also testing on other random rhythm sets would confirm results by eliminating the possibility of results being specific to this particular rhythm set.

This experiment is intended to be the first part of a two-phase exploration into human perception of rhythm. The second phase would consist of seeking out the neural basis for performance on multi-modality rhythm perception tasks. Neural synchronization in the EEG gamma-band has been shown [1,2] to have a direct correlation with musical perceptional acuity. Future experiments could compare the degree of neural synchronization in subjects observing audio and visual rhythms. A strong similarity of the degrees synchronization across rhythmic modalities could lend important neural support to current psychophysical results. The areas of synchronization for audio and visual rhythm perception, as well as their intersection, would also shed insight into the mechanisms involved in rhythmic perception.

TOPIC AREAS

Perception And Psychophysics, Rhythm, Meter And Timing, Memory And Music, Mixed Modality Rhythm Discrimination

REFERENCES

1. Bhattacharya J, And Petsche H (2001) "Musicians And The Gamma-Band – A Secret Affair?" *Neuroreport* 12: 371-374.
2. Bhattacharya J, Petsche H, Pereda J (2001) "Long-Range Synchrony In The Gamma Band: Role In The Music Perception." *Journal of Neuroscience* 15: 6329-6337.
3. Boltz, Marilyn G. (1999) "The Processing of Melodic and Temporal Information: Independent or Unified Dimensions ?" *Journal of New Music Research* 28, 1, 67-79
4. Decuir AA, Maranto CD (1979)"Analysis Of Visual And Auditory Rhythmic Perception Abilities Of Piano Majors." *Perceptual and Motor Skills,* Oct; 49(2):659-64
5. Fingelkurts An.A., Fingelkurts Al.A., Krause C.M., Möttönen R., Sams, M. Cortical "Operational Synchrony During Audio-Visual Speech Integration" *Brain And Language*. 2003, V. 85. No 2. P. 297-312.
6. Johnson-Laird, Philip N. (1991)." Rhythm And Meter: A Theory At The Computational Level." *Psychomusicology* 10, 88-106.
7. Large, E. W., And Palmer, C. (2002). "Temporal Responses To Music Performance: Perceiving Structure In Temporal Fluctuation." *Cognitive Science, 26, 1-37*.
8. McAngus N, Todd D, O'Boyle J, Lee C (1999) "A Sensory-Motor Theory of Rhythm, Time Perception and Beat Induction" *Journal of New Music Research* 28, 1, 5-28
9. Parncutt, Richard (1994), "Template-Matching Models of Musical Pitch and Rhythm Perception", *Journal of New Music Research* 23 145-168