Rhythm perception through different modalities

Katri Kosonen[1]
University of Tampere

Roope Raisamo[2]
University of Tampere

ABSTRACT

Rhythms are an everyday phenomenon. The most common source of rhythm is music. Rhythm perception studies in the past have mostly concentrated on the auditory and visual modalities. For example, loud music can also be felt as vibrations. Previous studies have showed that audio dominates rhythm perception. Our experiment was conducted to study rhythm perception through three different modalities. We had 12 subjects and 3 sessions with each. In each session there were rhythms through the auditory, tactile and visual modalities. Three different rhythm lengths were used: 4, 5 and 6 beats. In each category there were 14 rhythms. The tasks were to perceive and reproduce these rhythms as accurately as possible. The results showed that audio dominated the tactile and visual modalities. There were more correctly reproduced rhythms with the tactile modality than with the visual modality.

Keywords: tactile, audio, visual, haptics, rhythm perception, rhythm reproduction

1 INTRODUCTION

Audio is the most natural way for people to hear music and rhythms. With rhythms and temporal patterns many kind of information can be provided to people. Rhythms and temporal patterns could be used more than they presently are. To make the use of the rhythms basic research is needed to uncover the capabilities and limitations of people in perceiving the rhythms.

There are several comparative studies about perceiving rhythms given through audio and vision. These are not the only modalities that can be used use in rhythm perception. Haptics is a good possibility: rhythms can be also provided through vibrotactile rhythmic stimuli.

The sound is based on similar waves as the vibrotactile effects. Thus, vibrotactile effects could be good means to provide the users with rhythmic information.

Vibrotactile stimuli could benefit hearing impaired people by helping them to feel the rhythms. This would be essential, for example, when using messaging devices and other devices that can give information to the user about the different events and states of the machine.

Also blind and visually impaired people could get help with this kind of extension to the information channels. Previously the only actively used information channel between computers and blind users has been the audio channel. Because of that the audio channel easily runs out of capacity. There are too many things to tell to the user so that all information could be given through audio. If the users could be provided with rhythmic vibrotactile stimuli, the audio channel could be used for the other kind of information and this way the pressure on the audio channel could be eased up.

Vibrotactile rhythms can also be used to create tactons and for example vibrotactile tactons [3]. This way vibrotactile feedback could benefit the hearing impaired, visually impaired and possibly even people with no vision problems.

The aim of our experiment was to find out how rhythms are perceived with other senses than hearing. A comparison was made between audio, tactile and visual stimuli to find out if the tactile and the visual modality could be used when delivering rhythmic information. There were several research questions in this experiment.

1. Are there more mistakes in the rhythm reproduction, when the rhythm example is given in tactile or visual manner?
2. Do the results change when more experience is gained? Are there fewer mistakes, with more experience?
3. Are the rhythms easier to reproduce, if the rhythms used in the test are regular rhythms instead of irregular rhythms?
4. Does the quality of the rhythm example affect the beat length the user uses while reproducing the rhythm?
5. How do the users react to different modalities in the rhythm perception and reproduction?

The first research question was aimed to find out the overall rhythm perception capabilities through the modalities used. The second question was chosen to find out how learning affects to the results. Does experience in different modalities help in rhythm perception? What kind of rhythms should be used? Which are the most easily remembered regular rhythms and irregular ones? The goal of the fourth question was to find more precisely how the rhythms are perceived through different modalities. The fifth research question addressed the user's opinion and experiences while perceiving the rhythms. Users' opinions are very important. If there is too much work in perceiving the rhythms the users will not use that modality.

The main goal of this experiment was to get basic knowledge of facts that affect rhythm perception.

2 PREVIOUS WORK

Vibrotactile rhythms have already been found to be an effective and very useful method in, for example, Tactons by Brown *et al*. [3]. Still, there are no earlier studies on the rhythm perception in user interfaces making use of the tactile modality. Most of the experiments have compared the auditory and visual modalities in rhythm perception. Based on these studies it can be said that audio dominates rhythm perception and the finest differences in rhythm are perceived through audio [9]. Other researchers have tried to find the factor that would explain the audio dominance over the visual modality, and they have found out that, for example, the loudness or the visual disturbance doesn't affect the fact that audio dominates the cross-modal temporal coding [4]. There has also been a hypothesis that experience in audio rhythms would explain the difference, but that has been proven wrong [4].

It has also been studied if the rhythm reproductions methods used in the experiments would have explained the audio dominance. It was found that writing the rhythms with

[1] e-mail: katri@kososet.net
[2] e-mail: roope.raisamo@uta.fi

letters ‘S’ for short rhythms and ‘L’ for long rhythms, or rhythm reproduction through a piano keyboard doesn’t affect the audio dominance [4]. While either of those methods was used audio dominated the visual modality in rhythm reproduction.

It has been found that if the chunking of the beats is disturbed, there is very little difference between the auditory and visual modalities [5]. But when using rhythms usually it is not desirable to disturb the chunking effect, because it comes naturally from the rhythm structure. The natural temporal pattern of the rhythm would need to be altered.

Modality effects decrease also if the beat durations increase to about 2000 ms and over that [5]. With longer beat durations it becomes too hard to remember the beat lengths. So the human’s memory capacity sets the limit in that aspect. There is also an upper limit to the interval between two stimulus markers. The comprehension of a rhythm as a unit structure is impossible if the longest interval between two stimulus markers transcends 600-800 ms [8]. The pattern of the rhythm also has a time limit. In order to comprehend a rhythm as a whole the pattern shouldn’t be longer than 3 or 4 seconds [8].

In previous research it has been noticed that learning has some effect on the performance in the visual condition when comparing the first and the second trials [4]. In the present experiment, one of the goals was to find out how learning would affect tactile rhythm perception and reproduction.

A study also revealed that the optimal ration of the long to the short intervals for the perception and reproduction of the rhythm structures is about 2:1 [8].

3 EXPERIMENT

The experiment was conducted in a usability laboratory. The users were asked some background questions concerning their previous experience with music and with the devices used in the experiment. During the tactile phase of the experiment white noise was played to the subjects and they were asked to wear hearing protectors to eliminate the influence of the sounds the tactile device makes when tactile effects are played. After each modality the users evaluated their performance with that modality. After all the modalities and tasks were done, the final questionnaire was filled in. The questionnaire contained questions on subjects’ opinions on the devices and their preferences of modalities. The software saved log files for the analysis.

3.1 Participants

In the experiment the subjects were 12 normally sighted university students. Six of the subjects were female and six were male. Their age ranged from 19 to 24. Two male and four female subjects had experience on rhythms based on a musical hobby. Almost all subjects had tried some tactile devices in games, mostly game pads or steering wheels. Three subjects said that they have never tried these devices and 2 subjects told to use some of them regularly.

The subjects participated in three sessions in different days. One session included all three different modalities: auditory, tactile and visual. With each modality there were three different lengths of rhythms: 4, 5 and 6 beats.

In the following sessions the subjects made the same test tasks in a different order. The order of modalities and lengths of the rhythms were counterbalanced between the subjects. In one session with all three modalities the order of the rhythm lengths was always the same. The order of rhythms was randomized in each round so that the learning effect wouldn’t affect the results. One round included 14 rhythms. Overall, there were 9*14=126 rhythms in one session.

3.2 Apparatus

Logitech WingMan force feedback mouse shown in Figure 1 was used to create the rhythms. A java applet with the Immersion TouchSense SDK for Java [7] was used in this experiment. This experiment was designed for the sighted subjects so some instructions were given as text. The program included instructions as text and the rhythm played through three different modalities (Figure 2). A button was used to continue forward in the experiment.

Figure 1: The Logitech WingMan force feedback mouse

Auditory rhythms were played with loudspeakers to the user using simple tones. Tactile rhythms were created by vibrations with the Logitech WingMan mouse. Visual rhythms were shown to the user with a blinking object on the screen. This method was selected, because in this way the rhythms were presented in a similar manner as the auditory and tactile rhythms. The software recorded the rhythms based on mouse clicks when the user reproduced the rhythms. There wasn’t any feedback for the users while reproducing the rhythms.

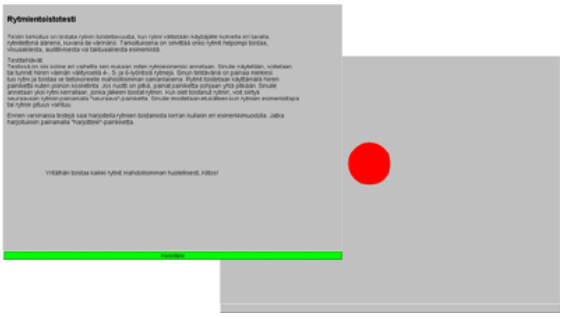

Figure 2: The prototype used in the experiment

In the first phase of the experiment the test procedure was introduced to the subject and there were two examples of each modality used in the test: auditory, tactile and visual. After this practice session the actual tests were done through each modality. The audio test played rhythms to the user through the loudspeakers. The tactile test produced the rhythms to the user through the rhythmic vibrations, and the visual test through the blinking visual object on the screen. After the user had heard or felt or seen the rhythm he or she was asked to reproduce it by clicking the mouse button according to the rhythm in the piano keyboard. The reproduced rhythm was recorded and the forward button was shown to the user. When the user was ready he or she could continue to the next rhythm. The prototype informed the user when the modality or the rhythm length was changed so that the user had a chance to focus on a new task.

3.3 Tasks

In the rhythm, the duration of pauses between the beats was constant (300 ms) and the duration of the beats was varied between short (300 ms) and long (600 ms). The lengths of beats were chosen so that the long beat was twice as long as the short beat. The interval between the beats was chosen to be constant in the whole experiment so that it would not make the task too difficult.

There were two kinds of rhythms in this experiment, irregular and regular rhythms. In irregular rhythms the order of the long and short beats can vary randomly. Regular rhythms used in this experiment follow the rule that short beats always appear in pairs. These are the most common and simplest rhythms that occur in the music: those easy rhythms taught in a musical playschool. There is an example of the graphical presentation of these kinds of rhythms in Figure 3.

Figure 3: Graphical presentation of simple regular rhythms.

It is interesting to find out how very simple rhythms are understood. Usually the simplest rhythms follow the rule of quavers always appearing in pairs. It was hoped that with this division of rhythms it could be found if the regular rhythms are easier to reproduce than rhythms with an irregular order of long and short beats.

With both kinds of rhythms the rhythm length was also varied. The number of beats in one rhythm varied between 4, 5 and 6. Previous research has proven that it is more difficult to reproduce longer rhythms. It was hoped that these longer rhythms would help to reveal if there is a real difference in the reproduction between regular and irregular rhythms.

The rhythms were selected so that they were as different from each other as possible. All the selected rhythms can be seen in Table 1. In the group of regular 4-beat rhythms there weren't enough different rhythms so two of the rhythms were repeated twice to match the number of rhythms in the other groups. The letter 'S' means a short beat and the letter 'L' means a long beat.

Table 1: The rhythms used in the experiment

Irregular rhythms			Regular rhythms		
4 beats	5 beats	6 beats	4 beats	5 beats	6 beats
LLLS	LSLLL	LLLLSL	SSSS	LSSSS	LLSSLL
LLSL	LLLLS	LSLLLS	LLSS	SSLSS	SSLLLL
LSLL	SLLSL	SLLSSL	LSSL	SSSSL	LLLSSL
LSLS	LLSLS	LLSSSL	SSLL	LLLSS	LLSSSS
SLSL	LSSLS	LSLSSL	LLLL	LLSSL	SSLLSS
SLSS	SSLLS	LSSSLS	LLSS	LSSLL	LSSSSL
SSLS	SSSLS	SSSSLS	LSSL	SSLLL	SSLSSL

3.4 Collected information

The main goal of the experiment was to find out whether the users were able to perceive the rhythms or not. In this experiment it was decided to reproduce the rhythms by using a mouse button as a piano key. This method was chosen so that the durations of the beats reproduced could be timed. These times were needed for finding out if the reproduced beat times varied depending on the rhythm modality used.

The program saved the information on the times spent in the reproduction of the rhythms and the lengths of the individual beats in each rhythm. From the beat times it was calculated if the beat was interpreted as a short or a long beat. The software looked for the shortest and the longest beat in each rhythm and counted the median of beat times. Each beat was then compared to this threshold value if the beat was longer or shorter than the median beat length. If the beat was shorter than median length it was marked as a short beat and if the beat was longer than the median it was marked as a long beat. The software counted which beats in each rhythm were wrong.

Information was also collected with a questionnaire. Some background information was asked: gender, age, experience in music and experience in using tactile devices. During the test the subjects were also asked to fill in a workload test of each modality. In the end the user were asked to tell their opinions about the modalities and to choose their preferred modality.

4 RESULTS

4.1 Overall difference between the modalities

There were more mistakes with the tactile and visual channel than with the auditory channel. A one-way ANOVA revealed that proportion of wrongly reproduced rhythms differed significantly as a function of different modalities ($F(2, 4533) = 31.22$, $p < 0.00001$). The Tukey post-hoc test revealed that differences between every modality were significant ($p < 0.05$). When rhythms were presented to the user through audio, 7.9% of the rhythms were reproduced incorrectly, while with the tactile channel the equivalent number was 14.0% and with the visual channel 17.5%.

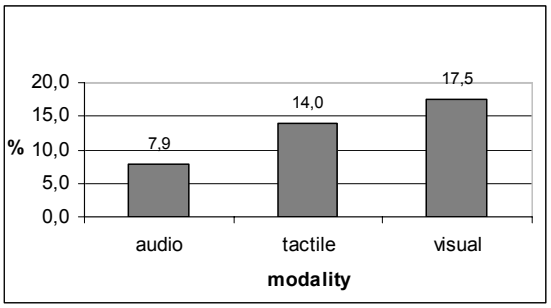

Figure 4: Proportion of wrongly reproduced rhythms

4.2 The effect of learning in rhythm reproduction

Differences between the modalities were the largest during the first time of using them. The one-way ANOVA revealed that the proportion of wrongly reproduced rhythms differed significantly as a function of different modalities ($F(2, 1509) = 17.88$, $p < 0.00001$). The Tukey post-hoc test revealed that differences between every modality were significant ($p < 0.05$). When rhythms were presented to the user using audio 12.3% of the reproduced rhythms were incorrect, while with the tactile channel the equivalent number was 21.0% and with the visual channel 27.2%.

During the second time of use the proportion of incorrectly reproduced rhythms reduced remarkably. The proportion of errors in reproduced rhythms differed significantly as a function of the different modalities (ANOVA $F(2, 1509) = 10.97$, $p < 0.00005$). The difference between the auditory and the tactile modality and the difference between the auditory and the visual modality were significant (Tukey

post-hoc test $p < 0.005$). The difference between the tactile and the visual modality was not significant ($p = 0.346$). When the rhythms were presented to the user using audio, 6.3% of the reproduced rhythms were incorrect, while with the tactile channel the equivalent number was 12.7% and with the visual channel 15.5%.

During the third time of use the differences in the proportion of errors were really small between the modalities. Proportions of error differed significantly (ANOVA $F(2, 1509) = 3.87$, $p < 0.05$). The difference between the auditory and the visual modality was significant (Tukey post-hoc test $p < 0.05$), but the differences between the tactile and the visual modality ($p = 0.687$) and between the auditory and the tactile modality were not significant ($p = 0.142$). When rhythms were presented to the user using audio, 5.2% of the reproduced rhythms were incorrect, while with the tactile channel the equivalent number was 8.3% and with the visual channel 9.7%.

Figure 5 represents the learning of perceiving and reproducing the rhythms through the three different modalities during the three times of use. Learning is statistically significant between the first and the second time of use in all the modalities ($p < 0.05$). There is statistically significant learning between the second and the third time of use only in the visual modality (ANOVA $F(2, 1509) = 28.68$, $p < 0.00001$, the Tukey post-hoc ($p < 0.05$)). In the learning curves it can be seen that the visual and tactile modalities approach each other when more experience is collected. The auditory modality is still the best of the modalities in spite of the learning effect.

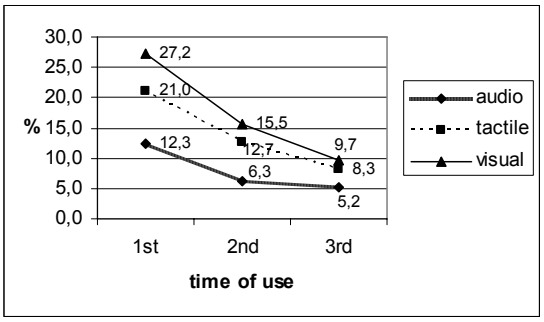

Figure 5: Learning effect in rhythm reproduction between the three modalities

4.3 Differences in reproduction between irregular and regular rhythms

Results of an independent samples t test for means calculated for the proportion of wrongly reproduced rhythms showed that the regular rhythms were significantly easier to reproduce ($t=7.2$, $p<0.0001$). The proportion of wrongly reproduced rhythms was 16.7% with the irregular rhythms and 9.6% with the regular rhythms.

4.4 Modality effect in the beat lengths in reproduced rhythms

The rhythm modality used in the experiment didn't affect the length of the beats in reproduced rhythms. The average length of a short beat (300 ms) the users used was 180 ms and the average length of a long beat (600 ms) was 680 ms. Differences between the average times of modalities were smaller than 30 ms.

4.5 Subjective preference about the modalities

The subjects were also asked about their subjective preference of the modalities. The preference scale was from -10 to 10. In the first time of use there were really big differences between the modalities but they were balanced a little bit when the experiment proceeded. In the third time of use, audio got 5.9 and tactile got 4.5 points. The difference between the auditory and tactile modalities was really small. The difference from the result of the visual modality was noticeable. Although the tactile and visual modalities had just 8.3% and 9.7% difference in incorrectly reproduced rhythms (Figure 5), there was a big 4.5 vs. 0.3 difference (Figure 6) in the average of subjective preferences. So the users strongly preferred tactile over the visual modality.

Several subjects commented that they kind of "heard" the vibrating rhythm in their heads even though they could not actually hear the sound of the vibrating mouse. One subject even said that the vibrating rhythms were easy to remember because of the muscle memory. It was easy to reproduce the rhythm with the hand when you felt the rhythm with the hand.

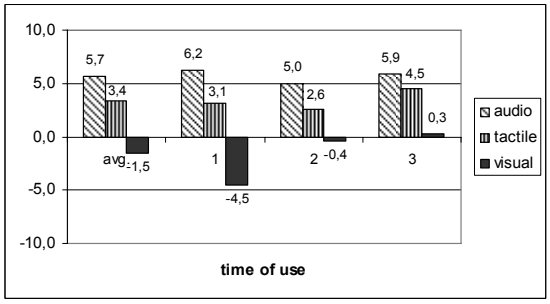

Figure 6: Subjective preference of modalities

5 DISCUSSION

The auditory and visual modalities were studied in rhythm perception and reproduction. For some reason, the tactile aspect of rhythm perception has, until now, been left uninvestigated. This is surprising, because people can often even feel the rhythm of the music through the baseline. People can feel the vibrations going through their body or the ground or chair vibrating in the rhythm of the music. In the experiment some subjects even said that they could "hear" the rhythms through vibrations, even though they wear hearing protectors and white noise was played on the background. The subjects said that they just changed the vibrations to sound.

Vibrotactile rhythms have been shown to be a very useful parameter in Tactons [3]. Tactons can be used to give information in mobile phones, PDAs, messaging devices and also in the desktop computers. For example, the tactile progress bars that make use of vibrotactile tactons, have been researched and found effective [8, 9]. The use of the tactile rhythms could also reduce the overuse of the auditory channel with blind users.

It would seem natural to use the tactile modality to give information about rhythms. In this study this possibility was investigated. It was interesting to find out how the tactile modality gives information about rhythms when compared to the auditory and visual modalities. Answers were found to all of the research questions.

The first research question was about the amount of mistakes in the reproduced rhythms, when the auditory, tactile and visual modalities were compared. The previous studies show that audio is the dominant modality in rhythm perception between audio and vision. This applies also when the auditory, tactile and visual modalities are compared. Quite often vision is the dominant modality of people and audio and tactile come after vision in this order. When we consider rhythm perception this order is turned around: then audio is

the most important sense and vision gives the weakest results, while the tactile modality is between audio and vision.

The previous research has proven that experience and learning does not affect audio dominance in rhythm perception [4]. There was significant learning only between the first and second time of use. These results were proven right also with the tactile modality The auditory modality dominated the rhythm perception even when more experience was collected. The order between the modalities didn't change because of learning. The dominance of the auditory modality significantly decreased between the first and the second time of use. There was also a statistically significant difference between the second and the third time of use, but it affected only visual the modality. It would require more sessions in a longitudinal study to find out how the learning effect behaves with more experience.

The third research question concerned the nature of the rhythm: whether the users make fewer mistakes with regular rhythms than with irregular rhythms. There was a clear result that regular rhythms were easier to perceive and reproduce than irregular rhythms. 9.6% of the regular rhythms and 16.7% of the irregular rhythms were incorrectly reproduced. This might be due to the possibility of chunking. In regular rhythms, there are always two short beats together, which makes it easy to chunk these beats together as a memory chunk. This makes the regular rhythms easier to remember. The subjects also commented that symmetric rhythms were easier to remember. There were more symmetric rhythms in regular rhythms than in irregular rhythms.

It was very clear from the results that the modality used to present the rhythms to the user had no effect to the beat length in the reproduced rhythms. The average beat lengths varied very little.

The last research question concerned the subjective preferences about the different modalities as a presentation channel of rhythms. From the user's point of view the tactile modality was almost as good as the auditory modality even though the tactile modality was closer to the visual modality in respect to the errors made in the reproduction. This was a significant result to show that the nature of the tactile modality is well suited to this kind of tasks from the user's point of view. The users' opinions and their comfort in using these modalities is an important issue.

In this experiment the users were not given any feedback intentionally while they were reproducing the rhythms. Some users commented that while pressing the mouse button they got some kind of feedback with the sound of button to be pushed. This sound however consists of two parts. There is one sound when the button is pushed down and another when the button is released. So this sound isn't similar to the original rhythm to be reproduced and it might have even made the rhythm reproduction even more complicated. The subjects also commented that they got feedback through the sensation that the button was pushed down. This feedback was also really different from the original vibrotactile stimulus and it was considered not to have really affected the results.

6 CONCLUSION AND FUTURE WORK

This paper described an experiment that was conducted to study rhythm perception through three different modalities. The goal in the experiment was to find out if the rhythms could be given to the user through the tactile or visual channels, instead of audio.

The results of the experiment revealed that audio still dominates the rhythm perception and the tactile and visual modalities are in a minority role. The tactile modality produced more correctly reproduced rhythms than the visual modality. In this way it can be positioned between the auditory and visual modalities. In their subjective preferences, several subjects thought that the tactile modality was almost as good as the auditory modality. The tactile modality was however closer to the visual modality if we compare the number of incorrectly reproduced rhythms.

In the future, the effect of different modality combinations in the rhythm perception can be studied. It would also be interesting to find out if a proper feedback while reproducing the rhythms would have an effect on the results. In the future it would also be interesting to find out if these results also apply to other user groups, for example children or disabled people.

7 ACKNOWLEDGEMENTS

We thank all the people who participated in the experiment. This research was funded by the European Union 6th framework programme, IST Specific Targeted Research Project "Multimodal Collaboration Environment for Inclusion of Visually Impaired Children" (MICOLE), IST-2003-511592 STP.

REFERENCES

[1] Stephen Brewster and Alison King. The design and evaluation of vibratactile proggress bar. In Proceedings of World Haptics 2005. IEEE Press: 499-500, 2005

[2] Stephen Brewster and Alison King. An investigation into the use of tactons to present Progress Information. In Proceedings of Interact 2005: 6-17, 2005

[3] Lorna Brown, Stephen Brewster and Helen Purchase. A First Investigation into the Effectiveness of Tactons. In Proceedings of WorldHaptics 2005. IEEE Press: 167-176, 2005.

[4] Arthur M. Glenberg, Stuart Mann, Lisa Altman, Tim Forman & Sean Procise. Modality effects in the coding and reproduction of rhythms. Memory & Cognition, 17(4): 373-383, 1989.

[5] Arthur M. Glenberg and Menachem Jona. Temporal coding in rhythm tasks revealed by modality effects. Memory & cognition, 19(5): 514-522, 1991.

[6] Sharon E. Guttman, Lee A. Gilroy and Randolph Blake. Hearing what the eyes see. Psychological Science, 16(3): 228-235, 2005.

[7] Immersion Corporation. http://www.immersion.com/

[8] L. Mavlov and I. Dascalov. Methods for intramodal and Cross-Modal Investigation of Rhythms (Temporal patterns) in audition, vision and touch. Acta Physiologiga et pharmacologica bulgarica, 7(1): 17-24, 1981.

[9] Gregg H. Recanzone. Auditoryinfluence on visual temporal rate perception. J neurophysiol 89: 1078-1093, 2003.