Sonification playback rates during matching tasks of visualised and sonified EEG data

MICHAEL GAVIN[1], ROKAIA JEDIR[2] AND FLAITHRI NEFF[3]

[1, 2, 3] Interactive Systems Research Group, Limerick Institute of Technology, Ireland
[1]e-mail: michael.gavin@lit.ie
[2] e-mail: k00194255@student.lit.ie
[3] e-mail: flaithri.neff@lit.ie

September 23rd 2016

Abstract

In this paper, the authors discuss a user study examining the role of sonification in electroencephalography (EEG) data presentation. Conventionally, EEG data are presented using visualisation techniques and incorporate multivariate, time-critical information. As the number of EEG channels increase, or when screen real-estate is reduced, visually-presented data can become cluttered and occluded. Our user study examined how accurately users could match visualised EEG data to sonic equivalents, and at what playback rate this was most effective. Accuracy and timing data were recorded, as well as task load index (TLX) questionnaires. Results show that faster playback rates of sonified EEG data yield more accurate results. However, matching accuracy of sonified EEG data in the form presented in this study was not sufficient to replace visualized EEG. Although presently sonified electroencephalograms are not a complete replacement, sonification has the potential to effectively represent aspects of EEG data when visualisation alone becomes challenging for the user. The authors therefore propose a multimodal approach to EEG data presentation aimed at reducing visual clutter and reducing the cognitive load experienced by users when presented with too many dynamic variables on screen.

1 Introduction

Electroencephalography (EEG) is a non-invasive method for monitoring electrical activity in the brain. Visualisation techniques are the most common ways of presenting EEG data to users who need to interpret the multivariate information being produced. However, in certain scenarios, such as when a large number of channels are being observed or other viewing options are simultaneously displayed, limited screen real estate means that visualised data can become cluttered or occluded [1]. In addition, the observer can only track a finite number of information channels using a single sensory modality [2].

The authors propose employing a multimodal approach to EEG data by representing some of the sensor output in the auditory domain using sonification techniques. This approach has the potential to reduce visual clutter and mitigate the cognitive load experienced by users when presented with too many dynamic data variables on screen.

Sonification is the use of non-speech sound for communicating and representing a wide range of dynamic multidimensional data. Additionally, it is frequently employed to inform computer users about real-time processes, notifications, system status, and system events. Sonification can be used to replace or augment information that is typically visualised, such as graphs or attributes of a dataset, as well as to relay alerts or progress information to the user. Popular sonification techniques include Auditory Icons, Earcons, Spearcons, Hybrids, and Audification.

Auditory icons mimic the sounds of real-world objects and activities that correspond to processes, data, and events in a computer system. Therefore, auditory icons rely on prerequisite experience and exposure to everyday sound sources, and attempt to initiate an instinctive link with the computer data or action that they represent. A simple example of an auditory icon is the action of deleting a file, resulting in a sound comparable to that of a crumpled piece of paper thrown into a trash can.

In cases where the function or object being represented has no intuitive real-world sonic basis, the use of Earcons may be more suitable. Earcons are abstract by design, incorporating musical tones and motifs to represent objects, events or functionality. Unlike auditory icons, there is no concrete relationship between the sound and the process. Hence, there is learning involved before an abstract sound becomes associated with a computer process, event or functionality. Earcons can be used to relay more complex information compared to auditory icons, since all of its parameters can be tightly controlled, such as dynamic pitch, loudness, and spectral content.

Spearcons (speech-based earcons) are created by using text-to-speech whereby the utterance is time compressed and is no longer recognisable as speech [3]. Hybrids are a combination of all of these sonic icons. In certain contexts, it is theorised that combining the strengths of various sonic icon types will compensate for the disadvantages that any one type may possess.

A more direct mapping of sound to data is called Audification. Dombois and Eckel define Audification as: “a technique of making sense of data by interpreting any kind of one-dimensional signal (or of a two-dimensional signal-like data set) as amplitude over time and playing it back on a loudspeaker for the purpose of listening” [4].

The authors believe that sonifying portions of EEG data will benefit users who are attempting to interpret real-time anomalies when presented with a large number of EEG channels on screen. However, careful consideration to how the sonification is designed is key so as not to elicit negative crossmodal interaction between auditory and visual streams when presented congruently [5]. Therefore, the design of the sonified stream needs to consider visual perception principles from the ground up, and at every stage of the sonification design process. In this study, the authors have implemented a simple direct audification of the EEG data stream, aimed at testing one characteristic - the effect of playback rate on the ability to match the sonified stream to the correct EEG visualisation.

2 Related Work

Medicine is increasingly incorporating sound and acoustic technology as a means for diagnostics and drug delivery. In 1816, French physician Hyacinth Laennec, initiated the use of sound in medicine. He developed the technique of auscultation through a stethoscope, facilitating the diagnosis of thorax and cardiac diseases by means of observing specific audible signs of irregularities [6]. This was an early demonstration of how sound was effective in revealing anatomical abnormalities that were visually inaccessible. The concept of converting data to sound in the field of neurophysiology was investigated and developed as early as 1883 by N. E. Wedenskii, as he derived features of excitation from properties of the sound produced by action potentials of nerve cells. These he observed when recording electrodes were connected directly to a speaker membrane [7].

The first EEG recording of the human brain was made by the German psychiatrist Hans Berger in 1924. In 1929, Berger revealed the generation of a noticeable rhythmic structure in electrical brain activity, particularly in the absence of visual stimuli [8]. In addition, technicians used chart recorders to listen to the rhythm of a pen scratching on paper in attempt to discover inconsistency in the rhythms and to recognise prominent spike-wave activity in long-term monitoring [7].

In 2007, Hermann, Baier and Stephani [9] proposed a multivariate Event-Based Sonification technique (EBS) whereby they used pitch and spatial location to provide cues about the location of specific events in brain activity. The EEG data is scanned for characteristics that were defined as events, which would later be used to trigger sound synthesis events. For diseases such as epilepsy, narcolepsy, Alzheimer’s and other conditions of which symptoms are observable through brain activity, Hermann argued that this technique provided an efficient method of revealing changes in rhythmic characteristics, allowing observers to recognise subtle differences between normal and abnormal rhythms.

Sonification research in general offers further practical value to disparate disciplines requiring alternative or complementary modes of relaying complex information streams to users. In particular, developments in the fields of auditory perception and auditory cognition are key to the effectiveness of these sonification techniques, and this has led to further maturation of auditory displays [10][11][12]. More recently, interactive sonification has become a more active research topic, allowing users to actively interpret, engage, and manipulate data using sound [13].

3 Methodology

Participants were tasked with matching sonified EEG data streams with visualised equivalents on screen. A simple audification process was applied to the EEG data streams aimed at assessing the effect of different auditory playback rates on matching accuracy and reaction times.

EEG source material was utilised from a previous study carried out by members of the Interactive Systems Research Group, LIT. These data, which were captured using a Neuroelectrics® Enobio® device, was saved in ASCII plain text format comprising data of 8 EEG channels, LSL marker data, and a UNIX timestamp. 30 seconds worth of EEG data was extracted from this dataset for use in this study (see figure 1).

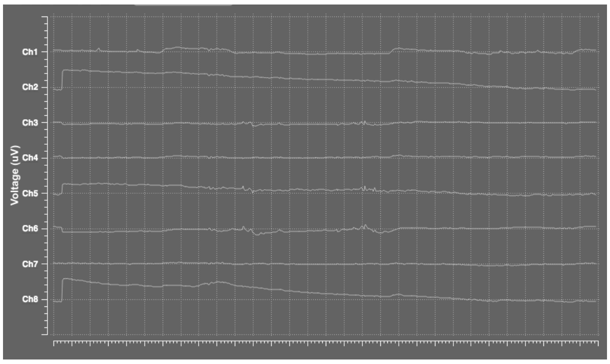

Figure 1: An example of the EEG data acquired from a previous study – 8 channels, 30 seconds in length.

The ASCII file was formatted and indexed using Microsoft® Excel® in preparation for use with the 'coll' object in the Max™ 7 programming environment. The 'coll' object "allows for the storage, organization, editing, and retrieval of different messages" [14], and was loaded with the EEG data. Each channel of the data was separated using the 'unpack' object in order to allow for the individual playback of EEG channels. For the purpose of audification, the datasets were scaled accordingly from 0 to 127 allowing for the generation of a MIDI note on/note off message using the Acoustic Grand Piano patch.

The Max™ patch was designed so that participants' response times and accuracy levels could be recorded within the programme for later analysis. To achieve this, a system of basic start/stop interaction was employed using the 'key' and 'select' objects in Max™ 7. When participants triggered the required input to begin (in this case, the spacebar), the sonified stream was initiated, simultaneously triggering a 'clocker' object to begin counting. When the participant decided to stop the playback, the 'clocker' was also triggered to stop counting. This timing data was captured and saved using the 'coll' object for later analysis in Microsoft® Excel®.

Five different playback rates were presented to participants: 2ms; 25ms; 50ms; 75ms; and 100ms. They were required to observe a static screenshot of 8 channels of EEG data (see figure 1) while listening to a sonified representation of one of the 8 channels. They were then tasked with accurately identifying what channel they were hearing. Each participant performed 15 identification tasks, where the order of the 8 channels was randomised. Each playback rate was presented 3 times over the course of the 15 identification tasks in random order.

31 participants volunteered to take part in the study. However, 3 participants were eliminated prior to analysis due to technical issues during the task process.

4 Experimental Results

A quartile function was implemented for each playback rate to identify any potential outliers. A number of outliers were identified in 4 of the 5 categories, which were then discounted from any further analysis (see table 1).

PLAYBACK RATE	NUMBER OF OUTLIERS
2ms	0
25ms	4
50ms	4
75ms	4
100ms	4

Table 1: The 2ms playback comprised 0 outliers, while all other playback rates each comprised 4, which were discounted from further analysis.

After initial analysis of the individual means for each playback rate, it became apparent that the 2ms playback category had more consistent clustering in terms of participant reaction times. The four remaining categories (25ms, 50ms, 75ms and 100ms) had less defined clustering in comparison. Figures 2 to 6 below show scatter plots of the individual means for each playback category.

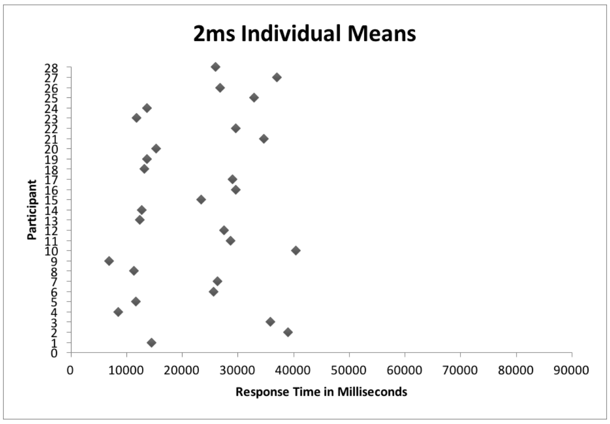

Figure 2: Scatter plot showing the individual means of each participant's reaction time in the 2ms playback category. Reaction times compared to the other playback rates are more consistent.

Figure 3: Scatter plot showing the individual means of each participant's reaction time in the 25ms playback category. Reaction times compared to the 2ms playback rate is less consistent, and follows a similar pattern as the 50ms, 75ms, and 100ms playback categories.

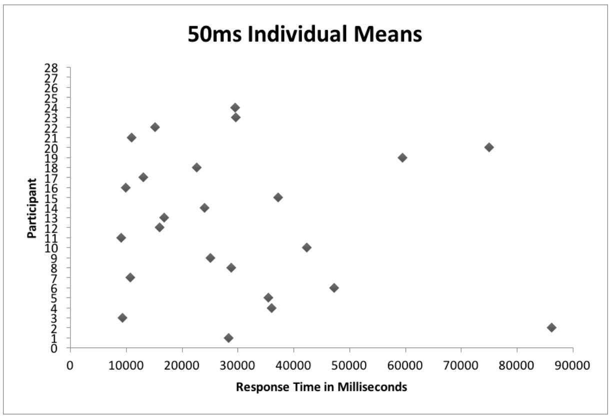

Figure 4: Scatter plot showing the individual means of each participant's reaction time in the 50ms playback category.

Figure 5: Scatter plot showing the individual means of each participant's reaction time in the 75ms playback category.

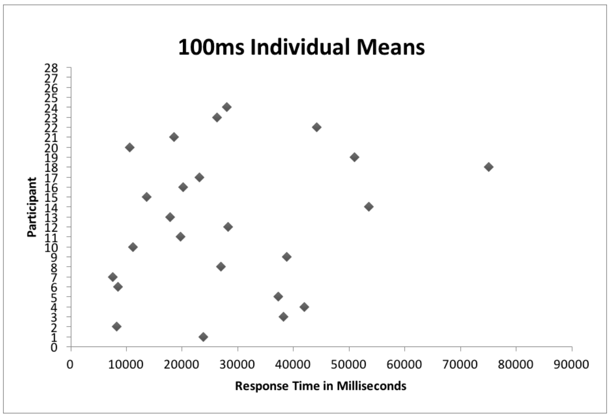

Figure 6: Scatter plot showing the individual means of each participant's reaction time in the 100ms playback category.

Although the 2ms playback category shows a trend towards more consistent reaction times amongst individual means, the overall means for the playback category indicates no statistically significant difference in reaction times between playback categories. Table 2 shows the overall means and standard deviation for each playback category, and although the 2ms category indicates faster reactions times, error bars do overlap between each (see figure 7).

	2ms	25ms	50ms	75ms	100ms
Mean (in ms)	22783	28492	29900	24611	28010
SD (in ms)	10273	28492	20458	13649	16830
CV	0.451	0.658	0.684	0.555	0.601
CI (95%)	3805	7501	8185	5461	6733
Correct (%)	23.6	17.9	12.4	15.1	5.5

Table 2: Overall reaction time means, standard deviations, and percentage of correct responses for each playback category. From these data, participants presented the fastest reaction times, the most consistent reactions times, and the most accurate response rates for the 2ms playback category.

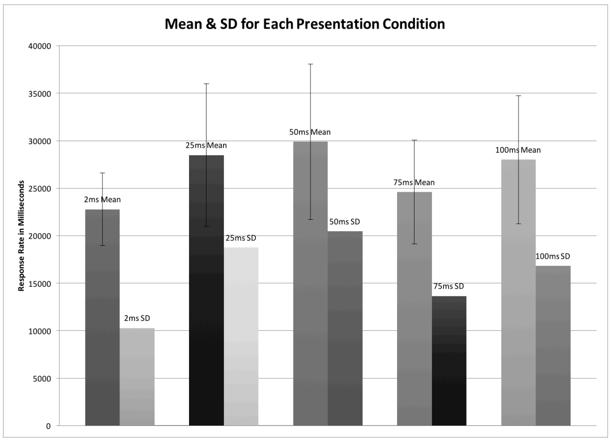

Figure 7: Overall mean and standard deviation for each playback category. Although the 2ms category provided the most consistent and fastest reaction times, error bars overlap between all playback categories, indicating that there is no statistically significant difference between any of the playback categories in terms of reaction times.

In relation to matching accuracy, error bars overlapped between each playback category (see figure 8). This was not the case, however, when comparing the 2ms and 100ms extremes. It is clear that the 2ms playback category outperforms the 100ms category in terms of matching accuracy, despite the fact that the 100ms playback rate offered participants more time to match the sonified and the visualised EEG data. Overall, the 2ms playback category was the most accurate, with c. 24% correct matches, followed by the 25ms category with c. 18% correct matches (see Table 2). All other playback categories performed more poorly with the 100ms playback category having a matching accuracy of only c. 6%. The trend seems to indicate that faster playback rates allow participants to gain an important auditory overview that is perhaps more akin with the visual snapshot of the EEG timeline on screen. However, further investigation is needed despite this trend, as only the 2ms and 100ms playback categories offered a statistically distinct difference in matching accuracy results.

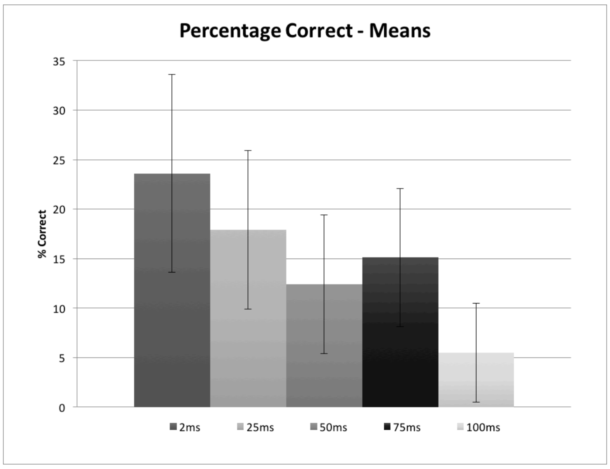

Figure 8: Accuracy for each playback condition. Error bars indicate a statistical difference between the 2ms rate and the 100ms only.

Each participant was asked to complete a NASA Task Load Index (TLX) questionnaire after they finished all study tasks. Across all participants, overall mental demand during tasks was high (85%), as were several other indicators with the exception of physical demand (17.1%) (see figure 9). Temporal demand had an overall mean of 59.3%, performance 63.2%, and frustration 53.6%. Participants considered that the effort required to perform the matching tasks was medium-high, with an overall mean of 72.5%. This indicates that the sonification design used in the study impacted negatively on the participant's cognitive load. However, the TLX questionnaire was given to participants after all tasks were completed, and not after each playback category. Therefore, the 2ms playback category, which had more favourable results in terms of accuracy and reaction times, can't yet be discounted in this regard until a more discrete TLX approach is employed in an upcoming study.

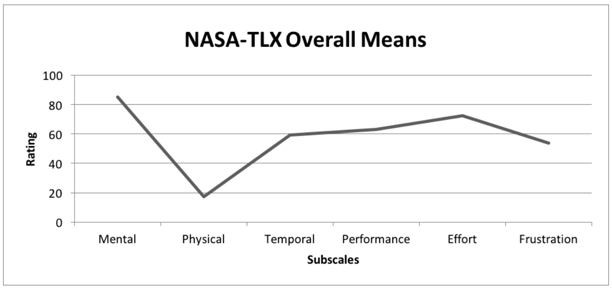

Figure 9: Results of a NASA TLX questionnaire after full study completion. Overall means reveal that participants considered mental demand and effort to be high. The TLX questionnaire results represent all playback categories accumulatively, and do not represent discrete playback categories. Further studies will implement TLX questionnaires that will focus on each playback category individually.

5 Conclusion

Conventional presentation of EEG data relies primarily on the visual domain. In circumstances where a large number of channels are presented on screen, visual clutter has the potential to impact negatively on the user's ability to track and recognise significant EEG anomalies in real-time. Therefore, relaying some of these data using the auditory domain may alleviate some of this cognitive overload. However, the auditory perceptual system is complex and very specific perceptual thresholds need to be established and tested before EEG data streams can be effectively presented this way. In this paper, the authors tested a simple audification mapping comprising 30 seconds of EEG data, presented to participants using five different playback rates – 2ms, 25ms, 50ms, 75ms, and 100ms. Participants were tasked with matching the sonified EEG stream with the correct visualised EEG graph.

In terms of reaction times, the 2ms playback rate provided the fastest response times once the file playback was complete. In addition, the 2ms playback rate had the most consistent reaction times across all participants, displaying a tighter cluster compared to the other playback rates. The 2ms playback rate also provided the best performance in terms of matching accuracy. This suggests that a rapid auditory presentation of EEG streams is similar to a user's quick visual overview of the EEG graph on screen. This perceptual "previewing" or "overviewing", where patterns in the information are quickly identified, is an important perceptual mechanism that allows the user to contextualise before more detailed processing of the information occurs [15]. In auditory terms, the notion of an "auditory glance" is recognised as being beneficial to users prior to them receiving more detailed information using the auditory domain [16] [17]. However, despite this trend showing advantages of presenting EEG streams using the fastest playback rate, statistical analysis did not show a significant difference compared to the other playback rates in terms of reaction times. Only matching accuracy between the 2ms and 100ms playback rates showed a statistically significant difference. Therefore, further evaluation is needed with perhaps the introduction of a faster playback rate and a more developed audification process.

The NASA TLX questionnaire provided some interesting feedback on the overall perception of mental load and effort experienced by the participants. The data showed that the matching task required pronounced mental concentration and effort by participants. However, this questionnaire was taken after all playback rates had been presented to participants, and is therefore not representative of each playback rate taken on its own merit. A follow-on study will evaluate each playback rate discretely in this regard to properly determine if there are differences in mental load and effort across individual playback rates. Biometric evaluation will also play a role in the follow-on study in order to gain a more comprehensive view of participant cognitive load while performing the matching tasks.

While the audification of EEG data streams in the form presented in this study does not accurately translate to visualised equivalents, the authors will evaluate faster playback rates with a more comprehensive audification design to determine if reaction time, reaction time consistency, and accuracy can be improved.

References

[1] M. ten Caat, N. M. Maurits, and J. B. T. M. Roerdink, "Data-Driven Visualization and Group Analysis of Multichannel EEG Coherence with Functional Units," *IEEE Transactions On Visualization And Computer Graphics*, vol. 14, no. 4, pp. 756–771, Jul. 2008.

[2] D. Fougnie and R. Marois, "Distinct capacity limits for attention and working memory: Evidence from attentive tracking and visual working memory paradigms," *Psychological Science,* vol. 17, no. 6, pp. 526-534, 2006.

[3] T. Dingler, J. Lindsay and B. N. Walker, "Learnability of Sound Cues for Environmental Features: Auditory Icons, Earcons, Spearcons, and Speech," in proceedings of the *14th International Conference on Auditory Display*, Paris, France, 2008.

[4] F. Dombois and G. Eckel, "Audification", in *The Sonification Handbook*, 1st ed., T. Hermann, A. Hunt and J. Neuhoff, Ed. Logos Verlag, 2011, pp. 301-324.

[5] V. Hulusic, C. Harvey, K, Debattista, N. Tsingos, S. Walker, D. Howard, and A. Chalmers, "Acoustic Rendering and Auditory–Visual Cross-Modal Perception and Interaction," *Computer Graphics Forum*, vol. 31, no. 1, pp. 102-131, Feb 2012.

[6] A. Roguin, "Rene Theophile Hyacinthe Laënnec (1781-1826): The Man Behind the Stethoscope", *Clinical Medicine and Research* vol. 4, no. 3, pp. 230-235, Sept. 2006.

[7] T. Hermann and G. Baier, "Sonification of the Human EEG", in *Sonic Interaction Design*, K. Franinović and S. Serafin, eds., Cambridge, USA: MIT Press, 2013, pp.285-297.

[8] E.D. Adrian and B.H.C. Matthews, “The Berger rhythm: potential changes from the occipital lobes in man”, *Brain; A Journal of Neurology,* vol. 57, no. 4, pp. 355-385, Dec. 1934.

[9] G. Baier, T. Hermann and U, Stephani, “Multi-Channel Sonification of Human EEG,” in the proceedings of the *13th International Conference on Auditory Display*, Montréal, Canada, 2007.

[10] D. Worrall, “Realtime Sonification And Visualisation Of Network Metadata (The NetSon Project),” in the proceedings of the *21st International Conference on Auditory Display*, Graz, Austria, 2015.

[11] L. Boschi, A. Pate, B. Holtzman, J-L. le Carrou, “Can Auditory Display Help Us Categorize Seismic Signals?,” in the proceedings of the *21st International Conference on Auditory Display*, Graz, Austria, 2015.

[12] R. McIlraith, P. Walton, J. Brereton, “The Spatialised Sonification of Drug-Enzyme Interactions,” in the proceedings of the *21st International Conference on Auditory Display*, Graz, Austria, 2015.

[13] T. Hermann and A. Hunt, "Guest Editors' Introduction: An Introduction to Interactive Sonification," in *IEEE MultiMedia*, vol. 12, no. 2, pp. 20-24, April-June 2005.

[14] Cycling ’74, “coll Reference,” [Online]. Available: https://docs.cycling74.com/max5/refpages/max-ref/coll.html. Accessed: Aug. 1, 2016

[15] S. K. Card, J. D. Mackinlay and B. Shneiderman, “*Readings in information visualization: Using vision to think.*” S. K. Card, J. D. Mackinlay and B. Shneiderman, Ed. San Francisco, CA: Morgan Kaufmann Publishers, 1999.

[16] R. D. Stevens, S. Brewster, P. C Wright and A. D. N. Edwards, “Design and evaluation of an auditory glance at algebra for blind readers,” in the proceedings of the *2nd International Conference on Auditory Display*, Santa Fe, New Mexico, United States of America, 1994.

[17] R. D. Stevens, P. C. Wright, A. D. N. Edwards, S.A. Brewster, “An audio glance at syntactic structure based on spoken form,” in the proceedings of the *International Conference on Computers Helping People with Special Needs*, Linz, Austria, 1996.