

**Remote interpreting:
Assessment of human factors and performance parameters**
**Joint project International Telecommunication Union (ITU)-Ecole de Traduction et d'Interpretation,
Université de Genève (ETI)**
Barbara Moser-Mercer¹

Abstract

This study seeks to assess, for the first time by way of a controlled experiment, the technical feasibility of remote interpreting and its impact on human factors such as psychological and physiological stress and fatigue. Overall acceptance of remote interpreting is still rather low, but this study distinguishes clearly between psychological/physiological and technical reasons for low acceptance and confirms that remote interpreting is more tiring. Interpreters tire significantly more quickly as evidenced by a faster decline in quality of performance over a 30-minute turn.

1. Introduction

Despite its air of novelty remote interpreting is not an entirely new idea. The first major experiments were carried out in the 1970s: the Paris-Nairobi ("Symphonie Satellite") experiment by UNESCO in 1976 and the New York-Buenos Aires experiment by the United Nations in 1978. More recently, a series of experiments was conducted by the European Commission in 1995 (Studio Beaulieu), and a pilot study on ISDN video telephony for conference interpreters was carried out by the European Telecommunications Standards Institute in 1993. The European Commission launched another test in 1997 (Zaremba, 1997) and yet another in 2000 (European Commission, 2000), the European Parliament launched one in 2001 (European Parliament, 2001), the United Nations explored the issue again in 1999 (United Nations, 1999) and in 2001 (no report available). Over the past decade video-conferencing has been used on and off for small meetings between a handful of delegates in a point-to-point video-conferencing set-up. Only the advent of N-ISDN and the H320 family of standards has reduced the complexity of video-conferencing set-ups and increased both sound and image quality.

As technology improves so does the technical feasibility of remote interpreting, defined as any form of simultaneous interpreting where the interpreter works away from the meeting room either through a video-conferencing set-up or through a cabled arrangement close to the meeting facilities, either in the same building or at a neighboring location. While cost and effort necessary to ensure high-quality remote interpreting set-ups are certainly still not negligible, they are likely to decrease as equipment becomes less expensive and technical support staff become more experienced. Human factors then emerge as one of the most important issue to be explored in remote interpreting: psychological aspects such as coping with the stress of a remote interpreting assignment, medical aspects such as having to rely on a screen to derive the visual support information necessary for carrying out the interpreting task, motivation, processing information from multiple sources, social isolation, operating multiple controls, and others.

2. Job design, task analysis and mental workload

Modern technology has raised important new issues for the design of jobs, which is concerned with the psychological and performance effects of general characteristics of jobs. Job design looks at the nature of jobs and their effect on employee performance and well-being, rather than the particular content of a job, which is more the focus of task analysis. While many jobs can be considered as having been designed at some point in time, interpreting falls more into the category of jobs that have evolved over time, with each major innovation

¹ The author wishes to acknowledge the collaboration of the members of the research team and the technicians: Hanne Laugesen, Marina Korac, Valérie Servant, Jaqui Jouffroy, William Ijeh, Craig Jones, Claude Lagrive; as well as the support of Prof. Hans Zeier, ETH Zürich for analysis of cortisol samples.

(such as the shift from consecutive to simultaneous successive interpretation and simultaneous reading of pre-translated texts, and then again to simultaneous interpreting in the late 1930s and mid-1940s (Gaiba, 1998)) requiring a major adjustment on part of those carrying out the job. The rationale for redesign has traditionally been an enhancement of productivity, simultaneous interpreting allows for real-time negotiations, with contentment of those carrying out the job taking a distant second.

Whether a job redesign enhances performance or not depends on a variety of circumstances. Not all job characteristics are likely to be important for performance, or feelings of satisfaction and well-being. Some of those characteristics that have been associated with occupational stress are lack of variety, absence of discretion and control, lack of contact with other people, and physical working conditions (e.g. ergonomically designed equipment) (Chmiel, 1998). One of the prerequisites to job design and redesign is a good task analysis which identifies what people are required to do and the constraints that are placed on them. Modern-day task analysis incorporates psychological factors and uses models of how people handle and process mental information; it considers factors such as memory, learning, attention, mental effort, and decision-making (Cassidy, 1999). With modern technology having increased the emphasis on mental, rather than physical work, and with complex technologies obliging operators to handle considerable amounts of information, it is only fitting that the most recent innovation in interpreting, remote interpreting, be subjected to careful analysis as regards its impact on the physical and psychological well-being of the interpreter and on his performance.

3. Objective of the ITU/ETI remote interpreting project

It was against this background that ITU and ETI launched the remote interpreting project. The objectives were to explore the feasibility and cost of remote interpreting for meetings of ITU and to evaluate the effect of this mode of interpreting on the quality of the service provided and on the human factors involved.

The research team's operating hypothesis was that there would be no difference in terms of perceived psychological and physiological stress, nor in terms of fatigue or quality of performance between live and remote interpreting. In addition, a large number of technical parameters were sampled in order to obtain as comprehensive a picture of the issues involved in remote interpreting as possible. The results of this part of the study can be found in Annex II.

The meeting used for this research project was a live meeting held from April 7 to April 9, 1999 at ITU headquarters in Geneva. For ITU, additional objectives particular to its needs were defined as follows:

- To provide remote interpretation in some of the conference rooms at headquarters.
- To reduce costs for the growing number of interpreted meetings of short duration held outside Geneva.
- To facilitate meetings outside Geneva by eliminating the need for interpretation booths, thus making it easier to find venues and less costly for host governments.
- To use ITU's edge in telecommunications for the benefit of the membership.

4. Logistics

4.1 Meeting type

ITU meeting of a general nature, held at ITU Geneva from April 7-9, 1999.

4.2 Languages

English, French and Spanish.

4.3 Team composition (Interpreters)

Standard team composition in conference room. Duplication of one (French) booth at remote site (ETI).

4.4 Technical support

A team of technical support staff were at the local site at ITU and one technician was available at remote site to handle all technical requirements.

4.5 *Documentation*

Interpreters in the conference room and at the remote site were given the same documents at the same time in advance in order to prepare the meeting. Separate phone, fax and e-mail numbers at the remote site were available for in-session document transmission and technical coordination during the meeting.

5. **Technical components**

5.1 *Equipment and Installation*

5.1.1 On the ETI site two interpretation booths were prepared for the test, equipped with a videoconference system, which would enable the interpreters to follow the conference live. One booth was used as the remote French booth, the second was designed as a back-up in case of technical difficulties. The picture on the interpreter's monitor showed an overall view of the conference room at ITU with an image of the speaker (delegate or chairperson) superimposed in the upper left-hand or lower right-hand corner of the screen.

5.1.2 The videoconference system used was a PictureTel video modem, series 4500, version 6.12, equipped with a Promptus IMX-1B Inverse-Multiplexer connected via ISDN links to the Swisscom switched digital network. The picture monitor in the booth was a 15" computer monitor, in which the video signal had been converted for computer display.

5.1.3 On the ITU site a different infrastructure was installed in order to provide the most faithful coverage of the conference environment. For this purpose, it was necessary to install three studio cameras and video facilities for processing the image. The audio signal from the conference room's interpretation equipment used other routes and interfaces. The two signals were then fed into a videoconference system similar in every way to the one used at ETI.

5.1.4 An identical videoconference system was used at ITU. The only difference was in the connections to the network: Bus SO (Hicom series 300) internal connection to the PABX were used.

5.1.5 Audio and video equipment used in ITU:

- 2 Sony DXC 325 professional video cameras;
- 1 Sony BVWV 507 professional video camera
- 3 video monitors;
- 1 digital video control unit;
- 1 audio mixer;
- 1 CODEC.

5.2 **Connections**

Throughout the experiment, all the installations were under permanent load. For the videoconference this meant a continuous link from 0845 hours to 1800 hours without interruption. H.320 standard connections were used with a bit rate of 384 kb/s and a coding algorithm H.261 for video and G.722 (48 kb) for audio, with a rate of 30 frames per second.

5.3 Video coverage

4.3.1 To cover the meeting room visually, three video cameras were used. These three cameras were operated by cameramen and all three were connected to a digital video control unit operated by an editor, who mixed the three signals and produced a composite image: almost the whole of the room as a background and an insert of the speaker with his place-card. The floor sound from the room was picked up with a clam shell socket. The composite image and the sound were sent to the CODEC. The CODEC, connected to the remote site (ETI) by four ISDN lines, transmitted the sound and the picture with good qualities. The French sound was received back from the ETI (the floor interpreted), and fed into an interpretation channel in the meeting room, using a mixer to adapt the impedance and the power level. (See Annex 1 for schematic of technical set-up)

6. Methods

6.1 Management

A coordination group was established in September 1998 to plan, experiment and evaluate the application of the concept of remote interpreting. Given the complexity of running a controlled experiment in a live international conference setting the group met regularly and agreed on a plan of action of which the main elements were: A method of evaluation covering equipment, transmission, and compatibility between human and equipment components; cost evaluation; and a schedule of tests on human factors covering stress level, concentration, distraction and adaptability to this new way of working.

6.2 Interpreters

Interpreters in the conference room constituted the control group, interpreters at the remote site the experimental group. The study uses a within-subject design to eliminate excessive variation in human factors measurements and in output quality. Within-subject designs are ideally suited for studies with small numbers of subjects, and experiments in conference interpreting usually fall in that category. Thus each of the six interpreters working in the French booth was his/her own control, working according to a prescribed rotation schedule both in the conference room and at the remote site. In addition, the 6 interpreters working in the English and Spanish booths at the conference site were included in part of the sampling (personality inventory, physiological and psychological stress measurements). There were 4 male and 8 female subjects, their average age was .

6.3 Quality

Interpreters' output from both French booths was recorded at both sites (dual track, original and interpretation) for the duration of the entire 3-day meeting. Interpreter output from both French booths was sampled at regular intervals (every 12 minutes), transcribed, evaluated according to a scale developed by ETI for comparable research projects (Moser-Mercer, Künzli and Korac, 1998) and compared. The interpreters' anonymity was ensured by assigning codes to each tape. Their names would never appear on any tape or document and were known only to a non-interpreter research assistant. This was important to ensure cooperation and to assuage any fears interpreters might have regarding their participation in the experiment.

6.4 Human factors

Baseline stress measurements (stress hormones, sIgA - salivettes, Sarstedt, Germany) were taken before the experiment (at the latest one day prior to the selected meeting), immediately before the selected meeting began, and at regular intervals during the experiment. ETH Zurich, Department of Behavioral Sciences (Prof. Hans Zeier), had kindly agreed to collaborate once again with ETI on analyzing the saliva samples and providing statistical analyses of stress hormones. Their experience in this area is well known from analogous studies on air traffic controllers (Zeier, Brauchli & Joller-Jemelka, 1996).

6.5 Questionnaires

The following standardized questionnaires were administered to interpreters before and after the experiment.

6.5.1 Eysenck Personality Questionnaire (Eysenck & Eysenck, 1975) - approx. 15 minutes to complete (administered before experiment to both experimental and control groups, as well as to other interpreters in English and Spanish booths).

6.5.2 State-Trait Anxiety Inventory (STAI) (Spielberger, 1983) - approx. 10 minutes to complete (general questionnaire administered before experiment to both experimental and control group and English and Spanish booths, situational questionnaire administered at the beginning of each meeting and at regular intervals throughout the experiment and to coincide with physiological stress measurements).

6.5.3 Technical questionnaire relating to technical aspects of the remote interpreting situation. Experimental and control groups in French booth filled in all parts, English and Spanish booths filled in general part only. (Annex 2)

These tests provide information on interpreters' basic personality profile and coping mechanisms and specific coping behavior on the job. The technical questionnaire was designed to capture specific (technical and other) problems interpreters in the experimental group may have experienced due to the novel technical arrangements.

7. Results

7.1 Technical findings

7.1.1 Connections

It is noteworthy that throughout the entire time period for which these connections were set up, there was no major hitch affecting the conference proceedings, CODEC, Hicom or network. On just one occasion it was necessary to re-establish the link with ETI as a precaution, because a greater than normal increase in the CODEC correction rate for the video frames transmitted from ITU had been noted.

7.1.2 Image

The image was acceptable (for details see Annex II) and further tests are being scheduled specifically for image and monitor evaluation. The screens used were not the most highly recommended for such tests, but were found to be manageable. Camera coordination was smooth; image transmission was consistent and was modified according to requests from the interpreters working remotely at ETI.

7.1.3 Sound transmission

The only real problems encountered by the technicians which could not be dealt with for lack of time were: hum (50 Hz) in the conference rooms' existing installations, and minor audio disturbances of various kinds and unknown origin.

7.1.4 Feedback

Delegates in the meeting room were not systematically polled but spontaneously expressed satisfaction with this experiment and some suggested that it would be interesting to conduct a meeting with full remote interpretation (i.e. for all three working languages). Delegates could not perceive any difference in quality between the on-site

and remote interpretation, except for the fact that the automatic floor relay function was not programmed for remote interpretation (the automatic floor relay function enables delegates to hear the floor language whenever the interpreter he/she is listening to turns off the microphone in the booth) as the ETI installation had been configured differently for educational purposes. This could not be rectified during the test.

7.2 Questionnaire on technical arrangements (see Annex 2)

Overall, interpreters seemed to be accepting remote interpreting. While certain dimensions of the technical arrangements (sound quality, image, screens) still needed to be improved, none of these emerged as a major factor in determining whether a participant would accept remote interpreting or not. The single most important factor appears to be the physical and perceived distance from the conference hall, the inability to be more closely involved with what is going on in the conference hall, which produces a feeling of “not being in control”. Only 50% of the respondents felt they were in the same hall as the speaker (Fig. 1) and 66% believed that the action outside their scope of vision (outside of what was visible on the screen) was important (Fig. 2).

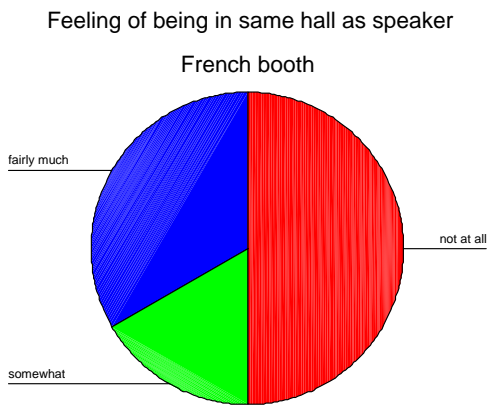


Fig. 1 French interpreters feeling of being in the same hall as the speaker

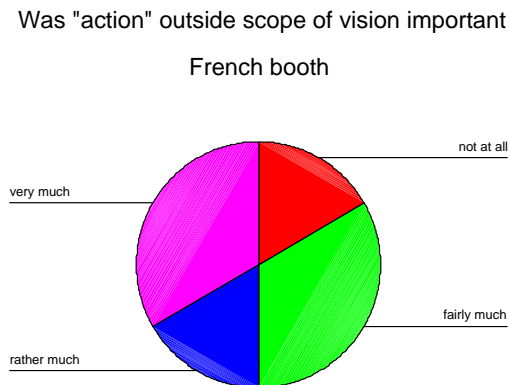


Fig. 2 Importance of action outside of scope of vision on screen for French booth

While the technical arrangements were such that it was probably easier and quicker to phone, fax or e-mail from the remote site to the conference site regarding any technical difficulty than going down into the conference hall to try and fix the problem, that did not change interpreters’ perception of being “remote”. 66% of those working

remote, and the same percentage of those working on site, felt that it is important to be able to speak to the chairman. It seems that the lack of proximity to clients and staff produces a feeling of alienation that ultimately results in lack of motivation and hence produces a decrease in interpreting quality (see 7.6).

It is obvious that some of the technical arrangements, such as for example an image that would allow the interpreter to clearly make out lip movements, facial expressions, etc., are being judged much more harshly under remote interpreting conditions than under normal conference conditions. The fact that 50% of the respondents in the remote condition could make out facial expressions compares well with 50% of them considering that aspect to be important. In other words, those who felt that these expressions were important were able to make them out on the screen, others were not or did not try to. Answers were similar for gestures and body movements: Those who felt gestures were important were able to make them out on the screen, others were not.

While participants in the French booth, who rotated between the conference hall and the remote location, were somewhat divided in their acceptance of remote interpreting (Fig. 3; 60% seemed favorably disposed, 40% much less so), no correlation could be established between their attitude towards remote interpreting and their general attitude towards new technologies, nor between their attitude towards remote interpreting and their desire to travel (a question to that effect had been included as a control question in the Technical Questionnaire), which would obviously be curbed if remote interpreting were to be used universally. 66% of interpreters questioned enjoy travel, which obviously includes those who were also accepting of remote interpreting. All agreed to some degree, however, that there were certain aspects of new technologies that worried them. It appears that the feeling of not being in control of the situation because one was physically remote from it is reflected in these replies.

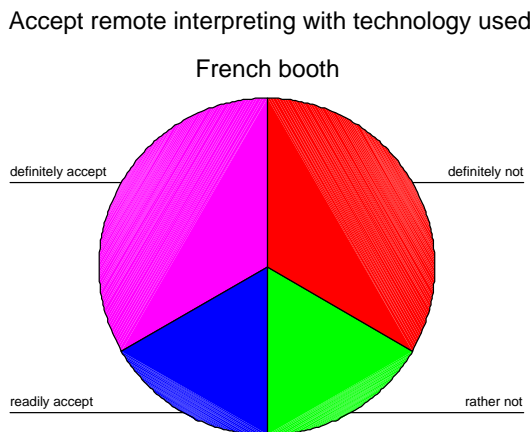


Fig. 3 Acceptance of remote interpreting by French booth

7.3 Standardized questionnaires

In order to obtain a comprehensive picture of interpreters participating in the experiment the design included the administration of the *Eysenck Personality Questionnaire - EPQ-R* as well as the administration of the *State-Trait-Anxiety Questionnaire* whose results are correlated with the physiological stress measurements (immunoglobulin-A - IgA, cortisol).

7.3.1 Eysenck Personality questionnaire - EPQ-R

The Eysenck Personality Scales embody the results of forty years of development, and many hundreds, if not thousands, of psychometric and experimental studies, carried out in many different countries. The Scales are

designed to measure the major dimensions of personality as they have emerged from self-ratings, ratings by friends or acquaintances, observational studies, experimental investigations, psychophysiological experiments and biochemical analyses. The major factors of personality here measured - *Psychoticism, Extraversion, Neuroticism, and Social desirability* - have achieved the widest consensus in this field and more is known about their psychological meaning and their importance in educational, clinical, industrial and other applied fields than could be said about any other personality factors.

Psychoticism (P): high scorers are known to be solitary, not caring for people, troublesome, not fitting in anywhere, hostile and aggressive. Socialization is a concept relatively alien to high P scorers.

Extraversion (E): high scorers are sociable persons, take chances, are impulsive, have a tendency to be aggressive and to lose their temper. They crave noise and excitement and have a low threshold for monotony and boredom. Low scorers are introverted, quiet, introspective, and known to plan ahead. Almost any stimulation is over-stimulation and is painful.

Neuroticism (N): high scorers are anxious, worrying, moody and frequently depressed, overly emotional, reacting too strongly to all sorts of stimuli.

Social desirability (L): high scorers have a tendency to fake well, give socially appropriate answers and tend towards social conformity.

The following tables give an overview of the results obtained both for interpreters in the French booth as well as for those working in the English and Spanish booths.

Table 1: Means of P, E, N and L for male participants

	N	reference values Mean	reference values SD	experimental values
Psychoticism	4	5.72	3.21	8.75
Extraversion	4	14.90	4.74	13.5
Neuroticism	4	10.55	5.49	8.75
Social desirability	4	6.22	3.79	7.5

Table 2: Means of P, E, N and L for female participants

	N	reference values Mean	Reference value SD	experimental values
Psychoticism	7	4.61	2.97	9
Extraversion	7	14.44	4.90	16.43
Neuroticism	7	13.66	5.49	10.72
Social desirability	7	7.62	3.90	9.43

Although the small number of participants does not enable us to make definitive statements regarding the population of interpreters as a whole, some clear indications emerge nevertheless:

Psychoticism - Values both for male and female participating interpreters are clearly above the norm. Perhaps the fact that interpreters, particularly free-lance interpreters, need to fit in with a particular team only for the

duration of a contract and then move on to another social grouping once the next contract begins, is responsible for interpreters being solitary beings, who are at times troublesome and don't fit in. Remarks regarding interpreters' elitist attitude are often heard outside the profession - the above results would lend some credence to that assessment insofar as the meaning of elitist includes being solitary and at times troublesome and aggressive.

Extraversion - Values of male participating interpreters are somewhat below the norm, those for female interpreters somewhat above the norm. Thus interpreters appear to be neither particularly impulsive, nor prone to taking too many risks, nor particularly introverted and overly pedantic. The relatively balanced scores appear to fit in well with the profile of an interpreter: he or she must be ready to take risks, but not excessively so, and he or she must be able to plan ahead (anticipate).

Neuroticism - Values both for male and female participating interpreters are below the norm. Thus interpreters do not appear to be anxious, worrying, moody or frequently depressed, nor do they react too strongly to all sorts of stimuli. Could one perhaps conclude from this that they are fairly well equipped for remote interpreting?

Social desirability - Values both for male and female participating interpreters are above the norm. Interpreters might want to look good and conform, and/or many years of conference interpreting and adapting to a large variety of speakers and situations have left their mark.

It is interesting to note in this context the answers to questions 2, 3 and 4 of the *Technical questionnaire - Part: General aspects* (Annex 2). When asked whether they would accept remote interpreting with the technology used in the experiment 50% of the respondents said they would accept. When asked whether they believe other interpreters would accept remote interpreting with the technology just used 70 % said they would. However, when asked whether the profession as a whole should accept remote interpreting only 33% said it should. It appears that respondents had an inkling that most of their colleagues were leaning towards accepting this new technology and that it was therefore socially desirable to join them, whereas they still harbored strong reservations about the technology: the profession (this abstract entity) should not accept it.

7.4 State-Trait Anxiety Questionnaires (STAI 1 & 2)(Spielberger)

Anxiety states are characterized by subjective feelings of tension, apprehension, nervousness, and worry and by activation or arousal of the autonomous nervous system. Scores increase in response to physical danger and psychological stress.

STAI (State Trait Anxiety Inventory) is a self-report scale, based on 3300 studies carried out in more than 30 languages. It reflects a person's subjective feeling of anxiety. The STAI-Y-2 is designed to capture a person's trait anxiety, i.e. his general level of anxiety independent of any particular anxiety-provoking event. The STAI-Y-1 is designed to capture a person's state anxiety, the level of anxiety experienced in a particular situation.

Spielberger (1966) believes there is a basic difference between trait anxiety and state anxiety. Trait anxiety tends to be relatively stable across time and place. People who are high in trait anxiety have a much greater tendency to be anxious whatever the situation and relatively more anxious all the time compared to those low in trait anxiety. For the high trait-anxious person, it takes relatively less external stress to trigger a stress reaction. On the other hand, people who are low in trait anxiety are more relaxed all the time, regardless of the situation. It takes relatively higher levels of stress to trigger an anxiety response in them.

In contrast to trait anxiety, state anxiety is specific to a situation, such as a driving test, a job interview, a novel job situation, etc. A person high in trait anxiety faced with such an event might be overwhelmed or panic. On the

other hand, someone with low trait anxiety might manage with no difficulty as long as state anxiety in any particular situation does not become extreme.

For the purposes of this study one STAI-Y-2 was filled in by each participating interpreter several days before the beginning of the experiment (together with the Eysenck Personality Questionnaire). Three STAI-Y-1 questionnaires were filled in by each interpreter during each morning and afternoon session, yielding 6 completed questionnaires per interpreter for each complete day in the booth. Frequent sampling was chosen in order to capture any variation in an interpreter’s state anxiety in the course of a workday.

A comparison of state anxiety adjusted for age and gender of all interpreters at the start of their first turn the first day showed no significant difference between interpreters working under live conditions and those working remote. Towards the end of the first turn a somewhat different picture emerges. Interpreters in the French booth experience more state anxiety when working at the remote site as compared to working under live conditions. This difference is noticeable, but did not reach statistical significance. Interpreters in the Spanish booth consistently had higher state anxiety values.

When looking at the start of the last turn and again comparing French interpreters working under live and remote conditions we observed an increase in state anxiety for those working under remote conditions. The increase was statistically not significant. Again, interpreters working in the Spanish booth consistently revealed higher levels of state anxiety.

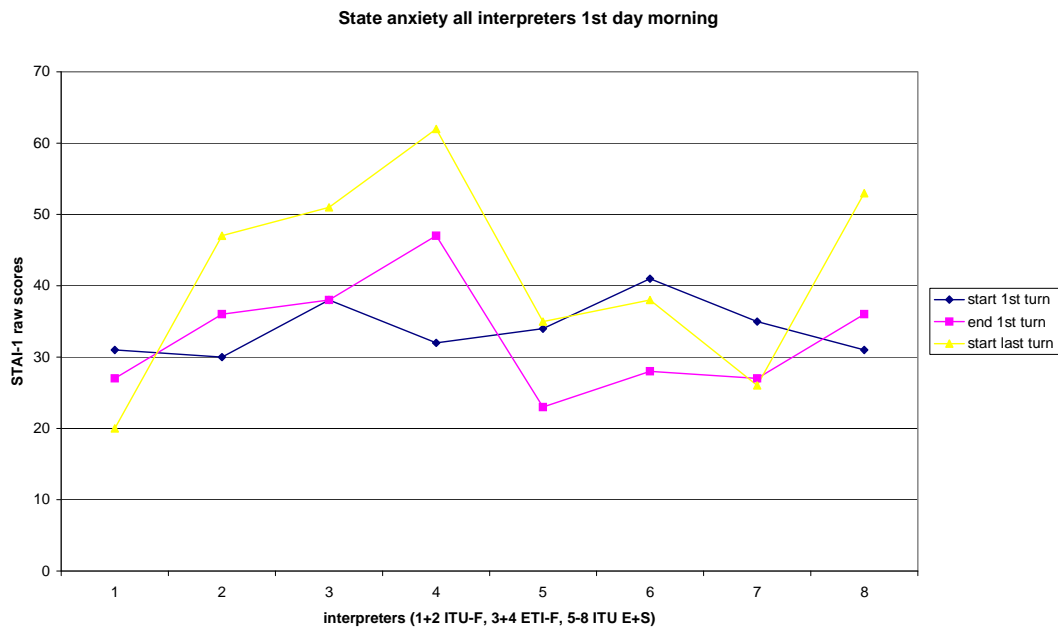


Fig. 4 State anxiety all interpreters first day

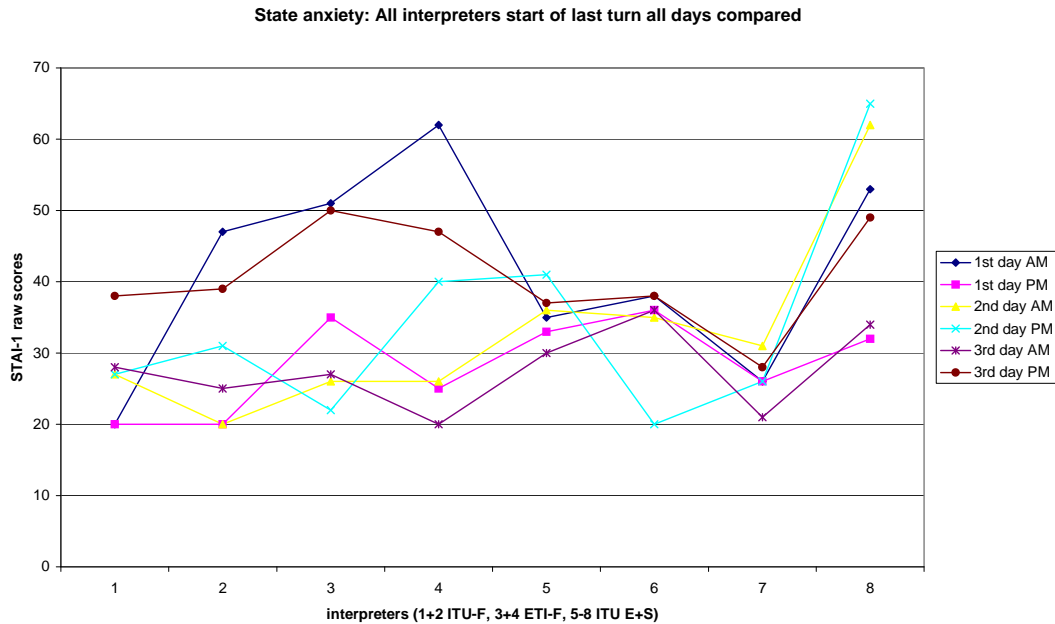


Fig. 5 State anxiety all interpreters start of last turns all days compared

On the whole, repeated psychological self-assessment by interpreters during the experiment indicated that they found working under remote conditions more stressful, although these results did not reach statistical significance.

7.5 Objective stress measurements – The measurement of cortisol and IgM levels

Many studies on occupational stress have investigated the effects of stress on various components of the immune system and have found immunosuppression. However, there are data on cellular and humoral immune markers showing that psychological stressors can also have stimulating effects on the immune system, such as increasing the number of natural killer cells or increasing the number of T-cells, the concentration of IgM and the complement component C3. Cortisol, one of the stress glucocorticoids, inhibits immune functions, especially the cellular immune response, but probably has no significant short-term effects on humoral immunity. Nevertheless, repeated and increased secretion of cortisol is assumed to have a negative long-term effect on humoral immunity. Prolonged stress floods the system with cortisol, which then suppresses the immune system and increases vulnerability to illnesses.

Cortisol secretion was measured by collecting saliva samples at regular intervals. Test tubes (salivettes, Sarstedt, Germany) marked with the personal code-number of each subject and containing a small sterile cotton roll were provided before each test. Subjects were instructed to remove the test tube stopper, take the rolls out of the test tubes, put them in their mouths and chew them slightly for exactly two minutes. Then they had to push the saliva-filled cotton rolls back into the test tubes, seal the tubes with the stoppers and return them to the investigator. They were immediately shipped to the Ecole Polytechnique Fédérale in Zürich (ETH, Department of Behavioral Sciences) in order to be frozen until all laboratory analyses could be performed. Instructions for taking the saliva samples were handed out in written form before the experiment. The subjects were instructed not to eat, and to rinse their mouths well with water 10 minutes before the first saliva collection. Drinks, sweets and chewing gum were allowed, smoking was not. Test tubes were later thawed in the laboratory of ETH Zürich and centrifuged in order to extract the saliva from the cotton. The volume of saliva in each vial was recorded,

thirteen test tubes (8.3%) out of a total of 144 samples taken over the three days of the test contained insufficient saliva volumes to be processed. Laboratory personnel, who had no knowledge of the experimental conditions of the analyzed samples, determined concentrations of cortisol. Concentrations of free cortisol were determined using the RIA kit Cortisol Coat-A-Count from Diagnostic Products and a gamma counter from Canberra-Packard using the RIA-CALC and 4PL programs (Zeier, 1994).

Contrary to interpreters' self-assessment, stress hormone values did not show much variation for French interpreters between their working under live and remote conditions. Again, interpreters in the Spanish booth revealed the highest levels of stress hormones; this correlates well with their self-reported stress on the STAI. Fig. 8 plots both stress hormone (cortisol) levels and state anxiety levels (expressed in percentile) for 6 different measure points distributed over the 6 half-days of the experiment for live (bold) and remote interpreting conditions. In reading the graph from left to right to obtain an idea of the evolution of anxiety over time we note that values tend to be higher at the beginning of the experiment and then settle into a regular, slightly lower pattern, without significant variation. Nevertheless, all values for remote are above values for live interpretation, without reaching significance, though. A larger sample of interpreters would most likely help establish a clearer, significant difference in anxiety between the two conditions.

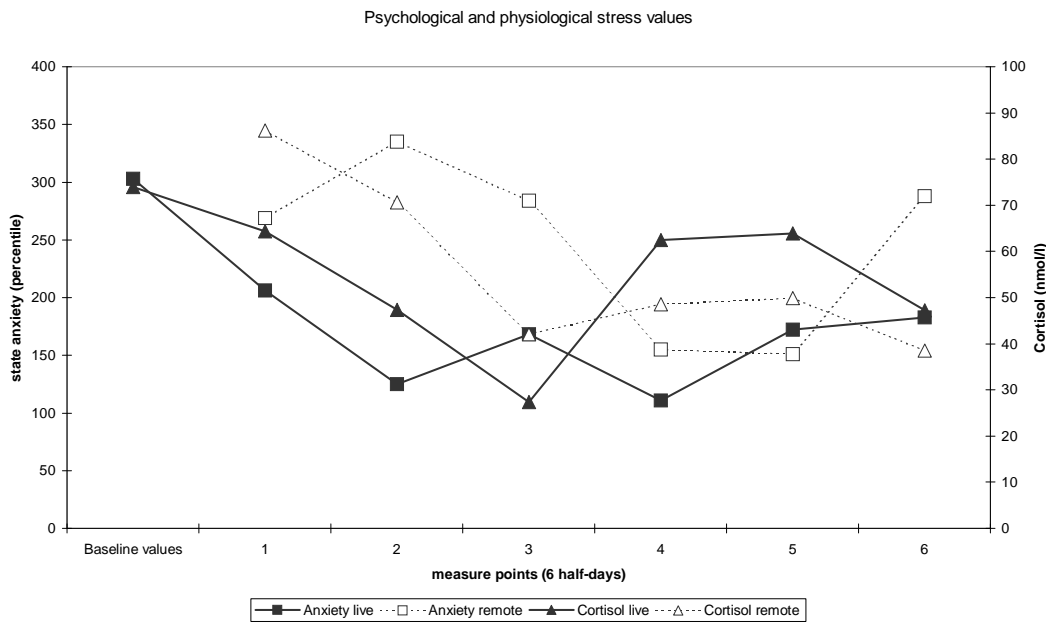


Fig. 6 Psychological (STAI-T) and physiological (cortisol) stress values compared

7.6 Fatigue

Fatigue on task can be assessed only indirectly. In order to study the effects of fatigue under live and remote conditions (French booth) interpreters' performance was audio-recorded, sampled at regular intervals, transcribed and assessed, using an error matrix specifically developed for calculating the effects of an experimental condition (independent variable) on quality as a dependent variable (see Moser-Mercer, Künzli & Korac, 1997). It has been found in earlier experiments that an interpreter's performance tends to degrade over time, especially with regard to completeness of information transfer. Other linguistic and prosodic aspects of interpreters' performance appear to be much less affected by time on task and other variations in external conditions (such as live versus remote interpreting).

Interpreters' output was sampled three times during each half-day, once in the beginning, once in the middle and once towards the end of a turn. It is important to use conference interpreters as jurors as only they can assess information loss correctly knowing that during simultaneous interpreting a variety of compression strategies are applied by the interpreter which do not necessarily lead to meaning/information loss, but might produce an "overt" deviation from the original for listeners/readers/jurors unfamiliar with interpreting. Given the quantity of discourse sampled over three days the authors decided against coding the transcribed data for propositional analysis (Tommola & Lindholm, 1995), but relied on an error coding system developed in earlier experiments (Moser-Mercer, Künzli and Korac, 1997). Two judges with the same language combination (French mother tongue, English as a passive language) and the same number of years of professional experience as conference interpreters (five years) scored each sample independently. Interjudge agreement was computed for the values returned by the two jurors for each sample; it was on average $P_r = 0.89$ ($p < 0.05$) for all samples.

When comparing the effects of fatigue on performance for the same interpreters working either live or in a remote condition we find significant differences ($t = 2.77$, two-tailed, $p < 0.05$). The same interpreter will be less tired, hence work at a higher level of quality, in live conference conditions as opposed to remote conditions. This significant difference holds for all but one of the sampled interpreters in the French booth.

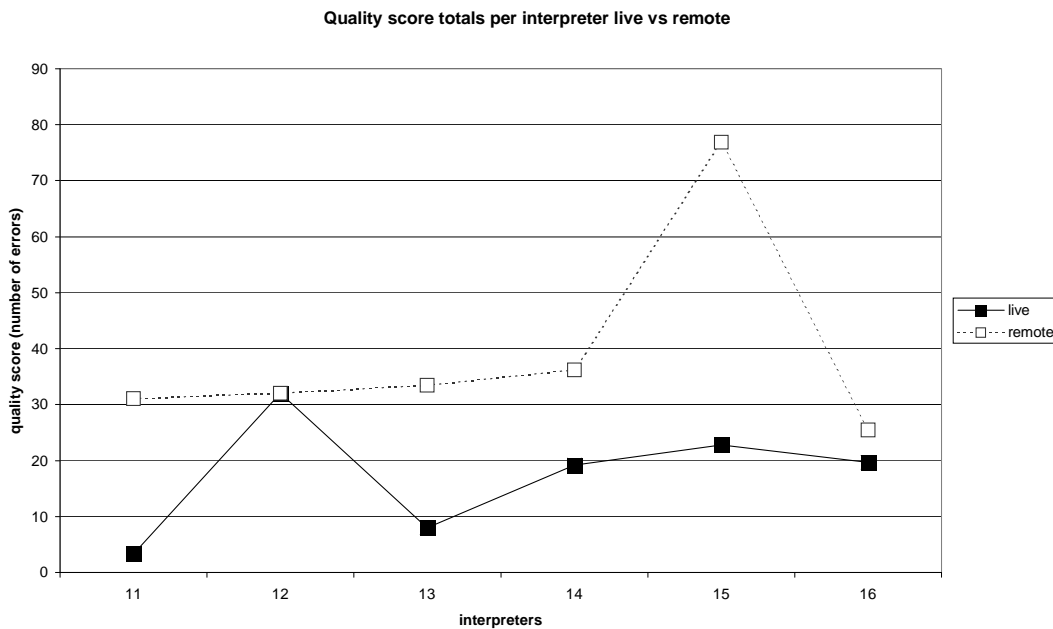


Fig. 7 Quality score totals per interpreter, live versus remote interpreting ($t=2.77$, two-tailed, $p<0.05$)

As to the onset of fatigue, hence the onset of decline in performance, we observed the same interpreter will tire faster (error rates increase) between the middle and the end of a turn in the remote condition as compared to the live condition, where error rate increases are known to occur past the normal end of turn time (past the 30 minute mark on average; see Moser-Mercer, Künzli & Korac, 1997). The onset of fatigue under remote conditions, as evidenced by a decrease in performance, appears to occur fairly soon after "half-time", i.e. somewhere between 15 and 18 minutes into a 30-minute turn. Quality of performance then declines consistently irrespective of time of day. Under live conditions variations in quality follow a very similar pattern throughout an interpreter's turn which confirms that a 30-minute turn corresponds largely to an interpreter's normal work span.

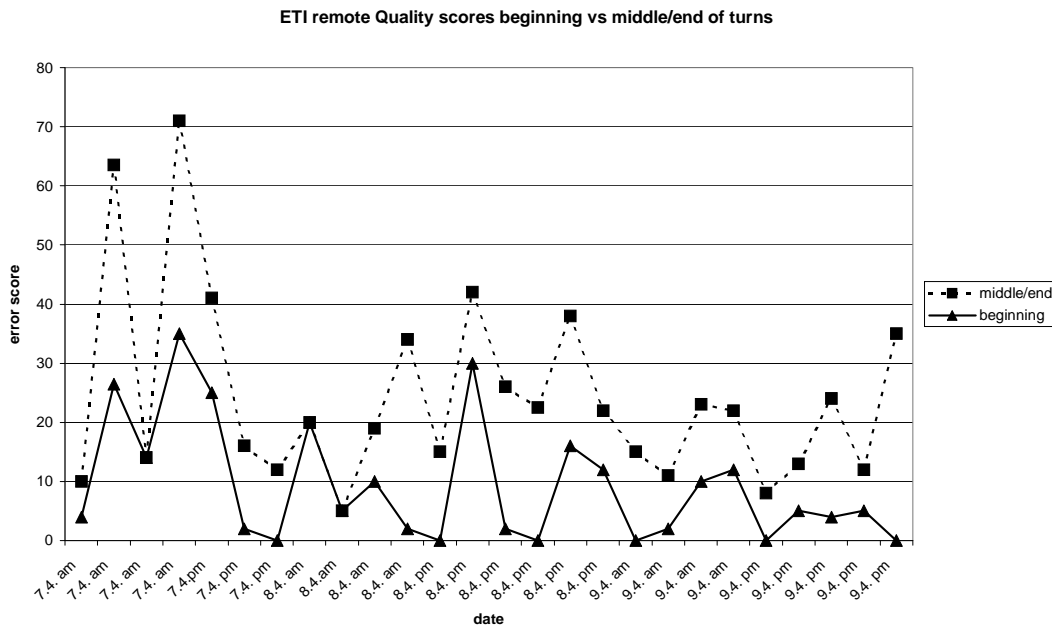


Fig. 9 Quality scores for beginning versus middle/end of turns in the remote condition (ETI)

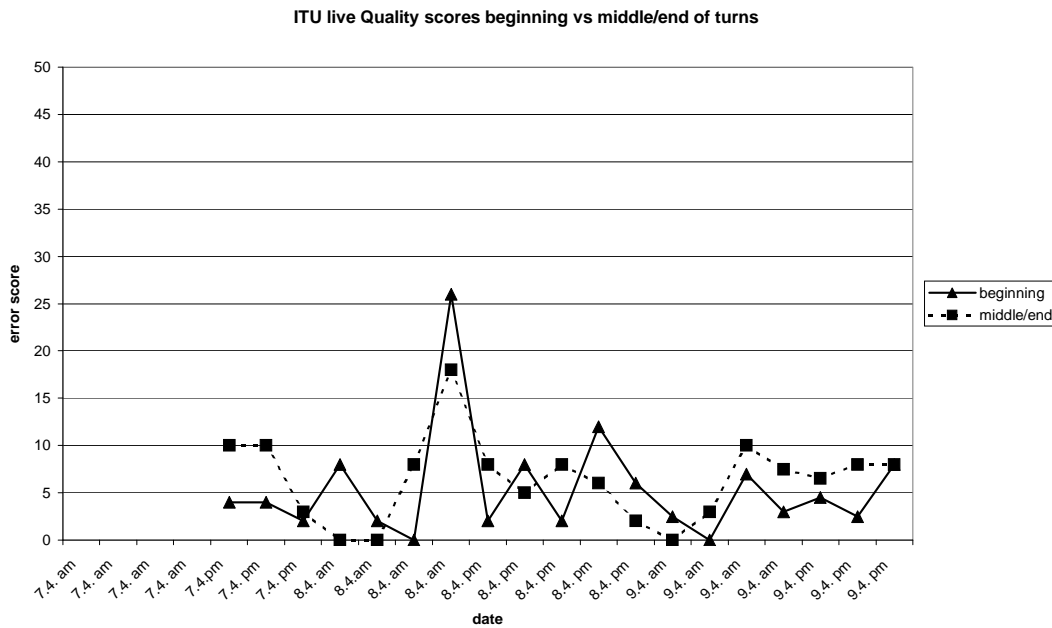


Fig. 10 Quality scores for beginning versus middle/end of turns in the live condition (ITU)²

We therefore need to conclude that remote interpretation increases an interpreter’s mental workload and leads to fatigue and decline in performance faster than live interpretation. These results have been obtained by controlling individual performance differences that are normal across the interpreting population, by choosing a within-subject design and comparing the performance of the same interpreters working in two conditions, live and remote, at the same conference, hence on the same technical subject material and often for the same speakers. Therefore, any difference in performance must be attributed to the condition the interpreter worked in and the effect it had on his or her output.

² Due to a technical error no recordings were available for the live condition on the morning of day 1.

8. Conclusions

The first controlled experiment to evaluate human factors and technical arrangements in remote interpreting has demonstrated that for the same group of interpreters working live in a conference room is psychologically less stressful (according to interpreters' self reports), less tiring as evaluated via performance indicators and conducive to better performance overall. The remote interpreting situation appears to represent not only a novel environment for interpreters in which they need to invoke more effortful problem-solving strategies, but seems to cause more than the usual physiological and psychological strain in that the coordination of image and sound, the piecing together of a reality far away and the concomitant feeling of lack of control all draw on mental resources already overcommitted in this highly complex skill.

The author has argued elsewhere (Moser-Mercer, 2002) that inference generation and the construction of situation models were crucial to discourse comprehension; we might conclude theoretically that the reason interpreters feel the need to be in control of the situation (see Technical questionnaire, Annex II) reflects their need to decide freely and quickly as to which contextual and extra-linguistic information is needed for successful comprehension to occur at high speed (Moser-Mercer, 2002). Interpreters seem to perform well under normal working conditions, but any change in these conditions has immediate repercussions on the efficiency and delicate balance of comprehension and production processes in short-term and long-term working memory (Ericsson & Kintsch, 1995). We must assume that remote interpreting as it is currently set up, and irrespective of technical parameters such as sound and picture quality, prevents interpreters from building up the requisite situation models in working memory that normally allow them to perform at a high level of quality throughout a regular 30-minute turn.

Based on the results obtained in this study and in order to guarantee commensurate quality we would need to recommend not only shorter turn times for interpreters working in a remote situation, but also a thorough analysis of interpreters' visual needs during time on task. This could be accomplished via sophisticated eye-tracking studies comparing different conference room seating arrangements which could provide important input for technical support staff (camera men, sound and image engineers) as to what images or what image selection to send on to interpreters working remote. In addition, it appears that although the present study could establish only trends, interpreters seem to be under increased psychological stress when working away from the conference room, mostly because they experience a lack of control of the situation. Only additional studies along the lines of the present one will allow us to corroborate these findings. Given the importance of establishing and guaranteeing correct working conditions for interpreters working remote so that the high level of performance interpreters are asked to provide can be maintained without undue psychological stress, no effort should be spared to continue to investigate human factors in remote interpreting.

Bibliography

Arocha, J.F., Patel, V.L. (1995). Construction-Integration theory and clinical reasoning. In Ch.A. Weaver, S. Mannes, & C.R. Fletcher (Eds.), *Discourse comprehension. Essays in honor of Walter Kintsch* (pp. 359-381). Hillsdale, NJ : LEA.

Böcker, M., & Anderson, D. (1993). Remote conference interpreting using ISDN videotelephony: A requirements analysis and feasibility study. In *Proceedings of the human factors and ergonomics society 37th annual meeting* (pp. 235-239). (no place)

Cassidy, T. (1999). *Stress, cognition and health*. London: Routledge.

Conway, M.A. (1997). *Cognitive models of memory*. Hove: Psychology Press.

- Chmiel, N. (1998). *Jobs, technology and people*. London: Routledge.
- Ericsson, K.A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102, 211-245.
- ETSI. (1993). *Study of ISDN videotelephony for conference interpreters. Draft final report*. Sophia Antipolis, France: European Telecommunications Standards Institute.
- European Commission. (2000). *Rapport concernant les tests de simulation de téléconférence au SCIC en janvier 2000*. Unpublished internal report. Brussels: EC/SCIC.
- European Parliament. (2001). *Report on remote interpretation test, 22-25 January, 2001, Brussels*. Unpublished internal report. Brussels: European Parliament, Working Party on New Technologies of the Interpretation Directorate.
- Eysenck, H.J., & Eysenck, S. B. G. (1994). *Eysenck personality questionnaire (EPQ-R Adult)*. San Diego, CA : EdITS.
- Gaiba, F. (1998). *Origins of simultaneous interpretation*. Ottawa: University of Ottawa Press.
- Gerver, D. (1969). The effects of source language presentation rate on the performance of simultaneous conference interpreters. In E. Foulke (Ed.), *Proceedings of the 2nd Louisville Conference on Rate and/or Frequency Controlled Speech* (pp. 162-184). Louisville, KY : University of Louisville.
- Greasser, A.C. (1981). *Prose comprehension beyond the word*. New York : Springer Verlag.
- Greasser, A.C., & Clark, L.F. (1985). *Structures and procedures of implicit knowledge*. Norwood, NJ : Ablex.
- Greasser, A.C., Lang, K.L., & Roberts, R.M. (1991). Question answering in the context of stories. *Journal of Experimental Psychology : General*, 120, 254-277.
- Graeser, A.C., & Zwaan, R.A. (1995). Inference generation and the construction of situation models. In Ch.A. Weaver, S. Mannes, & C.R. Fletcher (Eds.), *Discourse comprehension. Essays in honor of Walter Kintsch* (pp. 117-140). Hillsdale, NJ : LEA.
- ITU/ETI/Swisscom. (1999). Background paper to interim report on remote interpretation. International Telecommunications Union, Ecole de traduction et d'interprétation, Swisscom. Report published for IAMLAPD 1999, Rome, WFP. Geneva : ITU/ETI.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension. A construction-integration model. *Psychological Review*, 95, 163-182.
- Kintsch, W. (1998). *Comprehension. A paradigm for cognition*. Cambridge : Cambridge University Press.
- Moser-Mercer, B. (1997). *Visual contact in simultaneous interpretation*. Unpublished manuscript. Geneva: Ecole de traduction et d'interprétation.

Moser-Mercer, B. (2000). Cognitive issues in interpreting research. Paper presented at the colloquium *Linguistik und Translation* (21.-22.11.2000). Hamburg : Universität Hamburg.

Moser-Mercer, B. (2002). Situation models : The cognitive relation between interpreter, speaker and audience. In F. Israël (Ed.), *Identité, altérité, équivalence? La traduction comme relation. Actes du Colloque International tenu à l'ESIT les 24, 25 et 26 mai 2000 en hommage à Marianne Lederer* (pp. 163-187). Paris : Lettres Modernes Minard.

Moser-Mercer, B., Künzli, A., & Korac, M. (1998). Prolonged turns in interpreting : Effects on quality, physiological and psychological stress (Pilot study). *Interpreting* 3 :1,47-64.

Spielberger, C.D. (1966). Theory and research on anxiety. In C.D. Spielberger (Ed.), *Anxiety and behavior*. New York: Academic Press.

Spielberger, C.D. (1983). *State-Trait Anxiety Inventory (Form Y)*. Redwood City, CA: Mind Garden.

Tommola, J., & Lindholm, J. (1995). Experimental research on interpreting: Which dependent variable? In J. Tommola (Ed.), *Topics in interpreting research* (pp. 121-133). Turku: University of Turku, Centre for Translation and Interpreting.

United Nations. (1999). *A joint experiment in remote interpretation. UNHQ-UNOG-UNOV*. Geneva: United Nations, Department of General Assembly Affairs and Conference Services.

Van Dijk, T.A. (1994). Discourse and cognition in society. In D. Crowley & M. Mitchell (Eds.), *Communication theory today* (pp. 107-126). Oxford: Pergamon Press.

Van Dijk, T.A. (1995). On macrostructures, mental models and other inventions : A brief personal history of the Kintsch-van Dijk theory. In Ch.A. Weaver, S. Mannes, & C.R. Fletcher (Eds.), *Discourse comprehension. Essays in honor of Walter Kintsch* (pp. 383-410). Hillsdale, NJ : LEA.

Van Dijk, T.A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York : Academic Press.

Zaremba, H.-D. (1997). *Evaluation des questionnaires Beaulieu. Expériences en automne 1995*. Unpublished internal report. Brussels : European Commission - SCIC.

Zeier, H. (1994). Workload and psychophysiological stress reactions in air traffic controllers. *Ergonomics*, 37(3), 525-539.

Zeier, H., Brauchli, P., & Joller-Jemelka, H.I. (1996). Effects of work demands on immunoglobulin A and cortisol in air traffic controllers. *Biological Psychology* 42 (1996), 413-423.