



ETSI
TECHNICAL
REPORT

ETR 261-7

October 1996

Source: ETSI TC-HF

Reference: DTR/HF-01028-7

ICS: 33.020

Key words: Keypad, MMI, supplementary service

**Human Factors (HF);
Assessment and definition of a harmonized minimum
man-machine interface (MMI) for accessing and controlling
public network based supplementary services;
Part 7: Experimental evaluation of draft ETS 300 738**

ETSI

European Telecommunications Standards Institute

ETSI Secretariat

Postal address: F-06921 Sophia Antipolis CEDEX - FRANCE

Office address: 650 Route des Lucioles - Sophia Antipolis - Valbonne - FRANCE

X.400: c=fr, a=atlas, p=etsi, s=secretariat - **Internet:** secretariat@etsi.fr

Tel.: +33 4 92 94 42 00 - Fax: +33 4 93 65 47 16

Copyright Notification: No part may be reproduced except as authorized by written permission. The copyright and the foregoing restriction extend to reproduction in all media.

© European Telecommunications Standards Institute 1996. All rights reserved.

Contents

Foreword	5
Introduction	5
1 Scope	7
2 References	7
3 Definitions, symbols and abbreviations	8
4 Introduction to the experiment.....	8
5 Method	9
5.1 Subjects	9
5.2 Experimental procedure.....	9
5.3 Equipment.....	11
5.4 Data capture	12
6 Results	12
6.1 Personal data.....	12
6.2 Performance	12
6.3 Results for the "Assisted mode" (Turin sample)	16
6.4 Results for the "Learning mode" (Berlin sample).....	17
6.5 Overall analysis.....	18
7 Detailed analysis	20
7.1 Error analysis.....	20
7.1.1 Cross-service errors.....	20
7.1.2 Service-specific errors.....	21
7.2 Timing	21
7.3 Knowledge transfer	23
7.4 Short questionnaire.....	23
7.4.1 Assisted Mode.....	23
7.4.2 Learning Mode	24
7.5 Long questionnaire	25
7.5.1 Assisted Mode.....	25
7.5.2 Learning Mode	26
8 Discussion of results and conclusion	27
8.1 Effects of the experimental implementation.....	27
8.1.1 Meaning of tones.....	27
8.1.2 Sub-optimum implementation of the display	27
8.1.3 Sub-optimum instructions for experimental tasks	27
8.2 Critical issues identified during the study.....	27
8.2.1 Learning complex command strings	27
8.2.2 Service definition	28
8.2.3 User models in 3-Party Conferencing	28
8.3 Conclusions	28
8.3.1 Is PBI++ better than PBI+, and PBI+ better than PBI?.....	28
8.3.2 What effect does the information before the control action have on usability?.....	28
8.3.3 What effect do feedback or prompts during and as a result of the control action have on usability?	29
8.3.4 What effect does the increase in syntax complexity from "activate" and "de-activate" to including "interrogate" have on usability?	29

Annex A: Personal Data.....	30
A.1 Personal data tables for the German and Italian subjects.....	30
Annex B: General Instructions	34
Annex C: Description of the Services.....	35
Annex D: Tasks and Questionnaires	42
Annex E: Error Analysis	54
E.1 Types of errors vs. Tasks for PBI, Sessions 1 and 2, "Assisted Mode"	54
E.2 Types of errors vs. Tasks for PBI+, Sessions 1 and 2 ("Assisted Mode").....	55
E.3 Types of errors vs. Tasks for PBI++, Sessions 1 and 2 ("Assisted Mode").....	56
E.4 Types of errors vs. Tasks for PBI, Session 1 ("Learning Mode")	57
E.5 Types of errors vs. Tasks for PBI, Session 2 ("Learning Mode")	58
E.6 Types of errors vs. Tasks for PBI+, Session 1 ("Learning Mode")	59
E.7 Types of errors vs. Tasks for PBI+, Session 2 ("Learning Mode")	60
E.8 Types of errors vs. Tasks for PBI++, Session 1 ("Learning Mode")	61
E.9 Types of errors vs. Tasks for PBI++, Session 2 ("Learning Mode")	62
Annex F: Responses to Long Questionnaire	63
History	65

Foreword

This ETSI Technical Report (ETR) has been produced by the Human Factors (HF) Technical Committee of the European Telecommunications Standards Institute (ETSI).

ETRs are informative documents resulting from ETSI studies which are not appropriate for European Telecommunication Standard (ETS) or Interim European Telecommunication Standard (I-ETS) status. An ETR may be used to publish material which is either of an informative nature, relating to the use or the application of ETSs or I-ETSs, or which is immature and not yet suitable for formal adoption as an ETS or an I-ETS.

Introduction

The Technical Committee for Human Factors has prepared this ETSI Technical Report to report publicly its work on the assessment and definition of a harmonized minimum man-machine interface for the access and control of public network based supplementary services. It is intended to complement ETS 300 738 [3].

This ETR constitutes part 7 of a multi-part ETR ("Assessment and definition of a harmonized minimum man-machine interface (MMI) for accessing and controlling public network based supplementary services"), whose parts have the following titles:

- Part 1: "General approach and summary of findings";
- Part 2: "Literature review - Memory and related issues for dialling supplementary services using number codes";
- Part 3: "Experimental comparison of two MMIs - Simulated UPT access and prototype ISDN supplementary services";
- Part 4: "Experimental comparison of the effect of categorized and non-categorized formats within user instructions";
- Part 5: "Experimental comparison of the CEPT and GSM codes schemes";
- Part 6: "Survey of existing PSTN, ISDN and mobile networks, and a user survey of supplementary service use within Centrex and PBX environments";
- Part 7: "Experimental evaluation of draft ETS 300 738".**

Blank page

1 Scope

This multi-part ETSI Technical Report (ETR) presents the results of the research work conducted to develop a European Telecommunication Standard (ETS) defining a harmonized minimum man-machine interface (MMI) for the access and control of public network based telecommunications services, and in particular supplementary services.

This part 7 of the ETR presents the results of an experimental evaluation of the harmonized minimum man-machine interface as defined in the draft ETS being produced under ETSI Work Item to define a minimum harmonized MMI for the access and control of public network based supplementary services [3].

The report describes the methodology, subjects, and a simulation developed to support the MMI, as the common agreed experimental approach to the evaluation. The results and conclusions discussed are based on two sets of experimental data.

2 References

For the purposes of this ETR, the following references apply:

- [1] Allison, G and Hewson, T: "Usability evaluation of 3 protocols for conference calling (3-way calling)". Proc. 15th HFT Symposium, Melbourne, 6-10 March 1995.
- [2] CEPT: T/CAC 02: "Subscriber control procedures for supplementary services in modern telecommunication system".
- [3] ETS 300 738: "Human Factors (HF); Minimum Man Machine Interface (MMI) to public network based supplementary services".
- [4] ETR 261-3: "Human Factors (HF); Assessment and definition of a harmonized minimum man-machine interface (MMI) for accessing and controlling public network based supplementary services; Part 3: Experimental comparison of two MMIs - Simulated UPT access and prototype ISDN supplementary services".
- [5] ETR 261-4: "Human Factors (HF); Assessment and definition of a harmonized minimum man-machine interface (MMI) for accessing and controlling public network based supplementary services; Part 4: Experimental comparison of the effect of categorized and non-categorized formats within user instructions".
- [6] ETR 261-5: "Human Factors (HF); Assessment and definition of a harmonized minimum man-machine interface (MMI) for accessing and controlling public network based supplementary services; Part 5: Experimental comparison of the CEPT and GSM codes schemes".
- [7] ETR 261-6: "Human Factors (HF); Assessment and definition of a harmonized minimum man-machine interface (MMI) for accessing and controlling public network based supplementary services; Part 6: Survey of existing PSTN, ISDN and mobile networks, and a user survey of supplementary service use within Centrex and PBX environments".
- [8] ETS 300 511: "European digital cellular telecommunications system (Phase 2); Man Machine interface (MMI) of the mobile station (MS) (GSM 02.30)".
- [9] ITU-T Recommendation E.131: "Subscriber control procedures for supplementary telephone services".
- [10] ITU-T Recommendation E.161: "Arrangement of figures, letters and symbols on telephones and other devices that can be used for gaining access to a telephone network".

- [11] ETR 261-1: "Human Factors (HF); Assessment and definition of a harmonized minimum man-machine interface (MMI) for accessing and controlling public network based supplementary services; Part 1: General approach and summary of findings".

3 Definitions, symbols and abbreviations

For the purposes of this part of the ETR, the definitions, symbols and abbreviations given in part 1 [11] of the ETR apply.

4 Introduction to the experiment

The ETR is complementary to the ETS being drafted in parallel on a harmonized minimum MMI for accessing and controlling supplementary services [3], and reports a study that was conducted in order to support empirically some of the recommendations made in the draft ETS. In particular, it was intended to provide data to support the approach taken in the draft ETS of focusing on three elements as primary factors in usability. These comprise the information needed prior to the control action (e.g. handbooks, prompts, etc.), the control action itself (e.g. a command for activating a service), and the information after the control action (e.g. feedback). The issue addressed with the present study is whether the level of information (in terms of quantity and modality) specified in the draft ETS is indeed required for the MMI to be usable, or whether less information presented in a single mode only (e.g. through tones) is sufficient.

The results of an earlier survey (part 6 of the present ETR [7]) showed widespread ignorance about the existence or use of supplementary services. A pilot experimental study (part 3 of the present ETR [4]) also showed that the information given prior to the control action is critical to being able to control supplementary services successfully using the minimum MMI implemented through the 12-key telephone keypad. Other data collected during user experiments showed that using an MMI becomes almost impossible if no or inadequate information or handbook is available. Finally, the study results (part 4 of the present ETR [5]) showed that information provided in handbooks should be carefully structured in order to make explicit the logic underlying the syntax of the MMI.

The present study examined three interfaces that varied in terms of the amount of information and the modes in which information was presented prior to and after a control action. All three implementations were compliant with the recommendations given in the draft ETS (with the exception of PBI, see below). The reason for only testing compliant implementations lay in the fact that the previous experimental work reported in part 5 of the present report [6] indicated that the choice of the code scheme (i.e. CEPT or GSM) did not matter for the user as much as the instruction and feedback information, which seemed to be far more critical. It should also be added that the CEPT Service Codes used in the present study (e.g. "21" for Call Forwarding) were not changed in the different experimental conditions.

Three different man-machine interfaces were used:

- PBI the feedback consisted of network tones only;
- PBI+ the feedback consisted of tones and verbal announcements;
- PBI++ the feedback consisted of tones, announcements and visual messages.

Three supplementary services were tested in the present experiment: Call Forwarding Unconditional (CFU), Call Barring (CB) and 3-Party Call (3PTY). They were chosen because they are representative of the various categories of supplementary services. In particular, 3PTY is switching order based, whereas CFU and CB are code scheme based, and CB is defined in a negative manner (i.e. switching on the service bars the outgoing calls) and requires entry of supplementary information (e.g. PIN) in addition to the service code.

In addition to studying the effects of different feedback modalities as provided by the three interfaces (PBI, PBI+, and PBI++), the issue of syntax complexity was addressed. The absolute minimum functionality to be provided by a phone-based interface consists of control sequences for activating and de-activating a service. Any additional functionality increases the syntax in terms of further commands to be learnt by the user. Interrogating the system on whether a particular service is currently active or not ("interrogate"-function) is from a human factors point of view an additional desirable function that should be implemented as part of the MMI. In order to be able to assess the impact of increasing the minimum functionality (and consequently the syntax complexity) from "activate" and "de-activate" to include

"interrogate", two implementations of the PBI condition were tested only one of which included the specific command sequence for "interrogate".

Finally, the question was addressed whether, after sufficient learning time, the services could be used without supporting instructional material.

The experiments were conducted in two separate locations, at CSELT in Turin, Italy, and the Heinrich Hertz Institute in Berlin, Germany, using a common agreed experimental design.

Procedurally, each subject was asked to attempt to gain access to a service, interrogate it or deactivate it, on one of the three interfaces, according to specific user instructions supplied.

The aim of the experiment was to provide data to help answer five questions:

- 1) Is PBI++ better than PBI+, and PBI+ better than PBI?
- 2) What effect does the information before the control action have on usability?
- 3) What effect do feedback or prompts during and as a result of the control action have on usability?
- 4) What effect does the increase in syntax complexity from "activate" and "de-activate" to including "interrogate" have on usability?
- 5) Can the users learn the procedures after making use of the services twice, or do they have to rely on printed material for a longer period of time?

5 Method

5.1 Subjects

A total of 54 subjects, 30 men and 24 women, took part in the experiment, 18 subjects in Turin and 36 subjects in Berlin. In both cases, subjects were randomly assigned to one of the three different interfaces: PBI, PBI+, and PBI++, and the three subject-groups were balanced as far as possible across age, education level and type. Three categories of age were considered (<35, 36-49, and >50), and two education levels (medium and high). The type of education was further distinguished between technical (computers, telecommunications, etc.) and non-technical (secretarial, administration, medicine, etc.). For further data on the subject samples, see annex A.

5.2 Experimental procedure

Each subject was given the following material in his national language:

- general instructions about the scope and duration of the experiment, and the confidentiality of data (see annex B);
- description of the services (text instructions - see annex C);
- description of the services (graphical instructions - see annex C);
- scenarios (see annex D);
- short questionnaires (see annex D);
- long questionnaire (see annex D);
- questionnaires on personal data (age, education, profession, experience with telecommunications, attitude towards technology).

The text instructions (see annex B) contained a definition of the services, the commands available and an explanation of the logic underlying their syntax. This was done with the aim of helping the users to remember the commands. The instructions were also worded in a task-oriented rather than a service-oriented manner, according to the good practice of user-centred design.

The graphical instructions (annex C) consisted of as many sets of pictograms as the number of service options (switching on, off, interrogating, etc.). They illustrated the user procedure and included commands, prompt and feedback indications.

The instructions were presented in two alternative modes - Assisted Mode and Learning Mode.

In Assisted Mode, the subjects were presented with the textual and graphic instructional material, but the textual material was removed for the experimental sessions, leaving only the graphical material on the

table. The subjects could request assistance from the experimenter when they felt they needed it. The textual material was available on demand. The Turin group subjects used the Assisted Mode.

In Learning Mode, the same textual and graphical material was presented to the subjects who were expected to read, assimilate and remember the material, all of which was then removed for the experimental sessions. This left the subjects reliant solely on their memory, although, as in Assisted Mode, the subjects could ask for the material if they felt it necessary (in both modes subjects rarely took advantage of this). In this mode, the subjects were allowed to make more than one attempt at each task. All subjects of the Berlin group used Learning mode.

The tasks to be performed were developed into scenarios. They were chosen so as to cover all the possible options offered by the services, and are listed below:

CFU:

- switching the service on (CFA - Call Forwarding Activate);
- making an outgoing call (CFB - Call Forwarding Basic call);
- switching the service off (CFD - Call Forwarding De-activate);
- interrogating the service (CFI - Call Forwarding Interrogate).

CB:

- switching the service on (CBA - Call Barring Activate);
- trying to make an outgoing call (CBB - Call Barring Basic call);
- switching the service off on a per call basis (CBC - Call Barring Cond. de-activate);
- switching the service off (CBD - Call Barring De-activate);
- interrogating the service (CBI - Call Barring Interrogate).

3PTY:

- calling user A (3P1 - 3-Party Basic Call 1);
- putting call on hold and calling user B (3P2 - 3-Party Call 2);
- recalling user A (3PA - 3-Party Alternate);
- setting up the conference, i.e. connect with A and B (3PE - 3-Party Establish Conference Call).

Since it was hypothesized that a more complex syntax would lead to more errors, the interrogation tasks CFI (Call Forwarding Interrogate) and CBI (Call Barring Interrogate) were implemented under the PBI condition in two versions: the "Assisted Mode"-subjects (Turin sample) were instructed to dial the relevant command to interrogate (e.g. "*#21#"), whilst in "Learning Mode", (Berlin sample) the subjects were told to pick up the receiver and listen to the dial tone which indicated the state of the service (Special Dial tone: a service is activated, Normal Dial tone: no service is activated).

A Short Questionnaire consisting of only three questions on the user procedures and the adequacy of the feedback was administered after completion of all the tasks of each service. The three statements to be rated with a 5-point rating scale ranging from "agree completely" to "disagree completely" were:

- 1) The service commands were complicated;
- 2) The information conveyed by the tones and vocal messages (PBI+/PBI++) or solely by the tones (PBI) was insufficient, and;
- 3) The information conveyed by the visual messages was insufficient (PBI++ only).

A Long Questionnaire consisting of 15 questions on various issues (instructions, commands, tones, messages) was presented after all sessions at the end of the experiment.

The experiment was conducted over two sessions in which the subjects had to perform a number of tasks on each of the three services. These tasks concerned switching on and off and interrogating the services. The time taken for each session ranged from 50 minutes to over 180 minutes, with the majority of subjects completing a session in about 60 minutes.

The three services were tested in the six possible orders of presentation (A to F) to compensate for possible learning effects.

The performance data captured during the experiment consisted of the overall performance time, keystrokes, number and types of errors, the first two being automatically stored into log-files. The times

were calculated from the instant when the subjects picked up the receiver to the instant when it was replaced.

Since the services tested by the users were simulated and not accessed via a network, setting up calls did not result in speaking to another party, but a recorded message was heard instead. The subjects were informed of this prior to the start of the experiment.

5.3 Equipment

The telephone service offering the supplementary services for the experiment was simulated, using a program prepared by the Norwegian Telecom Research Institute, based on a specification developed by the project team. Essentially, the simulation reproduced the appropriate network responses, tones, voice messages and visual indications, for each MMI and responded to user input in a way that was as similar as possible to network supplementary services designed to comply with the draft ETS [3]. It must, however, be noted that the Special Dial tone was known to the Turin group (Assisted Mode) with a different meaning: the Special Dial tone (Continuous 425 Hz and 350 Hz tone) is known in Italy as a congestion tone.

The simulation ran under a Microsoft Windows 3.1 environment, implemented on standard PCs (Italy: Compaq Prolinea 425S, 486SX processor at 25 MHz, 120 Mbyte hard disk, 8 Mbyte memory and VGA colour monitor; Germany unnamed-PC 386SX processor at 25 MHz, 40 Mbyte hard disk, 4 Mbyte memory and VGA colour monitor). The user interface was implemented both on a screen (for the experimenter) and via a standard telephone set (for the subject). The telephone set was supplemented by a small LCD screen to display the visual messages in PBI++.

The experimental set-up is illustrated in figure 1 below.

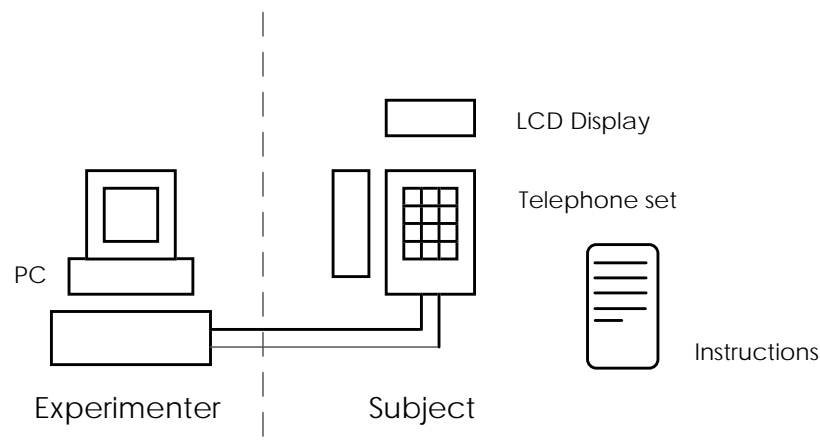


Figure 1: Schematic of experimental set-up

Unfortunately, before the experiment could be started on the Turin site, a breakdown occurred to the LCD of the telephone set used. It was not possible to repair or replace it in time, so that a display was presented on the screen of a monitor instead. Care was taken to adjust the relative position of the monitor and the phone so as to ensure that the display would appear as close as possible to the telephone keypad, as in a real terminal. It should be noted, however, that contrast and legibility of the display on the CRT screen were much better than those of an LCD.

At the Italian site, an additional application was prepared to capture the personal data of the subjects and their answers to the questionnaires. The environment used was a Sun IPX workstation with 525 Mbyte hard disk and 32 Mbyte memory. The graphical application was developed with the software tool Interface Architect 2.0, resident on a HP 725/50 workstation and then transferred to Sun/Unix environment.

5.4 Data capture

The following data were captured:

- personal data;
- keystrokes;
- performance times;
- number of failures;
- types of errors;
- answers to the questionnaires;
- subjects' comments.

Two separate questionnaires were employed, both based on a 5-point Likert scale:

- a) six short questionnaires (2 - 3 questions each) presented after completion of the tasks relevant to each service of each session;
- b) one long questionnaire (11 - 15 questions for each interface) presented after both sessions had been completed.

The procedure used to capture this data was:

- from a VDU (in Turin) for the personal data and the long questionnaire, (paper and pencil in Berlin);
- automatic logging for keystrokes and performance times;
- direct observation for errors and consultation of instructions;
- paper and pencil for the short questionnaires and the subjects' comments.

6 Results

6.1 Personal data

Subjects were carefully selected and allocated to conditions in a fair and as balanced a way as possible, but it nevertheless emerged that some bias appeared in both groups in certain attributes. For instance, in the Italian group (tables A3 and A4), the average age of subjects allocated to PBI, PBI+, and PBI++ was well balanced at a mean of 41, 43, and 45 respectively, but their educational levels were not. Two thirds of the group allocated to PBI and PBI++ were of a medium level of educational attainment and two thirds of those for PBI+ at a high level. The division between the sexes was also biased, with only 20 % of the subjects allocated to PBI being female, but two thirds allocated to PBI++ being male. For the German group (tables A1 and A2), PBI++ had twice as many university graduates allocated than to PBI, whereas twice as many subjects considered as potential graduates (i.e. would have had access to University) were allocated to the PBI interface. Further, some imbalance occurred in subjects allocated to interfaces in the Berlin group with experience with a personal computer and also with having had prior usage of the supplementary service Call Forwarding Unconditional.

In all other respects, the subject groups could be considered homogeneous, and although earlier work (part 4 of the ETR [5]) suggested a correlation between educational level and some aspects of performance, it is not thought that this apparent bias in itself had any significant effect on the results. An analysis of variance (ANOVA) performed on the data showed that the subjects classified into the three educational categories (medium, higher, and university) did not differ significantly in terms of their overall achievement scores in the experimental tasks. It must, however, also be noted that 19 out of the total of 54 subjects refused to answer the question on their educational background.

6.2 Performance

Both separate and combined analyses were conducted on the two modes.

In "Assisted Mode", when a subject did not perform a task correctly, he was invited to proceed to the next one and an error was counted. In contrast, the subjects under the "Learning Mode" condition were allowed to try to perform the task as often as they wanted, although only the first two attempts were considered in the analysis. Sometimes the subjects realised that they had made a mistake, restarted the task and completed it successfully: this further attempt was also ignored by the experimenter. When the subjects were not aware they had failed the task, they were informed of this and invited to read carefully the task and service description again, but were not permitted to make another attempted call. No explicit solution was suggested.

Missed tasks were counted as errors (e.g. some subjects forgot to perform a given task or intentionally did not perform it because they felt it unnecessary or because they had not read carefully the task instructions). Misdialling of directory numbers was classified as a "supplementary information" error (see below). Hesitations resulting in time outs expiring and consequent need to restart were not counted as errors, so that keystrokes contained in the log files needed a certain degree of interpretation.

Correctly achieving a target in a way not foreseen by the task assigned was considered as an error. A typical example is provided by the task requiring the subjects to interrogate the Call Barring service. Some subjects (maybe to avoid dialling the relevant command) achieved the target by trying to make a call; others listened to the modified dial tone after picking up the receiver.

In the following tables (tables 1 to 8), the percentage of successes (i.e. tasks completed correctly) is reported as a function of the variables of the experiment (interfaces, services, sessions, tasks). Two versions are provided for each table, a and b, containing the successes of the Turin sample (Assisted Mode) and of the Berlin sample (Learning Mode) respectively. In tables 3, 4, 7, and 8, performance figures are given including and excluding (in brackets) "basic call set-up" tasks in order to give a truer picture of errors caused solely by the supplementary service access and control procedures rather than by a misunderstanding of the task instructions.

Table 1: Percentage task performed successfully by Interfaces (INT), Services (SERV), Sessions (S) and Tasks: Turin sample, "Assisted Mode" n = 18

SERV	INT	S	CFU				CB					3P			
			Tasks				Tasks					Tasks			
			CFA	CFB	CFD	CFI	CBA	CBB	CBC	CBD	CBI	3P1	3P2	3PA	3PE
PBI		1	100	50	100	83	67	83	50	83	83	83	100	50	100
		2	100	67	100	100	83	100	83	83	83	100	83	67	83
PBI+		1	100	67	83	100	67	83	67	67	67	100	100	17	100
		2	100	83	83	83	83	100	83	83	83	100	100	83	100
PBI++		1	83	83	83	100	50	67	67	100	67	83	83	33	100
		2	67	100	100	100	83	100	100	100	100	83	100	83	67

Table 2: Percentage tasks performed successfully by Interfaces (INT), Services (SERV), Sessions (S) and Tasks: Berlin sample, "Learning Mode" n = 36

SERV	INT	S	CFU				CB					3P			
			Tasks				Tasks					Tasks			
			CFA	CFB	CFD	CFI	CBA	CBB	CBC	CBD	CBI	3P1	3P2	3PA	3PE
PBI		1	92	92	83	67	92	92	75	92	83	100	100	83	92
		2	91	91	91	100	82	100	82	100	91	100	100	100	100
PBI+		1	83	75	58	75	92	92	8	33	50	92	92	50	83
		2	92	92	58	92	75	100	17	67	58	92	92	92	100
PBI++		1	75	92	83	58	67	92	25	67	67	92	100	58	100
		2	75	100	83	83	83	92	25	67	42	100	100	83	100

KEY to tables 1 and 2 for Tasks. (see also subclause 3.2)

CFA: switching on the service: activate	CBA: switching on the service: activate	3P1: call party A (basic call)
CFB: make outgoing call (basic call)	CBB: try outgoing call (basic call)	3P2: put A on hold, call user B
CFD: switch off the service: de-activate	CBC: switch off per call	3PA: recall user A
CFI: interrogate service	CBD: switch service off: de-activate	3PE: set up conference call
	CB I: interrogate service	

Table 3: Mean percentage of tasks performed successfully by Interfaces and Services. Assisted Mode sample, n = 18. Figures in brackets [] represent mean percentage for all tasks excluding the basic call tasks

INTERFACES	CFU		CB		3PTY		MEAN	
PBI	88	[97]	80	[77]	83	[81]	84	[85]
PBI+	88	[92]	78	[75]	88	[83]	84	[83]
PBI++	90	[89]	83	[83]	79	[78]	84	[83]
MEAN	89	[93]	80	[78]	83	[81]	84	[84]

Table 4: Mean percentage of tasks performed successfully by Interfaces and Services. Learning Mode sample, n = 36. Figures in brackets [] represent mean percentage for all tasks excluding the basic call tasks

INTERFACES	CFU		CB		3PTY		MEAN	
PBI	90	[87]	89	[87]	97	[96]	92	[90]
PBI+	78	[76]	59	[50]	86	[85]	74	[70]
PBI++	81	[76]	63	[55]	92	[90]	79	[74]
MEAN	83	[80]	70	[64]	92	[90]	82	[78]

Table 5: Mean percentage task performed successfully by Tasks and Services. Assisted Mode sample, n = 18

Tasks	CFU		CB		3PTY	
1	CFA	92	CBA	72	3P2	94
2	CFD	92	CBC	75	3PA	56
3	CFI	94	CBD	86	3PE	92
4			CFI	81		

Table 6: Mean percentage task performed successfully by Tasks and Services. Learning Mode sample, n = 36

Tasks	CFU		CB		3PTY	
1	CFA	85	CBA	82	3P2	97
2	CFD	76	CBC	39	3PA	78
3	CFI	79	CBD	71	3PE	96
4			CFI	65		

Table 7: Mean percentage task performed successfully by Interface and Sessions and by Services and Sessions. Assisted Mode, n = 18

Session	PBI		PBI+		PBI++		CFU		CB		3PTY	
1	79	[82]	78	[77]	77	[77]	86	[92]	71	[70]	79	[76]
2	87	[87]	90	[88]	91	[90]	90	[93]	90	[87]	88	[85]

Table 8: Mean percentage task performed successfully by Interface and Sessions and by Services and Sessions. Learning Mode, n = 36

Session	PBI		PBI+		PBI++		CFU		CB		3PTY	
1	89	[86]	68	[62]	75	[70]	78	[75]	68	[63]	87	[84]
2	94	[94]	79	[74]	80	[74]	87	[85]	71	[66]	96	[96]

6.3 Results for the "Assisted mode" (Turin sample)

Effects of interfaces. Table 7 shows that, generally speaking, the subjects under all three interface conditions PBI, PBI+, and PBI++ performed fairly well: under conditions PBI+ and PBI++, nine out of ten subjects were able to complete successfully their tasks already at the second attempt. The same table also shows that the subjects under the three conditions did not differ much in terms of their overall performance (an analysis of variance confirmed that the differences were not significant). This result is surprising, since the feedback provided in PBI+ (verbal announcements) was more meaningful than that of PBI (tones only) and, hence, should have improved the performance. The same consideration applies to PBI++ which provided prompt and feedback messages in a redundant way, that is both verbally and visually. This inconsistency may only be an apparent one. One possible explanation may be the fact that the modified dial tone used in the experiment to indicate that a service (CF or CB) was active was the continuous tone at 425 Hz, which has been standardised only recently, and was already in use in the Italian network with the meaning of "line congestion" or "line out of order". Therefore, the Italian subjects under all three conditions may have interpreted this tone as a feedback indicating that dialling is not possible, and the verbal and visual messages may have confused the PBI+ and PBI++ subjects because they were (apparently) conflicting with the information conveyed by the tones.

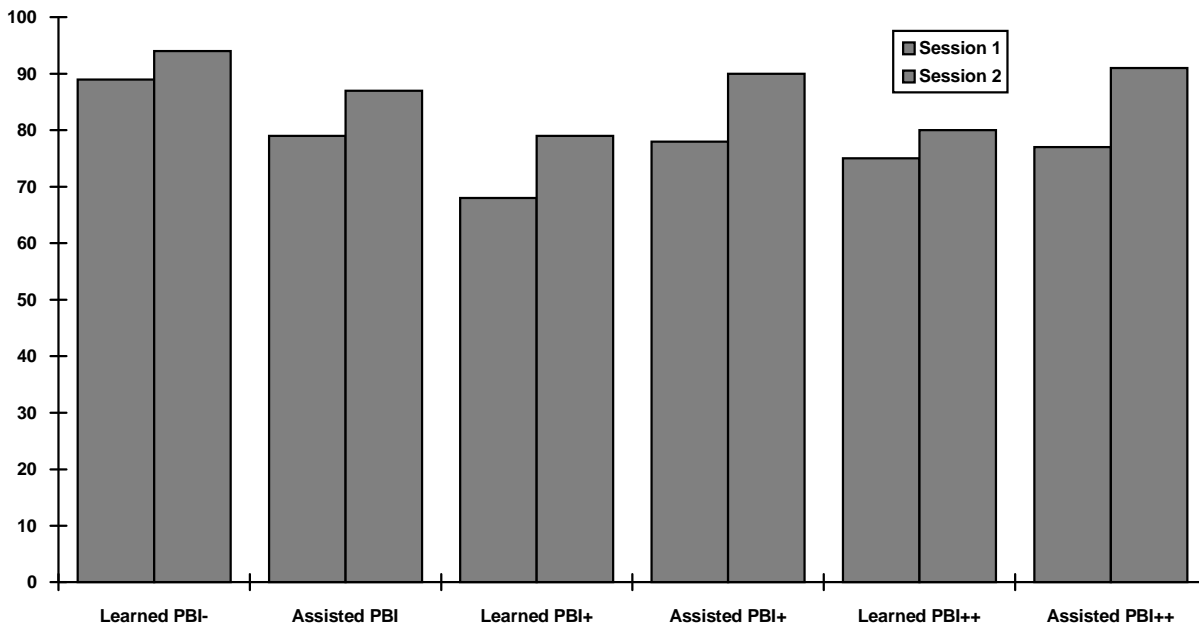


Figure 2: Success rates for different experimental conditions

Effect of services and tasks. Table 7 also shows that the lowest success rate was achieved with the CB service. The simplest user procedures were those of the 3PTY service (very short commands without

* and #), but this was not reflected in a better performance because the users expected a service model other than the implemented one, as is explained in clause 7.

Table 5 shows that the three most critical tasks turned out to be the Activation task (CBA) in the CB service, the de-activation on a per-call basis (CBC) in the CB service, and the re-connecting task (3PA) in the 3PTY service. The first case can be explained by the "negative" definition of the CB service (the service was defined in a negative way in that to switch the line *on* one has to switch the service *off*), and by the complex command syntax of the de-activation on a per-call basis task (need to enter the PIN) in the second case. As to the third case, the final target was to set up a 3-party call; therefore the subjects felt it more natural, after dialling the second party, to add that party to the conversation rather than checking again whether the first party was still on hold.

For the 3PTY service, the errors are focused on the task requesting the subject to put the first user on hold. Since in the previous task, the second user had been put on hold by pressing "2", many subjects guessed that pressing "1" would have put on hold the first user. Implementing this service option with a toggle key violates the expectations of a number of users.

Effects of learning. Finally, table 7 and figure 2 show that a learning process appeared to take place for all interfaces and services (the success rate improved from the first to the second session). An ANOVA showed, however, that the improvement of the performance was significant only for the CB service ($p < 0,04$).

6.4 Results for the "Learning mode" (Berlin sample)

Effects of interfaces. When considering the results of the subjects under the "Learning Mode" condition, it has to be kept in mind that in this mode the PBI commands were implemented using a simpler syntax for the interrogation command (hence PBI-). Generally speaking, most tasks could be well performed using any of the three interfaces (table 1b). Table 4b shows that the subjects performed best under PBI- followed by PBI++ and PBI+: Under PBI-, 94 % of the tasks were completed successfully after the second session, while with the two other interfaces, the performance rate only went up to 80 %. An analysis of variance confirmed that PBI-performance was significantly better than PBI+ and PBI++. Even though PBI++ performance was consistently better than PBI+, these differences were not statistically significant.

PBI- subjects thus performed better than any other Learning Mode or Assisted Mode subject group. The reason for the excellent performance of this group clearly lies in the fact that they had to learn only very simple commands (activate and de-activate while interrogation was based on listening to the dial tone). The simple syntax resulted in less material to learn and consequently in a smaller risk of confusion (see also clause 7).

The fact that under Learning Mode, PBI+ and PBI++ subjects performed less well than the corresponding groups under Assisted Mode (table 2b) is not surprising given the fact that under Learning Mode, all information on the syntax had to be memorised because any syntax information was available only after request (on average, one subject per group and per task requested to see the information again).

Effects of services and tasks. Most subjects (90-100 %) were able during the second session to perform the tasks of the 3-Party conference scenario with the exception of the 3PA task (going back to user A after having connected to B and before establishing the conference call). As will be shown later (clause 7), almost all of those subjects who did not perform this task correctly failed because they skipped the task, i.e. it was a case of misunderstanding the task description rather than of not knowing the command syntax. This interpretation is also supported by the fact that performance went up from Session 1 (after being told that they had skipped a task) to Session 2 (PBI- 83 → 100, PBI+ 50 → 92, PBI++ 58 → 83).

As regards the CFU and CB scenarios, PBI+ and PBI++ subjects had difficulties with the interrogation tasks (CFI and CBI), with the de-activation tasks (CFD and CBD), and, as in the case of the Assisted Mode subjects, with de-activation on a per-call basis (CBC). The error analysis (clause 7) shows that the reason for these problems is to be found in the complexity of the syntax leading to difficulties in memorising the commands and consequently to confusion errors. The results suggest that CBC is so complicated a command that its syntax was not being remembered even after having been used twice (success rate in Session 2 PBI+: 17 % and PBI++: 25 %).

Effects of learning: As in the case of the Assisted Mode sample, performance significantly improved from Session 1 to Session 2.

6.5 Overall analysis

In order to compare the two instruction modes (Learning Mode vs. Assisted Mode), a series of (repeated measures) analyses of variance were conducted.

Effect of interfaces (three levels between-subjects): An ANOVA was performed predicting total performance scores by Interfaces and Modes. The results showed no significant effects for the Instruction Mode factor but significant differences between the three Interfaces ($p < 0,05$) in that PBI scored better than PBI+ and PBI++ (figure 3). The reason for this result has already been discussed above: PBI- (= Learning Mode PBI) led to a much better performance because of its low-complexity syntax.

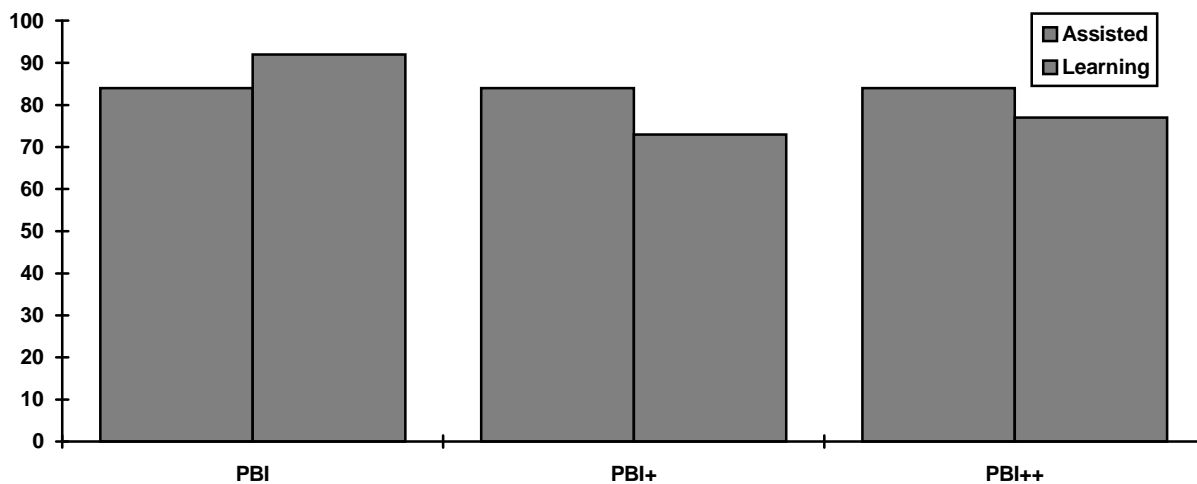


Figure 3: Performance by interface and presentation mode

Effect of sessions (two levels within-subjects): A repeated measures analysis of variance on the factors Session (within-subjects), Interface (between-subjects), and Mode (between-subjects) resulted in a significant main effect of the factor Session with the results of Session 2 being better than those of Session 1 ($p < 0,001$). No other main effect (Factor Interface, Factor Mode) nor any of the interactions (Interface by Mode, Interface by Session, Mode by Session, or Interface by Mode by Session) reached significance.

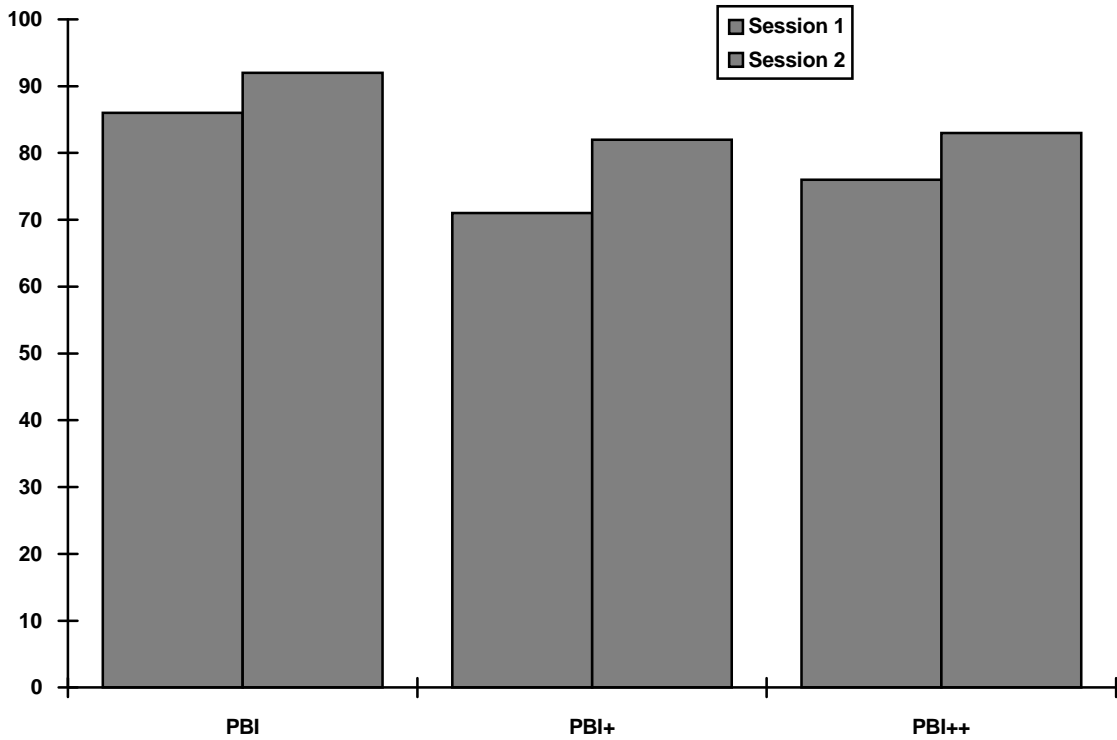


Figure 4: Performance by Interface and Sessions

Effect of services (three levels within-subjects): A repeated measures ANOVA on the factors Service (within-subjects), Session (within-subjects), Interface (between-subjects) and Mode (between-subjects) was conducted. Significant main effects were found for both factors Session (Session 2 > Session 1, $p < 0,001$) and Service (CFU and 3-Party > CB, $p < 0,001$). The interaction Mode by Service was also significant (CFU Assisted > CFU Learned, 3-Party Learned > 3-Party Assisted, $p < 0,005$). No other main effect or interaction was significant.

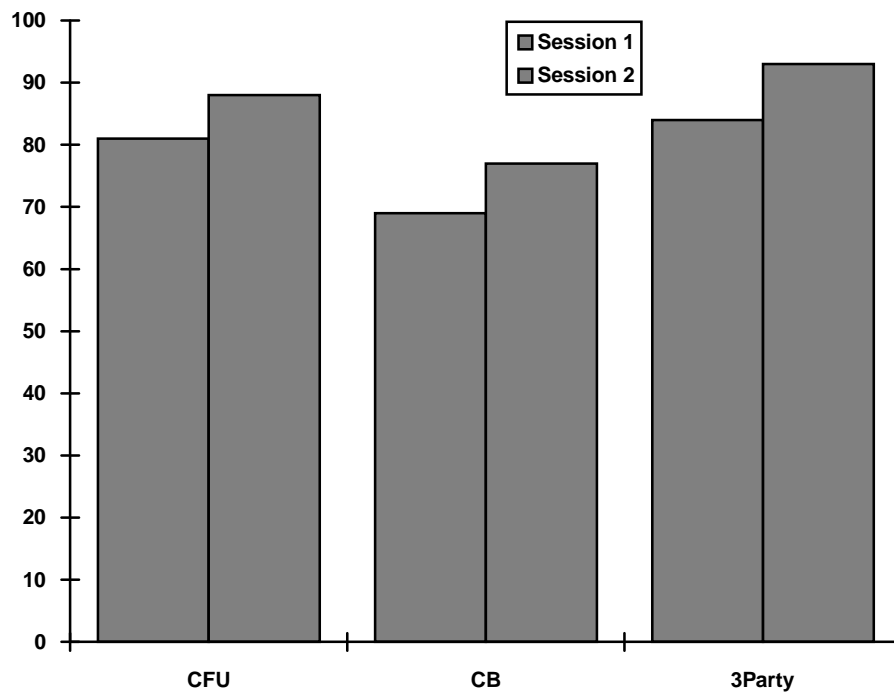


Figure 5: Performance by Services and Sessions

7 Detailed analysis

7.1 Error analysis

The errors which occurred during the experiment are classified in tables E1 to E6 ("Assisted mode") and E7 to E18 ("Learning mode") in annex E.

For the "Assisted Mode" experiment, tables E1 and E2 refer to PBI (1st and 2nd session), tables E3 and E4 to PBI+, and tables E5 and E6 to PBI++.

For the "Learning Mode" experiment, tables E7 to E10 refer to PBI (Session 1 Attempt 1; Session 1 Attempt 2; Session 2 Attempt 1; and Session 2 Attempt 2), tables E11 to E14 to PBI+, and tables E15 to E18 to PBI++.

Errors are classified as:

- Prefix errors (wrong prefix);
- Code/Switch order (wrong service code or wrong switching order in 3-Party);
- Separator error (using any other key than "*");
- Supplementary information error (wrong directory number or PIN);
- Suffix error (wrong or omitted suffix "#");
- Wrong task error (performing a different task);
- Skip task error (omitting a task);
- Restart error;
- No-hang-up error (subject forgot to replace the receiver), and;
- No interaction.

As a first point, it is important to distinguish between the errors caused by the test procedure and those caused by the user interface.

To the first category belong errors such as:

- using inadvertently data relevant to a wrong task (a single sheet contained more than one task);
- not understanding a task;
- skipping intentionally a task (e.g. because it was felt unnecessary);
- achieving a goal in a way different than the one required by a task (e.g. checking the status of the CB service by trying to make a call instead of using the interrogation function).

Given the goal of this experiment (validating the ETS on the PBIs), in the following, the focus of the discussion will be on the errors caused by the user interface.

The tables show that there are two types of errors:

- cross-interface errors;
- service-specific errors.

7.1.1 Cross-service errors

In the Turin sample ("Assisted Mode"), some errors (needless repetitions of tasks) and most uncertainties (hesitations, lack of confidence about the results of performed actions) were caused by the continuous tone, which was interpreted as rejected action or breakdown (see clause 4.3). This happened especially with PBI, but even with PBI+: the beneficial effect of the meaningful vocal message was cancelled by the ambiguity of the tone.

An important source of errors was the interrogation tasks which were felt to be disrupting and confusing. Many subjects did not understand their purpose and omitted them (e.g. in 3PTY, the target was to set up a 3-party call, so many subjects went straight to establishing the conference call omitting task 3PA) or interrogated the service instead of making a regular call (e.g. in CFU and CB: in the latter, the effect of calling or interrogating was equivalent). This occurred with all the interfaces.

7.1.2 Service-specific errors

CFU and CB:

- a) There were many more prefix errors under PBI+ and PBI++ in "Learning Mode" than under any other condition. For those conditions, prefix errors accounted for more than half of the errors made. For those subjects, interrogate (CFI and CBI) and de-activate on a per-call basis (CBC) were particularly difficult. Many tried "#*" (which was not defined), "#", or "*" for interrogate;
- b) The few Separator errors that were made occurred almost exclusively during Session 1 indicating that the separator symbol is easy to learn.

CFU:

- a) Interrogate (CFI) and de-activate (CFD) were the most problematic commands (mostly due to the prefix).

CB:

- a) A number of command errors concerned the command syntax (PIN required for activating and de-activating), which is complicated and difficult to remember;
- b) Many errors were caused by the confusion between switching on and off the service. In fact, for many users, enabling the line and enabling the service were synonyms and the definition of the service violated the user expectations (to enable the line the service must be de-activated);
- c) One particular difficulty existed for CBA under PBI: if the users entered the wrong service code ("21" instead of "33"), they diverted the call to their PIN (e.g. "**21*4321#"). In this case, the PBI-feedback indicated that a service was indeed active, and the users were unlikely to notice that they had made a mistake.

3PTY:

- a) A common error was to press "1" (instead of "2") to put the first user on hold. The subjects who made this mistake were convinced that different keys were associated with the different parties to be put on hold: 1 for the first party, 2 for the second party. The implementation of this service option with a toggle key (the *same* key was used to put on hold both the 1st and the 2nd parties) violated the expectations of these subjects;
- b) A few errors consisted of pressing "2" *before* the user answered. This was due to the ambiguous interpretation of a comma in the instructions;
- c) As mentioned before, most 3PA errors turned out to be Skip Task errors.

7.2 Timing

Tables 9 and 10 contain information on the behaviour of the Italian subjects observed by the experimenter. The tables show that hesitation and perplexity were fairly frequent situations. They can be explained partially by the fact that some subjects read the instructions after picking up the receiver. The main responsibility, however, is to be attributed to the modified dial tone.

Table 9: Hesitations and expirations of time out

Task	PBI		PBI+		PBI++	
	1	2	1	2	1	2
CFA					5	4
CFB			6			
CFI					5	
CFD					5	
CBA						
CBB					3 4	
CBC		2 3				
CBD						
CBI						
3P1					5	
3P2				2		6
3PA						
3PE						

Table 10: Perplexity/uncertainty

Task	PBI		PBI+		PBI++	
	1	2	1	2	1	2
CFA	2	2 6				
CFB	1 2		3 5 6			
CFI	2 6	2				
CFD	2	2				5
CBA	2 5 6	2				
CBB	5	3	3	3	6	3
CBC			3			
CBD	2 5	3	3			
CBI	2					
3P1						
3P2						
3PA						
3PE						

7.3 Knowledge transfer

It could be expected that, given the similarity of the command syntax of the CF and CB services, an effect of the order of presentation of the services exists. In other words, there could be a difference in the performance of, say, service CF (or CB) when used first and when used after CB (or CF). On the other hand, such effect should not be detected when either CF or CB is preceded by 3P, and vice versa.

No such effect was detected (probably because of the small sample): whilst there were differences, they were not consistent. This result agrees with the outcome of another experiment carried out on these code schemes.

7.4 Short questionnaire

The results of the Short Questionnaires as a function of Interfaces, Services, and Sessions are reported in tables 11 and 14. The Short Questionnaire consisted of only three questions on the user procedures and the adequacy of the feedback, and it was administered after completion of all the tasks of one service. The three statements to be rated with a 5-point rating scale ranging from 1 "disagree completely" to 5 "agree completely" were:

- 1) "The service commands were complicated";
- 2) "The information conveyed by the tones and vocal messages (PBI+/PBI++) or by the tones (PBI) was insufficient"; and
- 3) "The information conveyed by the visual messages was insufficient" (PBI++ only).

The answers to the questions of the second session, which were worded in a negative form, have been reversed to facilitate the comparison of the results. The highest score is 5 and the lowest score is 1.

7.4.1 Assisted Mode

Interfaces: Generally speaking, the marks given by the subjects on the command structure, auditory and visual feedback were fairly positive (table 11). When implemented on a terminal with speech and text feedback (PBI++), the commands were rated as less difficult than under the two remaining interfaces. The tones were considered insufficient to a larger degree by PBI-users who had to rely on them exclusively than by PBI+ and PBI++ users who made use of other modalities as well. Very few PBI++ users considered the text as insufficient. In absolute terms, the messages are rated between "good" and "very good".

Table 11: Mean percentage responses to Short Questionnaire by Interfaces, Services, and Sessions (Sample "Assisted Mode", n = 18) (1 "disagree completely" - 5 "agree completely")

Quest.	Interfaces			Services			Sessions	
	PBI	PBI+	PBI++	CFU	CB	3PTY	1	2
Comm. diff.	2,06	2,42	1,87	1,97	2,50	1,86	2,24	1,98
Tones insuff.	2,28	1,86	1,53	2,11	1,94	1,61	1,98	1,80
Text insuff.	-	-	1,53	1,58	1,67	1,33	1,61	1,44

Services: Not surprisingly, CB commands were rated as being more difficult than CFU commands. Handling 3-Party calls was not considered difficult at all. Throughout the three services, the tones were considered sufficient, and the textual information was least insufficient in the context of CFU and most under CB.

Sessions: Complementing the performance results reported above, the subjects rated the command difficulty and the adequacy of audio and visual feedback more positive in Session 2 than in Session 1.

Interfaces and services: The ratings improve when moving from PBI to PBI+ and to PBI++, but only for CFU and CB. This confirms that the problems related to the 3PTY service do not depend on the feedback or the complexity of the commands, but possibly on a different user model.

A statistically significant difference between the ratings under the three interface conditions concerned the adequacy of the audio feedback under CB, as shown in table 12:

Table 12: Comparison of subjective assessment of audio feedback in Call Barring

Session	PBI	PBI+	PBI++
1	2,50	1,83	1,67
2	2,83	1,33	1,50

PBI-subjects considered the tones insufficient in Session 1 and even more so in Session 2, i.e. having encountered the tones before did not seem to have helped in Session 2. In interpreting these results, it has to be kept in mind that the Assisted Mode sample had to dissociate the special dial tone from its line-congestion meaning in Italy). PBI+ and PBI++ subjects were confused by this because of additional feedback, and by Session 2, audio feedback was rated as being quite appropriate.

A second significant difference between interfaces (albeit only in Session 2 results) concerns ratings of commands for 3-Party conferences. PBI+ users found those commands difficult in Session 1 and, unlike the other two groups, even more so after Session 2. There is no obvious and plausible explanation for this result shown in table 13:

Table 13: Comparison of subjective assessment of user commands for 3 Party Conference

Session	PBI	PBI+	PBI++
1	1,33	2,17	2,17
2	1,33	2,83	1,33

7.4.2 Learning Mode

Interfaces: PBI- users rated the commands as less difficult than did PBI+ and PBI++ subjects. This is not surprising considering that the command syntax of PBI- was less complex in that the subjects did not have to learn the "interrogate" commands. The fact that Learning Mode PBI+ and PBI++ subjects did not rate the commands as (significantly) more difficult indicates that the commands were not considered particularly difficult irrespective of whether they had to be applied with instructional material present or whether they had to be memorised, and the performance figures support this interpretation. The Learning Mode sample gave consistently better ratings to audio feedback which can be explained by the fact that, to them, the special dial tone was new (it is not being used in Germany) and not associated with a particular line state.

Table 14: Mean percentage responses to Short Questionnaire by Interface, Services, and Sessions (Sample "Learning Mode", n = 36)

Quest.	Interfaces			Services			Sessions	
	PBI	PBI+	PBI++	CFU	CB	3PTY	1	2
Comm. diff.	1,58	2,43	2,19	2,03	2,67	1,54	2,22	1,94
Tones insuff.	1,50	1,36	1,44	1,43	1,50	1,37	1,58	1,31
Text insuff.	-	-	1,93	1,63	2,13	2,04	2,14	1,72

Services: The rank order of services in terms of the difficulty of their commands is the same as for the Assisted Mode subjects: CB commands were rated as being more difficult than CFU commands. Handling 3-Party calls was again not considered difficult at all. Audio and visual feedback were considered appropriate throughout the three services. The results do, however, indicate that the textual information was less sufficient for CB than for CFU and 3-Party conferencing.

Sessions: As in Assisted-Mode, the subjects rate the command difficulty and the adequacy of audio and visual feedback more positive in Session 2 than in Session 1.

Interfaces and services: a statistically significant difference between the ratings under the three interface conditions concerned the CB-commands, as shown in table 15:

Table 15: Comparison of subjective assessment of user commands in Call Barring

Session	PBI	PBI+	PBI++
1	2,08	3,25	2,75
2	2,00	3,33	2,58

The fact that CB-commands are less difficult for PBI- subjects is most likely due to the less complex command structure of CB under PBI-. As the error analysis (clause 7) showed, most errors made by the Learning Mode subjects were prefix errors, and most of those occurred for CFI, CBI, and CBC. The fact that, as in Assisted Mode, PBI++ subjects found the CB commands less difficult than did PBI+ subjects indicates that the feedback augmented by text results in a perception of the commands as being less difficult in multi-mode feedback.

7.5 Long questionnaire

The Long questionnaire was filled in by the subjects following the completion of the last task. It consisted of several statements (13 PBI+ and 11 PBI) for each of the three services which were rated with the same rating scales already used for the Short Questionnaires (5-point scales ranging from 1 "disagree completely" to 5 "agree completely"). Half of the questions were worded in a negative way. The 15 questions can be clustered into the following categories:

Table 16: Long questionnaire showing categories of questions

Category	No.	Questions - Statements to be rated
general	2	I had to concentrate hard when performing the task.
	6	When I used the service, I was always sure that my actions were correct.
	7	My experiences with the service were unpleasant.
	8	I enjoyed using the service.
instructions	3	The handbook section on service x was useful.
	4	The handbook section on service x was too long.
procedures	1	The commands for service x were easy to remember.
	10	The procedure for using service x were confusing.
audio feedback tones	5	The tones (and spoken messages) did not indicate what should be done next.
	9	The feedback provided by the tones (and spoken messages) was useful.
	11	I found it difficult to understand the meaning of the tones (and spoken messages).
verbal	12	I hardly paid attention to the spoken messages.
	13	This service is not usable without spoken messages.
visual feedback text	14	I hardly paid attention to the messages on the display.
	15	This service is not usable without messages on a display.
NOTE:		Where x is the specified service.

The answers to the Long Questionnaire on the three interfaces and the three services, after averaging over subjects, are reported in tables F.1 to F.3 in annex F.

7.5.1 Assisted Mode

General: Generally speaking, the subjects enjoyed using the services, even if they had to concentrate and were not always confident that their actions were correct. The highest scores were given to PBI+ which was rated significantly better ($p < 0,009$) than PBI++ on the "concentration" scale. This can be explained by the fact that with PBI++ the subjects, beside reading the instructions, performing the tasks and listening to the verbal announcements (as with PBI+), had also to heed the display. PBI+ was also the interface which received the highest rating as to the "confidence about the correctness of the actions performed": whilst with the other two interfaces the subjects were "neutral", with PBI+ they were "sufficiently confident".

Instructions: The assessments of the usefulness and conciseness of the instructions were extremely positive. This was confirmed by the opinions expressed by the subjects while "thinking aloud". There were no significant differences between interface groups on their attitudes to the handbooks.

Procedure: The subjective assessments of the statements that the procedures are "easy to remember" and "confusing" were only partly favourable. As already observed in the Short Questionnaires, the subjects considered CFU and 3-Party conferencing easier to learn than the CB commands. There were no significant differences between the interface groups.

Tones: The subjects' opinions about the usefulness and meaningfulness of the audio feedback were more negative for PBI. With PBI+ the scores are very positive ("the tones did indicate what to do") and significantly lower (i.e. more positive) than with PBI and PBI++ ($p < 0,01$ Question 5 for CFU and 3P, and $p < 0,01$ Question 11 for CB). This can be explained by the clarifying effect of the accompanying messages.

Verbal announcements: The opinions expressed about the verbal announcements were very interesting. All the subjects heeded the spoken messages, but significantly (at least under CFU and 3P, $p < 0,01$) more in the PBI+ than in the PBI++, where the attention was shared between the spoken and visual messages. In addition, most PBI+ subjects agreed that the services would not have been usable without the verbal announcements. As to PBI++ the attitude was neutral, probably because the availability of the visual messages reduced the need for the spoken ones. Some subjects pointed out that the messages delivered after interrogating and after switching on a service should be different (e.g. "the service is active" in the former case, "the service has been activated" in the latter case).

Visual messages: The users' reactions to the statements on the visual messages ("heeding" and "needing" the messages) were neutral. There are several possible reasons for this finding:

- the users were not accustomed to look at a display on the telephone set (only 50 % of the subjects were observed to look at the display);
- the visual messages (displayed in scrolling mode) were as ephemeral as the verbal ones;
- the visual messages were worded differently (more concisely) to the spoken messages;
- the visual messages were played at a lower speed than the verbal ones;
- the visual messages were less necessary, being accompanied by the spoken ones.

7.5.2 Learning Mode

General: The Learning Mode subjects were even more positive about their experience than were the Assisted Mode subjects. In CB and 3-Party conferencing, the PBI- enjoyed using the system (Questions 7 and 8) significantly more ($p < 0,05$) than PBI+ and PBI++ subjects, presumably because of the simpler command syntax.

Instructions: The assessments of the usefulness and conciseness of the instructions were extremely positive (means around 4,8 on the 5-point scale) throughout the three interface groups.

Procedure: Learning mode subjects found the commands easy to learn - in particular the 3-Party conference and CFU and CB under PBI-, possibly because they actually *had to* learn the commands and experienced that after Session 2, they were, in most cases, able to remember the commands whereas for Assisted Mode subjects, there was no need to memorise because they were always able to rely on the graphical instructions.

Tones: In CFU, PBI+ subjects rated the usefulness of the audio feedback significantly ($p < 0,05$) more positive than did PBI- subjects for whom tones were the only feedback mode, and than PBI++ subjects who could also make use of visual messages.

Spoken and Visual messages: In terms of the usefulness as a means for providing feedback, spoken messages were rated as far more important (in terms of paying attention to them (Questions 12 and 14) and in terms of usability (Questions 13 and 15) than visual messages. Possible reasons for this result have already been discussed above in the context of the Assisted Mode results.

8 Discussion of results and conclusion

The evaluation study confirmed most, but not all, of the assumptions made by the draft ETS on MMIs for phone-based systems [3]. In addition, it provided valuable insights into some critical issues in terminal and service design of such interfaces. Most errors observed in the study were due to the objective difficulty of using the services. However, as already mentioned above, the evaluation study was hampered in a few minor ways which will now be briefly discussed.

8.1 Effects of the experimental implementation

8.1.1 Meaning of tones

The Assisted Mode subjects (Turin sample) were confused by the fact that they had previously known the Special Dial Tone as indicating "Line congestion". This is particularly unfortunate since the meaning of the two indications are contrary: The Special Dial Tone indicates that a service is active, but dialling is still possible, whereas congestion means that call completion is not possible. In the case of the PBI, this led to confusion and lack of confidence about the outcome of the performed action. With PBI+ and PBI++, the semantic content of the tone was in conflict with the vocal and visual messages. This, too, led to confusion which resulted in no action or interrupted action (hanging up before or during dialling) or repeated action (restarting). This is a temporary problem related to the lack of familiarity of the subjects with the new tone and is bound to disappear with practice. The obvious conclusion therefore is that, when choosing new tones, care should be taken to avoid as far as possible the ambiguous ones. If a new tone already has a different meaning, this fact must be explicitly made known to the users.

8.1.2 Sub-optimum implementation of the display

Interviews conducted with the subjects after the experimental session revealed that the implementation of the display was sub-optimum. When the system was designed, care was taken that the information provided in a scrolling mode would not be displayed too fast. As it turned out, for some subjects, the display rate was too *slow*, leading to the effect that scrolling was completed seconds after the spoken message was completed. In such a situation, the written displays seemed superfluous for some subjects.

8.1.3 Sub-optimum instructions for experimental tasks

As mentioned before, many subjects skipped the 3PA task because the instructions did not make it clear that there was another sub-task (see annex B, instruction for 3PTY). This problem caused the performance rate for the 3PTY conference service to decrease even though the service was not considered very problematic (this view is also supported by the results of the Short and Long Questionnaires).

In two cases, some syntax errors can be accounted for by an inadequacy of the graphical instructions (see annex B). The instructions concerning the dial step consisted of a sentence ("Press the buttons:") followed by a pictogram representing a hand indicating the string of digits to dial. Since the index was pointing to the central digit of the string, this was interpreted by some subjects as a prompt to dial the central digit only.

8.2 Critical issues identified during the study

Generally speaking, the subjects performed well after the second session with some performance rates reaching 90 % and more, and most subjects enjoyed using the services. However, the study identified a number of critical issues which are discussed below.

8.2.1 Learning complex command strings

The issue of command complexity has already been addressed above. One command which turned out to be particularly difficult was CBC, a command that requires not one but two strings of supplementary information ("#33*PIN*Directory Number#"). Only very few subjects under Learning Mode condition were able to remember the command string correctly *and indicated that they would probably not use this command*: their solution would be to turn off CB, make the call, and activate CB again. The consequence is that commands like CBC must be well documented in appropriate instruction material, but should not be among the most basic commands a user can be expected to learn after few usage situations only: those should be restricted to activate and de-activate.

Another way to overcome the problem caused by the command syntax is to transfer the burden of associating the commands for the desired service from the user to the terminal. This can be done, for example, by using intelligent user interfaces equipped with a screen on which the services available are listed and can be selected by means of soft-keys. If a limited number of services is available, a solution based on dedicated keys is also possible. The basic phone interfaces are inadequate to solve the aforementioned problem, albeit that the addition of a display allows the user to check for the correctness of the data entered. A solution to the problem is to use different buttons for different functions. Were this not possible, the instructions should indicate clearly which command is associated with which service function. The best solution is to automate enabling tasks whenever possible; see Allison and Hewson 1995 [1].

8.2.2 Service definition

Some errors which may have been caused by a confusing service definition were recorded in connection with the CB service. Several subjects activated the service when they were supposed to de-activate it, and vice versa. This behaviour stemmed from the fact that the service was defined in a negative way: switching the service on switches the calls off and vice versa. For the subjects, on the contrary, enabling the "service" is equivalent to enabling the line. This problem can be solved by designing the service from the user's point of view, i.e. by relating the commands to the status of the line rather than of the service as seen from the network provider's point of view.

8.2.3 User models in 3-Party Conferencing

The error rate for the 3-Party conference service decreased from Session 1 to Session 2, but even after the second attempt it was still fairly high, and this may have been due to a conflict between service implementation and user expectations. The user procedure to put the connected party on hold required a press on the "2" button, regardless of the party, but some subjects had a different model of the service: they believed that two different buttons were to be used to put on hold the two parties, namely "1" for party B and "2" for party C ("1" and "2" could be interpreted as "short numbers" for parties B and C respectively). Since this "faulty" user-model (faulty in terms of the implementation) is equally plausible, it is likely that many users would make mistakes of the kind observed in the study. The consequence is that the instruction material must make explicit the model on which the implementation is based.

8.3 Conclusions

At the end of clause 5, a number of questions were listed that it was hoped would be clarified by the experiments. To conclude this report each question is briefly discussed below:

8.3.1 Is PBI++ better than PBI+, and PBI+ better than PBI?

There is some support for the hypothesis that the more information the better, as indicated for the success rate for the Assisted Mode subjects for CFU and CB, although not 3PTY. Over both groups, no statistically significant differences obtain, and given the experimental latitude, there seems to be no reason not to recommend that PBI++ is to be the preferred implementation, but nevertheless the axiom would be not to make things more complicated than they need to be.

8.3.2 What effect does the information before the control action have on usability?

The response to the question in the long questionnaire concerning the handbook was extremely positive, and confirmed by verbal responses from the subjects. This result was found consistently across all three interfaces. The conclusion is that usability is improved with this provision. The importance of providing usable instructions on the services and interfaces has been well documented by this study. The quality of the material was such that in Learning Mode, the majority of the subjects were able to use the services after memorising the commands and in the absence of any instructional material. The material received very positive marks and was considered optimum in terms of its length. Having a graphical representation of the main commands present during the actual usage situation accounted for the better performance under Assisted Mode.

8.3.3 What effect do feedback or prompts during and as a result of the control action have on usability?

The findings of the different feedback modalities are not as clear-cut as anticipated. In Assisted Mode, providing spoken and visual feedback led to a better (but not significantly better) performance in Session 2. However, in terms of *subjective* parameters, spoken messages were rated very positively (Long Questionnaire items "Tones and spoken messages are useful", "I was sure that my actions were correct", "Tones were difficult to understand", and "The system would not be usable without spoken messages"), indicating that, if spoken messages are provided, they increase the degree of subjective confidence in using the system the right way (this applied also to a smaller degree to the visual messages). The subjective confidence of being able to use the system is an important factor for the uptake of services of this kind. The conclusion is that providing feedback via spoken messages improves usability.

8.3.4 What effect does the increase in syntax complexity from "activate" and "de-activate" to including "interrogate" have on usability?

Syntax complexity was studied by means of condition PBI-. Not surprisingly, subjects under condition PBI- felt more positive about using the services (they enjoyed their experience significantly more) than those who used the more complex syntax including the interrogation commands. Introducing the Interrogate-command "*#" caused a number of confusion errors that did not occur under PBI-. Interrogate and other commands obviously have to be implemented because some subscribers will want to use them. The answer in terms of instructing the users on how to use the MMI is to teach them a set of survival or essential commands, minimally consisting of "Activate" and "De-activate" a service which can be expected to be memorised, and to present all other commands in the form of written instructions of the kind used in the study. Means should be provided to allow the users to have these instructions always at hand (e.g. by providing stickers that can be attached to the telephone set). All commands going beyond the basic activate (*Code#) and de-activate (#Code#) commands should be supported by appropriate feedback. It may be concluded that syntax complexity does have an effect on usability.

Annex A: Personal Data

A.1 Personal data tables for the German and Italian subjects

Table A.1: Personal data of the German subjects

		Interface			Total
		PBI	PBI+	PBI++	
Gender of respondent:	Female	6	6	6	18
	Male	6	6	6	18
Education Level	Medium Level	2	2	2	6
	Higher Level (university access)	7	4	3	14
	University Degree	3	5	7	15
Job Category	Non-technical	7	5	7	19
	Technical	-	3	3	6
	In training	5	4	2	11
Phone at home	No phone	1	-	-	1
	Basic phone	6	5	6	17
	Modern w/o display	3	5	1	9
	Modern with display	2	1	5	8
	Other	-	1	-	1
Phone at work	No phone	3	4	2	9
	Basic phone	2	3	6	11
	Modern w/o display	2	1	3	6
	Modern with display	3	2	-	5
	Other	-	-	1	1
Daily no. of phone calls	< 2	1	1	-	2
	2-10	7	10	11	28
	> 10	4	1	1	6
CF used before	yes	7	2	6	15
	no	5	10	6	21
CB used before	yes	2	1	1	4
	no	10	11	11	32
3P used before	yes	2	1	-	3
	no	10	11	12	33
Experience with PC	yes	11	7	7	25
	no	1	5	5	11
Experience Videotex	yes	2	1	1	4
	no	6	11	6	23
Experience with technol.	very little	-	1	3	4
	2	3	4	1	8
	3	4	2	3	9
	4	3	3	-	6
	considerable	1	2	5	8
Interest in technology	very little	-	-	2	2
	2	1	2	2	5
	3	6	3	2	11
	4	4	4	2	10
	considerable	1	3	4	8

(continued)

Table A.1 (concluded): Personal data of the German subjects

	Interface			Total
	PBI	PBI+	PBI++	
Advantages of technology outweigh disadvantages				
don't agree at all	-	1	-	1
2	1	-	2	3
3	6	4	3	13
4	3	4	5	12
agree completely	2	3	2	7
Technology makes life easier				
don't agree at all	-	1	-	1
2	-	-	1	1
3	4	1	1	6
4	4	5	6	15
agree completely	4	5	4	13
Concerned about today's technological progress				
don't agree at all	-	2	2	4
2	3	4	3	10
3	6	4	4	14
4	1	2	3	6
agree completely	2	-	-	2

Table A.2: Personal data of the Italian subjects

		Interface			Total
		PBI	PBI+	PBI++	
Gender of respondent	Female	1	2	3	6
	Male	5	4	3	12
Education Level	Medium Level	2	-	-	2
	Higher Level (university access)	2	2	4	8
	University Degree	2	4	2	8
Job Category	Non-technical	3	3	3	9
	Technical	3	3	3	9
	In training	-	-	-	-
Phone at home	No phone	-	-	-	-
	Basic phone	6	6	5	17
	Modern w/o display	-	-	1	1
	Modern with display	-	-	-	-
	Other	-	-	-	-
Phone at work	No phone	3	2	3	8
	Basic phone	-	3	1	4
	Modern w/o display	3	1	2	6
	Modern with display	-	-	-	-
	Other	-	-	-	-
Daily n. of phone calls	< 2	1	1	1	3
	2 - 10	3	4	5	12
	> 10	2	1	-	3
CF used before	yes	2	1	1	4
	no	4	5	5	14
CB used before	yes	-	-	1	1
	no	6	6	5	17
3P used before	yes	1	-	-	1
	no	5	6	6	17
Experience with PC	yes	4	5	3	12
	no	2	1	3	6
Experience Videotex	yes	-	-	-	-
	no	-	-	-	-
Experience with technol.	very little	-	-	1	1
	2	1	1	1	3
	3	3	3	1	7
	4	2	2	1	5
	considerable	-	-	2	2
Interest in technology	very little	1	-	1	2
	2	-	1	2	3
	3	1	2	3	6
	4	3	2	-	5
	considerable	1	1	-	2
Advantages of technology outweigh disadvantages	don't agree at all	3	-	-	3
	2	3	6	5	14
	3	-	-	1	1
	4	-	-	-	-
	agree completely	-	-	-	-

(continued)

Table A.2 (concluded): Personal data of the Italian subjects

	Interface			Total
	PBI	PBI+	PBI++	
Technology makes life easier	3	-	1	4
don't agree at all				
2	3	4	5	12
3	-	1	-	1
4	-	1	-	1
agree completely	-	-	-	-
Concerned about today's technological progress				
don't agree at all	4	-	4	8
2	2	4	-	6
3	-	-	1	1
4	-	2	1	3
agree completely	-	-	-	-

Annex B: General Instructions

GENERAL INSTRUCTIONS

Welcome to CSELT !

You are going to take part in an experiment on new telecommunications services. The purpose is to evaluate the effectiveness of and standardise user-to-telephone dialogues.

We are by no means going to test your abilities.

During the experiment, which will last about 2 hours, you will be asked to use a telephone set to access the following three services: Call Forwarding, Call Barring and 3-Party Call.

We shall give you two booklets: one contains the description of the 3 services, the other contains the tasks you have to perform

The experiment consists of 2 sessions, which are identical, with exception of the order in which the tasks are presented.

After completing the tasks relevant to each of the 3 services you will be asked to answer a few questions. At the end of each session you will be asked to fill-in a larger questionnaire.

Although names appear on the papers we give you, this is for our use only. All the papers will be treated in the strictest confidence and will be destroyed once the data have been processed.

Now, just before you begin working through the booklet, I would like to point out that there is no time limit: please work at your own pace. If you make an error, proceed to the next task.

Now, if you do not have questions, please begin.

Annex C: Description of the Services

DESCRIPTION OF THE SERVICES

This booklet contains a description of the following three services:

- three-party call** during a telephone conversation you can call another subscriber and talk simultaneously to both of them.
- call forwarding:** calls to your telephone can be temporarily diverted to another number
- call barring:** unauthorised people are stopped from using your telephone for outgoing calls; you can still receive incoming calls

Please read this booklet very carefully and try to memorise the commands because this booklet will not be available during the experiment.

Three-Party Call

What is Three-Party Call for?

Three-party Call allows you, during a telephone conversation, to call another subscriber and talk simultaneously to both of them.

How do I tell the telephone to set up a 3-party call?

To set up a 3-party call:

- 1) Lift the handset and dial the telephone number of user B;
- 2) Ask B to wait and press 2;
- 3) Dial the telephone number of user C;
- 4) When C answers press 3.


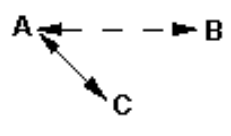


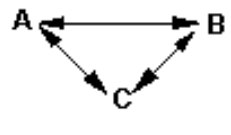

Why do I press 2 during the conversation with user B?

- 2 Pressing 2 tells the telephone that you want to put on hold the user you are talking to, in order to be able to dial user C.

Why do I press 3 during the conversation with user C (while B is on hold)?

- 3 Code 3 tells the telephone that you want user B (on hold) to join you in the conversation with C.

THREE-PARTY CALL

HOW TO USE THE SERVICE		
①	A ←————→ B	You are the subscriber A, in conversation with subscriber B
②	A ← - - - - -> B	Ask B to wait and press the button 
③		Dial-tone  Dial the telephone number of subscriber C 
④		When C answers press the button 
If C is busy or absent, to talk to B, press 1 To put C on hold and talk to B only, press 2		

Call Forwarding

What is Call Forwarding for?

If you want to be sure that you don't miss calls to your telephone while you are away from it, you can redirect your calls to a telephone with another number. This gives you a great deal of freedom of movement.

This service is also useful in other circumstances, such as when you don't want to be disturbed and have arranged for someone else to take your calls, or simply when your telephone is switched off.

How do I tell the telephone to forward my calls?

To forward calls you need to make use of the * and # buttons on your telephone as well as the number buttons you use to dial ordinary calls. They are located in the bottom left and bottom right corner of the telephone keypad, respectively.

To forward your calls to a different telephone number, say 898989:

- 1) Lift the handset;
- 2) Press *21*898989#;
- 3) Replace the handset.

Why do I press *21*898989#?

- * First you press the * button. This tells the telephone that you want to switch on one of its services. If you start typing a number without pressing the * first, the telephone thinks you are trying to dial an ordinary call.
- 21 There are several services available, and typing code 21 tells the telephone that you want Call Forwarding rather than a service with a different code number.
- * The second * tells the telephone that you have finished typing one number (21) and are about to type another (898989). Otherwise the telephone would think it was all one long number (21898989). In other words, * is used to separate two numbers, rather like a space is used to separate words printed on a page.

898989 This is the telephone directory number (DN) you want your calls be forwarded to.

Typing # tells the telephone you have finished typing the command.

Other commands




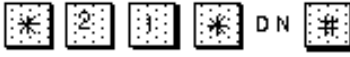





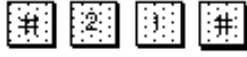






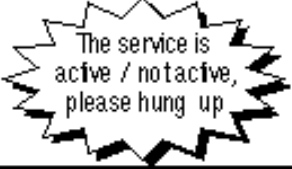

- To switch Call Forwarding off the command is

#21#.

- To ask whether Call Forwarding is switched on or off the command is

***#21#.**

CALL FORWARDING

SWITCH ON				
① 	Dial-tone 	② Press the buttons   D N is the telephone number calls are forwarded to	Audio message  The service is active, please hang up	③ 
SWITCH OFF				
① 	Special dial-tone 	② Press the buttons  	Audio message  The service is not active, please hang up	③ 
CHECK				
① 	Dial- or special dial-tone 	② Press the buttons  	Audio message  The service is active / not active, please hang up	③ 

Call Barring

What is Call Barring for?

Call barring allows you to stop outgoing calls from your telephone. In this way you can stop unauthorised people from adding to your telephone bill.

What stops other people from switching Call Barring on and off?

You identify yourself to the call barring system with a Personal Identification Number (or PIN), rather like the ones people use to obtain money from cash dispensing machines. Keep this number to yourself, unless you want other people to be able to switch call barring on and off.

In this booklet we'll assume that your PIN is 4321.

How do I tell the telephone to switch on Call Barring?

To switch on Call Barring:

- 1) Lift the handset;
- 2) Press *33*4321#;
- 3) Replace the handset.

Why do I press *33*4321#?

- * Pressing * tells the telephone that you want to switch on one of its services.
- 33 Code 33 tells the telephone that you want Call Barring rather than a service with a different code number.
- * The second * tells the telephone that you have finished typing one number (33) and are about to type another (4321). Otherwise the telephone would think it was all one long number (334321).
- 4321 This is where you press your PIN to identify yourself as a person authorised to switch on Call Barring.
- # Typing # tells the telephone you have finished typing the command.

Other commands

- To switch Call Barring off, i.e. to unbar calls, the command is

#33*PIN#.

- To unbar calls on a per call basis the command is











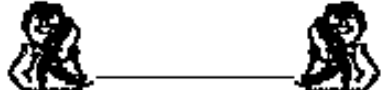


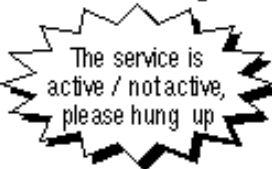

#33*PIN*DN#,

where DN is the directory number.

- To ask whether Call Barring is switched on the command is

***#33#.**

CALL BARRING

SWITCH ON				
①	Dial-tone 	② Press the buttons  * 3 3 * PIN #	Audio message 	③ 
SWITCH OFF				
①	Special dial-tone 	② Press the buttons  # 3 3 * PIN #	Audio message 	③ 
SWITCH OFF ON PER CALL BASIS				
①	Special dial-tone 	② Press the buttons  # 3 3 * PIN * DN # DN is the telephone number calls are forwarded to	③ 	
CHECK				
①	Dial- or special dial-tone 	② Press the buttons  * # 3 3 #	Audio message 	③ 

Structure Of Services Commands

* to switch on

to switch off

*# to interrogate

Please always end with: #

Annex D: Tasks and Questionnaires

PBI++

TASKS on PBI++

In this booklet you will find a number of tasks. Please perform all of them in the same order as they are presented in the booklet.

To do so you will use the keypad telephone set on the desk. Please note that this phone is equipped with a display. The information you will receive from the phone consists of network tones, auditory messages and visual messages presented on the display.

Basic call

To help you become familiar with the experimental set up we ask you to make a call to the following telephone number.

The telephone number you have to call is **33 44 55 66**.

PBI++

Session 1

Task session 1

CFU

- 1) You are leaving your home with your family for your country-house where you will be spending the weekend. You want your incoming calls to reach you there.
Therefore you use the appropriate service.
The telephone number of your country-house is **33 44 55 66**.
- 2) After having redirected your incoming calls to your country-house, you want to check whether you can still make calls from your phone.
Therefore you decide to call the number **88 77 66 55**.
- 3) After the week-end your family and you are now back home. You would like your incoming calls to reach you again at your usual telephone address, but your relatives might have done so already.
In the doubt you decide to ask the status of the telephone.
- 4) From the answer you receive from the telephone you understand that your incoming calls are still forwarded to your country-house.
Therefore you use the appropriate command to forward them again to your home.

PBI++

Session 1

SHORT QUESTIONNAIRE

This is a short questionnaire asking you to rate how you felt about the service commands and the messages, you have just used. The questionnaire presents a few statements of opinion. For each statement, please rate your degree of agreement or disagreement by ticking the appropriate box as in the example below.

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely.

Please do not use the central position, unless you really have no opinion on the statement.

We suggest you do not spend too long on each statement. First impressions are usually best.

Please consider each statement and rate your agreement or disagreement.

1) The service commands were complicated:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely.

2) The information conveyed by the tones and vocal messages was insufficient:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely.

3) The information conveyed by the visual messages was insufficient:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely.

PBI++

Session 1

CB

- 1) You are leaving for a business trip and, during your absence, you want to prevent your son from calling his friends living in another town.
Therefore you use the service that allows you to do so.
Your PIN is **4321**.
- 2) To make sure that your calls have really been barred you decide to try for yourself to make a long distance call.
The number you call is **02 33 44 55**.
- 3) You suddenly remember that you have to inform your business partner about the time of your arrival in order to allow him to pick you up at the airport.
You decide to use the unbar per call facility.
Your PIN is **4321**.
The telephone number of your partner is **02 88 77 66**.
- 4) Back home you want to re-enable your phone to the long distance calls.
Your PIN is **4321**.
- 5) Now you interrogate the telephone to check that calls have been unbarred.

followed by short questionnaire

PBI++

Session 1

3PTY

You want to organise an excursion with a couple of friends, Mary and John. The quickest way to do so is to have a phone conversation all together.

- 1) To do so you first call Mary.
- 2) Then you put Mary on hold and call John.
- 3) Before putting all three together you put John on hold and go back to Mary to check whether she is still there.
- 4) Finally, you set up a three-party conversation.

Mary's telephone number is **33 44 55 66**.

Paul's telephone number is **88 77 66 55**.

followed by short questionnaire

PBI++

Session 2

Task session 2

CFU

- 1) It is Friday and you have to write an urgent report before leaving for the week-end. Not to be disturbed, you ask your secretary to answer your incoming calls and therefore you forward them to her telephone number.
The telephone number of your secretary is **33 44 55 66**.
- 2) After having redirected your incoming calls to your secretary, you want to check whether you can still make calls from your phone.
Therefore you decide to call the number **88 77 66 55**.
- 3) After the week-end you cannot remember if your calls are still diverted. Therefore you decide to ask the status of the telephone.
- 4) Since your incoming calls are still forwarded to your secretary, you use the appropriate command, so that they can reach you again at your office.

PBI++

Session 2

SHORT QUESTIONNAIRE

Please consider each statement and rate your agreement or disagreement.

- 1 The service commands were easy to use
 - agree completely
 - agree mostly
 - undecided
 - disagree mostly
 - disagree completely.
- 2 The information conveyed by the tones and vocal messages was sufficient
 - agree completely
 - agree mostly
 - undecided
 - disagree mostly
 - disagree completely.
- 3 The information conveyed by the visual messages was sufficient
 - agree completely
 - agree mostly
 - undecided
 - disagree mostly
 - disagree completely.

PBI++

Session 2

CB

- 1) You have asked your neighbours to water your flowers during your summer holidays, but you suspect they could profit to use your phone for their long distance calls.
Therefore, before leaving you decide to disable your phone as for the long distance calls.
Your PIN is **4321**.
- 2) Just to make sure that your calls have really been barred you try to make a long distance call.
The number you call is **02 33 44 55**.
- 3) While waiting for the taxi for the airport you realise you forgot that your mother's birthday is next week. You pick up the phone and call her using the unbar per call facility.
Your PIN is **4321**.
Your mother's telephone number is **02 88 77 66**.
- 4) You are back home after your holidays and want to restore the possibility of making long distance calls from your phone.
Your PIN is **4321**.
- 5) Now you check that your phone allows you to make again long distance calls.

followed by short questionnaire

PBI++

Session 2

3PTY

You want to discuss some problems with Jane and Peter, two colleagues working in other branches of your firm. The best way to do so is to have a phone conversation all together.

- 1) Therefore you start calling Jane;
- 2) Then you put Jane on hold and call Peter,
- 3) While briefly talking to Peter and before setting up a three-party conversation, you decide to go back to Jane and check whether she is still on hold;
- 4) Finally, you put the three of you together.

Jane's telephone number is **33 44 55 66**.

Peters telephone number is **88 77 66 55**.

followed by short questionnaire

PBI++

LONG QUESTIONNAIRE

You have now completed all the tasks.

Before you leave, we would like still to ask you a few questions.

Please consider each statement and rate your agreement or disagreement.

We suggest you do not spend too long on each statement.

Call Forwarding

1a) The commands for the Call Forwarding service were easy to remember:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

2a) I had to concentrate hard when using the Call Forwarding service:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

3a) The section of the handbook describing the commands for the Call Forwarding service was useful:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

4a) The section of the handbook describing the commands for the Call Forwarding service was too long:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

5a) When using the Call Forwarding service, the tones did not suggest what to do next:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

6a) When I used the Call Forwarding service, I always felt confident I did the right actions:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

7a) The experience with the Call Forwarding service was unpleasant:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

8a) I found it enjoyable using the Call Forwarding service:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

9a) The feedback tones of the Call Forwarding service are useful:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

10a) The procedures for using the Call Forwarding service are confusing:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

11a) I found it difficult to understand the meaning of the tones of the Call Forwarding service:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

12a) I hardly paid attention to the vocal messages when I used the Call Forwarding service:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

13a) I would not have been able to use the Call Forwarding service if I had not listened to the vocal messages:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

14a) I hardly paid attention to the written messages when I used the Call Forwarding service:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

15a) I would not have been able to use the Call Forwarding service if I had not been able to read the written messages:

- agree completely
- agree mostly
- undecided
- disagree mostly
- disagree completely

followed by the same 15 questions on Call Barring and Three Party Conferencing

followed by

Any other comments ?

.....
.....
.....

Annex E: Error Analysis

E.1 Types of errors vs. Tasks for PBI, Sessions 1 and 2, "Assisted Mode"

Table E1: PBI Session 1

TASK	Prefix	Code/ Switch order	Separ.	Suppl. inform.	Suffix	Wrong task	Skip task	Re-start	No hang-up	No or inter. action
CFa										
CFb						3				
CFi										
CFd						1				
CBa	1	2						1		
CBb						1				1
CBs	1	1		1		1				
CBd						1				
CBi						1				
3Pb										
3P2										
3Pa							3			
3Pe										

Table E2: PBI Session 2

TASK	Prefix	Code/ Switch order	Separ.	Suppl. inform.	Suffix	Wrong task	Skip task	Re-start	No hang-up	No or inter. action
CFa										
CFb						2				
CFi							1			
CFd										
CBa						1		1		
CBb										
CBs			1							
CBd						1				
CBi						1				
3Pb										
3P2		1								
3Pa							2			
3Pe							1			

Annex F: Responses to Long Questionnaire

Table F.1: Total sample (n = 54) (5=agree completely; 1=disagree completely)

Question	PBI			PBI+			PBI++		
	CF	CB	3P	CF	CB	3P	CF	CB	3P
1 Command easy to remember	4,2	3,4	4,6	3,8	3,3	4,5	3,8	3,0	4,1
2 Had to concentrate	2,6	3,4	2,2	3,0	3,3	1,7	3,2	3,8	2,8
3 Handbook section useful	4,8	4,9	4,8	4,7	4,8	4,7	4,4	4,6	4,2
4 Handbook section too long	2,1	2,0	1,7	1,6	1,7	1,3	1,7	1,8	1,7
5 Tones / spok. mess. didn't indicate what to do	3,2	2,6	2,4	1,9	1,7	1,3	2,4	2,1	2,6
6 Sure actions correct	3,7	3,4	4,3	3,2	3,1	4,3	3,1	2,8	4,0
7 Experience unpleasant	1,7	1,7	1,3	1,9	2,4	1,3	1,9	2,0	2,0
8 Fun to use	4,8	4,6	4,8	3,8	3,6	4,7	4,1	4,0	4,1
9 Tones / spoken mess. useful	3,6	3,6	3,6	4,3	4,2	4,3	3,6	3,8	3,4
10 Procedure confusing	1,6	2,2	1,4	2,1	2,2	1,6	1,7	2,6	2,3
11 Tones / spok. mess. difficult to understand	2,7	2,7	2,0	1,7	1,4	1,4	2,1	2,3	2,4
12 Hardly paid attention to spoken messages	-	-	-	1,6	1,6	1,5	2,3	2,2	3,1
13 Not usable without spoken messages	-	-	-	2,7	3,2	2,7	2,3	2,6	2,2
14 Hardly paid attention to text	-	-	-	-	-	-	3,7	3,6	3,6
15 Not usable without text	-	-	-	-	-	-	2,0	2,1	1,8

Table F.2: Sample "Assisted Mode" (n = 18)

Question	PBI			PBI+			PBI++		
	CF	CB	3P	CF	CB	3P	CF	CB	3P
1 Command easy to remember	3,5	3,0	4,0	3,5	3,5	3,8	3,7	3,0	4,0
2 Had to concentrate	2,8	3,8	3,2	2,7	3,0	1,8	3,5	4,0	3,8
3 Handbook section useful	4,8	5,0	4,8	4,8	4,7	3,8	3,8	4,3	4,3
4 Handbook section too long	2,3	2,0	1,7	1,3	1,3	1,2	1,5	1,3	1,5
5 Tones / spok. mess. didn't indicate what to do	3,8	2,8	3,2	1,8	1,7	1,3	3,5	2,7	2,7
6 Sure actions correct	3,2	2,8	3,8	3,3	4,0	4,3	3,0	2,7	3,3
7 Experience unpleasant	1,7	2,2	1,7	1,3	2,0	1,3	1,7	2,0	2,3
8 Fun to use	4,7	4,5	4,5	4,0	4,0	4,7	4,5	4,3	4,5
9 Tones / spoken mess. useful	3,0	2,7	3,5	3,2	3,5	3,3	3,3	3,3	3,8
10 Procedure confusing	2,2	2,5	2,0	1,8	1,5	1,8	1,8	2,5	3,2
11 Tones / spok. mess. difficult to understand	3,5	3,8	2,3	1,8	1,5	2,0	2,3	3,2	3,2
12 Hardly paid attention to spoken messages	-	-	-	1,2	1,2	1,2	2,3	1,8	2,7
13 Not usable without spoken messages	-	-	-	4,0	4,0	3,5	3,0	3,0	2,8
14 Hardly paid attention to text	-	-	-	-	-	-	2,8	2,7	3,3
15 Not usable without text	-	-	-	-	-	-	2,8	3,2	3,2

Table F.3: Sample "Learning Mode" (n = 36)

Question	PBI			PBI+			PBI++		
	CF	CB	3P	CF	CB	3P	CF	CB	3P
1 Command easy to remember	4,6	3,7	4,8	3,9	3,2	4,8	3,8	3,0	4,2
2 Had to concentrate	2,4	3,3	1,7	3,2	3,5	1,7	3,0	3,8	2,3
3 Handbook section useful	4,8	4,9	4,8	4,7	4,8	4,6	4,8	4,7	4,1
4 Handbook section too long	2,0	2,0	1,8	1,8	1,9	1,4	1,8	2,1	1,8
5 Tones / spok, mess. didn't indicate what to do	2,9	2,4	2,1	1,9	1,8	1,3	1,8	1,8	2,6
6 Sure actions correct	4,0	3,7	4,6	3,1	2,7	4,3	3,1	2,8	4,3
7 Experience unpleasant	1,8	1,4	1,1	2,2	2,6	1,3	2,0	2,0	1,8
8 Fun to use	4,8	4,7	4,9	3,8	3,4	4,8	3,8	3,8	4,0
9 Tones / spoken mess, useful	3,9	4,0	3,7	4,9	4,6	4,8	3,8	4,0	3,7
10 Procedure confusing	1,3	2,0	1,1	2,2	2,5	1,4	1,7	2,7	1,9
11 Tones / spok, mess, difficult to understand	2,3	2,2	1,8	1,6	1,4	1,2	2,0	1,8	2,1
12 Hardly paid attention to spoken messages	-	-	-	1,8	1,8	1,7	2,3	2,3	3,3
13 Not usable without spoken messages	-	-	-	2,1	2,8	2,3	2,0	2,3	1,8
14 Hardly paid attention to text	-	-	-	-	-	-	4,1	4,0	3,8
15 Not usable without text	-	-	-	-	-	-	1,6	1,5	1,2

History

Document history	
October 1996	First Edition