ETSI

ETSI TECHNICAL REPORT

ETR 261-1

October 1996

Source: ETSI TC-HF

Reference: DTR/HF-01028-1

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 1: General approach and summary of findings

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

Contents

Blank page

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 1 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".

The intended users of this ETR are shown in table 1.

The work reported in this ETR was conducted in response to the rapid growth of telecommunications networks and services, and the increasing mobility of the human user, both within and between networks.

One of the missions for ETSI TC-HF is the recommendation of user control procedures for gaining access to and control of telecommunication services. This multi-part ETR describes the results of a programme of work conducted within ETSI STC HF1 concentrated on the user control procedures for supplementary services. The objective was to assess the user and technology requirements and to define a minimum man-machine interface (MMI) that can be applied consistently across networks. The definition of the MMI is presented separately in ETS 300 738 [3].

The key stimulus for the work came when ETSI TC-SMG chose to define the minimum MMI for controlling GSM supplementary services [2] with significant differences to the existing European (CEPT) [1] or International (ITU-T) [11] recommendations. Now, it seemed, that users would have to learn two very similar command languages to operate two very similar sets of supplementary services; and to distinguish between these depending upon which terminal / network system they were using. What is more, when the project surveyed European network operators, 17 out of 18 PSTN operators reported that they had implemented a third code scheme, ETR 261-6 [8], a cut down version of the CEPT recommendations [1]. For example:

- To switch on and register a number for Call Forward Unconditional:

  - CEPT requires *21*NNNNNN#
  - GSM requires **21*NNNNNN#
  - PSTN practice *21*NNNNNN#

- To switch off and erase a stored number from Call Forward Unconditional:

  - CEPT requires #21*1#
  - GSM requires ##21#
  - PSTN practice #21#

Quite obviously there was a case for harmonization within Europe.

Table 1: Intended users and potential benefits

	User	ETR used for	Potential Benefit
1	Service developers and providers	Access to background research complementary to an ETS, to help the development of usable supplementary and other services	Increased usability through harmonized and supportive services
2	User groups	Access and awareness of the background to the complementary ETS	Increased awareness by user groups of user requirements
3	Researchers	Access to the background research and data to the complementary ETS	Availability of data for other research groups
4	ETSI Technical committees	Access and awareness of the background to the complementary ETS	Increased usability through harmonized and supportive services
5	TC-HF	Access and awareness of the background to the complementary ETS	Improved usability through consistency and coherence of recommendations

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 1 of the ETR describes the approach taken to the work and presents a summary of the results obtained from the experimental, questionnaire and survey data collected.

2 References

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

[1] CEPT T/CAC 02: "Subscriber control procedures for supplementary services in modern telecommunication system".

[2] ETS 300 511: "European digital cellular telecommunications system (Phase 2); Man Machine interface (MMI) of the mobile station (MS) (GSM 02.30)".

[3] ETS 300 738: "Human Factors (HF); Minimum Man Machine Interface (MMI) to public network based supplementary services".

[4] ETR 261-2: "Human Factors (HF); Assessment and definition of a harmonized minimum man-machine interface (MMI) for accessing and controlling public network based supplementary services; Part 2: Literature review - Memory and related issues for dialling supplementary services using number codes ".

[5] 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".

[6] 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".

[7] 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".

[8] 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".

[9] ETR 261-7: "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".

[10] Israelski E (1988): "An experimental comparison of user performance with alternative access codes for PBX features." 12th Symposium on Human Factors in Telecommunications, The Hague.

[11] ITU-T Recommendation E.131: "Subscriber control procedures for supplementary telephone services".

[12] 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".

[13] Jones MLR (1990): "Making numeric command languages more usable." 13th Symposium on Human Factors in Telecommunications, Turin, pp 99-106.

[14] Lindgaard G (1993): "Wow - 568 smart features on your PABX: What really determines the uptake of technology?" 14th Symposium on Human Factors in Telecommunications, Darmstadt.

[15] Schwartz B (1993): "Advanced screen telephony - making today's services easier to use." 14th Symposium on Human Factors in Telecommunications, Darmstadt.

[16] Zeidler G 1970): "How to co-ordinate operational procedures for new telephone facilities." 5th Symposium on Human Factors in Telecommunications, London.

3 Definitions, symbols and abbreviations

NOTE: The definitions, symbols and abbreviations given below are also applicable to the other parts of the present ETR.

3.1 Definitions

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

analysis of variance: Statistical test to verify significant differences between two or more variables.

centrex: A Centralized PBX service provided by a local exchange for private (business) customers.

digit span: The length of a sequence of random digits which a person can reliably remember.

Likert scale: A common 5-point rating scale enabling the rating of agreement with a given statement.

Neuman-Keuls' test: Statistical test to verify significant differences.

PBI: The basic user interface or MMI provided between a telecommunications terminal and a supplementary service using only the twelve key keypad (0-9, plus * and #), and providing only auditory tones as feedback.

PBI+: The first level of improved MMI, which in addition to the use of the standard 12 keypad and auditory tones also provides spoken voice announcements as feedback.

PBI++: The second level of improved MMI, which in addition to the use of the standard 12 key keypad and auditory tones and voice announcements, also provides a visual display for text messages.

PBI-: A reduced version of PBI, which only provides for activate and deactivate; there is no interrogate command.

Pearsons *r*: Statistical test to verify correlations between two variables.

Tukey's test: Statistical test to verify significant differences.

3.2 Symbols and abbreviations

For the purposes of this ETR, the following symbols and abbreviations apply:

***** The Star on the standard telephone keypad arrays, see ITU-T Recommendation E.161 [12]. Also known as *asterisk*.

#	The Square on the standard telephone keypad arrays, see ITU-T Recommendation E.161 [12]. Also known as *hash*, *sharp*, or *number sign*.
3P1	Three Party Conference - Call 1
3P2	Three Party Conference - Call 2
3PA	Three Party Conference - Alternate (between Call 1 and 2)
3PE	Three Party Conference - Establish (join Call 1 and 2)
3PTY	Three Party Conference Service
I	> 95 percentile (Raven's score)
II+	90-95 percentile (Raven's score)
II	75-90 percentile (Raven's score)
III+	50-75 percentile (Raven's score)
III-	25-50 percentile (Raven's score)
IV	10-25 percentile (Raven's score)
IV-	5-10 percentile (Raven's score)
V	< 5 percentile (Raven's score)
--	Non-significant
✓	Significant at p = 0,05 - 0,01
✓✓	Significant at p > 0,01
A	General Certificate of Education, Advanced Level ("A-Level") (UK school leaving examination normally sat at age 18)
A5	International standard paper size
ANOVA	Analysis of Variance
AOC-D	Advice of Charges - During Call (Supplementary Service)
AOC-E	Advice Of Charge - End of call
AOC-S	Advice Of Charge - Set up stage of call
A-Party	the calling party (convention)
B-Party	the called party (convention)
C-Party	a second called party (convention)
ASR	Automatic Speech Recognition
ASS	"Set-up" function key
BBC	British Broadcasting Corporation
BT	British Telecom PLC
C	C computer programming language
CAT	Categories (form of manual)
chi-square	statistical computation (as in Kruskal-Wallis non-parametric analysis of variance)
CB	Call Barring Service
CBA	Call Barring - Activate
CBB	Call Barring - Basic Call
CBC	Call Barring - Conditional De-activate (on a per call basis)
CBD	Call Barring - De-activate
CBI	Call Barring - Interrogate
CB-I	Call Barring - Incoming (Supplementary Service)
CB-O	Call Barring - Outgoing (Supplementary Service)
CBACK NR	Call Back Number (fictitious service similar to Call Completion on Busy)
CEPT	Conférence European des Administrations de Postes et de Télécommunications
CCBS	Call Completion to Busy Subscriber (Supplementary Service)
CD	Call Deflection (Supplementary Service)
CF	Call Forwarding
CFA	Call Forwarding - Activate
CFB	Call Forwarding on Busy / Call Forwarding - Basic Call (part 7 only)
CFD	Call Forwarding - Deactivate
CFI	Call Forwarding - Interrogate
CFNR	Call Forwarding on No Reply
CFU	Call Forwarding Unconditional Service
CLIP	Calling Line Identification Presentation (Supplementary Service)
CLIR	Calling Line Identification Restriction (Supplementary Service)
COLP	Connected Line Identification Presentation (Supplementary Service)
COLR	Connected Line Identification Restriction (Supplementary Service)
CONF	Conference (as in Conference; Add on or 3-Party) (Supplementary Service)
CRT	Cathode Ray Tube
CSE	Certificate of Secondary Education (former UK school examination, lower academic level than O-Level, normally sat at age 16, superseded by GCSE)

CSELT	Centro Studi e Laboratori Telecomunicazioni, Turin, Italy
CT-E	Call Transfer - Explicit (Supplementary Service)
CT-S	Call Transfer - Simple (Supplementary Service)
CUG	Closed User Group (Supplementary Service)
CW	Call Waiting (Supplementary Service)
D	Degree (British education level)
D1 (2, 3, or 4)	Duration 1 (2, 3 or 4) see figure 16
DDI	Direct Dialling In (Supplementary Service)
df	degrees of freedom (statistical term)
DTMF	Dual Tone Multi Frequency
E	Errors (Summed for the 1st, 2nd, 3rd & 4th occurrence of the target services)
ECT	Explicit Call Transfer
F	Factor (statistical term, e.g. in an analysis of variance, or F-Test) or statistical value computed in ANOVA tests
FPH	Freephone (Supplementary Service)
G (or g)	Group (as in experimental group)
GCSE	General Certificate of Secondary Education (UK school examination normally sat at age 16, successor to CSE and O-Level)
GSM	Global System for Mobile communications
HCI	Human Computer Interaction
HHI	Heinrich Hertz Institute, Berlin, Germany
HOLD	Hold (Supplementary Service)
IDN	Integrated Digital Network
INT	Interface (as in PBI, PBI+, or PBI++)
ISDN	Integrated Services Digital Network
ITU-T	International Telecommunications Union - Telecommunication Standardization Sector
LCD	Liquid Crystal Display
m	minute(s)
max.	Maximum
MCID	Malicious Caller Identity (Supplementary Service)
MMC	Meet Me Conference (Supplementary Service)
MMI	Man-Machine Interface
MSN	Multiple Subscriber Number (Supplementary Service)
n	number (of subjects in a particular condition)
N	None recorded (British education level)
n.a.	not applicable
NA7	Subcommittee of ETSI TC-NA, responsible for UPT standardization
NL-PTT	Netherlands - PTT
NMT	Nordic Mobile Telephone System
NNN	Number (usually to show a string of digits)
NoCAT	No Categories (form of manual)
O	General Certificate of Education, Ordinary Level ("O-level") (former UK school examination, higher academic level than CSE, normally sat at age 16, superseded by GCSE)
OCB	Outgoing Call Barring
p	probability
PBI	Phone-Based Interface
PABX / PBX	Private (Automatic) Branch Exchange
PC	Personal Computer
PIN	Personal Identity (or Identification) Number
PN	Personal Number (also known as Personal User Identity - PUI)
Q	Question (as in Q1 in questionnaire)
r	Pearson's correlation coefficient (statistical term)
R	Register Recall, see E.161 [12].
®	The sign of a registered trade mark or name.
rho	Spearmans rank correlation coefficient (statistical term)
RIC	"Service access" function key
RIS	"Redial" function key
RS232	RS232 computer input / output port
S	Sessions (as in experimental session 1 or 2)
SC	Service Code
sd	Standard Deviation (statistics)

SERV	Service (as in CFU, CB, or 3PTY)
SI	Supplementary Information
SO	Switching Order
SPM	Standard Progressive Matrices (Raven's test of abstract reasoning)
SPSS	A PC based statistical package
STC	Sub Technical Committee
SUB	Subaddressing (Supplementary Service)
STM	Short Term Memory
TA	Technical Assembly
TACS / ETACS	(Extended) Total Access Communication Systems (Analogue Mobile Service)
TN	Telephone Number
TP	Terminal Portability (Supplementary Service)
TV	Television
UPT	Universal Personal Telecommunications
UUS	User to User Signalling (Supplementary Service)

4 Approach

The work required for this programme was split into two phases. Phase 1 was to enable the team to identify and examine the user and technology requirements. Phase 2 was for the development and testing of the proposed new standard.

4.1 Phase 1

One of the recurring themes within the human factors research in telecommunications revolves around the simple question: "How can we get more people to use more network services?" From Lindgaard [14] and Schwartz [15] at Darmstadt in 1993, via Jones [13] at Turin in 1990, Israelski [10] at The Hague in 1988, and many others to Zeidler [16] at London in 1970 the question has taken many forms but the underlying theme is the same - making services more accessible.

From these and many other researchers the answers seem to touch on four interacting elements:

1) the user's need for and awareness of the service;
2) the simplicity and structure of the information provided to the users;
3) the simplicity and structure of the controlling command language;
4) the type and content of the feedback and prompts.

The problem then was how to address these elements within a minimum man-machine interface standard which would offer a route to harmonization of the conflicting recommendations and enable future solutions like the developing idea of an Analogue Display Services Interface (ADSI) - see Schwartz 1993 [15].

However, before attempting to answer these questions in detail it, was obvious that more information was needed. For example, it was known that at least one European service provider had not implemented the full CEPT recommendations within its PSTN. Therefore, it was appropriate to find out just who had implemented what, and what alternatives if any were being considered. Hence a survey of European PSTN, ISDN, TACS/ETACS, and GSM Network Operators and other Service Providers was set up to ask for detailed information on their use of *# code schemes to access and control supplementary services. Also the survey attempted to get an idea on the type and content of the feedback provided to their users. At the same time it was decided to survey a sample of users within centrex and PBX environments to ensure that users' practice and experience in accessing these services had not radically changed since previously reported surveys, and to get a feel for the users' perceived need for and awareness of the services.

In addition to the survey work, it was decided that more detailed information was necessary on the past and present memory research. Consequently a broad ranging literature review was commissioned to look at previous memory and related areas of study to examine the directions open for possible solutions and what areas were likely to cause users most problems.

Finally, within this preliminary examination of the user and technology requirements, it was decided to look in more detail at user behaviour within two different technologies using the same broad type of interface but with differing levels of user feedback. The opportunity arose to test a simulated UPT access and to compare this with access to supplementary services within a prototype ISDN service. This opportunity

also facilitated the testing of experimental methodologies for examining user behaviour and experience with these type of services.

On completion of some of these preliminary studies it was decided to explore three of the four elements defined above (User Needs / Awareness, User Instructions, Command Languages and Feedback) in more detail. Three experimental situations were established to explore the effects on user performance of the selected three elements.

1 Users' needs and awareness of the services were partially addressed in the user survey, and no further work was undertaken on this topic as it was difficult to conceive how standardization could effect this element.

2 The effect of user instruction material was explored in a scenario based learning experiment (part 4 of the present ETR [6]). The experiment was designed to look at the effect of categorizing vs. non-categorizing the service codes and command syntax on user performance, and on user retention or recall. The experiment also allowed a comparison of the learning in younger vs. older users.

3 The effect of the different command languages CEPT vs. GSM was explored in a paper-based learning experiment (part 5 of the present ETR [7]). The experiment was set up to consider whether one code scheme is better than the other, whether there were elements of either code scheme that caused specific problems, and whether learning the two slightly different code schemes effected performance.

4 The effect of differential types of feedback tones, announcements and visual displays was explored in the final experiment which was undertaken in phase 2. The experiment also evaluated the draft ETS.

4.2 Phase 2

The principal objectives of the second phase of the work were to draft a new standard to define the minimum man-machine interface for public network-based supplementary services, and to evaluate this draft experimentally to demonstrate the potential level of usability.

Following the results of phase 1, the draft ETS [3] was to try to establish, for the first time, three component elements to a minimum man-machine interface:

1) the information given to a user before he can use the service;
2) the controlling command language; and
3) the feedback provided in response to a control action.

The ETS should also offer two approaches to service providers for enabling control of the service, either through the traditional command language style of *# code scheme or through an interactive dialogue. If the command dialogue approach is chosen, the final choice of command format should broadly be independent of the type of service being offered. However, it was recognized that some services were radically different to others and required a different command language set, e.g. the switching orders required to control user options when a Call Waiting signal is presented. If the interactive dialogue approach is selected, there should be no specific dialogue or command set defined, but compliance could be met by demonstrating an equivalent usability level.

The ETS should also set out to define a minimum standard for the information content provided to users both within the user instructions and within the feedback provided. The research demonstrated that there were a number of potential errors that could be avoided if users had proper feedback. For example, if users are told the number they have just diverted their calls to, they will be able to check whether it is the same as the number they intended to divert them to.

The ETS should conclude with a listing of the current recommendations for service codes as defined by CEPT [1].

The various drafts of the ETS were also to be subjected to an extensive consultation process. Liaison was established with a number of other international standards bodies with interests in this area, including: ITU-T SG1 SWG-HF, ISO/IEC JCT1 SC18 WG9, and ECMA TC32, as well as with ETSI TCs NA, BTC, SPS, TE, and SMG.

Finally within this phase, the experimental evaluation (part 7 of the present ETR [9]) was set up to explore two points:

- the effect of different forms of feedback (tones, voice announcements, and text display) on user performance, and
- the potential level of usability.

The results confirmed that a high level of usability can be achieved by compliance with the draft ETS. The usability target set was for naïve subjects to achieve a 75 % success rate on their first attempt and 90 % success on the second. The results also suggested that, at least for a command based dialogue format, feedback based on a combination of tones and voice announcements was preferred and offered the most consistent improvement between successive attempts.

5 Summary Results

5.1 Introduction

The summary results are subdivided according to the work being reported. Subclause 5.2 presents the outline findings from the review of the existing literature, which may be relevant. In particular this was concentrated on memory research. Subclause 5.3 presents the results of the surveys and experiments, as follows:

5.3.1: Summarizes the survey of existing networks and user practice, reported in part 6 of the present ETR [8];

5.3.2: Summarizes the experimental comparison of two user interfaces, a simulated UPT interface and a prototype ISDN supplementary service interface, reported in part 3 of the present ETR [5];

5.3.3: Summarizes the results of the experiment comparing categorized and non-categorized user instruction material, reported in part 4 of the present ETR [6];

5.3.4: Summarizes the results of the experiment comparing the CEPT and GSM code schemes, reported in part 5 of the present ETR [7];

5.3.5: Summarizes the results of the experiment comparing feedback and evaluating the usability of the proposed draft standard, reported in part 7 of the present ETR [9].

5.2 Memory Research

Before getting to deeply involved in the intricacies of * vs. ** etc., it was decided to take a step back and to review the existing literature with respect to the learning, remembering and using of such abstract command strings.

5.2.1 Type of material and practice

Perhaps the first major conclusion that can be drawn from this review is that the material (abstract symbol / number combinations) and the context (infrequent, irregular usage) really does not help the user's learning task. Many studies have demonstrated that the simple replication of all sorts of material without some form of mental processing will invariable result in a very poor recall performance. A pragmatic example might be made by asking how many people can accurately recall their credit card number. Yet no doubt they read and perhaps write it sometimes several times a month. Typically what is required for the material to become committed to long term memory is some form of semantic processing.

Perhaps equally as many studies have shown the power of the distributed practice effect. From mice learning mazes to humans learning to type and children learning to read or spell, the effect is the same. Probably every music pupil will have been told at some time that 15 minutes a day is much better than 90 minutes at the weekend. Given that the typical supplementary service may initially be used once or twice a week and then perhaps only once a month or less, it is not surprising that there is poor recall.

5.2.2 Memory aids

The second major conclusion from the review is that there is evidence that people's success rate at learning abstract material should be open to improvement if cognizance is taken of some of the things they can find helpful.

5.2.2.1 Learning the services

Users first need to learn what the services are and how they can be useful within their business and/or domestic life. Users need to be able personally to relate the service's function to their own real life scenarios. They have to perceive that they own the problem the service is attempting to address. Once they have done that, they may find they need to learn several interacting sub-services. For example the Call Waiting service may expect users to take on-board Call Waiting, Hold, Swap (between callers), Release (one of the callers), 3-Party Conference, and Transfer. If this is the case, then it can help if there is some systematic way to structure the group of services, and to explain this structure to the users.

The names assigned to the different services should be carefully designed to act as mediators between the service's functions and the various commands needed. The names should be descriptive of the service and indicate its goals. "Call Waiting", for example, probably meets this criterion, but "Hold" and "Release" may be perceived as telephony terms which have meaning only once they are learnt. Ideally, of course, once they have been learnt, the names should not change as users move between networks. However, the reality is that competing networks often chose different names to emphasise their individuality. In practice it maybe even worse, as different network standardization groups can subtly differentiate their services but still use a common name (e.g. "Conference, Add-on" in Euro-ISDN and in GSM).

5.2.2.2 Learning the codes

Assigning code numbers to services should take account of the grouping or structure of services so that the structure is reflected in the assigned code number in a systematic way. It does not help when the three common Call Forward services have the codes 21, 67, and 61 and general cancel for all Call Forward services is 002. To be fair to CEPT, who assigned these codes in Europe, significant effort was taken to try to keep the services and their respective numbers grouped, but the rapid growth of the number of services in both the private and public sectors meant it proved impossible.

The syntax of the command and the ordering of the arguments should be generated according to a set of rules that can be consistently applied. These rules should also be carefully explained to the users so that they can comprehend why the commands are constructed as they are. In that way the users have a chance to semantically process the command elements, and by understanding be able to extrapolate the commands from one service to another.

5.3 The surveys and experiments

In parallel with the literature review, discussions in TC-HF concluded that some of the limitations on the usability of public network services could be attributed to poor user guidance and feedback. For example, one network simply states its services and lists the relevant codes on some introductory pages within their directory, and many networks provide no voice messages as prompts or feedback. The team listed four key questions:

1) What has been implemented, e.g. within the PSTN, and how usable is this implementation?

2) What effect is there on user performance if the information presented to the user is better structured?

3) What differences are there in user performance between the two recommended code schemes (CEPT and GSM), and what transfer effects are there, if more than one scheme is learnt?

4) What effect is there on user performance from different levels of feedback, e.g. tones only; tones and voice messages; tones, voice messages and displayed text?

5.3.1 Current implementations

A questionnaire was sent to all the European network operators (fixed networks). Replies from 18 PSTN operators, representing probably 80 % to 90 % of European subscribers, were analysed. The key questions related to the detailed format of the command dialogue (stimulus protocol) used, and the type of feedback provided.

5.3.1.1 Command dialogue format

Table 2 shows the collated results from the 18 operators. In effect 17 out of the 18 had implemented cut down versions of the CEPT recommendations [1]. Only one operator reported a facility to store data without switching on, and even its erase command differed from CEPT's proposal.

Table 2: Command dialogue implemented by European PSTN operators

Function	Command	% Using
Switch on (basic)	*SC#	100 %
Switch on with data	*SC*SI#	100 %
Switch off (basic)	#SC#	100 %
Switch off erases data stored	-	67 %
Check status/data	*#SC#	83 %
Store data only	*SC*SI*0#	6 % (1 operator)
Switch on using stored data	*SC#	6 % (1 operator)
Switch off and Erase data	#SC*0# (CEPT proposed #SC*1#)	6 % (1 operator)
Where: SC = Service Code, e.g. 21 for Call Forward Unconditional; and SI = Supplementary Information, e.g. a subscriber number		

5.3.1.2 Type of feedback

The feedback offered by most operators was information tones with no supporting voice announcements. Figure 1 shows the type of feedback offered (tones vs. voice) within European PSTNs for a range of code scheme error conditions. At best, of the 16 operators completing this question, only about 50 % offered any form of voice announcement. Perhaps of even greater concern was that there was no consistent meaning that could be assigned to the tones used.

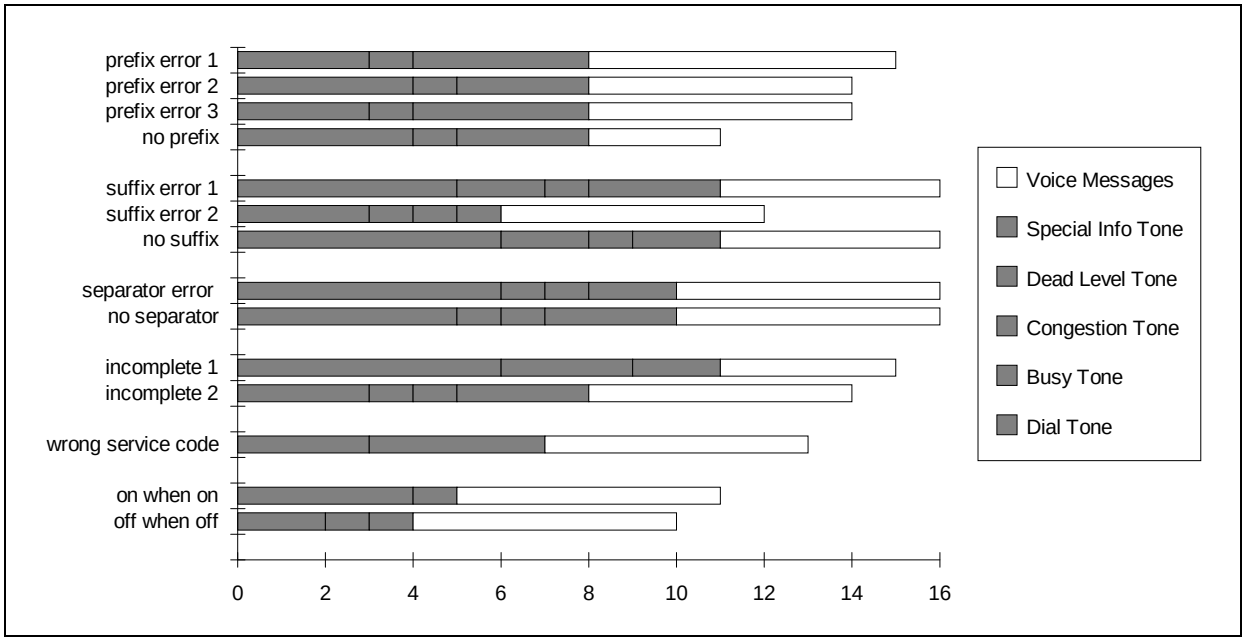

Figure 1: Responses to dialling errors in service control - PSTN Fixed Networks

5.3.1.3 How usable?

The question of how usable the current implementations are can only be addressed by example. A parallel survey compared users from two different environments. User group 1 were all research scientists and engineers working at a telecommunications research centre: the services were provided by Centrex. User group 2 were also research scientists, this time within a non-technical university department, where the services were provided by PABX. The target services (Call Forward, Call Barring, Call Waiting and Conference) were accessed and controlled in a similar way in both groups. Table 3 shows the reported frequency of use.

Table 3: Reported frequency of use of 4 services between two subject groups

	Group 1: Telephone aware (14 subjects)			**Group 2:** Less aware (18 subjects)			
	Daily	Weekly	Monthly	Daily	Weekly	Monthly	total
Call Forward	28 %	57 %	7 %	6 %	6 %	17 %	56 %
Call Barring						6 %	3 %
Call Waiting	14 %		7 %				9 %
Conference			28 %				12 %

Group 1 were also asked to expound on specific difficulties in relation to each service. Table 4 shows their reported level of difficulty for each service on the scale 0 (very difficult) to 10 (very easy). Note that not all questions were asked for each service.

Table 4: Areas of reported difficulties with Centrex services

Remembering...	**CFU**	**CB**	**CW**	**CONF**
...the service is available	8,5	**2,4**	5,9	6,7-
...the service code	**5,4**	**3,4**	-	**4,0**
...to switch on	**5,4**	**3,0**	**4,9**	**3,3**
...to switch off	**4,5**	**2,8**	-	**-**
...how to switch on	6,2	-	-	**3,0**
...how to switch off	-	-	-	7,3
...other numbers needed (subscriber, PIN, etc.)	7,7	-	-	-

The problems areas - at least for these users - were remembering the code and remembering to switch the service on and off. These current users seem to be little different to many other groups that have been reported previously.

5.3.2 Comparing UPT vs. ISDN interfaces

The surveys of European network operators and selected users of supplementary services within a PABX and a Centrex environment gave some insights into the usability of these services. In parallel, two experiments were conducted in CSELT, Italy, to give more detailed information within two specific interfaces to two different services: a UPT phase 1 simulation, and a prototype ISDN supplementary service.

5.3.2.1 The simulated UPT phase 1 interface

The simulation was created by connecting a basic telephone subset to a Sun© workstation. The simulation software responded to user actions on the telephone (handset on- / off-hook plus the 12 keys 0-9, * and #) with either tones or recorded announcements, so presenting a PBI+ interface.

The simulation was developed to support the three basic UPT phase 1 services:

- Incoming-call Registration;
- Incoming-call Deregistration;
- Single Outgoing-call.

The relevant codes for access and control of these services have been defined provisionally and unilaterally by the Italian public network operator (see table 5).

Table 5: UPT codes

Service	Code
Registration	1484 XXXXX YYYYY 1
Deregistration	1484 XXXXX YYYYY 0
OutCall	1484 XXXXX YYYYY 8
Where: 1484 is the service access code XXXXX indicates the UPT number YYYYY indicates the PIN the last digit indicates the procedure.	

The recorded announcements (prompts, feedback, status information, etc.) for these services were specified directly for this experiment.

5.3.2.2 The prototype ISDN user interface

The ISDN prototype service has been implemented within the Turin area for trial purposes. Access was provided via a FACE-Alcatel© telephone subset consisting of:

- a 12-push-button keypad (digits 0 to 9, * and #);
- 3 editing keys (←, Cancel, →);
- 4 function keys labelled with abbreviations or acronyms to access the following functions: Setup (ASS), Suspend/Resume (S/R), Redial (RIS), Access by menu to specific ISDN supplementary services (e.g. 3-Party Conference);
- a 2-line by 20-character LCD display;
- 2 small lamps (LEDs) associated with the ASS and RIC keys.

Prompting and feedback messages were presented on the display, the only auditory indications being the tones (dial, ringing, busy tones) and one or two recorded announcements.

Within the range of services available, the following three were considered representative:

- Call Barring (CB);
- Call Forwarding Unconditional (CFU);
- 3-Party Conference (3PTY).

For two of these services (CB and CFU) three functions were possible, for the other service (3PTY) only activation was tested. See table 6 for the relevant codes:

- activation;
- deactivation;
- interrogation of the status of the service.

Table 6: ISDN services and relevant codes

Service	Activation	Deactivation	Interrogation
CB	*33#	#33*PIN#	*#33#
CFU	*21*TN#	#21#	*#21#
3 PTY	TN1 RIC 1 TN2 1		
Where:	PIN = Personal Identity No. (to enable calls) TN = Telephone Number (new destination) RIC = "Service access" function key		

5.3.2.3 The experimental design

Both experiments used a scenario approach to present the experimental task to the user. For example the scenario for the UPT Incoming Call Registration was as follows:

UPT Registration

You are waiting for an important call but you have been invited by friends to dinner.

As soon as you arrive at your friends' home you register at their phone for the incoming calls: you use the PERSONAL MOBILITY SERVICE to communicate the telephone number at which you wish to be reached from that moment on.

To do so you dial the Service Access Code, 1484, and follow the instructions given through the telephone.

Please note that:

- your Personal Number is 56 789;
- your Personal Identification Number is 12 345;
- the function code is 1;
- the Telephone Number of your friends is 011/22 86 166;
- after dialling the last digit press #.

On lifting the handset and dialling the UPT access number the user would then be prompted by the recorded announcements and would be required to complete the task.

The scenario for ISDN Call Forwarding was similar:

ISDN Call Forwarding- Activation

You are leaving for your country house where you will be spending the weekend. You want your incoming calls to reach you there. Therefore you dial *21* followed by the telephone number of your country-house and #. Then you hang up.

The telephone number of your country-house is 228 6166.

Deactivation

Back home from your weekend in the country-side you want your incoming calls to reach you again at your usual telephone address. Therefore you pick up the receiver and, upon hearing a continuous tone (replacing the dial tone to remind you that the service is active), dial #21#. Then hang up.

Following completion of each experiment the subjects were given a simple questionnaire, and asked to rate (on a 5 point Likert scale) different aspects of the interface.

5.3.2.4 Results

Both performance data (time and errors) and subjective data were collected in both experiments.

5.3.2.4.1 UPT

No errors were recorded for any of the scenarios by any of the subjects. The only errors were due to subjects exceeding time out.

The mean performance times varied according to the task (see table 7), and the number of keystrokes.

Table 7: Mean performance time in UPT

Service	Code
Registration	66,4 s
Deregistration	49,0 s
OutCall	67,3 s

The scores given on the Likert scales are shown in table 8.

Table 8: Mean subjective reactions in UPT

Attributes	Registration	Deregistration	OutCall
Overall satisfaction	4,5	4,5	4,6
Length of the procedures	4,1	4,4	4,2
Ease of use of the procedures	4,7	4,8	4,8
Understandability of the messages	4,7	4,8	4,8

The subjective data indicates very favourable attitudes, which in at least one case (length of the procedures) seem to be overoptimistic.

5.3.2.4.2 ISDN

The performance data collected in the ISDN experiment was slightly different from the UPT data:

- the failure rate;
- the number of attempts made before performing correctly the tasks concerning the Call Barring and Call Forwarding services;
- the number of unnecessary steps made to complete successfully the task on the 3-Party Conference service;
- the time spent to perform correctly a given task (time of the last attempt, in case of more trials).

Table 9: Mean performance data in ISDN

ISDN Services	Functions	Failures (%)	Unnecessary Trials/Steps	Performance Time (sec)
Call Barring	Activation	0 %	1,07	28,6 s
	Deactivation	0 %	0,19	43,4 s
	Interrogation	0 %	0,18	28,8 s
Call Forwarding	Activation	0 %	0,52	56,2 s
	Deactivation	0 %	0,15	34,1 s
3 Pty Conference	Set-up	37 %	3,2	506,9 s

The scores given on the Likert scales are shown in table 10.

Table 10: Subjective reactions in ISDN

Attributes	Conference	Other services
Overall satisfaction	3,0	4,1
Ease of use of the procedures	2,9	4,3
Understandability of the messages	2,3	3,9

5.3.2.4.3 Comparing UPT and ISDN results

The two sets of results are compared in table 11. It can be inferred that there was more uncertainty within the ISDN subjects from there increased average time per key-stroke and the reduced average rating (of satisfaction, ease of use, etc.). The comparison also highlights the significant problem that the users had with the current implementation of the 3-Party Conference service.

Table 11: Comparison between UPT and ISDN

Data Measured	UPT Registration Deregistration	ISDN Activation Deactivation Interrogation	ISDN 3-Party Conference
Failures	-	-	37 %
Errors	-	0,42	3,2
Time per keystroke	4 s	6 s	508 s
Satisfaction	4,5	4,1	3,0
Length of procedures	4,2	-	-
Ease of procedures	4,8	4,3	2,9
Understandability of messages	4,8	3,9	2,3

5.3.2.5 UPT vs. ISDN conclusions

It could be argued that the difference between the two services was caused by the introduction of * and # into the codes used to access and control the service. However, from the subjects' feedback after the experiment, this is probably too simplistic a statement. There was considerable criticism also of the recorded announcement in UPT and the visual messages in ISDN.

Finally, these two experiments also brought attention to a ceiling effect. The subjects' task, in simply reading and interpreting a typed scenario into a set of sequences to be keyed into a telephone, was probably too easy. The lack of errors within the UPT experiment and in two thirds of the ISDN experiment meant that the subjects did not really test any shortcomings in the PBI+ or PBI++ design. The subsequent experiments (see subclauses 5.3.3, 5.3.4 and 5.5.5) were much more searching.

5.3.3 Structured user information

Both Israelski [10] and Jones [13] showed that it is more helpful to users if the service codes can be structured into meaningful groups. Unfortunately current European implementations have a problem: there are only a few services on offer, and for these few, the service code allocation must seem random, even though CEPT has applied a structure [1]. For example:

21 = Call Forward Unconditional (CFU)
33 = Outgoing Call Barring (OCB)
43 = Call Waiting (CW)
61 = Call Forward No Reply (CFNR)
67 = Call Forward on Busy (CFB)

The memory research also suggested that a user's retention is improved if some form of cognitive processing is involved. Consequently, an experiment was set-up to test the effect on user performance of two different types of user manual (part 4 of the present ETR [6]). Table 12 shows an example of part of the Categorized Manual. The services were presented in distinct groups with simple direct names for each group, and sub-names for each service within a group. To ensure there was some cognitive processing the users had to work out for themselves what the relevant service code was. They were told the syntax was, for example, * "group number" " service number" *, etc. In contrast, the Non-Categorized Manual tried to copy the format and content of the local network operator's manual (see table 13).

Table 12: A typical section from the Categorized Manual

GROUP 2	**Send on**
Service 0	not in use
Service 1	**Send on - "all calls"** This service allows you to send on *everyone* who calls you up to another telephone automatically. For that purpose you must enter the number of that telephone (*the destination*), after activating the service. *How to use it?* Activation and deactivation is possible only if you hear the dialling tone.
Service 2	**Send on - " when busy"** This service allows you to send on *everyone* who calls you up to another telephone automatically. For that purpose you must enter the number of that telephone (*the destination*), after activating the service. Phone calls are sent on ONLY if your telephone is BUSY. *How to use it?* Activation and deactivation is possible only if you hear the dialling tone.

Table 13: A typical section from a Non-Categorized Manual

Send on - "all calls" This service allows you to send on *everyone* who calls you up to another telephone automatically. For that purpose you must enter the number of that telephone (*the destination*), after activating the service. *How to use it?* Activation is possible only if you hear the dialling tone. > *21* **Stop sending on - "all calls"** This service allows you to stop sending on phone calls by means of the "Send on - all calls". *How to use it?* Deactivation is possible only if you hear the dialling tone. >#21#
Send on - " when busy" This service allows you to send on *everyone* who calls you up to another telephone automatically. For that purpose you must enter the number of that telephone (*the destination*), after activating the service. Phone calls are sent on ONLY if your telephone is BUSY. *How to use it?* Activation is possible only if you hear the dialling tone. >*22* **Stop sending on - "when busy"** This service allows you to stop sending on phone calls by means of the "Send on - when busy". *How to use it?* Deactivation is possible only if you hear the dialling tone. >#22#

Following an initial familiarization with the relevant manual, subjects were then asked to complete a series of 25 scenarios each of which would require them to identify and use one of the services described in the manual. The subjects used a telephone user interface (simulated on HyperCard™) which provided them with a 12 key keypad and full feedback (tones, voice announcements and displayed text).

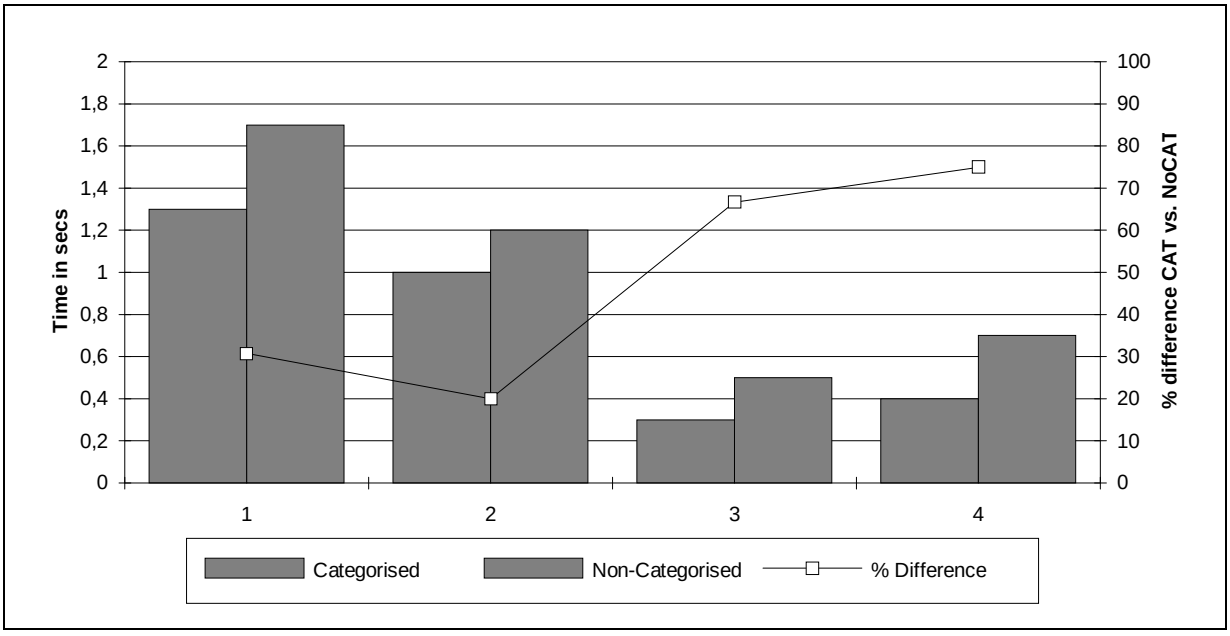

Figure 2: Mean duration of manual consultations for each scenario replication

Figure 2 shows a comparison of the manual consultations the subjects made between the Categorized and Non-Categorized manuals. Clearly the subjects spent significantly less time looking in the Categorized manual, and interestingly this difference increased as the experiment progressed (key scenarios were each repeated four times).

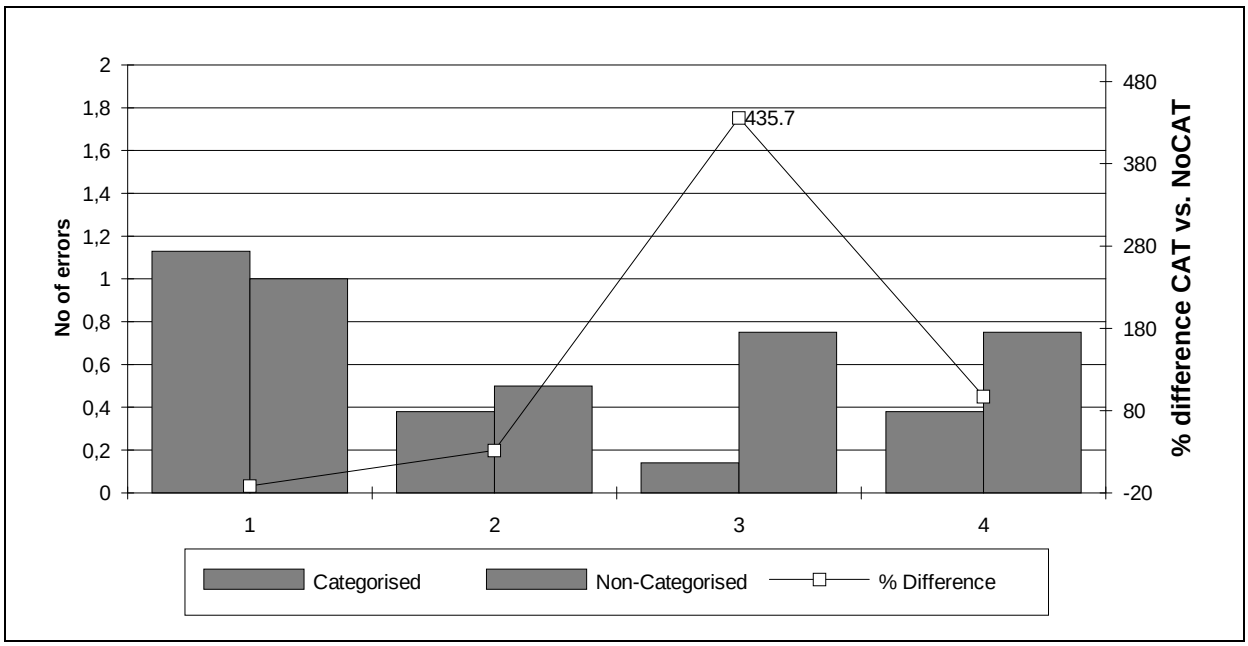

Figure 3: Mean number of errors per scenario replication

Figure 3 shows a similar comparison of the number of errors the subjects made, and the same effect is evident. Subjects were making significantly fewer errors if they used the Categorized manual (after the first replication) than if they had used the Non-Categorized manual. It is felt that further replications would be needed to understand why the slight increases in replication 4, seen in figures 2 and 3, occurred.

Looking back at the two tables 4 and 5 which show examples of the two manuals used, it is certain that every effort was taken to keep the "information content" and the "text" as near as possible the same. The difference between the manuals was, as far as possible, kept to the structuring or categorization of the information. After the subjects had completed all the experimental scenarios, they were asked to complete an effort scale, an acceptance questionnaire and a set of progressive matrices, they were then tested on the services and codes they could recall. The Category manual subjects scored 72,5 % and the Non-Category subject 53,75 %, again confirming that cognitive processing seems to assist long-term recall.

5.3.4 Code scheme differences

A paper based learning experiment (part 5 of the present ETR [7]) was set-up to compare the two recommended code schemes, CEPT [1] and GSM [2]. Four fairly matched groups of subjects attended one session a week for three consecutive weeks. In week 1 they learnt a code scheme from a prepared booklet, which trained them on the commands for three services (CFU, CB and 3PTY). During the training they had the opportunity to practice the commands on dummy telephones. After a brief questionnaire on their opinions of the training and code scheme, they were tested. In week 2, they were either retested on the learnt scheme, or learnt and tested the alternative scheme. In week 3, they were retested on one or both schemes. Table 14 summarizes the experimental design.

Table 14: CEPT vs. GSM learning experiment.

	Week 1		**Week 2**		**Week 3**		
	Learn	**Test**	**Learn**	**Test**	**Test**	**Test**	**Test**
Group 1	CEPT	CEPT	GSM	GSM	CEPT	GSM	Ravens
Group 2	CEPT	CEPT	-	CEPT	CEPT	-	Ravens
Group 3	GSM	GSM	CEPT	CEPT	GSM	CEPT	Ravens
Group 4	GSM	GSM	-	GSM	GSM	-	Ravens

The design allows direct comparison between the groups learning CEPT and GSM. It also allows assessment of any transfer of training if two schemes operated in parallel. The Ravens test (Progressive Matrices) was given to all subjects to assess performance in relation to abstract problem solving.

5.3.4.1 CEPT vs. GSM

Figure 4 compares the mean and standard deviation (sd) of the errors between the two groups learning only one scheme. Analysis of variance confirmed that there was no significant difference between the two schemes, even though the GSM group appears to show consistently less variation.

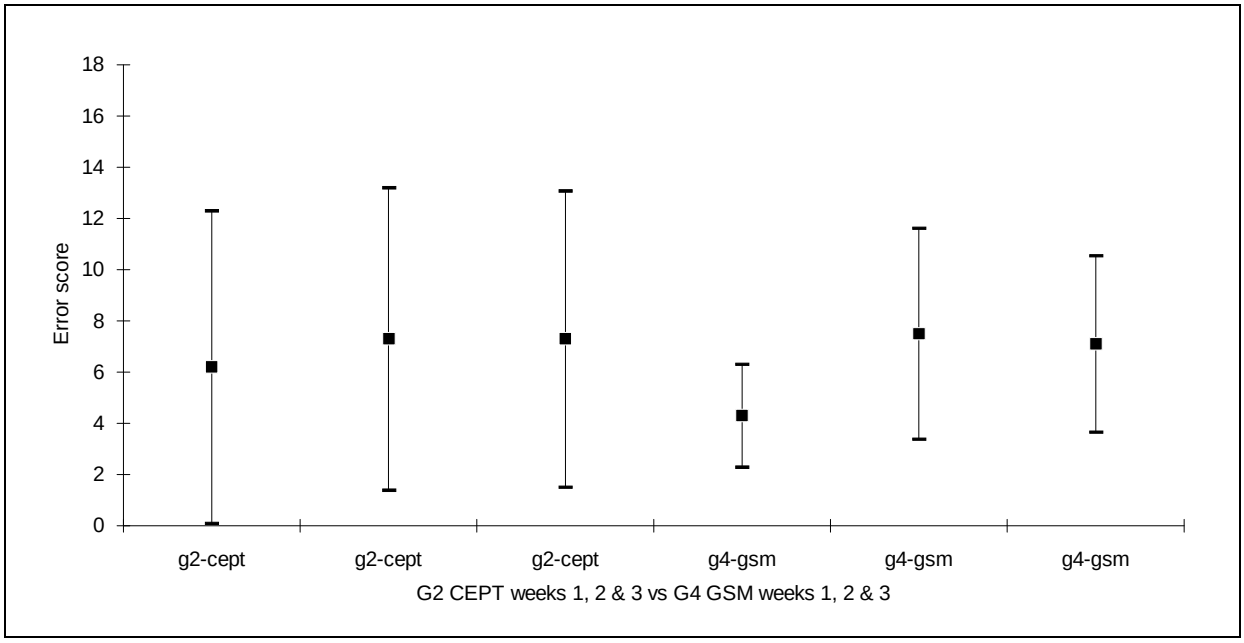

Figure 4: Groups 2 (CEPT) & 4 (GSM) - mean and ± 1 sd of raw error scores (max. 18)

Figure 5, compares the two groups learning both schemes. Analysis of variance again showed there was no significant difference between the two schemes, nor was there any apparent transfer of training (positive or negative). However detailed analysis of individual subjects error rates suggested that the individual effects were masking each other (about 50 % of subjects showed some degree of positive effect and 50 % a negative effect). From these broad analyses it would seem that both schemes are as good or as bad as each other.

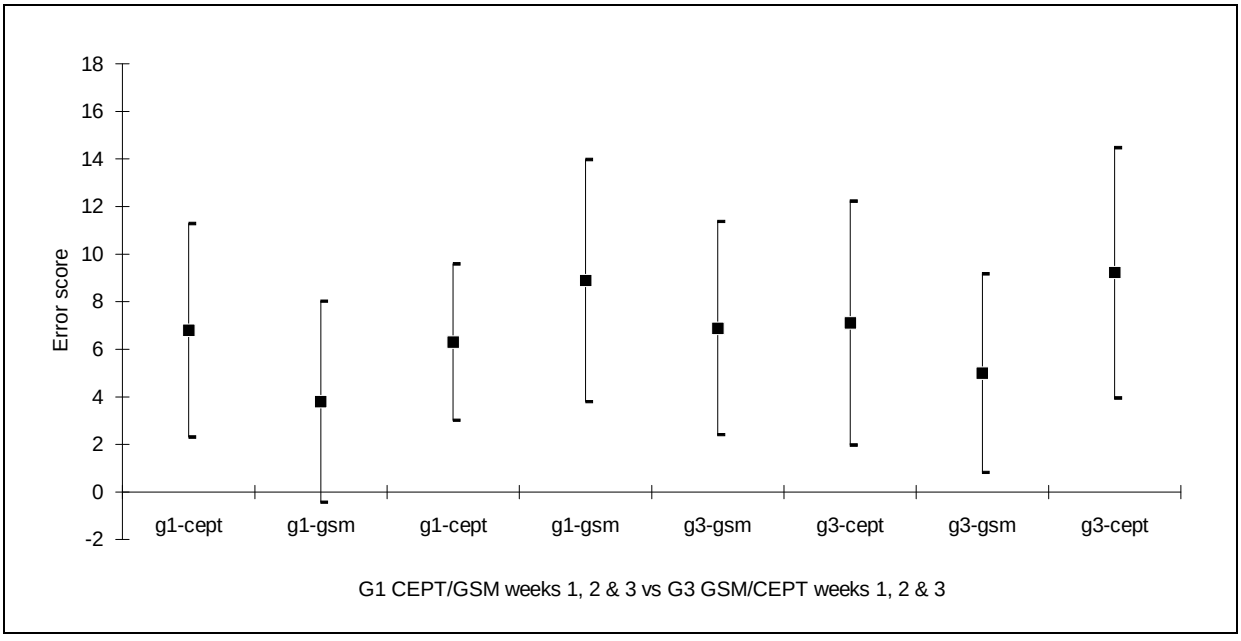

Figure 5: Groups 1 (CEPT/GSM) & 3 (GSM/CEPT) - mean and ± 1 sd of raw error scores (max. 18)

5.3.4.2 Error analysis

From the standards development perspective, a detailed analysis of the errors was more important. Which particular aspects of the two code schemes were giving the most problems. Figures 6 and 7 show the error scores for each question within the CEPT and GSM tests. The first point is the remarkable consistency in the patterns, irrespective of when the test was done. The most difficult questions for both schemes were:

Q.5	Call Forwarding	- How to erase the stored number (GSM is easier than CEPT)
Q.7	"	- How to store a number for later use (Register only)
Q.9	Call Barring	- How to switch off (subjects forgot their PINs)
Q.11	"	- How to suspend call barring for one call.

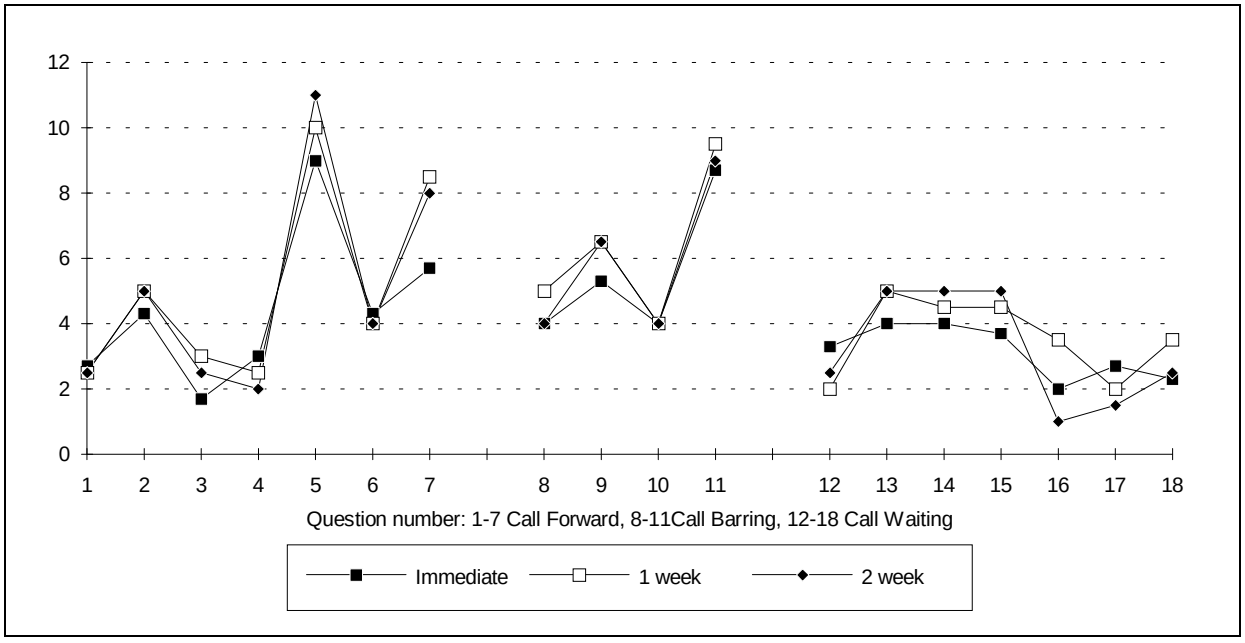

Figure 6: CEPT mean errors for each question, immediately after learning, 1 week and 2 weeks later

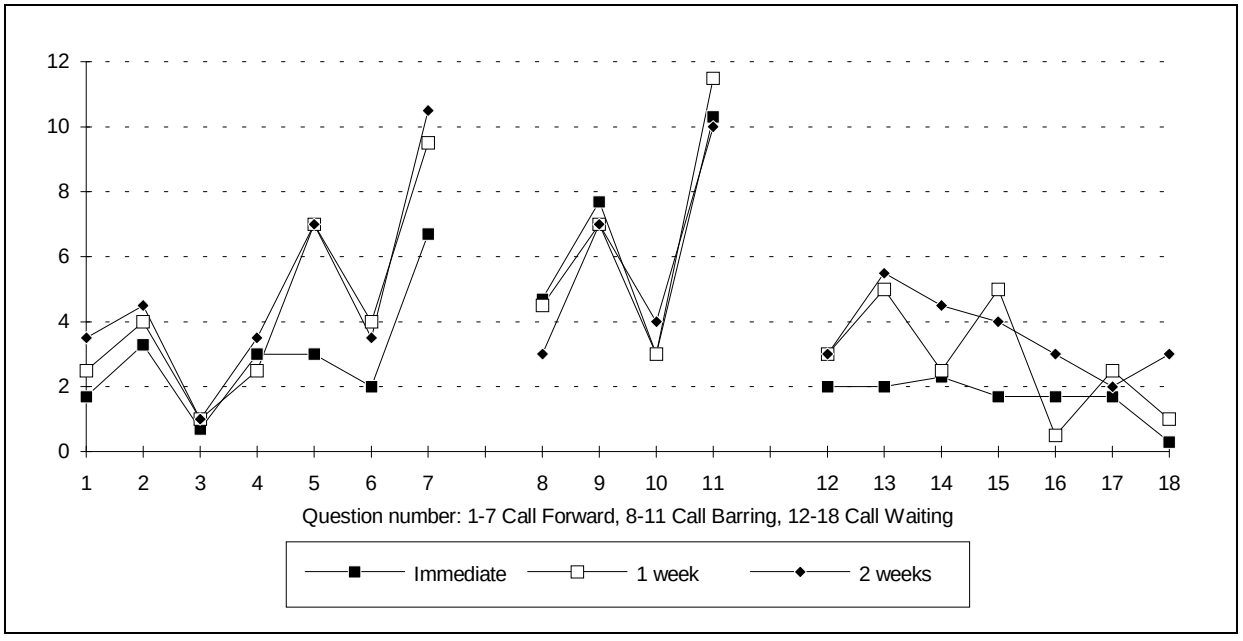

Figure 7: GSM Mean errors for each question, immediately after learning, 1 week and 2 weeks later

Table 15 compares the percentage of errors for each type of error between the two schemes. Clearly, the CEPT use of supplementary information fields to control register only and erase commands causes significant problems, and should perhaps be discontinued. The errors showing as "incomplete command set" and the "wrong or missing PIN" could probably best be solved by appropriate feedback. Whereas the switching order errors may be due to an inappropriate user model. Other experiments have reported users as thinking 1 controls call 1, 2 controls call 2, etc. One improvement maybe to present the correct model in the user manual, but a context sensitive display of options would obviously be better.

Table 16 zooms in for a closer look at the prefix errors, and the simplest comment is that problems arise as soon as the prefix involves more than one key press (e.g. *#, ** or ## induce errors). The solution can only be to keep a code scheme as simple as possible and if *# combinations are required for further prefix commands then the user requires a better than minimum error support facility.

Table 15: Analysis of Error Types

Type of Error Analysis percentage of possible errors	**CEPT**		**GSM**	
	Errors	Percentage	Errors	Percentage
Wrong or No CEPT Supp. Info.	74	54,4 %	-	-
Incomplete Sequence of Commands	15	22,0 %	50	36,8 %
Wrong or No PIN No.	72	26,5 %	87	32,0 %
Wrong Switch Order	73	26,4 %	61	22,4 %
Wrong Prefix	100	9,8 %	124	11,4 %
Wrong or No Suffix (Terminator)	61	5,9 %	74	6,8 %
No Answer offered	111	9,1 %	42	3,4 %
Adding PIN No.	11	1,5 %	30	4,0 %
Wrong or No Separator	18	3,3 %	9	1,4 %
CEPT/GSM Confusion	8	3,5 %	0	0,0 %
Wrong Service Code	15	1,5 %	3	0,3 %
Wrong Telephone No.	2	0,6 %	6	1,8 %
Wrong Syntax (Code sequence)	0	0,0 %	0	0,0 %

Table 16: Analysis of Prefix Errors

Wrong Prefix Analysis percentage of possible errors	**CEPT**			**GSM**		
	Errors	%	Rank	Errors	%	Rank
Activate/Deactivate - Swap */#	31	4,5 %	4	17	2,3 %	4
Activate/Interrogate - Swap */*#	34	5,0 %	3	58	8,5 %	1
Interrogate/Deactivate - Swap *#/#	21	5,1 %	2	16	3,4 %	3
Interrogate - Swap *#/#*	13	6,3 %	1	0	0,0 %	5
Erasure/Deactivate - Swap ##/#	1	0,3 %	5	33	8,1 %	2

5.3.4.3 Abstract problem solving

The correlation between a subject's Ravens score and his errors scores was highly significant (CEPT = 0,000 09 and GSM = 0,000 3). The only conclusion can be that those people who can learn command based code schemes are the same people that do well at abstract problem solving. Which means that achieving a reasonable and consistent level of usability for anything other than a simple code scheme may be restricted to the "top 25 %" of the population.

5.3.5 Feedback - tones, announcements and visual messages

The third experiment (part 7 of the present ETR [9]) had two objectives:

- to assess the effect of different levels of feedback; and
- to evaluate a compliant implementation of the proposed draft standard.

A VisualBasic™ simulation was developed that would drive a simple DTMF telephone and a two line LCD display, to present each of three types of interface:

PBI = a 12 key keypad with tones only as feedback;
PBI+ = a 12 key keypad with tones and voice announcements;
PBI++ = a 12 key keypad with tones, announcements and visual messages.

A generic user manual was also developed to present the descriptive and procedural information on the three test services (CFU, CB and 3PTY). Two experimental situations were envisaged: Learning and Assisted Mode. In the Learning Mode, subjects had to learn the necessary services and commands, then without the manual they completed 26 "mini" scenarios designed to test all the possible service commands twice. In the Assisted Mode subjects had a reminder sheet beside them throughout the experiment. Fifty-four subjects were assigned to the six experimental conditions, Learning Mode: PBI- 12, PBI+ 12, PBI++ 12; Assisted Mode: PBI 6, PBI+ 6, PBI++ 6. Table 17 shows the correct service code and switching order commands used to access and control the services. The scenario description always gave the relevant subscriber number and PIN.

Table 17: Example of service code and switching order commands used

Task	CFU	CB	3PTY
Switch on (and store data)	*21*NNNNNN#	*33*PIN#	
Dial number (basic call)	NNNNNN	NNNNNN (fails)	NNNNNN (1st call)
Check status	*#21#	*#33#	
Switch off	#21#	#33*PIN#	
Suspend CB per call		#33*PIN*NNNNNN#	
Hold first call			2
Dial 2nd call			NNNNNN (2nd call)
Swap between 1st and 2nd calls			2
Join conference			3

In order to recheck the effect of syntax simplicity, the PBI interface in the learning mode was further simplified by removal of the check status or interrogation command (Prefix *#), and was then labelled PBI-.

5.3.5.1 Usability levels

As part of the compliance clause of the proposed draft standard, TC-HF set a target usability level:

> "Using only the information provided by the service provider, 75 % of a sample of the intended target users of the service should be able to access and control the service successfully at the first attempt; and 90 % of the same sample should be able to access and control the service successfully at the second attempt."

Figure 8 shows the relative success rates (percentage of successful tasks) for the first and second attempt, for each type of interface and across both learning and assisted modes. Interestingly the simplest interface, the non-compliant PBI-, showed the greatest level of success. Which perhaps confirms that the more complex commands cause users more problems.

Each of the compliant assisted mode interfaces (PBI+ and PBI++) would then seem to just about meet the predetermined usability target. The assisted mode interface that only just fails to achieve 90 % at the second attempt is the non-compliant PBI. Unfortunately the story is not so good in the compliant PBI+ and PBI++ learned mode. In particular, subjects had very significant problems with the "Suspend Call Barring" task and several forgot to complete the "Swap" task before "Joining a Conference". If these tasks are extracted, the success rates become Learned PBI+ 1st = 75 % and 2nd = 83 %, and for PBI++ 1st = 81 % and 2nd = 84 %. Obviously these tasks are stopping the subjects from reaching the first attempt criteria point, but other tasks (mainly call barring switch off and interrogate) are also preventing subjects from reaching criteria on the second attempt.

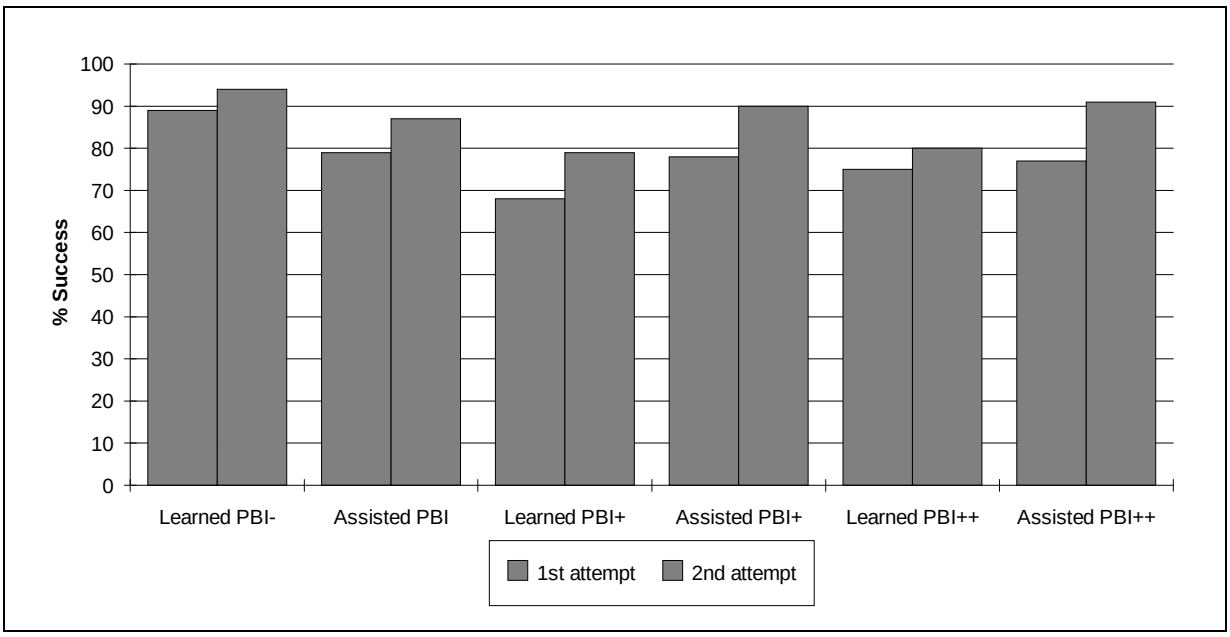

Figure 8: Success rates for different experimental conditions Learned vs. Assisted (PBI-, PBI, PBI+ and PBI++)

5.3.5.2 Tones, announcements and visual displays

To consider the effect of different levels of feedback, another parameter may be more appropriate than the overall success rates shown in figure 8. If a particular form of feedback has been more useful than another, then it might be expected that subjects experiencing that feedback type might improve more between 1st and 2nd attempt than on the other. Figure 9 looks at the comparative percentage improvements (between attempt 1 and 2) in success rates across the three feedback conditions.

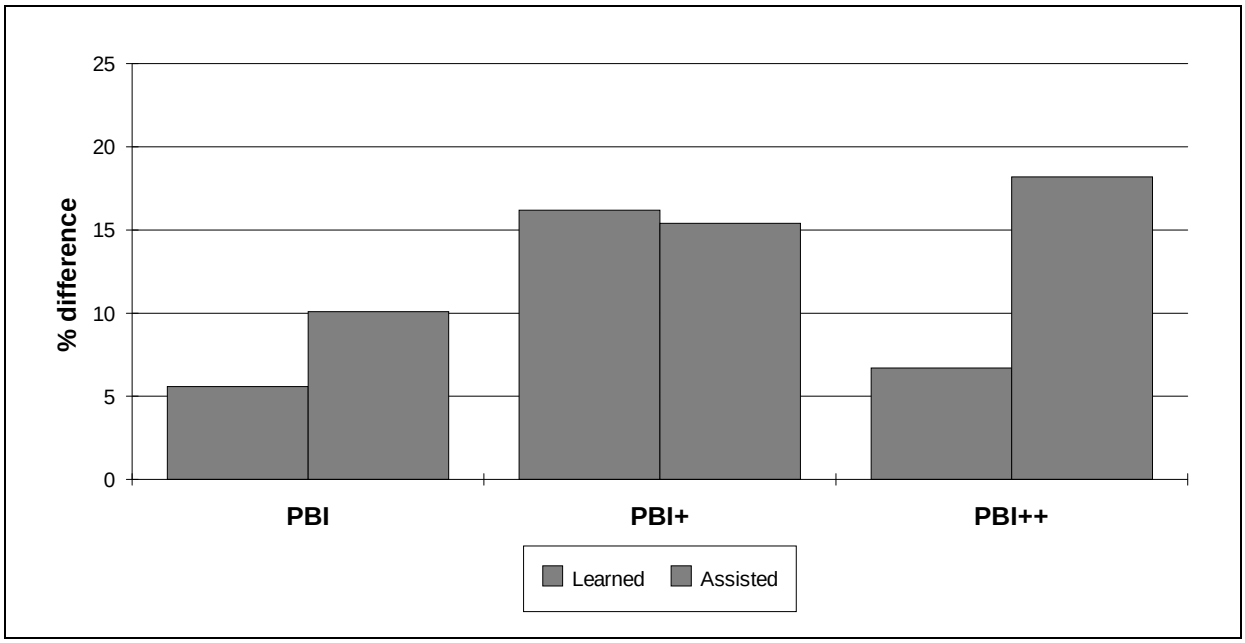

Figure 9: Percentage improvement between replications for different types of feedback (Learned vs. Assisted mode)

From this it can be seen that PBI+ shows a consistent improvement rate, and that PBI++ only betters this in the Assisted mode. Of course this comparison has also to be tempered by the fact that these two conditions had more room to improve than the PBI condition. To get a fairer picture it would need more replications. However, the announcement feedback taught at least one subject where he went wrong; he was expecting to switch on call barring rather than forward his calls to his PIN number!

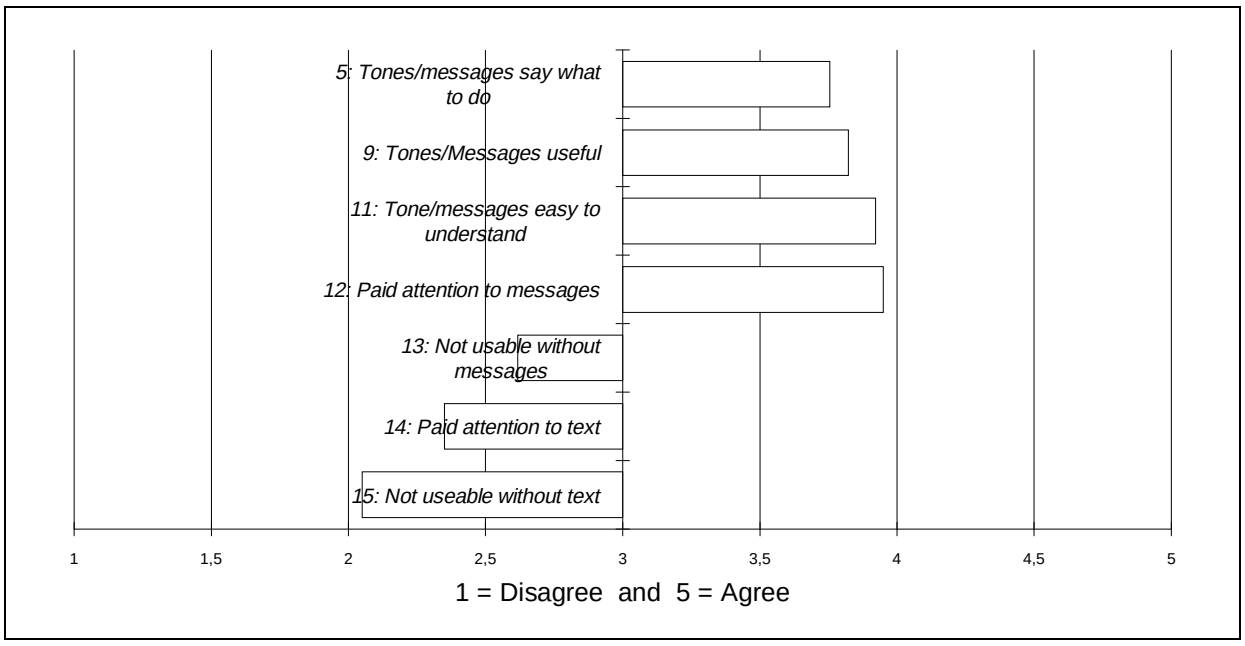

Figure 10: User perceptions of tones, announcements and visual messages (Learned and Assisted modes combined)

The differences between the feedback level is more starkly shown in the subjects' opinion responses shown in figure 10. Five statements related to the tones and voice announcements and four received a quite strong positive rating, the fifth pushed the subjects to say whether the services were "Unusable without the announcements" and this received a moderate negative rating; which is not unreasonable. Perhaps more disappointing were the ratings for the visual messages, these were distinctly negative. However as the experimenters discovered too late, the simulation did not show these visual messages in an optimum way. They tended to be displayed after the announcement had started and were scrolled. Which only goes to prove that there is no substitute for careful design and two or more pre-tests when simulations are involved!

Finally, it is worth stating the obvious. Where learning is involved, usability is directly related to syntax complexity. PBI- subjects consistently performed better than the other groups. Therefore the experimenters believe that it would be better if the service instructions only attempted to teach the two basic commands "Switch on" and "Switch off", and to offer prompting material so that users could learn the additional facilities over an extended period.

6 Conclusions and recommendations

The proposed ETS [3] should, as far as is practical, reflect the collated research and experimental evidence. It should set out to harmonize the existing code schemes and to establish a new minimum level of usability for supplementary service control. For the first time, an ETS should define a minimum man-machine interface (MMI) for public network based supplementary services to include:

1) the information provided to a user before a control action is performed;
2) the control actions a user performs to gain access to and control a service;
3) the information provided to a user after a control action has been performed.

These three elements of the MMI should then be defined in detail. The information provided before a control action, for example, should define some minimum contents for both the descriptive and procedural information given to users of the service, as well as for the prompting information that a service is active or invoked. Equally, the information provided after a control action should define separately the information content required as a response for each type of control action.

The control actions may be command or interactive dialogue based. Where a command dialogue is used, the syntax (prefixes, separator and suffix) should be defined for Service Code, Switching Order, Abbreviated Dialling and Alphanumeric commands. Where an interactive dialogue is used the dialogue may use any sequence of control actions and feedback/prompts commensurate with achieving an acceptable level of usability.

The draft ETS should conclude with an informative annex listing the existing CEPT recommendations on the allocation of service codes.

History

Document history	
October 1996	First Edition

ISBN 2-7437-0411-X - Edition complète
ISBN 2-7437-0862-X - Partie 1
Dépôt légal : Octobre 1996