



ETSI
TECHNICAL
REPORT

ETR 261-5

October 1996

Source: ETSI TC-HF

Reference: DTR/HF-01028-5

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 5: Experimental evaluation of the
CEPT and GSM code schemes**

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	7
4 Experimental design	7
4.1 Introduction	7
4.2 Subjects	8
4.3 Materials	9
4.4 Context	10
4.5 The Pilot Study	10
4.6 The Main Study	10
5 Results	11
5.1 Introduction	11
5.2 Were the subject groups homogeneous?	12
5.3 What performance and preference differences were there, if any, between the CEPT and GSM schemes?	12
5.3.1 Comparison of errors between Group 2 (CEPT) and Group 4 (GSM)	12
5.3.2 Comparison of errors between Group 1 (CEPT/GSM) and Group 3 (GSM/CEPT)	13
5.3.3 Comparison of errors across all four groups, CEPT vs. GSM	14
5.3.4 Comparison of Opinion Ratings	14
5.4 Is there any effect, positive or negative, from learning both schemes?	16
5.4.1 Individual subject differences	16
5.5 What type of errors were subjects making?	19
5.5.1 Which questions caused problems?	19
5.5.2 What type of errors?	21
5.6 What level of usability / learnability is achievable with the current code schemes?	22
5.6.1 Is there a link between a subject's Raven's grade and his error scores?	22
5.6.2 Are some subjects consistently better than others?	23
5.6.3 Are some questions easier to answer?	25
5.6.4 Does the opinion data give any indications of which aspects of the commands are more usable?	26
5.6.5 Is there a relationship between a subject's performance and his opinion ratings?	29
5.6.6 Answering the question	30
6 Conclusions	30
6.1 General	30
6.2 Training materials	30
6.3 Test materials	30
6.4 The results	31
Annex A: Sample Training Booklet for CEPT scheme	32
Annex B: Opinion Questionnaire	53
Annex C: Sample Quiz for CEPT Code Scheme	56
Annex D: General Instructions	60

History 61

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 5 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 5 of the ETR describes the experimental comparison of two code schemes, CEPT and GSM, used to access and control supplementary services provided by a telecommunications network. The UK-based experiment compared user performance and preferences within a learning task between the current CEPT [1] and GSM [2] recommended code schemes.

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 261-4: "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".
- [5] ITU-T Recommendation E.131: "Subscriber control procedures for supplementary telephone services".
- [6] 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".
- [7] Raven JC, Court JH & Raven J (1992): "Manual for Raven's progressive matrices and vocabulary scales. Section 3: Standard progressive matrices ", Oxford: Oxford Psychologists Press Ltd.
- [8] 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 [8] of the ETR apply.

4 Experimental design

4.1 Introduction

A simple paper-based learning experiment was chosen to assess people's performance and preferences on the two code schemes. The main reason for this approach was to get as large a sample as possible within a limited time-scale.

Four groups of subjects attended one session a week for three consecutive weeks:

In the first session, two of the groups (Groups 2 and 4) learnt one of the code schemes from a prepared booklet, then answered a simple opinion questionnaire, and was then tested on the learned code scheme using a quiz with a standardized set of questions. On the following two weeks both groups were retested on the same scheme using the same quiz. On the third week after the quiz, the subjects were asked to complete Raven's Standard Progressive Matrices (SPM).

In the first session, the other two groups (Groups 1 and 3) each learnt one of the code schemes, answered the questionnaire and completed the quiz, in the same way as Groups 2 and 4. In the second session the two groups each learnt the alternative code scheme, answered the questionnaire and completed the quiz. In the third session both groups (1 and 3) were tested on both code schemes, in the same order as they learned them. They were then asked to complete Raven's Standard Progressive Matrices. The experimental design is summarized in table 2.

Table 2: Summary of experimental design

	Week 1		Week 2		Week 3		
	Learn	Test	Learn	Test	Test	Test	Test
Group 1	CEPT	CEPT	GSM	GSM	CEPT	GSM	Raven's
Group 2	CEPT	CEPT	-	CEPT	CEPT	-	Raven's
Group 3	GSM	GSM	CEPT	CEPT	GSM	CEPT	Raven's
Group 4	GSM	GSM	-	GSM	GSM	-	Raven's

This structure provides a well balanced experiment and enables basic learning to be compared across the two schemes, as well enabling the effects of transfer (negative or positive) to be considered. It also approximates to a real world scenario, where a user has learnt to use one set of services, for example on his office telephone system, and may wish to start using them on his new mobile phone.

4.2 Subjects

Thirty nine subjects (23 female and 16 male) within the age range 25 - 52 volunteered to take part in the main study. Six different subjects (3 female, 3 male) also volunteered for a pilot study. The subjects all came from a white middle class professional background and recorded a higher proportion educated to A-level and above than might be expected in the general population. This subject pool may be seen as particularly valid because it is frequently regarded as a typical target market for supplementary services both fixed and mobile.

Subjects self-selected themselves into one of four groups, depending on their preferred time and day. The experiment was scheduled over three consecutive Tuesday and Thursday evenings, starting at either 19.00 or 20.30. The subjects assigned to each experimental group were therefore not random; for executive reasons it was decided that the two early evening groups would be groups 1 and 3.

As far as possible, it was attempted to keep the groups evenly balanced for sex, approximate age spread and approximate education level spread. This proved to be difficult due to the work patterns of the local male population (a high proportion commuted daily to London and were not available in the early evening). The original target was for twelve subjects in each group, six of each sex, but this proved impossible within sensible time scales. Table 3 summarizes the subject characteristics for each experimental group.

Table 3: Summary of subject characteristics for each experimental group

	Number	Mean age	Age range	Female : Male	Education level N : O : A : D	Mean Raven's grade (I, II+, II, III+, III-, IV, IV-, or V)
Group 1	10	43,9	38 - 52	4 : 6	1 : 5 : 2 : 2	III+ (I to IV-)
Group 2	10	38,4	25 - 45	5 : 5	1 : 2 : 3 : 4	III+ (I to IV)
Group 3	9	43,5	38 - 49	3 : 6	- : 7 : - : 2	III+ (II to IV)
Group 4	10	43,0	35 - 50	4 : 6	- : 2 : 2 : 6	III+ (II to III-)

Key: Education Level Raven's
 N = None recorded I = > 95 percentile II+ = 90 - 95 %
 O = O-Level/GCSEs II = 75 - 90 % III+ = 50 - 75 %
 A = A-Level III- = 25 - 50 % IV = 10 - 25 %
 D = Degree IV- = 5 - 10 % V = < 5 percentile

The key differences that might effect the data are:

- Group 2 is younger than the other three groups (there is one 25 year old; without him the mean age is 39,9, and the range is 35 - 45);
- Groups 2 and 4 include the majority of degree level subjects;
- Group 4 has the narrowest range of Raven's scores (II to III-).

4.3 Materials

A training booklet was prepared for each code scheme, (see annex A for the CEPT example). The booklet introduced three supplementary services: Call Barring, Call Waiting and Call Forwarding; and used these to present the basic range of commands for the subjects to learn.

These were:

	CEPT	GSM	
Register and Activate	*SC*SI#	*SC*SI#	same
Register Only	*SC*SI*0#	*SC*SI#, #SC#	different
Activate	*SC#	*SC#	same
Interrogate	*#SC#	*#SC#	same
Deactivate	#SC#	#SC#	same
Erasure	#SC*1#	##SC#	different
Switching Order	(R)SO	SO	different

Where: SC is Service Code, a two or three digit code (e.g. 21, 43, etc.)
 SI is Supplementary Information (e.g. a telephone number)
 SO is Switching Order Number (e.g. 1, 2, 3, etc.)
 R is Register Recall

The two booklets, one for CEPT and one for GSM, were identical in style and content for all aspects except the necessary command sequences. The booklets presented each supplementary service in turn and the commands required within a service in a set pattern: Description of the service, What to do, Why to do it. The sequence of the commands slightly varied from service to service but remained logical, i.e. Switch Off always came after Switch On. After the description, what to do and why, the booklets included a brief exercise to reinforce the command just learnt. Subjects were asked to write out the relevant service command for one or two specific scenario(s), and then to go through a complete keying sequence on the telephone provided. The telephones used, fixed and mobile, remained unconnected to the network and consequently offered no additional form of feedback as to the correctness or otherwise of the keyed command. After presenting the commands used for all three services, the booklet presented a general summary of the command syntax rules.

A common opinion questionnaire was prepared asking subjects to rate their level of agreement/disagreement on a five point scale for ten statements (annex B). The statements addressed five topics: the booklet, the services in general, the commands in general, and the use of service codes (21, 33, 43, etc.) and symbols (* and #) in particular. A negative and positive statement was presented for

each of the topics. The statements sequence was structured to present simple questions about the booklet and services first, before presenting the more complex statements about the service commands, etc. The negative/positive statements were separated and distributed evenly throughout the list. The instructions and question format were copied from a previously validated satisfaction inventory.

A common test paper or quiz was prepared for both code schemes, (annex C). The quiz had 18 questions covering the range of tasks possible for each service: 7 on Call Forwarding, 4 on Call Barring and 7 on Call Waiting. The questions were presented in their groups and asked subjects to record the *# and digit command sequences required for each task. A list of the required Service Codes was presented on the first page of the quiz, together with the PIN required for Call Barring. Any necessary telephone numbers were also included in the question. Thus the only elements being tested were the subjects' memories of the *# command sequences and the command syntax. The CEPT and GSM quizzes were identical except for the context information (Company Logo) at the top of the first page, and the colour of the paper.

4.4 Context

In order to help the subjects remember and separate the code schemes they were learning, each scheme was always presented within a "Context". For example, the CEPT scheme was introduced as coming from "Star Telephones Ltd.". The experimenter and assistant had name badges for this fictitious company, on the wall of the experimental room there were Posters declaring "Star Telephones Ltd" and showing the "Company Logo". The name badges, poster printing, the booklet, the quiz and the telephones (BT Viscount) were all red in colour. For the GSM scheme the fictitious company was "Mobile Communications plc", the colour was blue and the mobile telephones were black. The opinion questionnaire, being common to both code schemes, was yellow.

4.5 The Pilot Study

Before conducting the main experiment, a pilot study was conducted to verify the reliability of the learning and test materials and of the experimental design. Six subjects took part, drawn from the same population as for the main study. Two subjects were assigned to each of conditions 1 and 3 and one subject each to conditions 2 and 4. The pilot experiment was conducted at Ferris Associates offices over one week, e.g. day 1 was session 1, day 3 was session 2, and day 5 was session 3. Subjects completed the training booklets, opinion questionnaire and test papers necessary for each condition. They did not complete the Raven's exercise.

The results from the pilot confirmed several points:

- The materials had some errors; these were corrected before the full test set were reproduced.
- The methodology would appear to be acceptable to the subjects, and was successful in collecting the data required.
- An initial concern that the task may be too simple to obtain meaningful data, i.e. the data collected would present a ceiling effect; this was resolved.

There was also significant variance between subjects, from 72 % to 5 % errors, irrespective of condition; and significant variance between questions, from 81 % to 0 % errors, irrespective of condition. The individual data from the pilot experiment was not included in the main study.

4.6 The Main Study

The main experiment was conducted in one of the classrooms of a local school. For each experimental session the tables were evenly spaced about the classroom and the context material mounted on the walls. Each subject sat at a separate table. The relevant telephone and training booklet were placed on each table together with a pen or pencil.

At the start of the first session for each group of subjects, they were greeted at the entrance to the building, directed to the classroom and asked to sit at any desk. Before starting, the experimenter introduced herself and her colleague and gave a brief introduction on the purposes of the study and the order of events for the experiment (annex D).

At the start of the second session the room and tables were laid out as before. Groups 1 and 3 were reminded of the purposes of the study (to test the material and the commands, and not the subject) and of the order of events (training booklet, questionnaire, quiz) and of any special controls/keys necessary

(the "Recall" key on the telephone and the "Send" and "End" keys on the mobile were not obvious). Groups 2 and 4 were simply given the quiz for their particular condition/code scheme. The telephone/mobile was also placed on the desk for reference/context.

At the start of the third session the room and desks were laid out as above. Groups 1 and 3 were given their first quiz (for the code scheme they learnt in session 1) and after completion given their second quiz (for the code scheme they learnt in session 2). The telephone/mobile on the desk was also changed. The posters on the wall were not. Group 2 and 4 were again given their respective quiz. Following completion of the quizzes each group of subjects were asked to complete the Raven's Standard Progressive Matrices. This exercise was introduced using the standard set of instructions provided.

Finally at the end of the third session there was a short debriefing, covering the following points:

- Statement of thanks for the time and effort subjects had made.
- Clarification for the reason for the Raven's exercise (to enable the subject groups to be compared and to enable comparison with subjects used in the other studies).
- Invitation to question or comment on the study, and to have their individual results.
- Payment of the reward (two bottles of wine per subject).

Following the requests made during this debriefing, a short report was made back to all subjects providing feedback on the general results of the study; in addition a few individual subjects were given their individual Raven's scores/grade.

5 Results

5.1 Introduction

Three sets of data were obtained during the experiment:

- 1) Error scores from each of the quizzes completed in each experimental session.
- 2) Ratings from each of the opinion questionnaires for each subject on each code scheme learnt.
- 3) Raven's SPM scores for each subject.

The error scores used were raw error scores obtained simply by marking the subjects answers in the relevant quiz as right or wrong. The error score for any particular quiz was the sum total of wrong answers. To have the answer recorded as correct, the subject had to record the complete command sequence correctly. There were three questions, 1, 6 and 16, to which the subject could get a correct result with a different answer from the target answer. Within the analysis these pseudo-errors were counted as correct. These were:

	Target answer	Pseudo-error
Question 1	*21#	*21*566788#
Question 6	*21*0104969566788#	*#21#, to check then if necessary *21*0104969566788#
Question 16	0 or R0	"Ignore" or "Do nothing"

NOTE: The analysis of error types used a slightly different approach, see subclause 3.5

This data has been used to address five questions:

- 1) Were the subject groups homogeneous?
- 2) What performance and preference differences were there, if any, between the CEPT and GSM schemes?
- 3) Was there any effect, positive or negative, from learning both schemes?
- 4) What type of errors were subjects making?
- 5) What level of usability / learnability was achievable with the current code schemes?

Each question is considered separately.

5.2 Were the subject groups homogeneous?

The basic characteristics for the four groups of subjects were quoted above (see subclause 4.2: Subjects). The main data collected for comparing the groups was the Raven's SPM. Scores for each subject were converted to a Raven's Grade depending on the subject's age, from tables compiled for the UK population. The range of grades is I, II+, II, III+, III-, IV, IV-, and V, and these relate to 95+, 90-95, 75-90, 50-75, 25-50, 10-25, 10-5 and 5- percentile of the population respectively. The mean grade for each group of subjects was III+. If each subject's grade was converted to a ranking, 1 to 8 respectively, any differences between the groups could be tested by the Kruskal-Wallis non-parametric one way analysis of variance. A one-tailed probability of 0,617 confirmed that the groups did not differ significantly in respect of their Raven's scores. Therefore, for the purposes of the experiment, they could be considered homogeneous.

5.3 What performance and preference differences were there, if any, between the CEPT and GSM schemes?

The comparison of subject performance and preference between the two code schemes was made possible by the degree of homogeneity within the experimental groups. The comparison could be viewed at many levels, including:

- Comparison of errors between Groups 2 and 4, and between Groups 1 and 3;
- Comparison of errors between all four groups;
- Comparison of Opinion Ratings between the four groups.

5.3.1 Comparison of errors between Group 2 (CEPT) and Group 4 (GSM)

The raw error scores for Group 2 can be directly compared with those for Group 4 by using the mean and standard deviation for each experimental session.

Figure 1 shows this data for the two groups, comparing CEPT vs. GSM over the three consecutive weeks of the experiment. It demonstrates the broad spread of errors recorded in each group, and the similarity between the means. The main differences are that the GSM figures for week 1 are much lower than for any other session, also that the GSM sd (approx. ± 4) is generally narrower than the CEPT sd (approx. ± 6).

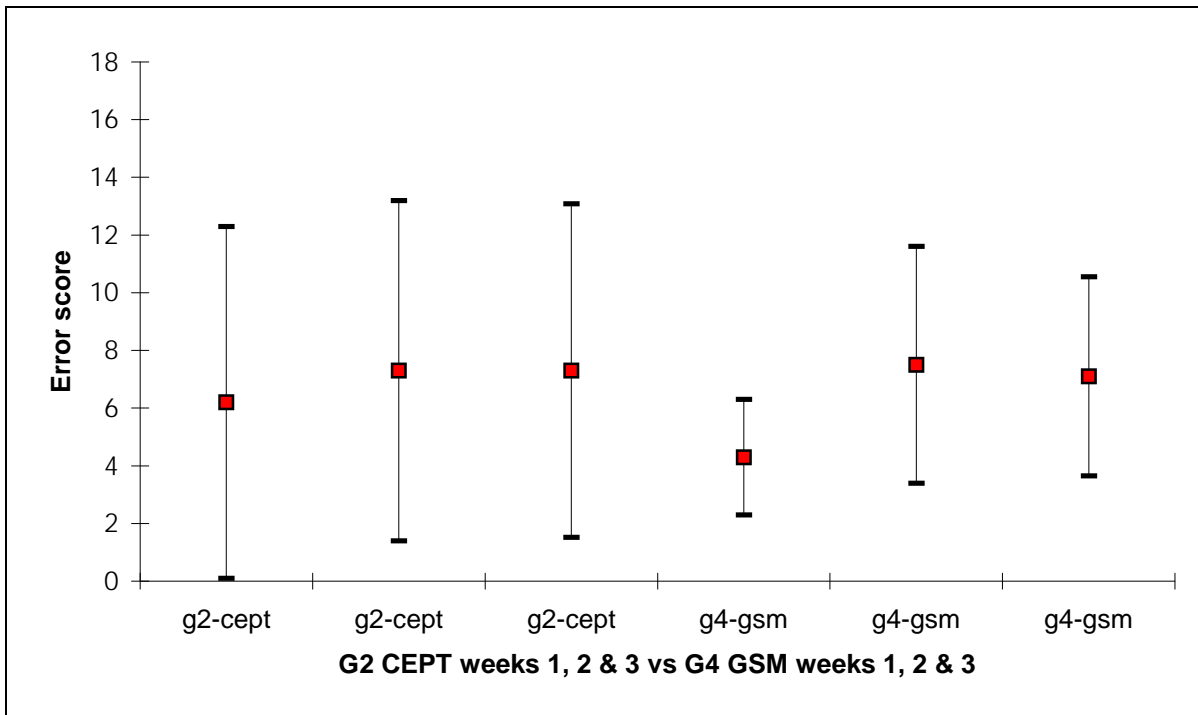


Figure 1: Groups 2 (CEPT) & 4 (GSM) - Mean and ± 1 sd of raw error scores (max 18)

Statistically, a parametric two way analysis of variance gave the following figures:

	F	Probability	
CEPT vs. GSM	0,122	0,73	
Session 1 vs. 2 vs. 3	4,466	0,019	<i>Significant > 0,05</i>
CEPT/GSM vs. Session	0,878	0,424	

Therefore, although figure 1 demonstrates the GSM data as having a smaller variance than CEPT, this is not significant. A separate F-test highlighted that what is significant between sessions is the error performance of Group 4 in week 1 with that in week 2, although the significance is not strong.

	F	Probability	
G4 GSM wk.1 vs. G4 GSM wk.2	4,2	0,02	<i>Significant > 0,05</i>

The indication is that there is no difference between the CEPT and GSM code schemes.

5.3.2 Comparison of errors between Group 1 (CEPT/GSM) and Group 3 (GSM/CEPT)

Figure 2 shows the mean and standard deviation for these two groups over the three experimental weeks. (NB: Both groups were tested on both code schemes in week 3). Again it demonstrates the broad spread of errors recorded for each session, irrespective of the code scheme being tested. Also again, the means are all very similar, most varying between an average of 30 to 40 % error. The differences are that the error rate for Group 1 goes down in week 2, but returns to the "normal" level in week 3. Also, that the error rate is slightly broader and higher in both of the second tests in week 3. (NB: For Group 1 this is GSM and for Group 3 it is CEPT).

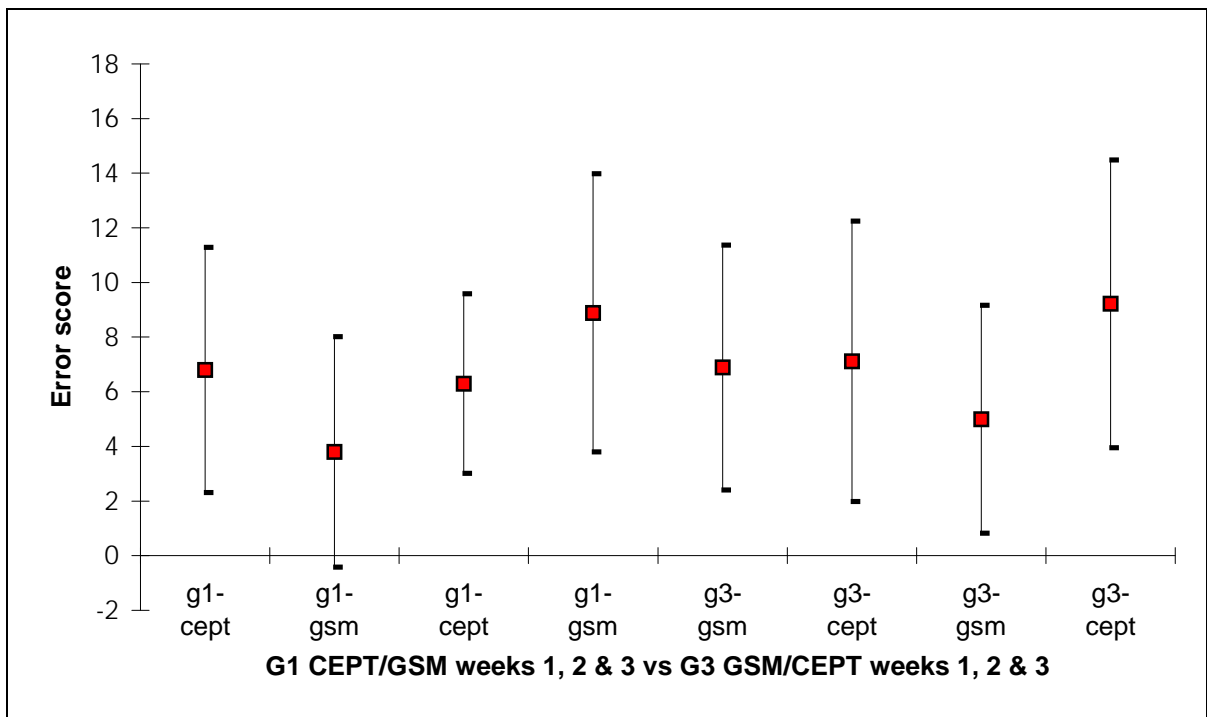


Figure 2: Groups 1 (CEPT/GSM) & 3 (GSM/CEPT) - Mean and ± 1 sd of raw error scores (max 18)

Statistically a parametric two way analysis of variance (unbalanced data) gave the following figures:

	F	Probability	
Main effects	1,49	0,22	
CEPT vs. GSM	1,39	0,24	
Session 1 vs. 2 vs. 3	1,53	0,22	
CEPT/GSM vs. Session	1,32	0,27	
Explained	1,42	0,22	<i>All non-significant</i>

These figures suggest that there was no consistent significant differences between any of the groups of data. Again, implying there is no difference between CEPT and GSM.

5.3.3 Comparison of errors across all four groups, CEPT vs. GSM

Comparing figures 1 and 2, graphically two points suggest themselves. The Group 4 week 1 mean and sd are very different to all other groups data; and Group 2 data is more spread than most of the other groups.

Statistically a parametric analysis of variance is not possible across all four groups. However it is possible to construct a table of F-tests comparing each experimental session with each other session. There is, however, a danger that this analysis has the possibility of increasing type 1 errors (i.e. rejecting the null hypothesis when it is true). This can be balanced to some degree by increasing the level of significance required; this carries a risk of type 2 errors (i.e. accepting the null hypothesis when it is false). Table 4 shows the results of the series of tests.

Table 4: Cross-comparison of each experimental session, F-tests

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A: G1cept1	X	--	--	--	--	--	--	--	--	--	--	✓	--	--
B:G1gsm1		X	--	--	--	--	--	--	--	--	--	✓	--	--
C:G1cept2			X	--	✓	✓	--	--	--	--	--	--	--	--
D:G1gsm2				X	--	--	--	--	--	--	--	✓	--	--
E:G2cept1					X	--	--	--	--	--	--	✓✓	--	--
F:G2cept2						X	--	--	--	--	--	✓✓	--	--
G:G2cept3							X	--	--	--	--	✓✓	--	--
H:G3gsm1								X	--	--	--	✓	--	--
I:G3cept1									X	--	--	✓✓	--	--
J:G3gsm2										X	--	✓✓	--	--
K:G3cept2											X	✓✓	--	--
L:G4gsm1												X	✓	--
M:G4gsm2													X	--
N:G4gsm3														X

Key: -- = Non-significant
 ✓ = Significant 0,05 - 0,01
 ✓✓ = Significant > 0,01

Clearly there is a consistent pattern: the data from Group 4 week 1 is consistently significantly different to all other groups except Group 1 week 3 CEPT scores. This is unlikely to have happened by chance even with the increased possibility of type 1 errors. However, what the table also shows is that there is no consistent pattern of differences between CEPT and GSM, or between the sessions in weeks 1, 2 and 3.

5.3.4 Comparison of Opinion Ratings

If there is no evidence of performance differences between the two code schemes, the question arises: is there a difference in the opinion ratings?

The subjects' ratings for each questionnaire were separated into the positive and negative statements. The rating for each negative statement was then reversed (i.e. 6-rating = reversed rating, thus 5 becomes 6 - 5 = 1, etc.). The ratings and reversed ratings for each statement were then totalled together to reach a general figure for each subject on each code scheme he had learnt. These totals ranged from 10 = maximum positive to 50 = maximum negative opinion.

Figure 3 shows the mean and standard deviation for the opinion ratings for each of the groups. The Group 1 and Group 3 data, being the two groups that completed both schemes, show a slight increase in negative opinion from CEPT to GSM. This increase is irrespective of the order in which the two groups learnt the schemes. However, this difference is not consistent with the other two groups, 2 and 4. What is perhaps more striking is the higher average of Group 3 over Groups 1, 2 and 4, and the larger spread of Groups 1 and 3 over Groups 2 and 4, for both CEPT and GSM. (NB: All four groups completed week 1 without knowing in detail what they would be doing in weeks 2 and 3, it was only groups 1 and 3 which completed a questionnaire in week 2 after learning their second code schemes).

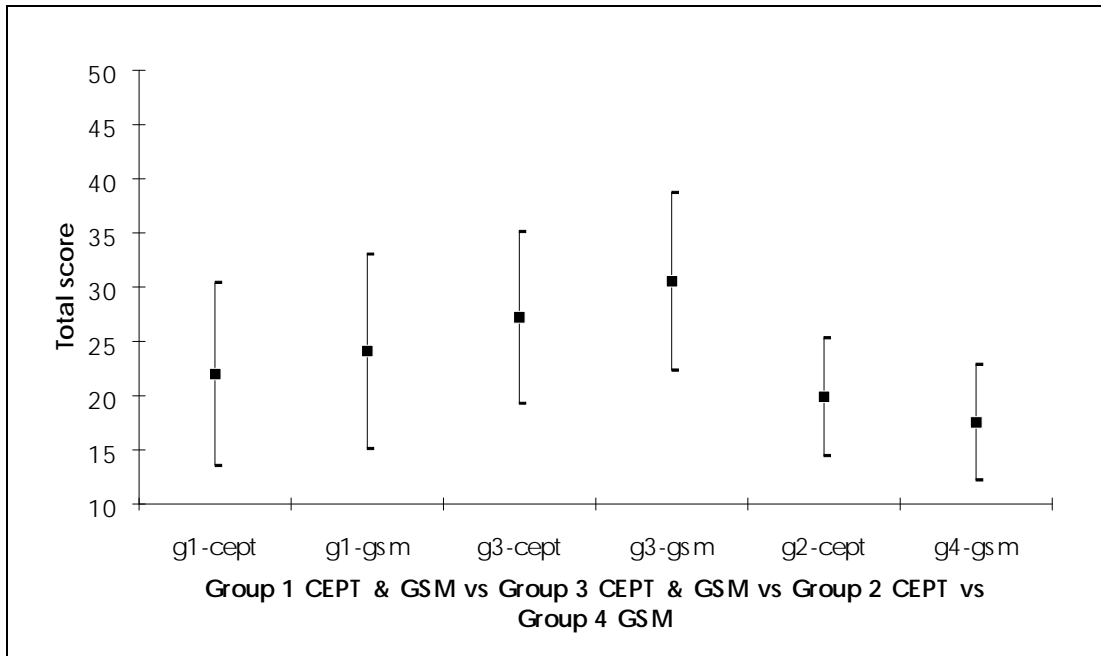


Figure 3: Opinion Questionnaire, Subject's total score per session - mean and standard deviation (min 10 = positive, max 50 = negative)

It might also be noted that figure 3 suggests a general positive opinion response to the code schemes. Almost all the means and standard deviations are less than 30, the mid point. However, it should not be forgotten that subjects invariably give a positive response in preference to a negative one in this type of situation. Whether this is due to a willingness to please the experimenter, or a partial justification for an evening's work, or any other combination of factors, remains untested.

Statistically, a Kruskal-Wallis non-parametric one way analysis of variance across the collective questionnaire scores for all the subjects in each group gave the following figures.

G1-Cept vs. G1-Gsm vs. G2-Cept vs. G3-Gsm vs. G3-Cept vs. G4-Gsm

chi-square	13,9	
df	5	
probability	0,016	<i>Significant 0,05 - 0,01</i>

Further analysis showed that the significant differences lay between the Group 3 and Groups 2 and 4 data.

G2-Cept vs. G3-Gsm vs. G3-Cept vs. G4-Gsm

chi-square	13,9	
df	3	
probability	0,0036	<i>Significant > 0,01</i>

This difference is not between CEPT and GSM, but rather identifies a possible difference between the groups.

From all of these comparisons graphical and statistical, error data and opinion data, it is clear that there is no distinction that can be assigned solely to the two code schemes.

5.4 Is there any effect, positive or negative, from learning both schemes?

The effect of learning both code schemes can be considered by graphical and statistical comparison of Groups 1 and 3 with Groups 2 and 4, and also by looking in more detail within Groups 1 and 3.

If there is an effect, positive or negative, from learning two code schemes, it could be expected that there would be differences between Groups 2 and 4 and Groups 1 and 3 with respect to their week 3 data. Figures 1 and 2 show that whilst there are differences, they are not consistent. For example, the group means are:

	CEPT	GSM
Group 1	6,3	5,0
Group 2	7,3	-
Group 3	8,9	9,2
Group 4	-	7,1

Thus it could be that Group 1 demonstrates an advantage, or positive effect, but Group 3 shows a disadvantage, but perhaps this is too simple.

Looking in more detail, figure 2 shows the collapsed data for Groups 1 and 3 in the sequence the codes were learnt and tested. A positive effect could be expected if after the retraining in week 2 there was a reduction in error rate over that for week 1. This reduction may also be expected to transfer to week 3. The reduction from week 1 to week 2 is shown in Group 1, but there is no continuity into week 3; and, unfortunately the effect is not repeated in Group 3. (NB: Perhaps the effect depends on the sequence the codes are learnt?) If, on the other hand, there is a negative effect, then the data from weeks 2 and 3 might be expected to be significantly worse than that for week 1. This is plainly not the case for either of the groups. Again, the overall spread of the data within each group and the similarity of the means across three weeks suggests that there is perhaps no real effect one way or the other.

One aspect that is interesting is that the error scores for the second test in week 3 tend to be higher than those scored for the first test, and this is irrespective of whether test 2 was CEPT or GSM. (NB: Perhaps subjects were inventing differences between the two schemes which did not exist?) Unfortunately, this difference is not statistically supported as can be seen from table 4.

Statistically, if there was any effect from learning both schemes, then it might be expected to show up within table 4, the F-test comparison across all groups. No such result was found.

5.4.1 Individual subject differences

One of the factors that a mean and standard deviation analysis can hide is a bimodal distribution. There is a possibility that some subjects were, for example, obtaining a positive transfer from learning two schemes, but that their benefit is being masked by others who are obtaining a negative transfer.

Figures 4a, 4b, 4c and 4d show the raw error scores for each subject in each group over the three weeks of the experiment. Week 1 in figure 4a, for example, might be thought to show three sub-groups within Group 2, e.g. those with < 2 errors, those with around 5 errors, and those with around 15 errors. To a large extent the relative error rates these subjects make in weeks 2 and 3 confirm their membership of a sub-grouping.

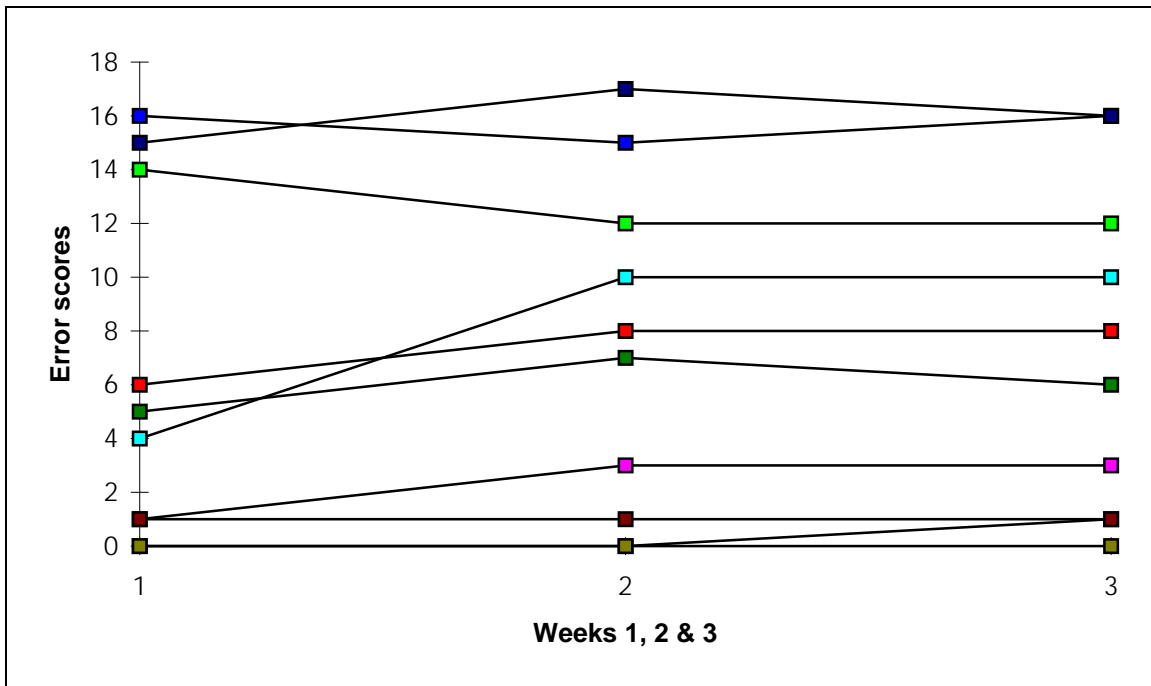


Figure 4a: Group 2 (CEPT), raw error scores for each subject in each session

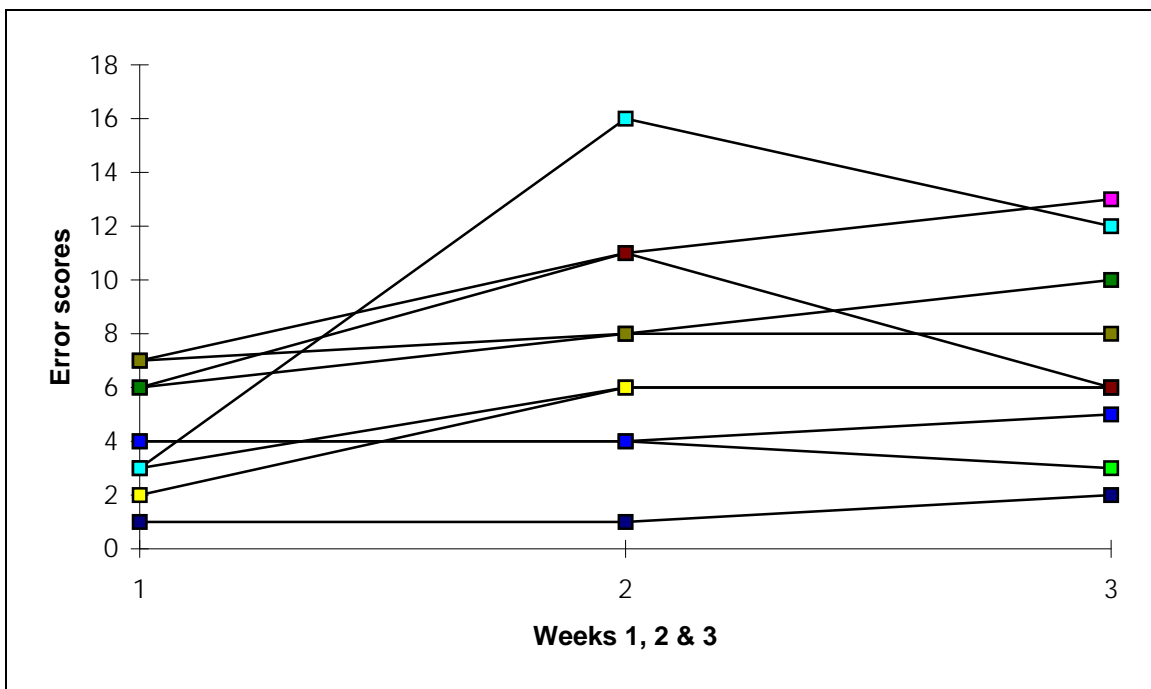


Figure 4b: Group 4 (GSM), raw error scores for each subject in each session

These sub-groupings can also be seen, but perhaps less clearly within each of the other subject groups, figures 4b, 4c and 4d. But what is also interesting is that there is a considerable cross-over. Where a subject performed well in week 1, it might be replicated in weeks 2 and 3, or it might deteriorate in weeks 2 and 3. Similarly, subjects that did less well in week 1 may do better in weeks 2 and 3. A visual check of the four figures (4a,b,c & d) confirms that all possible cases are represented.

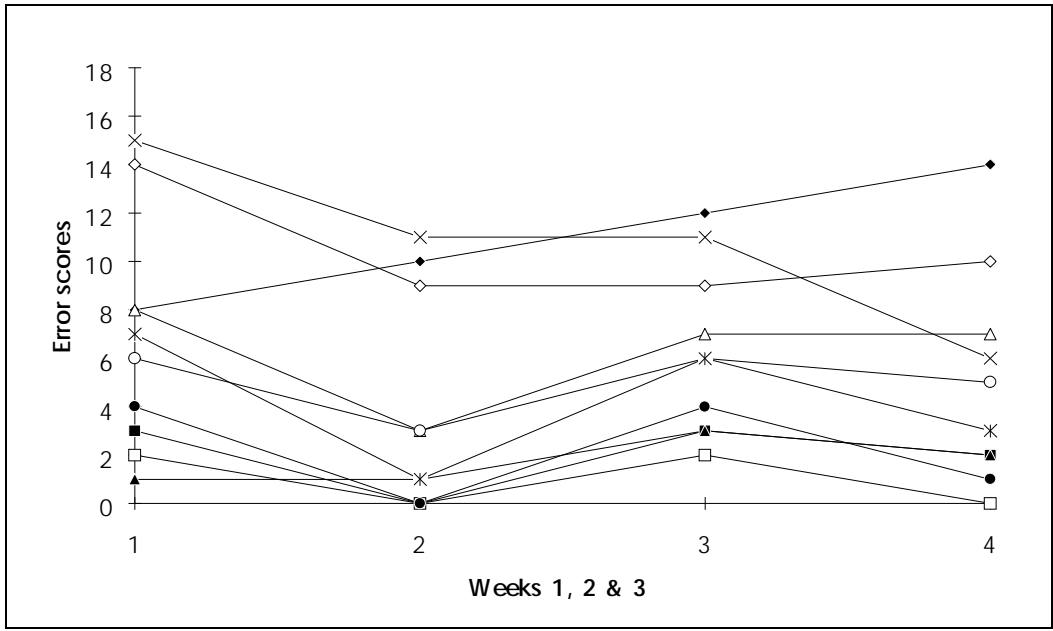


Figure 4c: Group 1 (CEPT/GSM), Raw error scores for each subject in each session

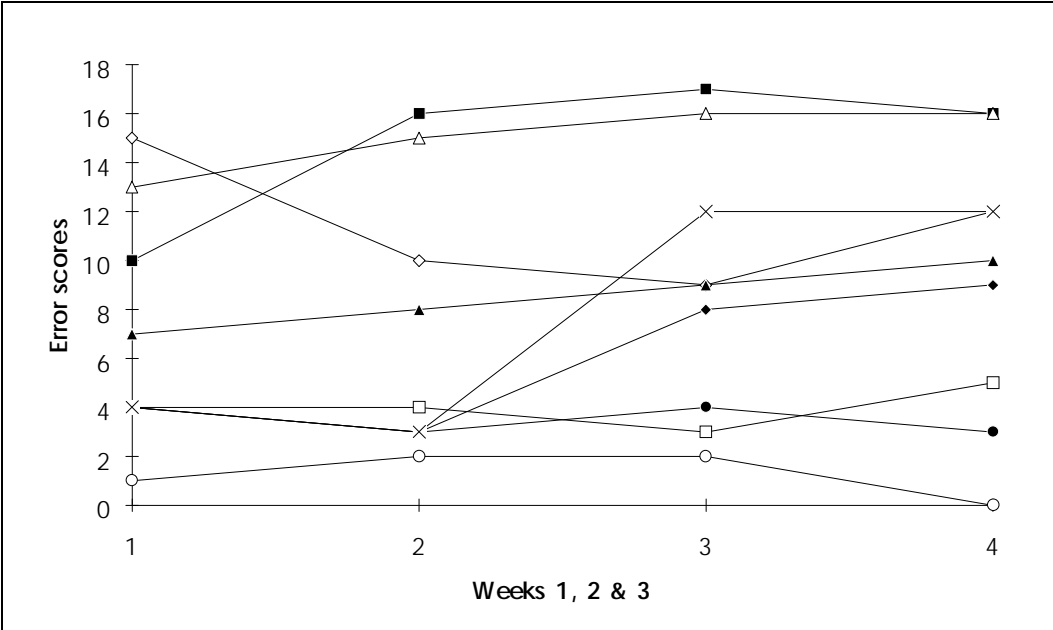


Figure 4d: Group 3 (GSM/CEPT), Raw error scores for each subject in each session

What can be stated from these four figures is that Groups 2 and 4 generally deteriorate slightly between weeks 1 and 2 and then seem to maintain a level for week 3. This is not the case for Groups 1 and 3. Group 1 (CEPT/GSM) subjects show a clear consistency for a positive improvement between weeks 1 and 2, and then deteriorated somewhat in their first test in week 3. However several subjects then improved their error rate in their second test (perhaps the first test acted as a reminder, or was it easier to remember one week ago rather than two?). Unfortunately Group 3 subjects showed no such general trend. Their data shows a complete mix, some got a similar score in all three weeks, some deteriorated and some improved.

From all these data it is evident that there is no common effect from learning both code schemes, either positive or negative, that is consistently transferred to the test in week 3. But there is evidence that subjects are behaving differently within their group and that each group contains two or more sub-groups of subjects who show similar behaviour.

5.5 What type of errors were subjects making?

For the detailed error analysis, all the data from all the subjects across all three weeks has been collapsed together. The only distinction that has been kept is to allow the comparison the CEPT and GSM schemes. The errors were scored not from the raw error data but from a re-scoring of the individual errors within the answer to any one question. For example a subject may have used the wrong prefix and the wrong service code, or may have neglected the PIN and the # terminator/suffix, within the same answer. For this analysis, both error types would be scored in each case.

5.5.1 Which questions caused problems?

Table 5 shows the number of errors logged against each question in the quiz, and therefore shows some error rates for each supplementary service. The analysis then computes the percentage error relative to all errors, and to all errors for that service, and then splits this latter one into errors from CEPT and from GSM tests.

The first point to note is the differences between the three services tested.

	% of total	mean errors per question
Call Forwarding	40,3 %	5,75
Call Barring	32,3 %	8,07
Call Waiting	27,3 %	3,9

Call Forwarding showed the highest percentage of errors, followed by Call Barring and then Call waiting. However, when the error rate is shared across the number of contributory questions the relative positions change: Call Barring shows an average of 8 errors per question against Call Forwarding 5,75 and Call Waiting 3,9. Call Barring obviously gave the subjects proportionally more problems.

Within the Call Forwarding service, the majority of errors occurred in two questions, "Erase" and "Register Only". The fewest errors were recorded for the simple commands "Activate" and "Deactivate". This difference was consistent also between CEPT and GSM. The greatest difference in percentage errors occurred where CEPT and GSM differed in their use of commands. For example, to erase CEPT requires the difficult construction #SC*SI#, whereas GSM uses the simpler ##SC#. The error rates were 16,4 % and 8,8 % respectively. Similarly, to register only CEPT again uses a difficult construction *SC*SI*SI#, but this time GSM requires two actions *SC*SI# to store and switch on the service and #SC# to switch it off again. Quite obviously with error rates of 11,9 % and 14,2 % the subjects were not comfortable with either solution.

Call Barring gave a high number of errors, basically because users kept forgetting one key part of the service, the PIN. Although subjects accepted the logical requirement for the PIN to both Activate and Deactivate Call Barring, they had significant problems, particularly when attempting to deactivate the service. However by far the highest number of errors (134) was logged for the question which required the most complex answer: To switch off Call Barring, To make a call, and then to reinstate Call Barring again. Many subjects assumed that suspending Call Barring on a per call basis would be possible. The command might be similar to that used in Call Forwarding, i.e. #SC*PIN*Tel_No#. Others may have thought more simply that once Call Barring is activated, it automatically reinstates even after a switch off command and a telephone call? As might be expected, there is no great difference between CEPT and GSM: 47,6 % and 52,4 % respectively.

Table 5: Number of errors per question

Call Forward - Unconditional					
	Errors	% of Total	% for Service	% CEPT	% GSM
1. Activate	35	3,3 %	8,3 %	4,3 %	4,0 %
2. Interrogate	60	5,7 %	14,2 %	7,8 %	6,4 %
3. De-activate	22	2,1 %	5,2 %.	3,8 %	1,4 %
4. Register and Activate	39	3,7 %	9,2 %	4,3 %	4,9 %
5. Erase	106	10,1 %	25,1 %	16,4 %	8,8 %
6. (Interrogate) Register and Activate	50	4,7 %	11,9 %	6,9 %	5,0 %
7. Register only	110	10,5 %	26,0 %	11,9 %	14,2 %
Totals	422	40,3 %	100 %	55,2 %	44,8 %

Call Barring					
	Errors	% of Total	% for Service	% CEPT	% GSM
8. Activate	60	5,7 %	17,8 %	9,2 %	8,5 %
9. De-activate	93	8,9 %	27,5 %	12,4 %	15,1 %
10. Interrogate	51	4,9 %	15,1 %.	8,2 %	6,8 %
11. Make a call De-activate, Call, Activate	134	12,8 %	39,6 %	17,8 %	21,9 %
Totals	338	32,3 %	100 %	47,6 %	52,4 %

Call Waiting					
	Errors	% of Total	% for Service	% CEPT	% GSM
12. Activate	38	3,6 %	13,3 %	7,0 %	6,3 %
13. Switch Order (R)2	58	5,5 %	20,3 %	11,5 %	8,7 %
14. Interrogate	52	5,0 %	18,2 %.	10,8 %	7,3 %
15. Switch Order (R)1	53	5,0 %	18,5 %	10,8 %	7,7 %
16. Switch Order (R)0 (or Ignore)	25	2,4 %	8,7 %	4,9 %	3,8 %
17. De-activate	30	2,9 %	10,5 %	5,6 %	4,9 %
18. Switch Order (R)2	30	2,9 %	10,5 %	7,3 %	3,1 %
Totals	286	27,3 %	100 %	58,0 %	42,0 %

Call Waiting had by far the fewest errors, only 3,9 per question. This is not surprising in this off-line situation, as it had the simplest command set. Activate, deactivate and interrogate did not require supplementary information and the switching orders are single digit commands. The true difference between CEPT and GSM was slightly exaggerated in this service, through the insistence for CEPT subjects to use the "R" (Register Recall) with the single digit for their switching orders. This explains some of the total error difference between CEPT at 58 % and GSM at 42 %, but does not explain the higher value for the interrogate command (10,8 % to 7,3 %).

One aspect which is worth noting is the degree of similarity between the error rates for the basic commands - Activate, Deactivate and Interrogate - across two of the three services tested. It again demonstrates the problem Call Barring has because of the need to include the supplementary information.

	Call Forward	Call Barring	Call Waiting
Activate	3,3	5,7	3,6
Deactivate	2,1	8,9	2,9
Interrogate	5,7	4,9	5,0

5.5.2 What type of errors?

Table 6 shows an analysis of the different error types, across all questions. The percentage error rates in this table have been calculated against the maximum possible number of errors of that type that could occur across all the completed quizzes. The table also compares the results for CEPT and GSM.

The table is presented in approximate order of relative error rates. For the CEPT data, the highest proportion of errors is for the questions requiring the special CEPT supplementary information codes. They are *0 for Register Only (e.g. *21*Tel_No*0#) and *1 for Erasure (e.g.#21*1#). This confirms the problems shown in table 4.

Table 6: Analysis of error types

Type of Error Analysis percentage of possible errors	CEPT		GSM	
	Error/Max.	Percentage	Errors/Max.	Percentage
Wrong or No CEPT Supp.Info.	74 / 136	54,4	0 / 0	0,0 %
Incomplete Sequence of Commands	15 / 68	22,0 %	50 / 136	36,8 %
Wrong or No PIN No.	72 / 272	26,5 %	87 / 272	32,0 %
Wrong Switch Order	73 / 272	26,4 %	61 / 272	22,4 %
Wrong Prefix	100 / 1 029	9,8 %	124 / 1 088	11,4 %
Wrong or No Suffix (Terminator)	61 / 1 020	5,9 %	74 / 1 088	6,8 %
No Answer offered	111 / 1 224	9,1 %	42 / 1 224	3,4 %
Adding PIN No.	11 / 748	1,5 %	30 / 748	4,0 %
Wrong or No Separator	18 / 544	3,3 %	9 / 612	1,4 %
CEPT/GSM Confusion	8 / 228	3,5 %	0 / 228	0,0 %
Wrong Service Code	15 / 1 020	1,5 %	3 / 1 088	0,3 %
Wrong Telephone No.	2 / 340	0,6 %	6 / 340	1,8 %
Wrong Syntax (Code sequence)	0 / 1 020	0,0 %	0 / 1 088	0,0 %

For both schemes the next three highest proportion of errors were very similar:

- Incomplete sequence of commands (mainly in Call Barring, bias in GSM caused by Register Only solution);
- Wrong or no PIN (only in Call Barring);
- Wrong switch order (only in Call Waiting).

The lowest proportion of errors were also very similar between CEPT and GSM:

- Wrong or no separator (possible in about half the questions);
- CEPT/GSM confusion (GSM answers to the Register Only and Erasure questions in a CEPT quiz);
- Wrong service code;
- Wrong telephone number;
- Wrong syntax.

Quite pleasingly, the data seems to confirm that people had no trouble in remembering the syntax of the command sequences. Also pleasing is the fact that very few people were making copying errors. Therefore it suggests that it is the learning of the abstract command codes (*, #, *#, etc.) and of the requirements for particular services (i.e. need for PINs, sequences of commands, etc.) which are difficult to learn.

Table 7 gives a detailed breakdown of the errors made in the use of the prefix, and gives some clue as to where confusions arose. The differences between CEPT and GSM are also included. Again the error rates are quoted as a proportion of the total possible.

Table 7: Analysis of Prefix Errors

Wrong Prefix Analysis percentage of possible errors	CEPT			GSM		
	Errors/Max.	%	Rank	Errors/Max.	%	Rank
Activate/Deactivate - Swap */#	31 / 680	4,5 %	4	17 / 748	2,3 %	4
Activate/Interrogate - Swap */*#	34 / 680	5,0 %	3	58 / 680	8,5 %	1
Interrogate/Deactivate - Swap *##	21 / 408	5,1 %	2	16 / 476	3,4 %	3
Interrogate - Swap */*#*	13 / 204	6,3 %	1	0 / 204	0,0 %	5
Erasure/Deactivate - Swap ###	1 / 340	0,3 %	5	33 / 408	8,1 %	2

The most striking aspect is the difference between CEPT and GSM for the interrogation command swap/confusion (i.e. #* instead of /*#). It occurs in 6,3 % of interrogate questions in CEPT yet not at all in GSM. No obvious explanation is available. On the other hand there is an explanation for the difference between CEPT and GSM for the Erasure/Deactivate confusion. The high GSM ranking is due to the fact that the ## Erasure command is solely a GSM command. The one example in CEPT for this confusion was in either Group 1 or 3.

The other three confusions Activate/Deactivate, Activate/Interrogate and Interrogate/Deactivate might all be seen to be similar even though they vary in error rate from 8,5 % to 2,3 %. For the most part these errors might be corrected quickly in practice if the terminal or network gives feedback on the command entered and its result.

5.6 What level of usability / learnability is achievable with the current code schemes?

This is perhaps the most difficult question to answer. The approach taken is to look at the data from three or four different angles.

5.6.1 Is there a link between a subject's Raven's grade and his error scores?

Figures 5 and 6 show scattergrams that plot all subjects error scores against their respective Raven's grades (figure 5 for CEPT and figure 6 for GSM). Each scattergram represents 68 data points (20 from Group 1, 18 from Group 3, and 30 from Group 2 or 4), and consequently the figures hide a considerable number of overlapping data points. The Raven's grades of I, II+, II, III+, etc. are represented by the rankings 1 to 8.

To test the idea that a person's capability with abstract constructs/problem solving is related to his ability to learn/remember supplementary service command codes, a Spearman's rank correlation coefficient, rho, was computed for each data set. The results show that the correlation was highly significant, for both CEPT and GSM.

CEPT rho = 0,695 *Significant > 0,00009*
 GSM rho = 0,410 *Significant > 0,0003*

These levels of significance suggest an almost causal relationship between the two parameters considered. The direction of the correlation was positive, which confirms that a better than average capability on the Raven's test will predict a better than average performance on the service codes, i.e. a low error score.

These data also suggest that only a minority of the population, say the 75th percentile and above on the Raven's test, will find the current supplementary service commands usable.

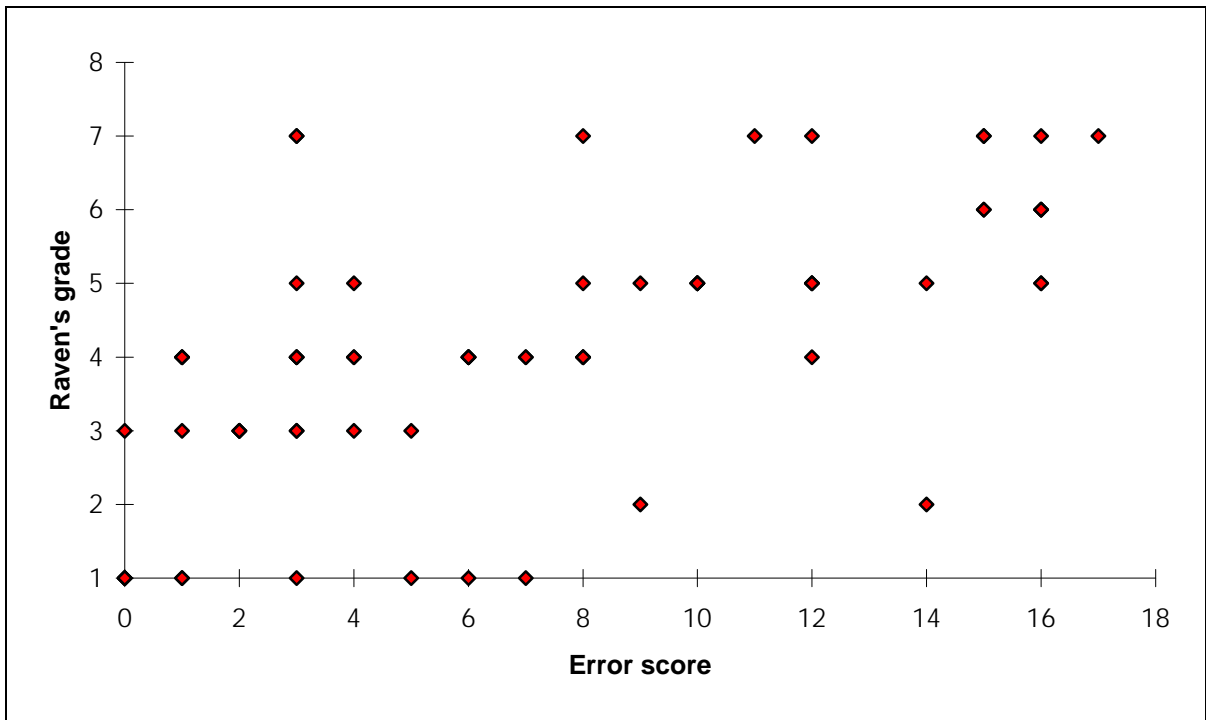


Figure 5: Correlation of Raven's grade vs. raw error scores, CEPT trials
 (Spearman's rho = 0,695, significant > 0,00009)

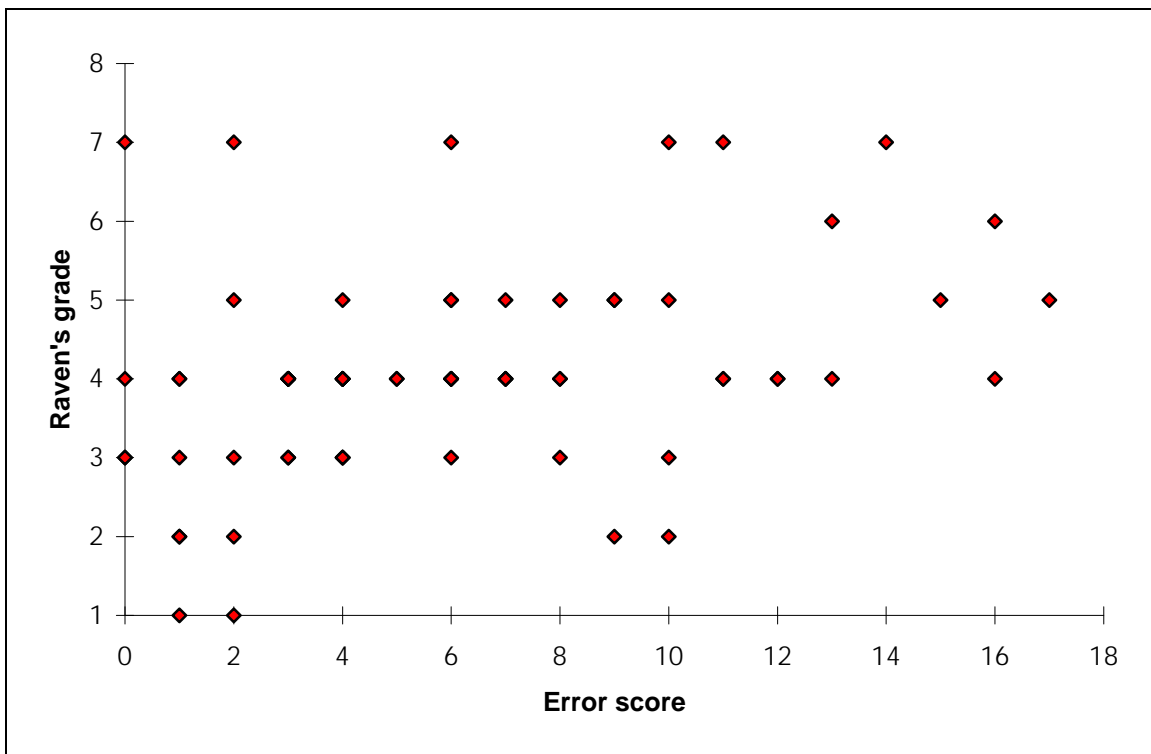


Figure 6: Correlation of Raven's grade vs. raw error scores, GSM trials
 (Spearman's rho = 0,410, significant 0,0003)

5.6.2 Are some subjects consistently better than others?

If the correlation evidence is true then it is reasonable to expect that some subjects may score consistently well or consistently poorly, irrespective of the experimental condition or session. To examine this, the data in figures 4 a-d can be re-drawn to compare the subjects separately. Figures 7 and 8 re-draw the data for figures 4c and 4d.

What these two figures demonstrate is that most subjects do tend to perform consistently across the three weeks of testing. For example, subjects 1, 2, 5, 7 and 8 in Group 1 all have error scores within four points over the four tests. Similarly, for subjects 2, 5, 6, 7, and 8 in Group 3.

This confirms that the service codes are only usable by some people and that this is irrespective of the amount or recency of training they have had.

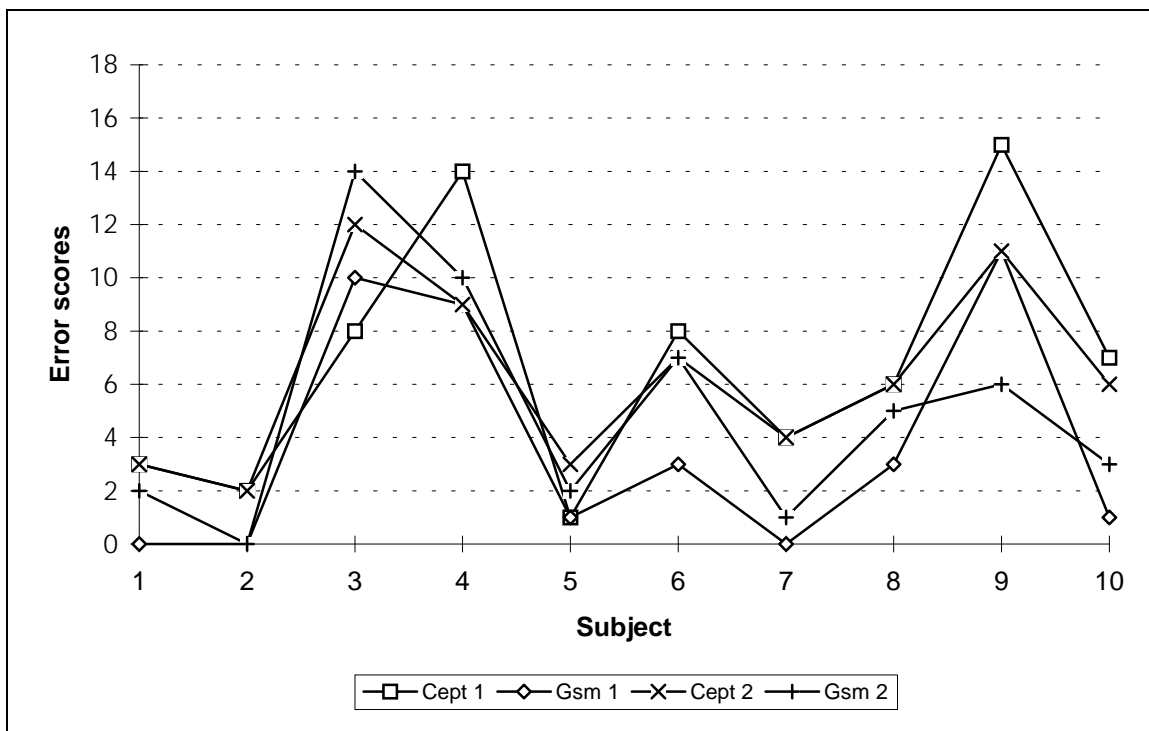


Figure 7: Group 1, comparison of error scores for each subject on four tests

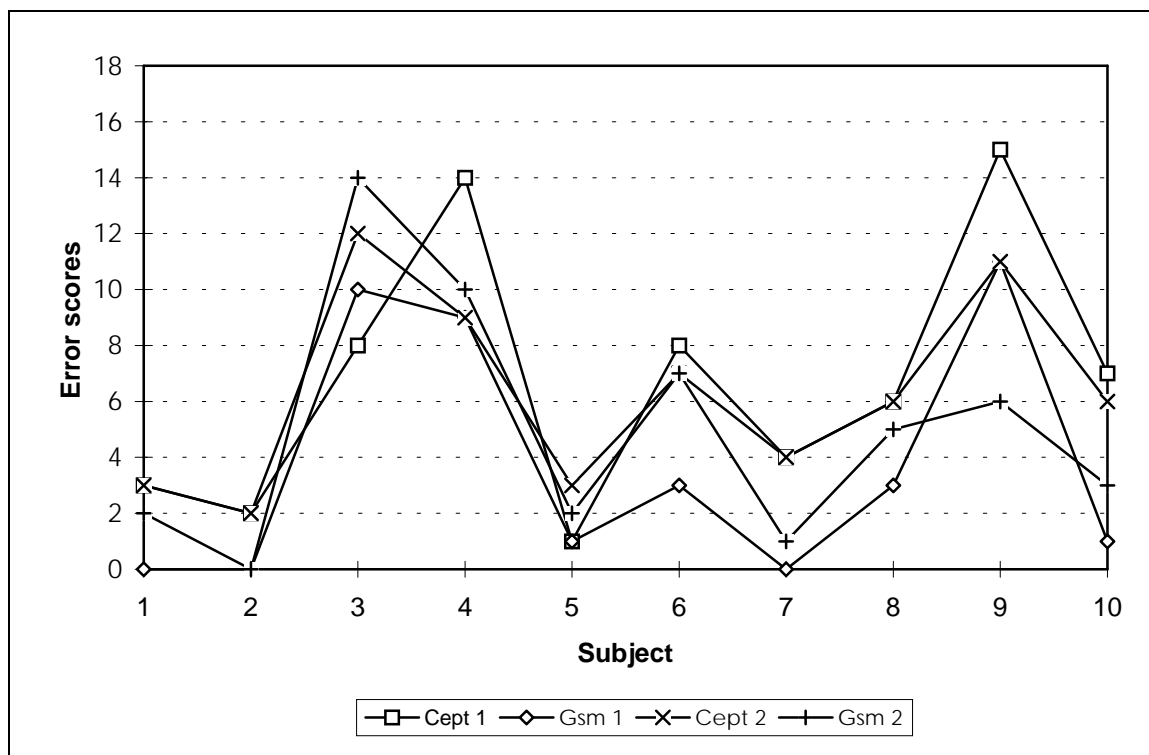


Figure 8: Group 3, Comparison of error scores for each subject on four tests

5.6.3 Are some questions easier to answer?

If only some people find the service codes usable, is it possible that some questions, from the quiz, were easier to answer? If so this might suggest that some command codes are more usable than others.

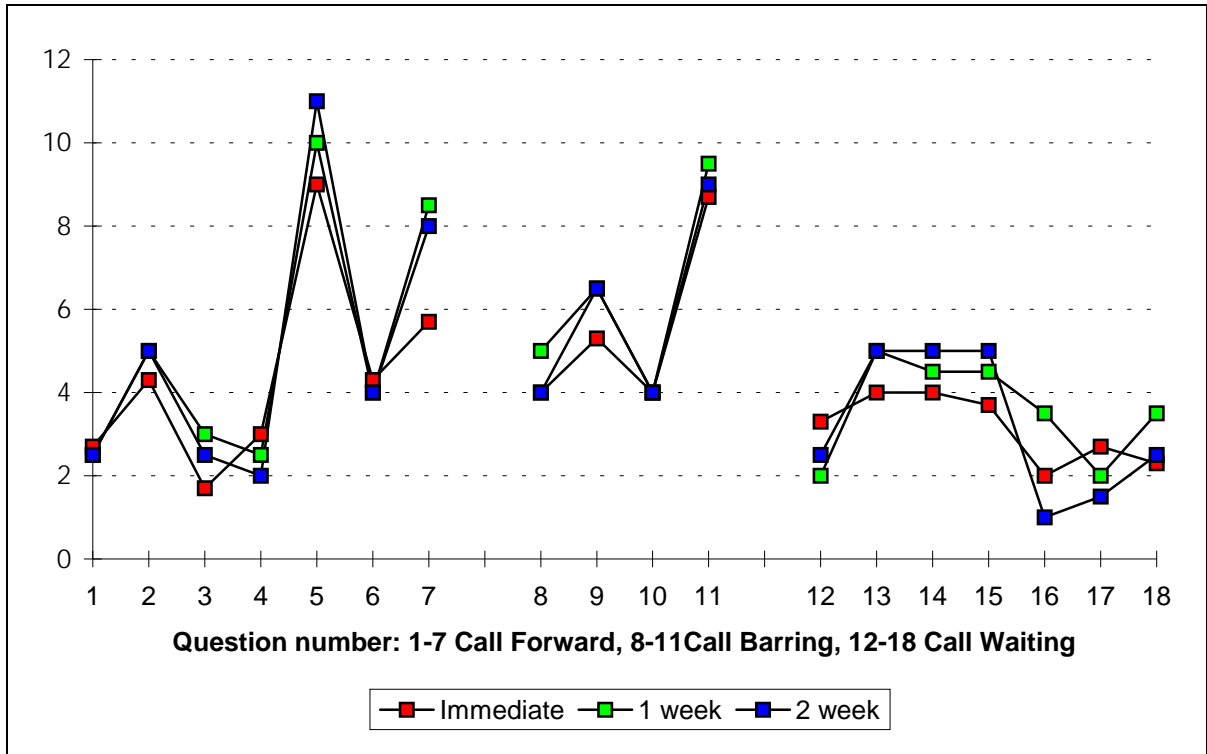


Figure 9: CEPT Mean errors for each question, immediately after learning, 1 week and 2 weeks later

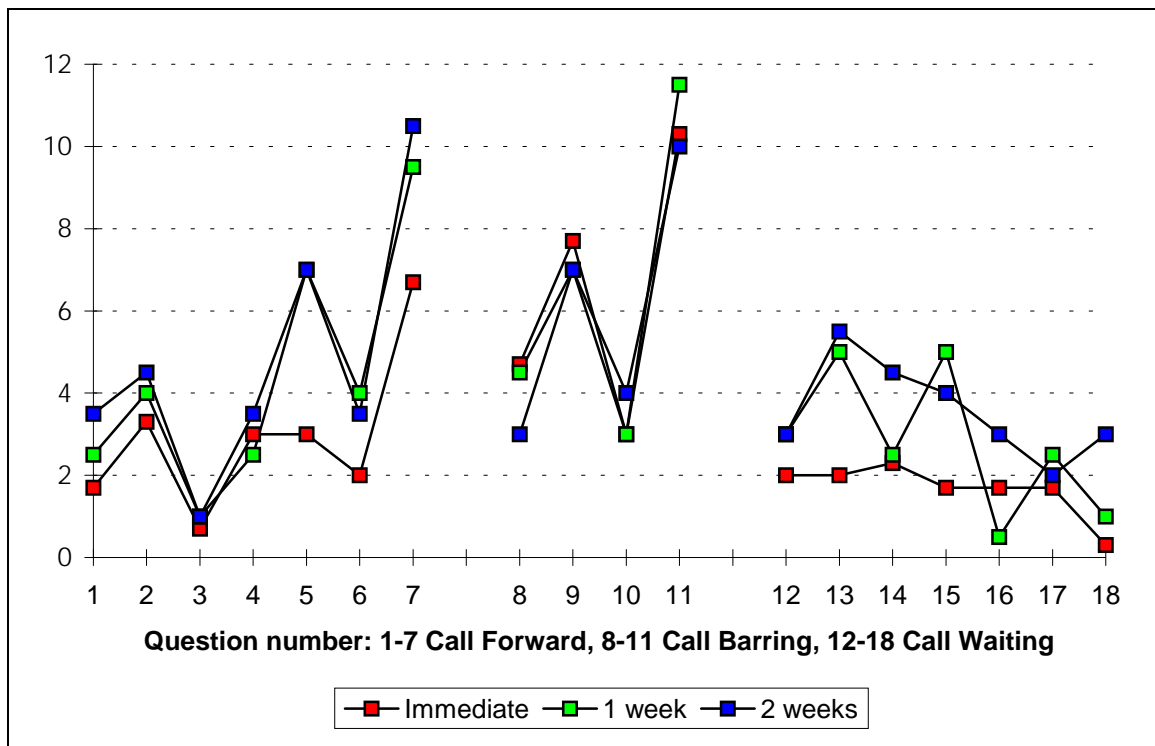


Figure 10: GSM mean errors for each question, immediately after learning, 1 week and 2 weeks later

Figures 9 and 10 show the respective means number of errors for each question from the CEPT and GSM quizzes. The figures show three curves. One for the quiz completed immediately after learning the commands; one for the quiz taken after a week; and one for the quiz taken after two weeks. The curves clearly demonstrate the consistency of errors between the three weeks of testing across the eighteen questions. It reinforces the data shown in table 5, and the problems subjects had with: Register Only, Erasure and Call Barring's PIN. Clearly some supplementary service commands are more usable than others.

Figures 9 and 10 also demonstrate that there is no significant change in the error rate for each question, for quizzes answered immediately after learning a code scheme, to that answered two weeks later. There is perhaps just about a general trend for the errors from immediately after to be slightly lower than those for later tests, but the difference is rarely more than two points. The exception might be question 5 in figure 10 (GSM), where probably the subject could remember the simple construction necessary to "Register Only" immediately after it was learnt, but one week later this unusual answer is forgotten

5.6.4 Does the opinion data give any indications of which aspects of the commands are more usable?

The data from the opinion questionnaires was analysed in terms of the spread of responses given for each topic and separated to highlight any differences between CEPT and GSM. Figures 11 to 15 show these data as simple histograms. The five topics are: the training booklet, the usefulness of the services, the complexity of the service command (syntax), the use of service codes (21, 33, etc.), and the use of symbol codes (*, #, *#, etc.). Subjects rated each statement against five response categories (1 to 5, strongly agree to strongly disagree). All five figure show that the subjects accepted the instructions and generally avoided the central neutral response category.

Figure 11 confirms that, for the most part, subjects claimed to feel comfortable with the training booklets: "They were easy to understand" and not "Too much to read". These statements scored the most positive response of the five, although a minority did think there was too much to read.

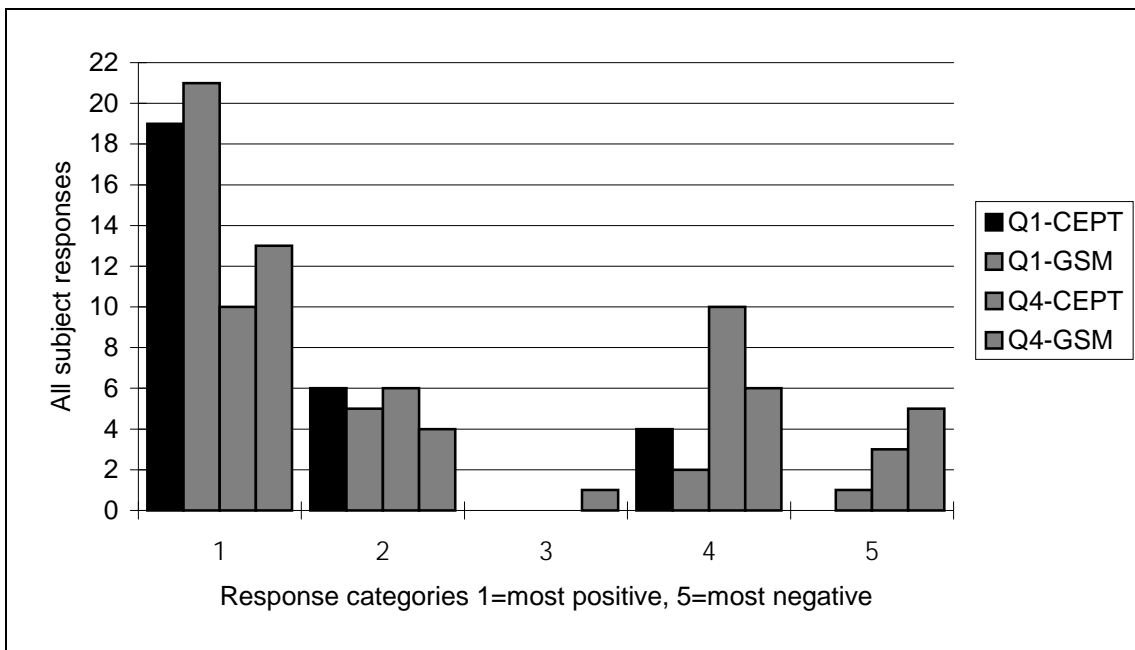


Figure 11: Questions 1 & 4 - Training Booklet (easy to understand vs. too much to read), spread of responses

Figure 12 shows that subjects gave a strong positive response to the usefulness of the supplementary services tested. Only a small minority, about 10 %, scored the services negatively. However, several subjects also commented that they thought they would be unlikely to use them within their everyday lives, i.e. they could foresee the usefulness for other people.

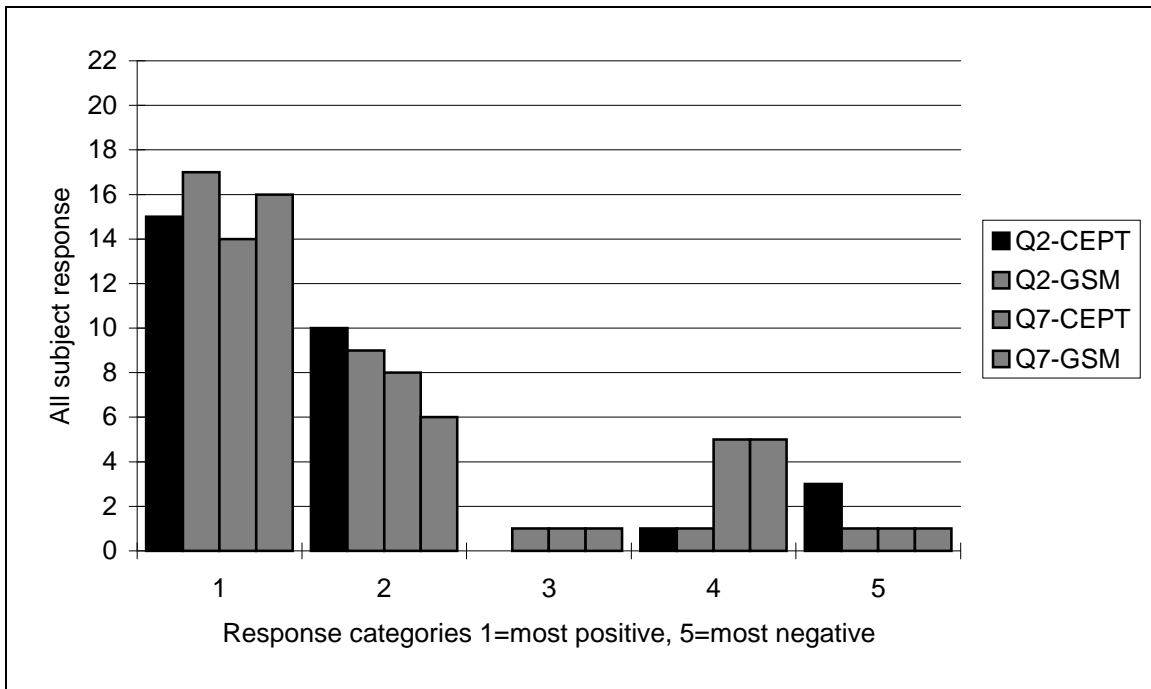


Figure 12: Questions 2 & 7 - Services (useful vs. difficult to understand use), spread of responses

Figure 13 shows a very strong bimodal spread of responses to the statements relating to the service commands: "They were too complicated" or "They were easy to learn". Although the mean response is still slightly positive, almost half the subjects gave a negative response. Confirming the fact, perhaps, that a significant proportion felt they had problems and found the commands to some degree unusable.

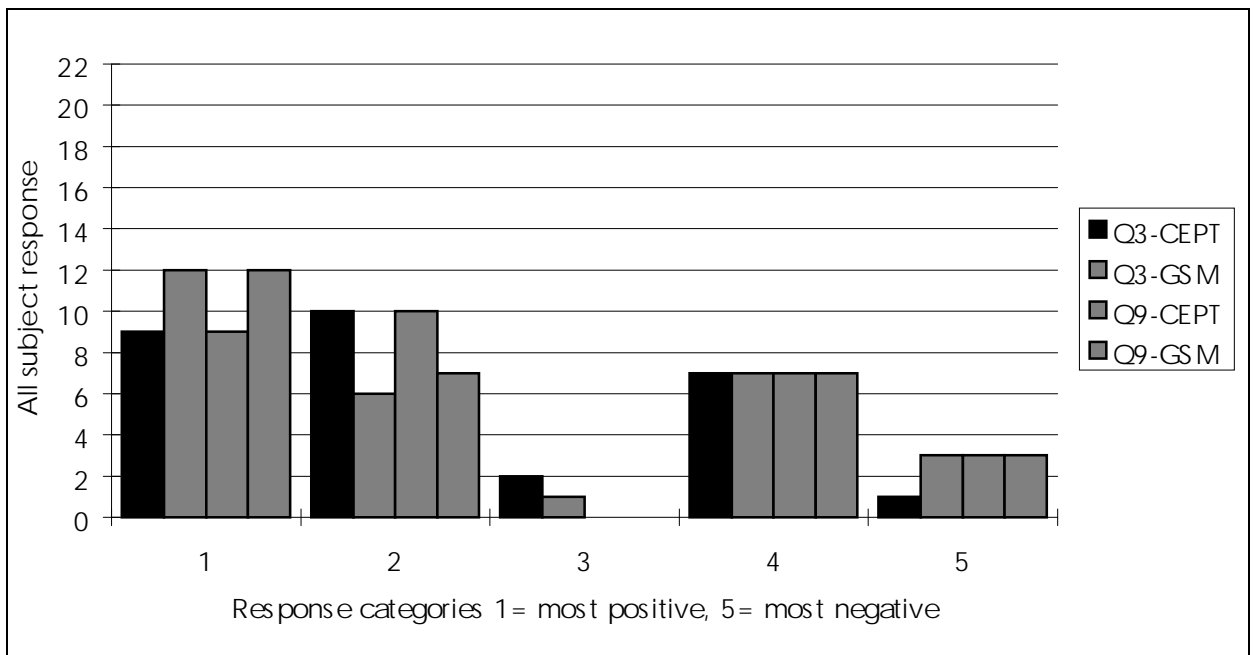


Figure 13: Questions 3 & 9 - Service Commands (too complicated vs. easy to learn), spread of responses

Figure 14 shows another very strong bimodal distribution. The statements this time related to the use of the service codes (21, 33, 43, etc.). Subjects were evenly split about whether these were, "Easy to remember" or "Confusing". This is interesting as the subjects were not asked to learn / remember the supplementary service codes. These were re-quoted on the front of the quiz paper. In general it seems that these data confirm the fact that people do not like or find it easy to learn / remember abstract numeric codes.

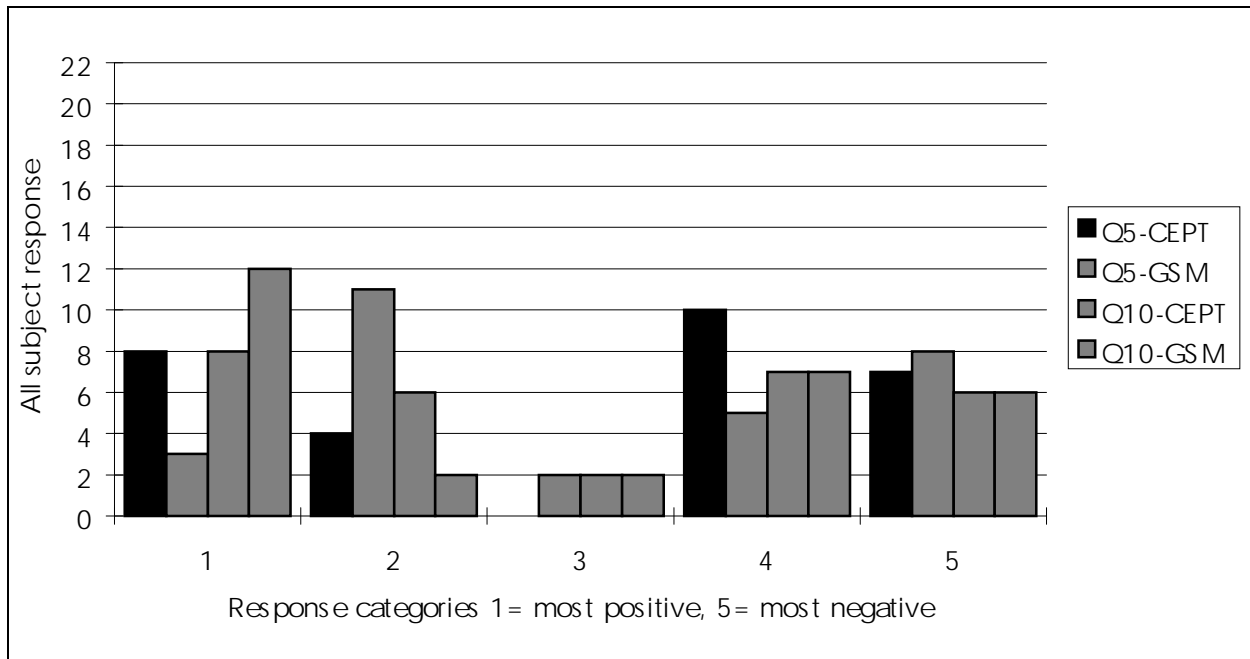


Figure 14: Questions 5 & 10 - Service Codes (easy to remember vs. confusing), spread of responses

Figure 15 shows the third significant bimodal distribution. The statements related to the ease of use of the symbol codes (*, #, *#, etc. sequences). Subjects were divided over whether these were "Easy or difficult to remember". Although the mean response towards these codes, about 28 % of responses were negative. This probably reflects the fact that some but not all the symbol codes were found to be unusable.

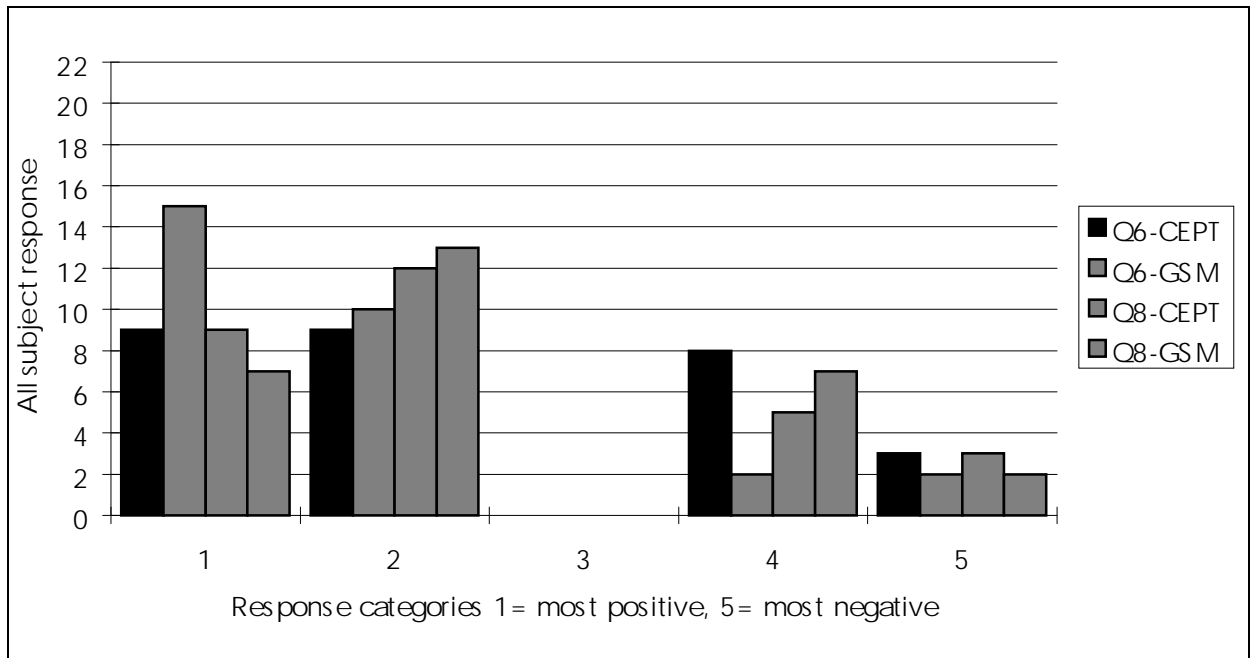


Figure 15: Questions 6 & 8 - Symbol Codes *, #, etc. (difficult to remember vs. easy to remember), spread of responses

These five figures (11-15) are able to underline some of the areas where subjects had problems. They show that it was probably some aspect of the service command structure, together with the use of abstract service codes which people disliked, more than the symbol codes or even the services themselves. This also reflects the error rate results where it was the more complicated command structures users had most difficulty with.

The five figures also confirm that subjects' opinions were similar, in respect of the negative to positive ratio of responses, between the two different code schemes. Where differences did occur, it was usually in the relative strength of the particular negative or positive response.

5.6.5 Is there a relationship between a subject's performance and his opinion ratings?

This final question is perhaps less about the usability of the supplementary services and more about the overall consistency of the data. It would be expected that subjects who expressed a negative opinion in the questionnaire would go on to perform less well in the quiz. Similarly, subjects who were positive in their opinion could be expected to perform better. Figures 16 and 17 show these two measures combined on the same graph for Groups, 1 and 3 (the other two were similar). They confirm that there is a relationship, but also show that some subjects had a more positive/negative attitude than their error rate would suggest.

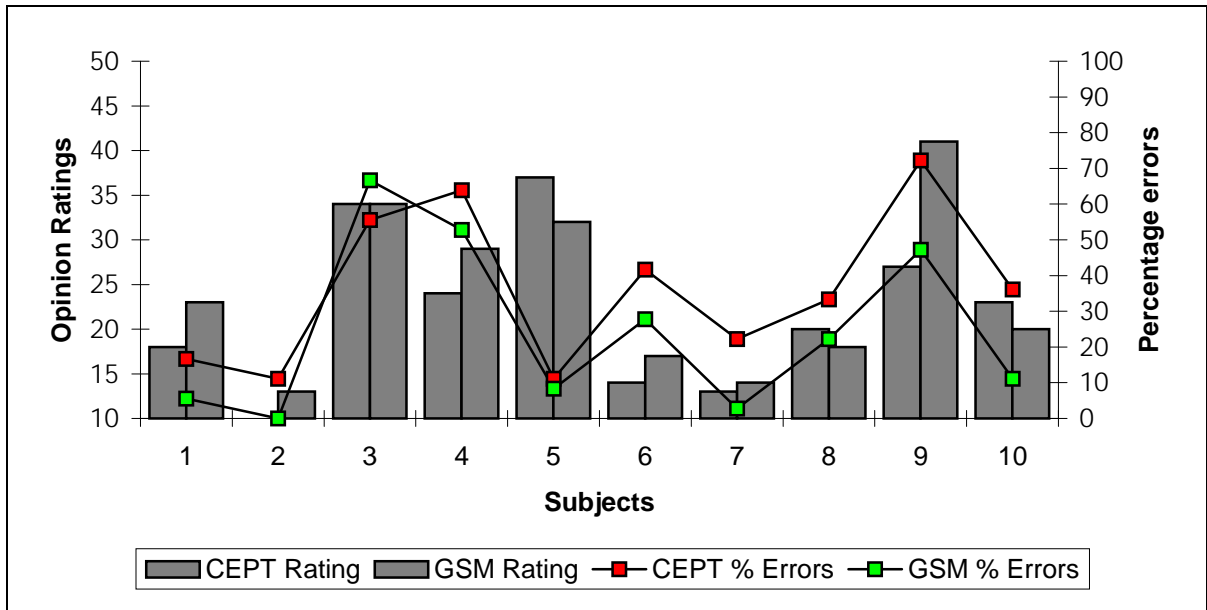


Figure 16: Group 1 (CEPT/GSM), mean percentage error scores and collective opinion ratings, CEPT vs. GSM for each subject

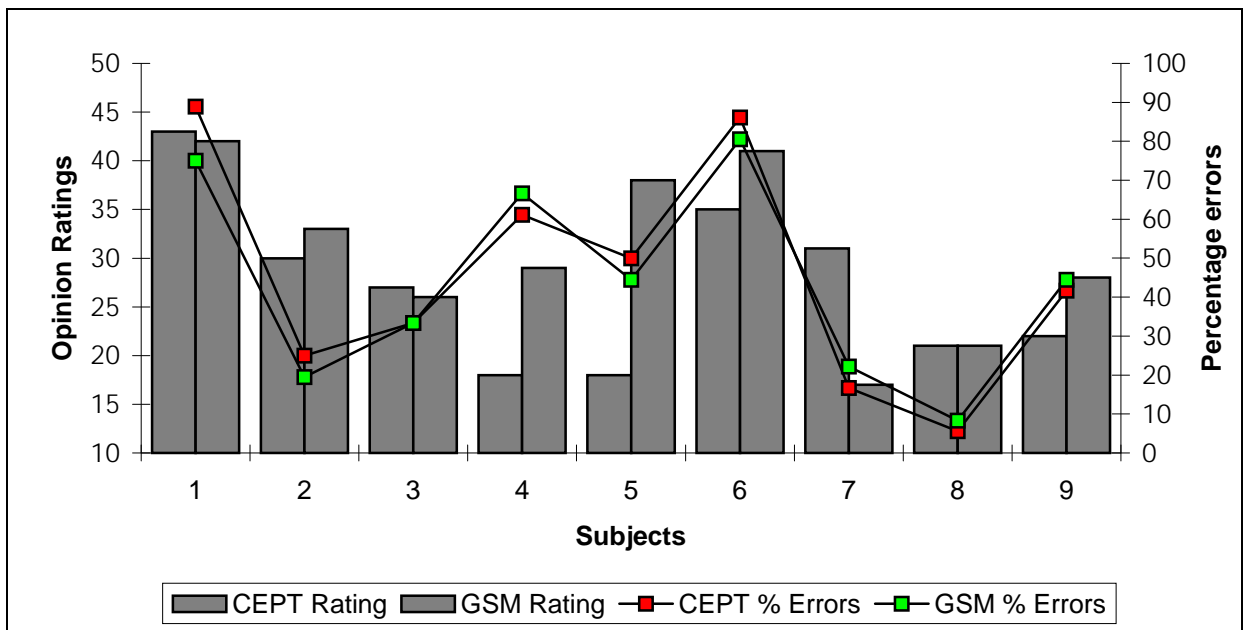


Figure 17: Group 3 (GSM/CEPT), mean percentage error scores and collective opinion ratings, GSM vs. CEPT for each subject

5.6.6 Answering the question

The answer to the question posed in subclause 5.6 "What level of usability is achievable?" may be summarized as:

- Those people who can do well with abstract problem solving can predictably perform better on supplementary service code learning tasks. But achieving a reasonable and consistent level of usability is certainly restricted to less than 50 % of the population, and may be restricted to the "top" 25 %;
- Some service commands within the code schemes tested consistently showed high error rates. Only the basic commands activate and deactivate had consistently lower error rates, and only then if no unexpected additional information had to be used, e.g. PINs. (Subjects performed better when adding six or ten digit telephone numbers to the activate command to invoke a Call Forwarding, than when adding a four digit PIN.)

6 Conclusions

6.1 General

The experiment fulfilled its objective and was successful in collecting a large amount of data, relatively quickly, from a subject population that can be regarded as typically a prime market for telecommunications services: the thirty to fifty year-old, well-educated professional.

The data confirmed that services based on these types of abstract command language are not highly usable, even by users of above average intelligence. There are significant proportions of the population sample which had difficulty learning and recalling many of the necessary commands. There are some commands that seem unusable to almost all subjects, even after some thirty minutes of dedicated training. What is more, this poor level of performance was maintained even if subjects re-learned many of the commands in a second training session.

6.2 Training materials

The training materials used and the training exercises given were considerably more detailed and intense than would normally be the case. It would be very unusual for anybody to receive 30 minutes concentrated learning on supplementary service codes in the "real world". Supplementary services are normally sold with only minimal end user documentation and support. For example, one country simply lists the services and access codes in the front of its directory, another advertises the services, but provides only a simple prompt card to tell the user which commands to use.

Where the "real world" users gain over the subjects in the experiment, is in the feedback and responsiveness of their telephone terminal or network. Within the experiment it was not possible to enable users to access a live terminal/network (the GSM mobile subjects would have potentially blocked the local "cell"). It would have also introduced a number of uncontrollable variables, i.e. the level and type of feedback given to any keying sequence (whether the phone was in the right state to send service code data, whether the service was in the right state to give the required feedback, etc.). Consequently it is accepted that the subjects' learning experience was not as rich as it would have been in reality. This limitation with the experiments' "feedback reality" should be balanced with the amount of time subjects dedicated to learning.

6.3 Test materials

The opinion questionnaire was based on a similar style satisfaction inventory which has been validated with a large subject pool. The specific statement set used, though, was not completely pre-tested for comprehension or consistency. There were only six subjects in the pilot study and no specific testing or examination of subjects' understanding of the questionnaire was executed. Consequently the data from this questionnaire should only be used indicatively. For example, there may well be problems in users' interpretation of the statements: did subjects appreciate the difference between the service commands, the service codes, and the symbol codes? did all subjects see these differences in the same way?

The test quiz questions were tested in a small pilot test and some changes made. However, it is possible that some subjects may have been confused by the wording of the question and gave the wrong answer not because they did not remember the correct one, but because they were answering a "different" question.

There is also the possibility that users were confused by the training booklet and learnt a command incorrectly. Each subject's training exercises were reviewed after the experiment to consider this point. In total ten subjects (four in Groups 1 and 3, and one each in Group 2 and 4) appeared to have made learning errors in one or other of the training booklets they used. Only one subject made a learning mistake in both booklets. In total there were a maximum of 30 of these training errors. This amounts to 3,3 % of the total number of errors made by all subjects, 906.

6.4 The results

The overall error rate for the experimental data, 37 %, shows that this subject pool had difficulties with this learning task. The individual error rate varied from 0 % to 89 %, which confirms the varying ability of people within a typical target market. The degree of difficulty any particular subject had depended on which command code he was asked to recall, and on his general ability with abstract construct problem solving. It was independent of which code scheme he learnt, how long he had to remember it, or whether he had learnt two code schemes.

Annex A: Sample Training Booklet for CEPT scheme



How to use the new services on your telephone

About this booklet

"Fictitious Company" are providing you with new services to help you make better use of your telephone.

This booklet combines instructions and exercises to help you learn to use the services provided. It is based on three typical services that will introduce you to the general rules that you can then apply to all the other services as well.

The three services the booklet is based on are:

- **call barring** stopping unauthorized people from using your telephone for outgoing calls, you can still receive incoming calls;
- **call waiting** when you're talking on your telephone and someone else tries to ring you, instead of hearing the engaged signal, the caller is held on the line while you're given an opportunity to speak to them;
- **call forwarding** calls to your telephone can be temporarily diverted to another number.

Please read this booklet very carefully and complete the practice exercises.

Call barring

What is call barring?

Call barring allows you to stop outgoing calls from your telephone. In this way you can stop unauthorized people from adding to your telephone bill. You can switch call barring on and off yourself whenever you want.

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.

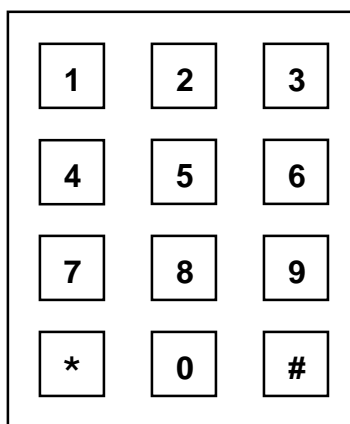
In this booklet we'll assume that your PIN is 2468

How do I tell the telephone to switch on call barring?

To bar 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.

To switch on call barring:

- 1 Lift the handset.
- 2 Type *33*2468#.
- 3 Replace the handset.



Why do I type *33*2468# ?

- * 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.
- 33 There are several services available, and typing 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 (2468). Otherwise the telephone would think it was all one long number (332468). In other words it is used to separate two numbers, rather like a space is used to separate words printed on a page.
- 2468 This is where you type your PIN to identify yourself as a person authorized to switch on call barring.
- # You type # to tell the telephone you've finished typing the command.

Exercise 1 Telling the telephone to bar outgoing calls

Now try the following exercise. You may look at the explanation and example above.

- A Write down what you would type to bar outgoing calls from your telephone, assuming your PIN is 2468:

.....

- B Using the telephone provided:
- 1 Lift the handset.
 - 2 Type the complete command you have written as answer to A above.
 - 3 Replace the handset.

How can I turn off call barring when I no longer need it?

To turn off call barring:

- 1 Lift the handset.
- 2 Type #33*2468#.
- 3 Replace the handset.

Why do I type #33*2468# ?

- # Starting with a # tells the telephone that you want to switch off a service.
- 33 Typing code 33 tells the telephone that your switch off instruction refers to call barring rather than a service with a different code number.
- * As before you use a * to separate the service code from the PIN.
- 2468 As before you type your PIN to identify yourself as a person authorized to switch off call barring. The same PIN is used to switch call barring on and off.
- # As before you type # to tell the telephone you've finished typing the command.

How can I make an outgoing call when call barring is on?

To make an outgoing call when call barring is switched on, you shall first switch call barring off, then make your call, then remember to switch call barring on again. The complete sequence would be:

- 1) Lift the handset;
- 2) Type #33*2468#;
- 3) Press the hook-switch, (to get a line again);
- 4) Dial number;

When conversation is completed:

- 5) Press the hook-switch again, (to terminate the call);
- 6) Type *33*2468#;
- 7) Replace the handset;

Exercise 2 Turning off call barring

Now try the following exercise. You may look at the explanation and example above.

- A Write down what you would type to stop barring outgoing calls from your telephone, assuming your PIN is 2468:

.....

- B Write down what you would type to make a call to 081 987 9877 if call barring is on, assuming your PIN is 2468, and that you wanted keep call barring on after the call. You can ignore the handset/hook-switch instructions:

.....

- C Using the telephone provided:
- 1) Lift the handset;
 - 2) Type the complete command you have written as answer to A above;
 - 3) Replace the handset.

How can I tell whether call barring is switched on or off?

There's a way of asking the telephone whether call barring is switched on or off.

- 1) Lift the handset;
- 2) Type `*#33#`;
- 3) You will be shown or told whether call barring is switched on or off. The information may be presented on the visual display and/or it may be spoken in a pre-recorded or synthetic voice;
- 4) Replace the handset.

Why do I type `*#33#` ?

- `*#` Starting with `*#` tells the telephone that you want information about a service. In this case the only relevant information is whether it is switched on or off;
- `33` Typing code `33` tells the telephone that your request for information refers to call barring rather than a service with a different code number.;
Notice that your PIN is not required for this command;
- `#` As before you type `#` to tell the telephone you've finished typing the command.

Exercise 3 Asking whether call barring is switched on or off

Now try the following exercise. You may look at the explanation and example above.

A Write down what you would type to find out if call barring is switched on:

.....

- B Using the telephone provided:
- 1) Lift the handset;
 - 2) Type the complete command you have written as answer to A above;
 - 3) Replace the handset.

Call waiting

What is call waiting

Call waiting gives you an opportunity to accept a call that arrives when you are already having a conversation on your telephone. Instead of hearing an engaged signal, new callers hear a message saying Trying to connect you. Please hold the line. At the same time you hear a brief tone over the top of your current conversation. You can then decide whether to talk to the new caller, and if so whether to put the existing call on hold until you have finished speaking to the new caller.

How do I tell the telephone to switch on call waiting?

Switching on and off the different services is similar, so that once you've learnt to use one, it is easier to learn another. Notice that for call waiting you do not need a PIN. To switch on call waiting:

- 1) Lift the handset;
- 2) Type *43#;
- 3) Replace the handset.

Why do I type *43# ?

- * As before, the * tells the telephone that you want to switch on one of its services.
- 43 Code 43 tells the telephone that you want call waiting rather than a service with a different code number.
- # You type # to tell the telephone you've finished typing the command.

Exercise 4 Telling the telephone to switch on call waiting

Now try the following exercise. You may look at the explanation and example above.

A Write down what you would type to switch on call waiting:

.....

B Using the telephone provided:

- 1) Lift the handset;
- 2) Type the complete command you have written as answer to A above;
- 3) Replace the handset.

How can I turn off call waiting when I no longer need it?

To turn off call waiting:

- 1) Lift the handset;
- 2) Type #43#;
- 3) Replace the handset.

Why do I type #43# ?

- # Starting with a # tells the telephone that you want to switch off a service.
- 43 Code 43 tells the telephone that your switch-off instruction refers to call waiting rather than a service with a different code number.
- # As always you type # to tell the telephone you've finished typing the command.

Exercise 5 Turning off call waiting

Now try the following exercise. You may look at the explanation and example above.

A Write down what you would type to switch off call waiting:

.....

B Using the telephone provided:

- 1) Lift the handset;
- 2) Type the complete command you have written as answer to A above;
- 3) Replace the handset.

How can I tell whether call waiting is switched on or off?

Just as with call barring, there's a way of asking whether call waiting is switched on or off.

- 1) Lift the handset;
- 2) Type *#43#;
- 3) You will be told or shown whether call waiting is switched on or off;
- 4) Replace the handset.

Why do I type *#43# ?

- *# Starting with *# tells the telephone that you want information about a service. As with call barring, the only relevant information in this case is whether it is switched on or off.
- 43 Code 43 tells the telephone that your request for information refers to call waiting rather than a service with a different code number.
- # As always, you type # to tell the telephone you've finished typing the command.

Exercise 6 Asking whether call waiting is switched on or off

Now try the following exercise. You may look at the explanation and example above.

- A Write down what you would type to find out if call waiting is switched on:

.....

- B Using the telephone provided:
- 1) Lift the handset
 - 2) Type the complete command you have written as answer to A above;
 - 3) Replace the handset.

What do I do when I hear a call-waiting tone?

When you hear a short beep or tone over your existing conversation, it means that a new caller is trying to get through to your telephone. The new caller is being held on the line to give you an opportunity to decide whether to interrupt your existing call to speak to the new caller. Your options are as follows:

Your action	The effect of your action
Ignore the tones and do nothing	(Reject new call) After a few moments, the call waiting tone stops and the new caller is told your number is currently engaged.
Press R button, then Type 0	(Reject new call) The call waiting tone stops immediately and the new caller is told your number is currently engaged.
Press R button, then Type 1	(Accept new call, finish existing call) Your existing call is ended and you are connected to the new caller.
Press R button, then Type 2	(Accept new call, hold existing call) You are connected to the new caller and the existing call is put on hold. You can then switch backwards and forwards between the two calls by typing 2 each time.

Note that when you hear the call waiting tone you can press the R button and then type 0, 1 or 2 without pressing the hook-switch or replacing the handset, and remember to tell your existing caller what you plan to do.

Exercise 7 What to do when a call is waiting

Now try the following exercise. You may look at the explanation and example above.

A Write down what you would type to continue your present call and reject the new one:

.....

B Write down what you would type to accept the new call and finish the existing one:

.....

C Write down what you would type to accept the new call and put the existing one on hold:

.....

- D Write down what you would type to switch backwards and forwards between the two calls:

.....

Call forwarding

What is call forwarding?

If you want to be sure that you do not miss calls to your telephone while you're away from it, you can redirect your calls to a telephone with another number. This service is also useful in other circumstances, such as when you do not 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 your calls to a different telephone number, say 898989, a number on your local exchange:

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

Why do I type *21*898989# ?

- * Pressing * tells the telephone that you want to switch on one of its services.
- 21 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).
- 898989 This is where you type the telephone number you want your calls forwarded to, including the STD or international codes, if necessary.
- # Typing # tells the telephone you've finished typing the command.

Entering telephone numbers as part of a command

You type the telephone numbers in the same way as if you were going to dial them from your telephone; in other words, you shall also include STD codes or international access and country codes as appropriate.

You shall not use the * key to separate parts of a telephone number, such as the STD code or the international code, from the rest of the number. For example:

Enter (0211) 678678 as 0211678678, or
Norway 34 30405 as 010473430405,
(where 010 is for international and 47 is for Norway)

Exercise 8 Telling the telephone to forward your calls

Now try the following exercise. You may look at the explanation and example above.

A Write down what you would type to forward your calls to these numbers:

Local number 765765:

STD number (0211) 321321:

International number 010 47 34 30405:

B Using the telephone provided:

- 1) Lift the handset;
- 2) Type one of the complete commands you have written as answer to A above;
- 3) Replace the handset.

How do I turn off call forwarding when I no longer need it?

When you get back to your telephone you will want to start answering your own calls again, you will need to turn off call forwarding.

To turn off call forwarding:

- 1) Lift the handset;
- 2) Type #21#;
- 3) Replace the handset.

Why do I type #21# ?

- # Starting with a # tells the telephone that you want to turn off a service.
- 21 Code 21 tells the telephone which service you want to turn off, in this case call forwarding.
- # Typing # tells the telephone you've finished the instruction.

You do not need to type any telephone number, as this instruction turns off call forwarding whatever number calls have previously been forwarded to.

Exercise 9 Turning off call forwarding

A Write down what you would type to stop forwarding your calls to these numbers:

STD number (0211) 321321:

International number 010 47 34 30405:

B Using the telephone provided:

- 1) Lift the handset;
- 2) Type one of the complete commands you have written as answer to A above;
- 3) Replace the handset.

How do I turn call forwarding on again?

Once you have used call forwarding but have turned it off, your telephone still remembers the number it forwarded calls to last time. This means that if you want to switch it on again, forwarding the calls to the same number as last time, all you have to type is *21# instead of the full *21*898989#. To turn on call forwarding again:

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

Notice that in the shortened command you leave out the telephone number and you do not need to use * as a separator because you only have one number, 21.

If you want to turn on call forwarding to a different number from last time, you shall type the whole sequence, including the new number preceded by a * to separate it from the service code 21. This new number will overwrite the previous one in the telephone's memory.

Exercise 10 Turning on call forwarding again

Now try the following exercise. You may look at the explanation and example above.

- A Write down what you would type to switch on forwarding your calls to STD number (0345) 30405 which is the same number as you used last time for call forwarding:

.....

- B Write down what you would type to switch on forwarding your calls to STD number (0211) 321321, which is not the same number as you used last time for call forwarding:

.....

- C Using the telephone provided:
- 1) Lift the handset;
 - 2) Type one of the complete commands you have written as answer to A or B above;

- 3) Replace the handset.

How can I wipe the telephone number from the memory when I turn off call forwarding?

There may be some occasions when you want to wipe out a telephone number stored as the call forwarding number in the telephone's memory. For example, you may want the number to remain confidential. To wipe out the stored telephone number:

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

Why do I type #21*1# ?

This is similar to switching off, but adds a 1 after the service code. You can think of it as switching off plus 1, a sort of double switching off, both the call forwarding service and the stored number.

Exercise 11 Wiping a number stored for use with call forwarding

Now try the following exercise. You may look at the explanation and example above.

A Write down what you would type to wipe a number stored for use with call forwarding:

.....

B Using the telephone provided:

- 1) Lift the handset;
- 2) Type the complete command you have written as answer to A above;
- 3) Replace the handset.

How can I enter a call-forwarding number without switching on call forwarding?

You may want to enter a call forwarding number so that it is ready for future use, without switching on the call forwarding service. For example you may want to enter it while you have it to hand, but not want to switch on call forwarding until some time later.

To store a number, in this case 898989, for future use with call forwarding.

- 1) Lift the handset,;
- 2) Type *21*898989*0#;
- 3) Replace the handset.

Why do I type *21*898989*0# ?

This is similar to switching on with a new number, but adds a 0 after the service code 21. You can think of it as switching on plus 0, or not quite switching on. You should also separate this 0 from the service code (21) and number (898989) with the usual *.

Exercise 12 Entering a call forwarding number for future use

Now try the following exercise. You may look at the explanation and example above.

A Write down what you would type to enter the telephone number 654321 for future use with call forwarding, without switching call forwarding on:

.....

B Using the telephone provided:

- 1) Lift the handset;
- 2) Type the complete command you have written as answer to A above,;
- 3) Replace the handset.

How can I tell whether call forwarding is switched on and what number, if any, is stored?

Just as with call barring and call waiting, you can ask the telephone whether call forwarding is switched on. Additionally the telephone will tell you what number, if any, is stored for call forwarding. You type:

*#21#

You will be told (or shown, if you have a display) whether call forwarding is switched on and any number that is stored for use with this service.

Exercise 13 Asking whether call forwarding is switched on and for any number stored for use with calling forwarding

Now try the following exercise. You may look at the explanation and example above.

A Write down what you would type to find out if call forwarding is switched on:

.....

B Write down what you would type to find out what number is stored, if any, for forwarding calls:

.....

C Using the telephone provided:

- 1) Lift the handset;
- 2) Type one of the complete commands you have written as answer to A or B above;
- 3) Replace the handset.

General rules

Now you have learned how to use call barring, call waiting and call forwarding. Use of other services is very similar, so you have also learned things you can apply to other services:

- 1) To switch on a service you always start with *, followed by the service code;
- 2) To switch off a service you always start with #, followed by the service code;
- 3) To ask for information about a service, such as whether it is switched on and what data (e.g. telephone number), if any, are stored with it, you always start with *#, followed by the service code;
- 4) If a command contains more than one number (e.g. service code number and telephone number) you always separate each number with a *;

This does not apply to the parts of a telephone number such as the STD code or the international code and the rest of the number; each whole telephone number should be typed just as you would dial it, i.e. with no * in the middle;

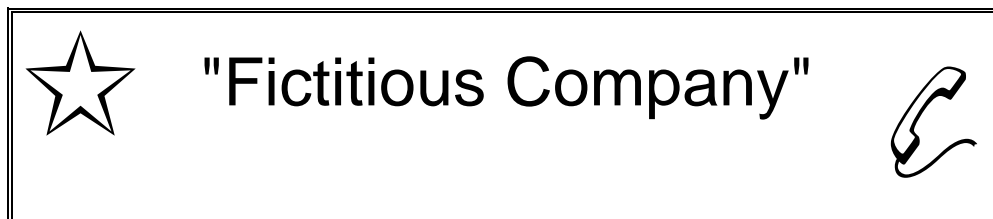
E.g. Enter (0211) 321321 as 0211321321

- 5) To enter data (e.g. a telephone number) for use with a particular service, but without switching the service on, you always start with a * and add a 0 as the last number before the # which ends the command;

E.g. In Call forwarding use: *21*0822898989*0#

- 6) To wipe out data (e.g. a telephone number) stored for use with a service, you always start with #, and add a 1 as the last number before the # which ends the command. If the service was switched on at the time this command will also switch it off;
- 7) Commands always finish with a #.

Thank you for reading these instructions and doing the exercises. You should now be able to get more out of the services on your telephone. We hope you find them useful.



Annex B: Opinion Questionnaire

Telephone Services Questionnaire

This is a short questionnaire asking you to rate how you felt about the service commands, you have just learnt. The questionnaire presents a number of statements of opinion. For each statement, please rate your degree of agreement or disagreement on the scale of 1 to 5.

- 1 Mostly agree with the statement.
- 2 Slightly agree with the statement.
- 3 Have no opinion either way.
- 4 Slightly disagree with the statement.
- 5 Mostly disagree with the statement.

Please do not use the central position "3", 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.

Name: **Group:**

Please consider each statement and circle the number you consider best matches your opinion.

- 1 The training booklet was easy to understand. 1 2 3 4 5
- 2 All of the services introduced could be very useful. 1 2 3 4 5
- 3 The service commands were more complicated than they needed to be. 1 2 3 4 5
- 4 There was too much to read when learning the service commands. 1 2 3 4 5
- 5 It is easy to remember which number codes (21, 33, 43, etc.) goes with which service. 1 2 3 4 5

- | | | |
|----|--|-----------|
| 6 | It is difficult to remember which symbol code (*, #, *#, etc.) is used to switch services on or off. | 1 2 3 4 5 |
| 7 | It was difficult to understand how the services should be used. | 1 2 3 4 5 |
| 8 | It is easy to remember which symbol (*, #, *#, etc.) is used at each point in a command sequence. | 1 2 3 4 5 |
| 9 | It was easy to learn the commands needed to use a service. | 1 2 3 4 5 |
| 10 | The assignment of number codes to services was confusing. | 1 2 3 4 5 |

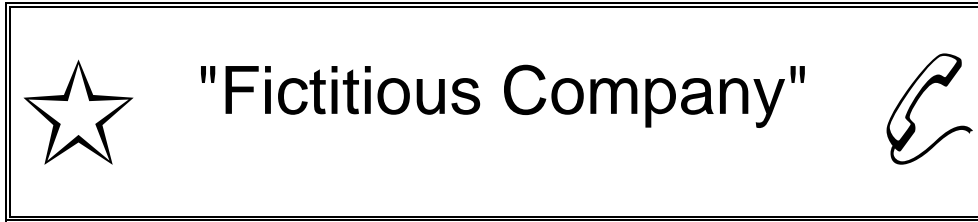
Any other comments:

.....

.....

.....

Annex C: Sample Quiz for CEPT Code Scheme



Quiz

Name:..... Education: CSE's/O-Levels
A-Levels
Group:..... Degree

To find out how helpful the booklet was, we should now like you to answer a quiz on what you may have learnt from it. Please try to answer each question correctly, but do not worry if you really cannot answer a question – it may mean that our booklet was just not good enough, or the codes were too confusing. You will only be asked about services you have read about and practised in the booklet.

To help you, here are the code numbers for the three services:

Call barring 33
Call waiting 43
Call forwarding 21, and
Your PIN No. is 2468.

Without looking at any of the examples or exercises, try to answer the following questions. Please write down the whole command sequence you need to type in each case. Don't worry about the lifting and replacing of the handset just ignore these. We are only interested in the *, #, and number command sequences.

- 1) How do you switch on call forwarding to number 566788, if this is the previous number used for call forwarding on your telephone?

.....

2) How do you ask the telephone what number is stored for use with call forwarding?

.....

3) How do you switch off call forwarding?

.....

4) How do you switch on call forwarding to number 0455 456456, if the last number it was set to was 0697 566788?

.....

5) How do you wipe out the number stored for call forwarding?

.....

- 6) How do you switch on call forwarding to number 010 49 69 566788, if you cannot remember the last number it was set to?
.....
- 7) How do you store the number 0505 678678 for later use on call forwarding?
.....
- 8) How do you bar outgoing calls?
.....
- 9) How do you stop barring outgoing calls?
.....
- 10) How do you check whether call barring is on or off?
.....
- 11) How do you make a call to 0211 898989, when call barring is on, and you want it left on?
.....
- 12) How do you switch on call waiting
.....
- 13) When you get a call waiting tone, how do you accept the new call and put the existing one on hold?
.....
- 14) How do you check if call waiting is on or off?
.....
- 15) When you get a call waiting tone, how do you accept the new call and finish the existing one?
.....
- 16) When you get a call waiting tone, how do you reject the new call and continue with your existing one?
.....
- 17) How do you switch call waiting off?
.....
- 18) When you have two calls, one on hold and one you are talking to, how do you switch between calls?
.....

Annex D: General Instructions

General Instructions

- 1 Welcome to everyone, on behalf of the fictitious company (e.g. "Star Telephones Ltd." or "Mobile Communications Plc.").
- 2 Introduce the experimenter and colleague. "We will be conducting the experiment over the next three weeks."
- 3 "On the desk in front of you should be a telephone (mobile phone), a pen and a booklet. Can I ask you at this stage not to touch anything until I ask you to start the experiment. First, I would like to tell you a little bit about what we are trying to achieve."

"You will be undertaking a learning exercise. We are attempting to test the effectiveness of the learning material, i.e. the test that you will undertake tonight and in the next two sessions are a measure of the material and not of you, the subjects. Although names appear on the papers, this is for our use only. All the papers will be treated in the strictest confidence, and will be destroyed once the results have been compiled."

"It is vital that you complete the three sessions. By way of a thank you, I will be handing out a form at the end of the evening, on which we would like you to indicate what wine you would like. Please make sure that you take one."
- 4 "Now just before you begin working through the booklet, I would like to point out that there are is no time limit on the exercise. Please work through at your own pace, and make as much use of the telephone (mobile) as you feel necessary. Work carefully through each of the exercises. For the purposes of the evening the important buttons to be aware of are *, #, R ("Send" and "End"). After everyone has completed the booklet there will be a short questionnaire, and then a test on the material you have learnt. Now please begin."
- 5 Before testing:
"Please can you note that the service codes appear on the front of the test as well as the PIN."
"Please read the questions carefully."

History

Document history	
October 1996	First Edition