CDI Input Devices: user test of prototypes and models

Ian McClelland
Bruce Thomas
Applied Ergonomics Group

July 1992

Philips Corporate Industrial Design

CDI Input Devices: user test of prototypes and models

Ian McClelland
Bruce Thomas

July 1992

Report No
3 CID AE 92-10

Philips
Corporate Industrial Design

CONTENTS

SUMMARY

A new CD-I remote control is to be developed as a replacement of the current CD-I remote control. A major goal is to improve on a number of perceived deficiencies in relation to the usability of the current remote control. As part of the development programme CID Applied Ergonomics carried out a previous usability test.

Following this study three prototype CD-I compatible input devices became availableand a number of design models were developed based on the results from the previous test by CID Design aimed at improving the handling aspects of the remote control.

Ten subjects evaluated the devices and models in a usability test. The tasks and method followed the same pattern of the previous test using the title Cool Oldies Juke Box.

The following recommendations are made for the design directions for the next generation CD-I general purpose input devices.

Handling the device

Fit to the hand
- anthropometric requirements need to be defined in order to encompass 95% of expected user population. Special consideration needs to be given to hand sizes of children. Minimum age groups need to be specified. Specific data requirements will depend on the particular device configurations proposed.

Physical relation between the primary control element and the action buttons
- position both on the top surface.
- place buttons adjacent to primary control element to enable easy coordinated operation,

Weight
- guidelines on weight and location of centre of gravity for device relative to grip position need to be determined.

Tactile quality
- smooth non-slip mat finishes especially for surfaces of control elements.

Grip and control
- user must be able to operate the primary control element and operate action buttons without significant change in grip,
- user must be able to operate the device with one hand, left or right
- do not compel the user to use two hands

Convenience
- user must be able to hold and carry the device in one hand,
- the device should be wireless

Controlling the cursor

Ease of control

- user must be able to position the cursor quickly and accurately
- enable a smooth control over change of cursor speed from zero to maximum speed at users discretion without perceptable thresholds
- the user should not perceive an absolute limit to maximum cursor speed that is determined by the system, eg Control Display ratio of the Yellow TBall, absolute speed limit imposed by Current RC.

Responsiveness of primary control element

- ensure direct response to user actions,
- ensure no perceptable delays in response to user actions,
- a balance is required between sensitivity (ratio of pixels to physical movement of device) and movement resistance in device mechanism. Specification method and recommendations need development.

General aesthetic quality of design

Style

- complement existing AV consumer electronics products
- colours, materials, and forms should complement each other.

Integrated design

- action buttons and CD player controls should not be 'add ons'

Shape

- provide rounded, smooth, slim shape.

Robustness and quality of fit

- provide 'solid feel' to device.
- housing parts should fit tightly together, no creaks or squeaks, no visible splits.
- primary control element and button housings should not feel loose in operation,

Usability measures

Subjective performance

The precise dimensions and the form of presentation used for collecting subjective data require review.

Objective performance

The objective data now available as a result of this test and the previous test require review in order to improve the accuracy of the data collected and enhance the long term benefit of usability testing of input devices.

Development of Usability test procedure

The tasks designed for usability tests must be reviewed. The cursor control tasks should reflect the range of tasks users of CD-I are likely to encounter when using CD-I with the general purpose device. The most notable cursor control tasks not covered in either this test or the previous test were system paced tracking tasks which require the user to respond rapidly.

The test procedure so far developed has proved to be successful to a point. However to improve the validity and value of the test procedure consideration must be given, amongst other aspects, to:

- how much experience should subjects have in the use of the input devices under test before reliable data can be measured?
- how many subjects are required in order to obtain a reasonable level of statistical reliability?
- what range of cursor control tasks are the most effective at discriminating between devices?
- can a set of 'Title independant' tasks be designed that provide a more objective basis for evaluating devices in the future?
- how should movement types and target sizes be best classified?

Acknowledgements

Thanks are due to a number of people who participated in running the test:

Govert de Vries, CID Applied Ergonomics group
Joyce Westerink, Instituut Perceptie Onderzoek
Karoliene vd Reek, Instituut Perceptie Onderzoek

As always thanks are due to the users who participated in the test.

1.0 INTRODUCTION

A new CD-I remote control is to be developed as a replacement of the current CD-I remote control. A major goal is to improve on a number of perceived deficiencies in relation to the usability of the current remote control.

As part of the development programme CID Applied Ergonomics carried out a previous usability test; CD-I Remote Control User test, Thomas and McClelland, April 1992, Report Number 92-09a. This study compared the following input devices:

- the current remote control
- the old joystick control used for the CD-I prototypes
- the Philips mouse
- the CD-I track ball

The previous test was designed as a first stage exploratory study which addressed three basic questions:

- is there a problem?
- what level of difficulty or ease is associated with using the current device?
- what level of convenience is associated with handling the current device?

The main results of the study showed that:

- the current remote control was perceived as being difficult in terms of the speed and accuracy with which the cursor could be positioned on a selected target.
- objective measures of time to target showed both joystick controls to be significantly slower than the track ball,
- the housing of the current remote control was generally liked, but that improvements could be made in terms of fit to hand, design and location of action buttons, appearance, and alignment of infra red beam with the player.

Therefore the main conclusion was that a 'problem' does exist.

Following this study three prototype CD-I compatible input devices became available. These input devices were two track balls and a pressure pad. In addition a number of design models were developed based on the results from the previous test by CID Design aimed at improving the handling aspects of the remote control. These models incorporated proposals based on the three technologies currently regarded as feasible options for use with CD-I; the joystick, the track ball, and the pressure pad. CID therefore took the initiative to run a second user test based on the previous test. This test had the following goals:

- to extend knowledge of the important usability characteristics for CD-I control devices,
- compare possible control technologies with the current RC and the current Track Ball,
- to provide early advice on shape/form studies for next generation devices,

- to ensure improvement over the current RC is achieved.

There were a number of significant constraints which restricted the scope of what the test could achieve.

The prototype devices

The prototype devices were only available in radically different different housings. They could not be dismantled and built into a common housing, eg the current remote control housing. The devices could not be modified in order to match their performance chartacteristics; eg Control Display ratios. The devices had to be tested as given.

The design models

The models were only shape/form dummies with no moving or operational parts.

Consequently the prototype devices could only be tested and compared as independent devices. In other words the test reported here cannot be used as a basis for judging one technology against another. We can only compare one device against the other. In relation to the models we know from the previous test, as well as similar tests carried out elsewhere, that the handling characteristics of a device in relation to controling the cursor is highly interdependant with the dynamic aspects of using one control technology or another. Therefore the result of the study was intended to:

- identify usability measures for CD-I input devices
- identify desirable and undesirable design attributes

2.0 METHOD

2.1 Test Situation

The usability test carried out in the CID Applied Ergonomics Usability Laboratory. Figure 1 illustrates the test situation used in the previous test. In test reported here camera three was not used. In all other respects, apart from minor details in terms of furniture layout, the setup was essentially the same.

Subjects were seated on a sofa facing the TV at a viewing distance of approximately 4m. Two CD-I players were used together with the TV monitor. One player was positioned to the right of the TV for use with the Infra-Red remote control. The other devices were wired devices and were supported by the second player positioned close to the subject.

The same views were recorded on video as in the previous test with the exception of camera three. Figure 2 shows the mixed view of a subject recorded during one of the test sessions. In addition a direct recording was taken of the TV image. The objective data was recorded from the video recordings.

2.1.1 Test subjects

Ten people took part in the test. All were drawn from the CID and IPO subject pools. None of the subjects had participated in the previous test. The profile setup for subject selection was:

- adults within the age range 18-45yrs
- no previous experience of CD-I
- users of CD audio
- user of common audio-video products
- not professional users of computers; SW engineers, programmers etc. Users of common software applications such as wordprocessing, spreadsheets etc. were accepted.

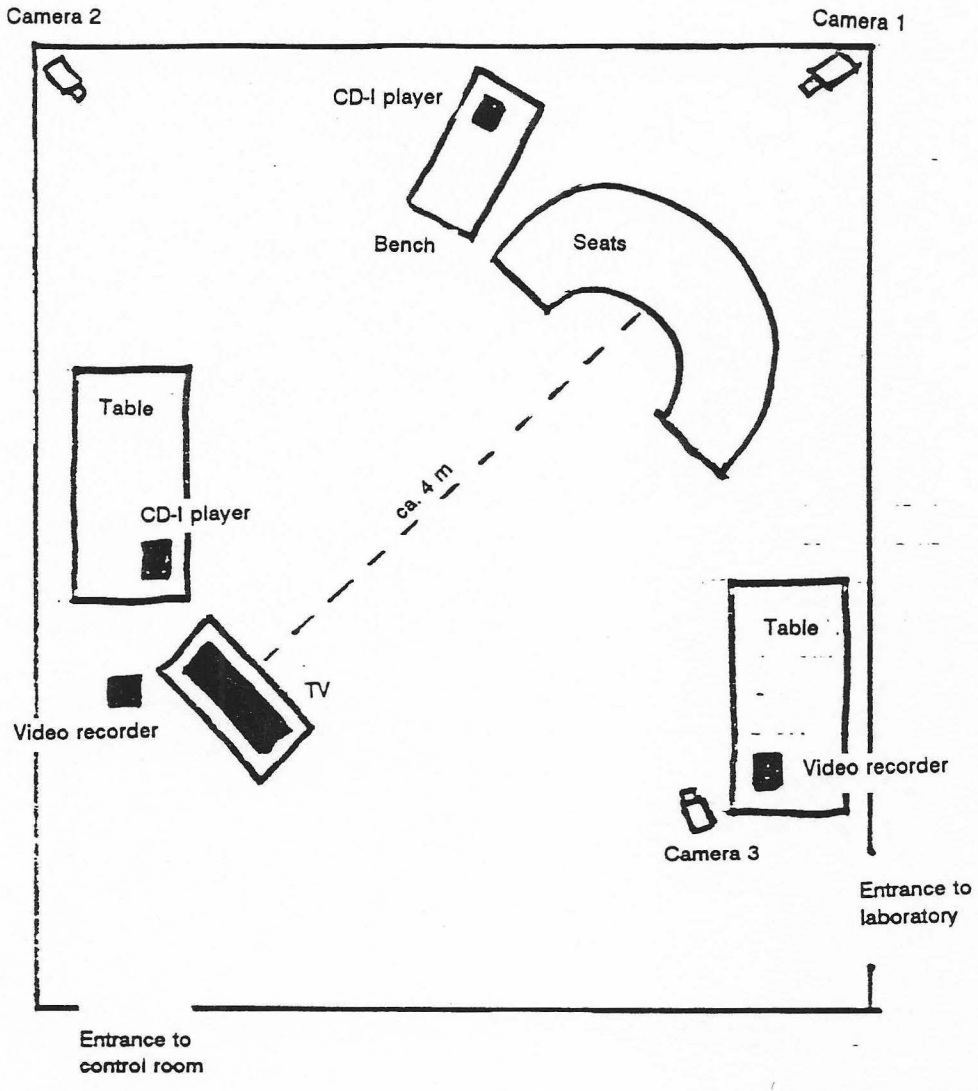


Figure 1 - The test situation used in the previous test.

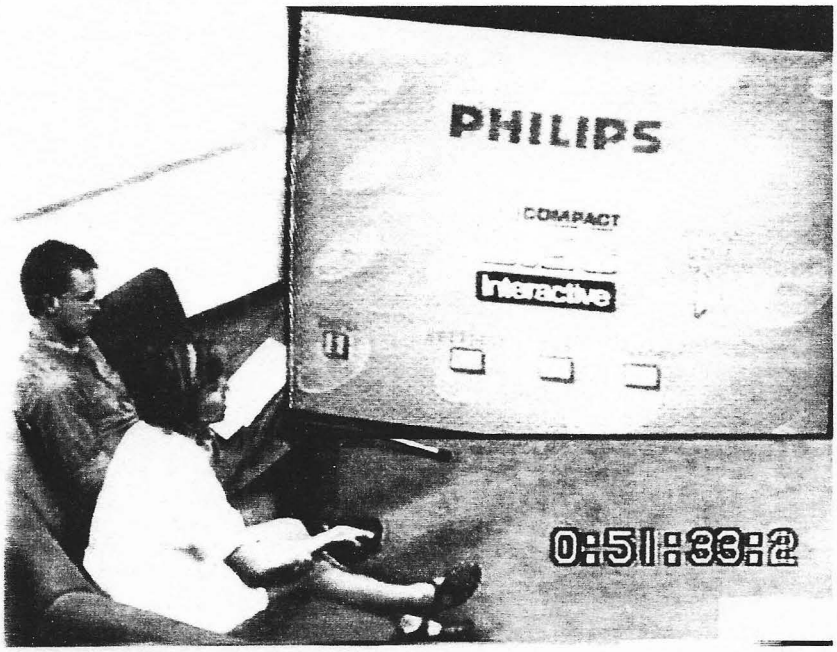


Figure 2 - The mixed view recorded for a subject during one of the test sessions.

2.1.2 The input devices

Five working input devices evaluated in usability trial. In this report the devices are referred to as follows:

- The Practice TBall Figure 3
- The Current RC Figure 4
- The Large TBall Figure 5
- TheYellow TBall Figure 6
- The Pressure Pad Figure 7

The first device, the Practice TBall, was designated as such because ther were known 'bugs' in the software which caused certain unreliabilities. Therefore this device was used to introduce subjects to the concept of CD-I and to the procedure to be followed during the rest of the test. The objective data collected from this device was not used in the results. The subjective data for this device is included for general interest.

2.1.3 The design models

The subjects evaluated eight models by inspection in discussion with the interviewer. The eight models were annotated as follows: FK, FG, RCB, RFB, FT, FP, RCP, RFP. The models are illustrated in the discussion, Section 5.4.

Figure 3 - The Practice TBall

Figure 4 - The Current RC

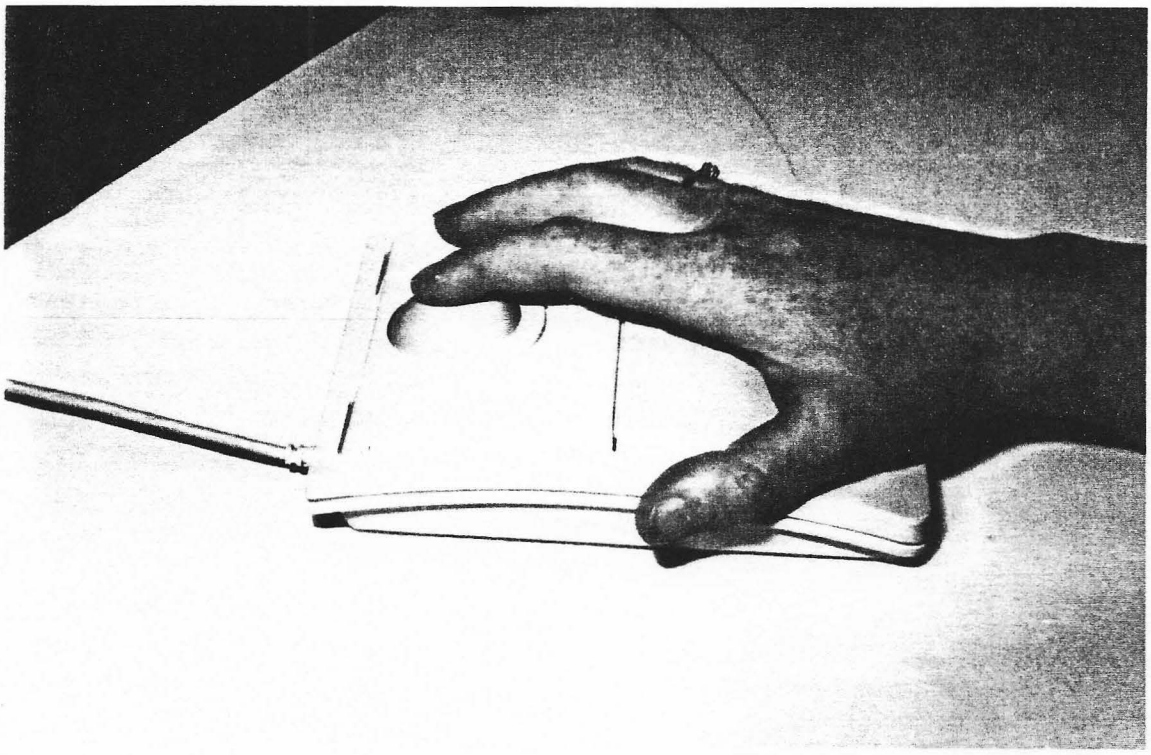


Figure 5 - The Large TBall

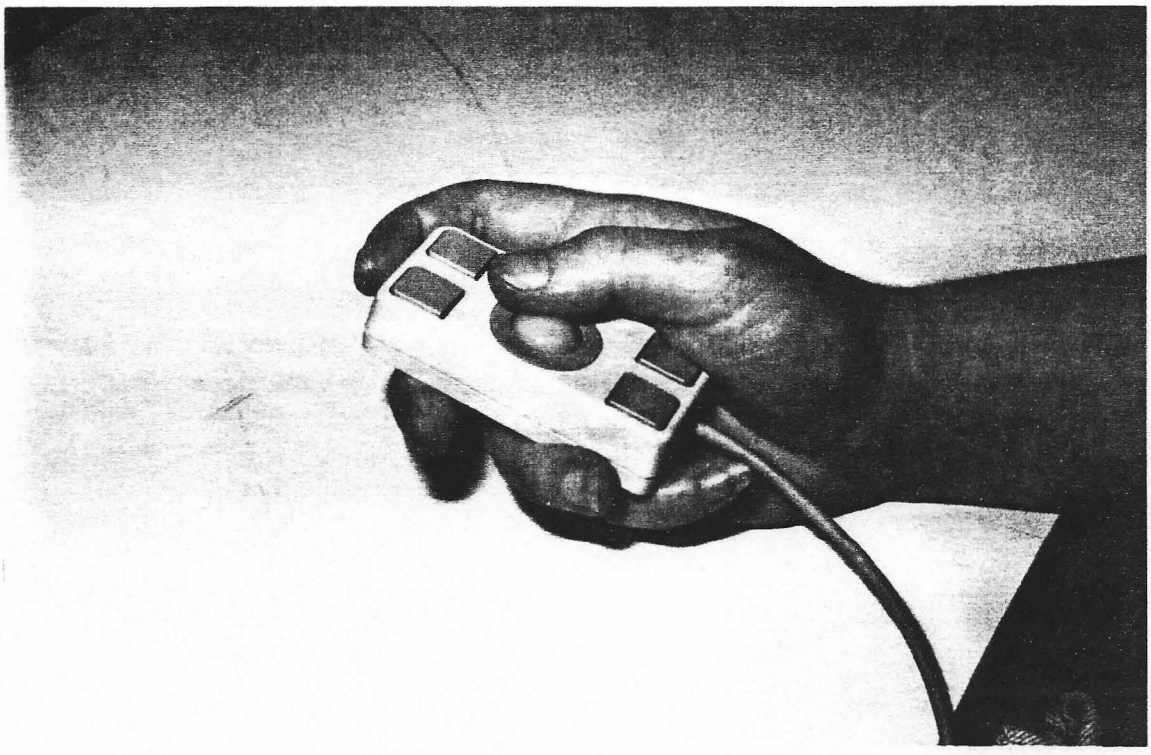


Figure 6 - TheYellow TBall

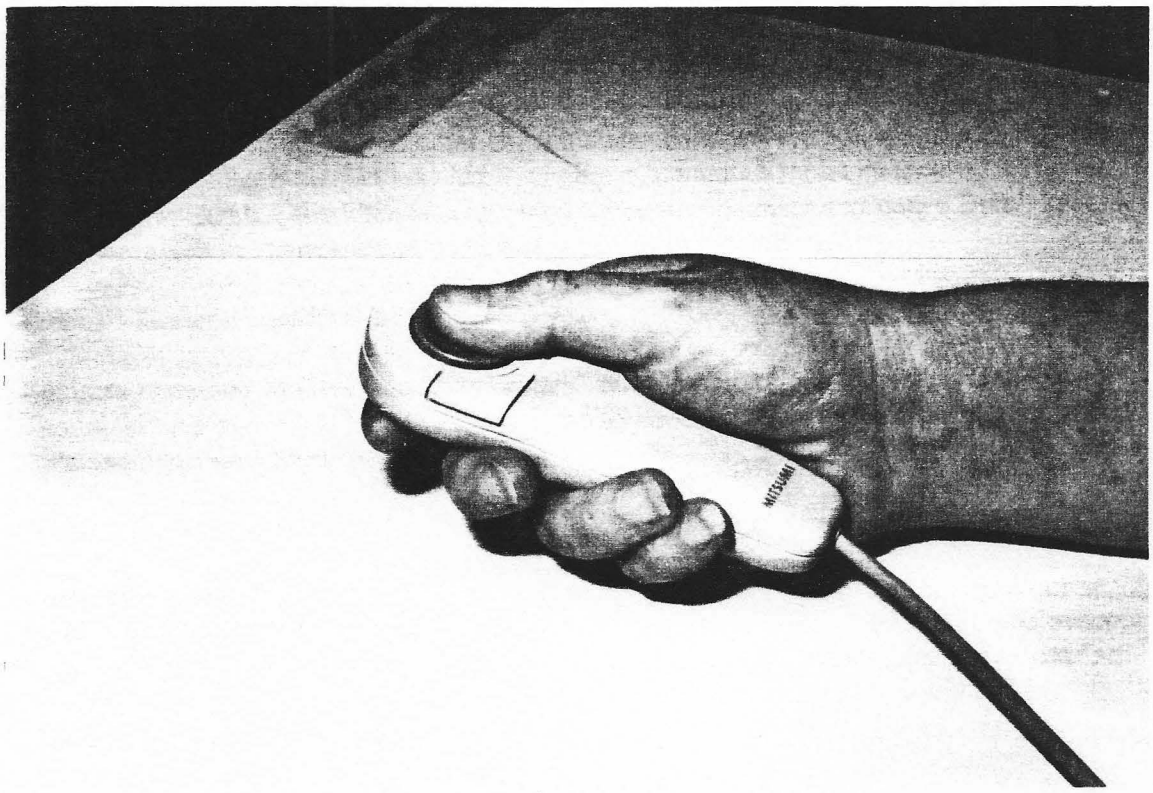


Figure 7 - The Pressure Pad

2.1.4 Procedure

The procedure followed in this test for the devices was basically the same as the previous test. The primary tasks were based on the title 'Cool Oldies Juke-Box'. The Player Shell and the Palm Springs Golf title were not used. Some other small modifications were made in order to improve efficiency in data collection. Each subject was led through the procedure by an interviewer. The test was conducted in Dutch. The interviewers were trained ergonomists and native Dutch speakers. An interviewer was in attendance at all times.

Stage1 - General introduction

Subjects were shown the usability laboratory, video recording equipment, and the observation room. The basic purpose of the test, the procedures to be followed, and the type of information to be collected were explained.

Stage 2 - Introduction to CD-I

Subjects use the Practice Track Ball, complete the tasks set and respond to the debriefing questions. The purpose of this stage was to ensure that subjects were familiar with the basic concept of CD-I and the details of the procedures they would have to follow during the course of the test.

Stage 3 - Use of the test devices

The four remaining input devices were used in sequence. The order of presentation for each subject was ordered systematically to counterbalance learning effects. The tasks each subject carried out with each device was the same, and were as follows:

Task	Button activated (IN CAPITALS) and movement
1	PLAY
2	STOP
3	PLAY to right arrow (repeat to Jimmy Jones)
4	right arrow to SEQUENCE
5	SEQUENCE to left arrow (repeat to James Brown)
6	left arrow to SEQUENCE
7	SEQUENCE to PLAY
8	go to next song
9	STOP
10	PLAY to FEATURES
11	BACK to SAVE SEQuence
12	SAVE SEQuence to PLAY
13	STOP
14	PLAY to FEATURES
15	BACK to CLEAR SEQuence
16	CLEAR SEQuence to left arrow (repeat to Janis Ian)

17 left arrow to PLAY
18 STOP
19 PLAY to FEATURES
20 BACK to GET SEQuence
21 GET SEQuence to PLAY
22 STOP
23 PLAY to FEATURES
24 BACK to EXIT (twice)

Following the use of each device subjects were asked to rate the qualities of each device and comment on the reason for their rating according to a number of criteria. The criteria used are given in the section on data collection.

Stage 4 - Comparing the devices

After the subjects had used all the devices each subject was asked to rank the devices from best to worst.

Stage 5 - Model evaluation

Subjects were tasked to inspect and describe the 8 models. Each subject was asked to choose the two best and the two worst cases according to a number of criteria. The criteria used are given in the section on data collection.

Stage 6 - Debrief and departure

Each subject was asked to comment generally on CD-I and the test. Each subject was thanked for their assistance and departed.

3.0 DATA COLLECTION

3.1 Subjective data - devices

Comments and subjective data was collected on the basis of a structured interview. Responses at the end of each session with each device were recorded using 1 - 5 rating scales. Both the rating data and the ranking data for each device were in terms of the following criteria:

General Appearance
General comfort
Adaptability to grip
Accuracy
Speed control
Access to buttons
Shape/form
Surface/finish

The best models and worst models were selected using the criteria above, except speed of control and accuracy.

3.2 Objective data - devices

For each device objective data on Time to Target for cursor positioning was collected for all tasks except tasks 1 and 2. Time to Target was defined as the time between the point at which the cursor begins to move and the point at which it comes to rest on the specified target. The data was collected from the video recordings.

3.2.1 Classification of movement types and target sizes

To classify the movement types the screen was divided into a 3x3 matrix having 3 rows; top, middle and bottom, and 3 columns; left, middle and right. The classification was based on the ideal path between target 'A' and target 'B', not paths actually taken by subjects.

Movement	Cursor moves
L: long	left to right column, or top to bottom row (or vice versa)
M: medium	from one row or column into adjacent row or column
S: short,	within one cell
D: diagonal,	at 45° +/- 10° to horizontal or vertical
U: up D: down,	
R: left to right	
L: right to left,	
V: vertical,	
H: horizontal	

Target size was based on the size of the target in the same dimension as the ideal approach path of the cursor to the target. Cursor size was defined as the maximum dimension of the cursor mark.

Target size	
L: large	greater than or equal to twice cursor size
M: Medium	greater than cursor size to less than twice cursor size
S: small	less than or equal to cursor size

3.3 Subjective data - models

The best models and worst models were selected using the same criteria as for the devices with the exception of speed of control and accuracy. The criteria were:

General Appearance
General comfort
Adaptability to grip
Access to buttons
Shape/form
Surface/finish

4.0 RESULTS

4.1 Devices - subjective data

The conclusions drawn from these results do not take account of the Practice TBall. However the results from this device are included for interest. Of note is that, in spite of the unreliability of the device, it still scored better than some other devices under certain categories. The data given in the body of the two tables below are mean values.

Note: B = Best score, W = Worst score

4.1.1 Device Ratings

Table 1 - Device Ratings

Device	Current RC	Lg TBall	Yell TBall	Pres Pad	Practice
Dimension					
Appearance	2.5	3.4	3.5 W	2.3 B	2.8
Comfort	3.2 W	2.6	2.8	2.1 B	3.1
Ease of grip	2.9	3.5	3.6 W	2.2 B	2.8
Accuracy	3.0 W	2.3 B	2.5	2.4	—
Speed Control	3.1 W	2.5	2.7	2.4 B	—
Move to buttons	2.2	2.0 B	2.8 W	2.3	2.9
Form	2.5	3.1	4.0 W	2.0 B	3.0
Surface finish	2.3	2.4	2.9 W	2.2 B	3.0
Overall preference	2.6	2.7	3.1 W	1.8 B	2.8

4.1.2 Device Preferences

Table 2 - Device Preferences

Device	Current RC	Lg TBall	Yell TBall	Pres Pad	Practice
Dimension					
Appearance	3.0	3.6	4.4 W	1.7 B	2.3
Comfort	3.1	2.9	3.3 W	2.1 B	3.6
Ease of grip	3.1	3.1	3.9 W	1.9 B	3.0
Accuracy	2.7	2.6 B	3.2 W	2.6 B	—
Speed Control	3.4 W	2.3 B	3.1	3.0	—
Move to buttons	2.4 B	2.6	3.9 W	3.0	3.1
Form	3.1	3.6	3.9 W	1.7 B	2.7
Surface finish	2.7	3.3	4.0 W	1.7 B	3.3
Overall preference	3.0	3.9 W	3.4	1.9 B	2.9

4.2 Devices - objective data

The table below gives the mean values in seconds for Time to Target scores. The movement number corresponds with the Task Number described in Section... The best and worst times are indicated. Where a statistically significant difference exists between two or more devices is also indicated.

4.2.1 Device Comparisons - Time to Target

Table 3 - Device Comparisons - Time to Target (mean values in seconds)

Movement type	Target size	Current RC	Lg TBall	Yell TBall	Pres Pad	Significant Difference
3	L,DU,R	L	7.2 W	3.0	2.7 B	3.5	N
4	M,VD	S	4.3 W	2.0 B	3.2	3.0	Y
5	L,DU,L	L	4.6 W	2.4 B	3.2	2.6	Y
6	L,DD,R	S	4.8 W	2.9 B	3.6	4.0	Y
7	L,HL	M	3.1 W	2.4	2.0 B	3.1 W	Y
8	M,HR	L	1.5 W	1.5 W	1.3 B	1.5 W	N
9	M,HL	L	1.2 B	1.5	1.9 W	1.9 W	N
10	L,DD,R	M	2.8 W	1.4 B	1.5	2.8 W	Y
11	M,VU	S	2.3 W	1.4	1.4	1.3 B	Y
12	M,HL	M	1.5	1.0 B	1.3	1.7 W	N
13	M,HL	L	1.9 W	1.3	1.1 B	1.8	N
14	L,DD,R	M	3.4 W	1.7	1.4 B	2.5	Y
15	S,VU	S	1.2 W	0.9 B	1.1	0.9 B	N
16	L,DU,L	L	4.1 W	2.3 B	2.6	3.9	Y
17	M,VD	S	1.7 W	1.5 B	1.5 B	1.7 W	N
18	M,HL	L	1.4	1.1 B	1.2	1.7 W	N
19	L,DD,R	M	3.0 W	1.4 B	1.5	2.2	N
20	S,VU	S	2.2 W	1.1	1.4	0.7 B	N
21	M,HL	M	2.3 W	1.3	1.1B	1.9	Y
22	M,HL	L	1.2 B	1.2 B	1.3	1.7 W	N
23	L,DD,R	M	2.2	1.4 B	1.6	2.9 W	Y
24	M,HL	M	1.2	1.2	1.0 B	1.8 W	Y

Movement type:
L: long M: Medium S: short, D: diagonal, U: up D: down, R: left to right L: right to left,
L: long M: Medium S: short, V: vertical, U: up or D: down,
L: long M: Medium S: short, H: horizontal, R: left to right L: right to left,
Target size:
L: large, M: Medium, S: small
Annotation:
B = Best score, W = Worst score; Y = Yes, N = No

4.3 Model Preferences

The table below summarises the preferences expressed by the subjects for the models. For each criterion and each model the total number of positive and negative votes are given. The positive vote is the number of occasions the model was regarded as the best for the particular criterion. The negative vote gives the number of occasions the model was regarded as the worst.

Table 4 - Model preferences

Model	FK	FG	FT	FP	RCB	RFB	RCP	RFP
Dimension								
Appearance	2+1-	8-	2+3-	3+1-	5+2-	2+1-	5+1-	3+1-
Comfort	2+3-	4+2-	6-	3-	3+1-	3+1-	3+	1+1-
Ease of grip	2+3-	2+4-	8-	2+4-	4+1-	2+1-	7+	2+1-
Move to buttons	1+1-	3+1-	2-	1+5-	2+2-	2+1-	3+	3+
Form	1+1-	8-	2+4-	3+	5+1-	---	4+	1+1-
Surface finish	1+2-	2+5-	1+4-	3+	2+2-	3+1-	3+1-	2+1-
Overall preference	1+1-	6-	4-	2+	3+1-	1+1-	6+	3+
Totals								
- Positive votes	10	11	5	14	24	13	31	15
- Negative votes	12	34	31	13	10	6	2	5
- Balance	2-	23-	26-	1+	14+	7+	29+	10+
Rank order	6	7	8	5	2	4	1	3

5.0 DISCUSSION

5.1 The Devices - comments and subjective data

The comments made by each subject on each device were collated by dimension and device. The comments were open ended and were invited as an 'explanation' for why the subject had given the particular device a particular rating or preference. The aim was to establish which apsects were of importance to the subject in coming to their judgement. The comments obtained were diverse. Comments in answer to one question often were related to other questions. Therefore in attempting to present and discuss the results in some useful way the following format has been adopted. For each dimension used to rate the device:

- the comments made have been interpreted in terms of the attention points which should lead to a successful design. The sequence of comment is in order of approximate priority; the most important issues first. The frequency of occurence of a phrase or term was used as a basis for judging priority.
- the device which scored best and the device which scored worst in terms of the rating and preference scores is given.
- where appropriate additonal notes relating to the devices are given.

A word of caution to the reader. It is extremely difficult to convey the richness of the information collected during the tests through only the printed word. This report can only give a partial insight into the strengths and weaknesses of the devices tested. In some cases comment was interpreted by using only key words and therefore rather cryptic. It is strongly recommended that any detailed interpretation of the test results is based on examining the video recorded during the test sessions.

5.1.1 Dimension - Appearance

Handling the device - the appearance of the device should clearly indicate how the device should be picked up and handled. Single handed use. Not two handed use. Must fit the hand.
Aestheic quality - provide rounded, smooth, slim shape.
Colour - complement current domestic AV equipment
Integrated design - buttons and controls should not be 'add ons'. Colours, materials, and forms should complement each other.
Simplicity - limit the number of buttons; too many appears complicated.
(note subjects were not required to use the CD player buttons on the current RC. They were not relevant to this test)

Rating scores:

Best device	Pres Pad	Worst device	Yell TBall
Mean Score	2.3	Mean Score	3.5

Preference scores

Best device	Pres Pad	Worst device	Yell TBall
Mean Score	1.7	Mean Score	4.4

The Pressure Pad; ease of handling and fluency of design.
The Yellow Track Ball; toy like appearance and crude shape.

5.1.2 Dimension - General Comfort

Handling the device - Single handed use. Not two handed use. Must fit the hand. The most important issue.
Handling and operation of primary control element - user must be able to hold device and operate the primary control element without changing grip.
Physical relationship between primary control element and action buttons - user must be able to operate the primary control element and operate action buttons without changing grip.
Portable - must be able to hold the device in one hand and carry device.
Cable - device should be wireless

Rating scores

Best device	Pres Pad	Worst device	Current RC
Mean Score	2.1	Mean Score	3.2

Preference scores

Best device	Pres Pad	Worst device	Yell TBall
Mean Score	2.1	Mean Score	3.3

The Pressure Pad; Pad very easy to operate without change of grip. Action buttons awkward to reach without grip change.
The Current RC; Operation of joystick required too much pressure to ensure desired response. Shape of housing generally liked.
The Yell TBall; Poor fit to hand, need to change grip to operate, number of buttons confusing.

5.1.3 Dimension - Ease of Grip

Handling the device - must fit the hand. The most important issue.
Handling the device - use with one hand, not two handed use.
Physical relationship between primary control element and action buttons - user must be able to operate the primary control element and operate action buttons without changing grip.

Rating scores

Best device	Pres Pad	Worst device	Yell TBall
Mean Score	2.2	Mean Score	3.6

Preference scores

Best device	Pres Pad	Worst device	Yell TBall
Mean Score	1.9	Mean Score	3.9

The Pressure Pad; as for General Comfort
The Yell TBall; mismatch between shape and size of the device and users hand. Positive attributes; small and light.

5.1.4 Dimension - Accuracy

Response characteristics - a balance is required between sensitivity (pixels/physical movement of device ratio) and movement resistance in device mechanism. Subjects seemed to find it difficult to distinguish between speed aspects and accuracy of positioning. With the Current RC subjects commented on having difficulty in positioning the cursor on small targets, and overshooting targets.

Rating scores			
Best device	Large TBall	Worst device	Current RC
Mean Score	2.3	Mean Score	3.0

Preference scores			
Best device	Large TBall	Worst device	Yell TBall
Mean Score	2.6	Mean Score	3.2

The Large TBall; responds immediately, fast and accurate.
The Current RC; slow response, target overshoots, fatiguing in use.

5.1.5 Dimension - Speed control

Ease of speed control - subjects preferred a smooth transition between stationary and maximum speed. The device should not impose limits to cursor speed, eg Control Display ratio of the Yellow TBall, absolute speed limit imposed by Current RC.
Responsiveness of primary control element - ensure direct response to user actions. Avoid perceptable lags in response.
Sensitivity - ensure appropriate level of sensitivity in device.

Rating scores			
Best device	Press Pad	Worst device	Current RC
Mean Score	2.4	Mean Score	3.1

Preference scores			
Best device	Large TBall	Worst device	Current RC
Mean Score	2.3	Mean Score	3.4

The Press Pad; specific comments inconclusive
The Large TBall; direct response, accurate and quick, precise
The Current RC; slow response, slow speed, target overshoots.

5.1.6 Dimension - Access to buttons

Button position relative to primary control element - place buttons on same surface as primary control element, place buttons on top surface, place adjacent to primary control element, enable

operation of buttons without significant change in grip.
Holding device - enable one handed use
Button size - big enough to distinguish easily without looking
Sensitivity - ensure appropriate level of sensitivity in device.

Rating scores			
Best device	Large TBall	Worst device	Yellow TBall
Mean Score	2.0	Mean Score	2.8

Preference scores			
Best device	Current RC	Worst device	Yellow TBall
Mean Score	2.4	Mean Score	3.9

The Large TBall; buttons large, easily operated, adjacent to TBall
The Current RC; easy to move from primary control element and buttons
The Yellow TBall; must change grip to reach buttons, too many buttons - confusing

5.1.7 Dimension - Shape and Form

Aesthetically pleasing - an attractive pleasing object that matches associated equipment
Fits to the hand - see comments on handling under Appearance above.
Resembles AV remote controls - communicates function and method of operation. Reflects experience with current AV remote controls; one handed use, no cable, light and portable.
Contempory - CD-I is a new generation of technology. The control device should reflect this.

Rating scores			
Best device	Pres Pad	Worst device	Yellow TBall
Mean Score	2.0	Mean Score	4.0

Preference scores			
Best device	Pres Pad	Worst device	Yellow TBall
Mean Score	1.7	Mean Score	3.9

The Pres Pad; an attractive pleasing object that fits to the hand.
The Yell TBall; no style, not attractive, not designed to be held.

5.1.8 Dimension - Surface Finish

Colour - colours should be combined in complementary ways. Dramatic contrasts to be avoided.
Primary Control element and Button housings - good fit into housing, not loose, smooth guides.
Grip and tactile quality - smooth non slip mat finishes especially for surfaces of control elements.
Robustness - provide 'solid feel' to device.
Quality of fit - housing parts should fit tightly together, no creaks or squeaks, no visible splits.

Rating scores			
Best device	Pres Pad	Worst device	Yellow TBall
Mean Score	2.2	Mean Score	2.9

Preference scores			
Best device	Pres Pad	Worst device	Yellow TBall
Mean Score	1.7	Mean Score	4.0

The Pres Pad; smooth housing, fluent design, good fit of control elements.
The Yellow TBall; unprofessional appearance, TBall slippery, poor fit to hand.

5.1.9 Dimension - Overall Preference

Ease of handling - device must be good fit to the hand, one handed operation, operate without changing grip,
Easy to control primary control element - positioning cursor quickly and accurately
Style - complement existing AV equipment,

Rating scores			
Best device	Pres Pad	Worst device	Yellow TBall
Mean Score	1.8	Mean Score	3.1
Preference scores			
Best device	Pres Pad	Worst device	Large TBall
Mean Score	1.9	Mean Score	3.9

The Pres Pad; good fit to the hand, one handed use, fluent design.
The Yellow TBall; poor fit to the hand, too small.
The Large TBall; must use surface, cannot hold and control with one hand.

5.2 The devices - objective data

5.2.1 Statistical analysis

The objective data was examined for statisitical significance using a 't' test. The purpose was to determine whether there were significant differences between the devices, and if so which devices were the worst and which the best. The table and the comments below summarise the result of this data analysis. The Practice TBall is excluded from this comparison because it was known beforehand to be unreliable.

The most significant differences in performance were related to the distance the cursor had to travel from target 'A' to 'B'. Secondly results also were influenced by whether the movement was diagonal rather than vertical or horizontal. The size of the target also influenced the performance recorded.

5.2.2 Long cursor movements

Of the nine long movements (3,5,6,7,10,14,16,19,23) performance with one of the TBalls was best on all occasions. Only for movements 3 and 19 were there non-significant differences between the devices. On six occasions (5,6,10,16,19,23) the Large TBall was best and on five occasions it was significantly better than at least one other device. For the other three occasions (3,7,14) the Yellow TBall was best and for movements 7 and 14 it was significantly better than at least one other device. In one case (5) the Pressure Pad was a close second to the Large TBall. In all cases except movement 23 the Current RC was the worst. All the long movements except 7 involved diagonal movements. This may have contributed to the poor performance of the Current RC because of the way in which subjects tended to operate the thumb cap with the side of their thumb.

5.2.3 Medium cursor movements

Of the eleven medium movements (4,8,9,11,12,13,17,18,21,22,24) results were more varied. In only four movements (4,11,21,24) was there a significant difference between devices. The Yellow TBall was significantly better than the other devices twice (21,24), the Large TBall was significantly better than the other devices once (4), and the Pressure Pad was also best once (11) although in this case the difference between it and the two Track Balls was insifgnificant. The Current RC was best on two occasions but the difference was not significant.

5.2.4 Small targets

In all movements that involved a small target (4,6,11,15,17,20,) the Current RC performed worst. But only in movements 4, 6, and 11 were the differences significant. The Large TBall was best in movements 4 and 6. The Pressure Pad was significantly better in movement 11.

The statisical analysis leads to the following conclusions:

No differences in the devices were found for:
- horizontal and vertical movements where the target was large.
- horizontal and vertical movements where the distance was short.

All track balls were better than the RC and the Pad for diagonal movements.
The Yellow track ball was better than all other devices for short diagonal movements.
The Large Track Ball was better than all other devices for long distances.
The Current RC was worse than all other devices for small targets.

In the table below the devices are compared based on the number of occasions that one device was significantly better than another device. From this table it can be seen that performance with the Large TBall was generally best. For only one comparison is another device better; the Yellow TBall. This confirms the trend found in the previous test.

Table 5 - Device Comparisons - Time to Target

	Better Device			
	Current RC	Large TBall	Yell TBall	Pres Pad
Worse Device				
Current RC	X	4	3	2
Large TBall	0	X	1	0
Yell TBall	0	2	X	0
Pres Pad	0	3	4	X

The test and analysis was based on a particular set of tasks for a particular title, The Golde Oldies. This is a significant limitation. Input devices should be evaluated on the basis of a range of cursor control tasks which reflect the full range of cursor control tasks which CD-I users are likely to encounter,.

It should be noted that for both the Current RC and the Pressure Pad the velocity of cursor movement is fixed by the system. It was not clear from the behaviour of subjects whether or not they were concerned by the limitation. However raising the upper limit on cursor velocity particularly in the case of the Current RC may have enabled quicker times to be recorded.

5.3 Setting Usability measures

The purpose of setting usability measures is to make the quality of the interaction with CD-I measurable in human performance terms. We therefore examined the objective data and the subjective data from this test and the previous test to see whether reliable estimates could be established as targets for the evaluation of the next generation of CD-I input devices.

5.3.1 How might a usability measure be expressed?

In the case of the objective data a specific time to target estimate, eg 2.5 secs, would be made for a specific cursor movement task. The estimate of time should be based on the data derived from our usability tests. The specification of task would take account of the distance to be covered, the target size, and the trajectory, eg Long, Diaganol, Up, Small target. Devices would be judged as satisfactory if users could consistently achieve or improve on this target in a usability test.

In the case of the subjective data target rating scores would be specified, eg 2, for a set of specified questions. The estimates and questions should be based on the data derived from our usability

tests.

In the tables below we compare data from this test and the previous test for the Current RC and the Large TBall, (the only devices common to both tests). The comparisons are based on only the questions asked in both tests, Table 6,and on the Time to Target data for the common cursor movements tasks, Table 7.

Table 6 - Comparison of This Test with Previous Test - Device Rating Means

	This test		Previous test	
Dimension	Current RC	Large TBall	Current RC	Large TBall
General appearance	2.5	3.4	1.8	2.5
Accuracy	3.0	2.3	2.7	1.3
Speed	3.1	2.5	2.2	1.8
Form	2.5	3.1	1.7	2.6
Surface Finish	2.3	2.4	1.9	2.2

In the case of the subjective data, Table 6, a number of observations can be made. Consistent trends can be seen within and between the data sets. For Generalappearance, Form and Surface Finish the Current RC always scores better than the Large Track Ball. For Accuracy and Speed the Large Track Ball always scores better than the Current RC. In all cases the scores recorded in this test are always worse than the scores scored in the previous test. What is inconsistent is the magnitude of the differences between the scores for the Current RC and the Large TBall in the previous test compared with this test. In each test the Current RC and the Large TBall were being used in conjunction with different devices which is likely to be a contributory factor in causing these differences.

Table 7 - Comparison of This Test with Previous Test - Time to Target Means

	This test		Previous test	
Movement This test/ Prev test	Current RC	Large TBall	Current RC	Large TBall
3/1b	7.2	3.0	3.9	4.8
4/2b	4.3	2.0	2.8	1.6
5/3b	4.6	2.4	3.3	4.2
6/4b	4.8	2.9	4.5	1.7
7/5	3.1	2.4	3.2	1.7
8/2	1.5	1.5	1.9	1.2
9/3	1.2	1.5	—	1.1
10/6	2.8	1.4	4.4	1.9
11/7	2.3	1.4	1.8	0.9
12/8	1.5	1.0	1.4	1.2
13/4	1.9	1.3	1.2	1.3
14/9	3.4	1.7	3.2	1.8
15/10	1.2	0.9	1.0	0.4
16/11a	4.1	2.3	4.2	1.3
17/12a	1.7	1.5	5.6	1.3
18/5	1.4	1.1	2.1	1.1
19/13	3.0	1.4	2.3	2.5
20/14	2.2	1.1	1.8	1.3
21/15	2.3	1.3	2.2	1.0
22/6	1.2	1.2	1.5	1.2
23/16	2.2	1.4	3.5	1.7
24/17	1.2	1.2	1.0	0.4

In the case of the objective data presented in Table 7 above the following observations are made. In this test the Large TBall is always quicker or as quick as the Current RC except for movement 9/3. In the previous test the Current RC was quicker on four occasions (3/1b, 5/3b, 13/4, 19/13). Within each test the magnitude of the differences between the devices for each movement is inconsistent which may be a function of the particular movement task. However we might have expected the same difference to exist for the corresponding task in this test. This was not generally the case. In addition the absolute mean times recorded for the corresponding tasks in the two tests are different. Examination of the raw data shows that the distribution of individual times was large and the differences in the means referred to above are the result.

The variations in the data referred to above do not enable us to make reliable estimates of usability measures. Further analysis of the data and a review of the test procedures for both tests is required which lies outside the scope of this study. In relation to a further examination of the data and reviewing the test procedures attention will be focussed on the following:

- the task presentation and structure of the test
- the number of subjects required
- the effect of learning on subject performance particularly in relation to the objective data
- the formulation of the questions and the method of presentation to subjects

5.4 The Models - comments and subjective data

In the following section the comments and the preferences expressed are presented. Each model is discussed individually. As with the data for the devices the raw information has been interpreted. It is strongly advised that detail interpretation of the comments are based on viewing the video recordings made during the tests.

5.4.1 Model FK

Positive votes	10
Negative votes	12
Balance	2-
Rank order	6

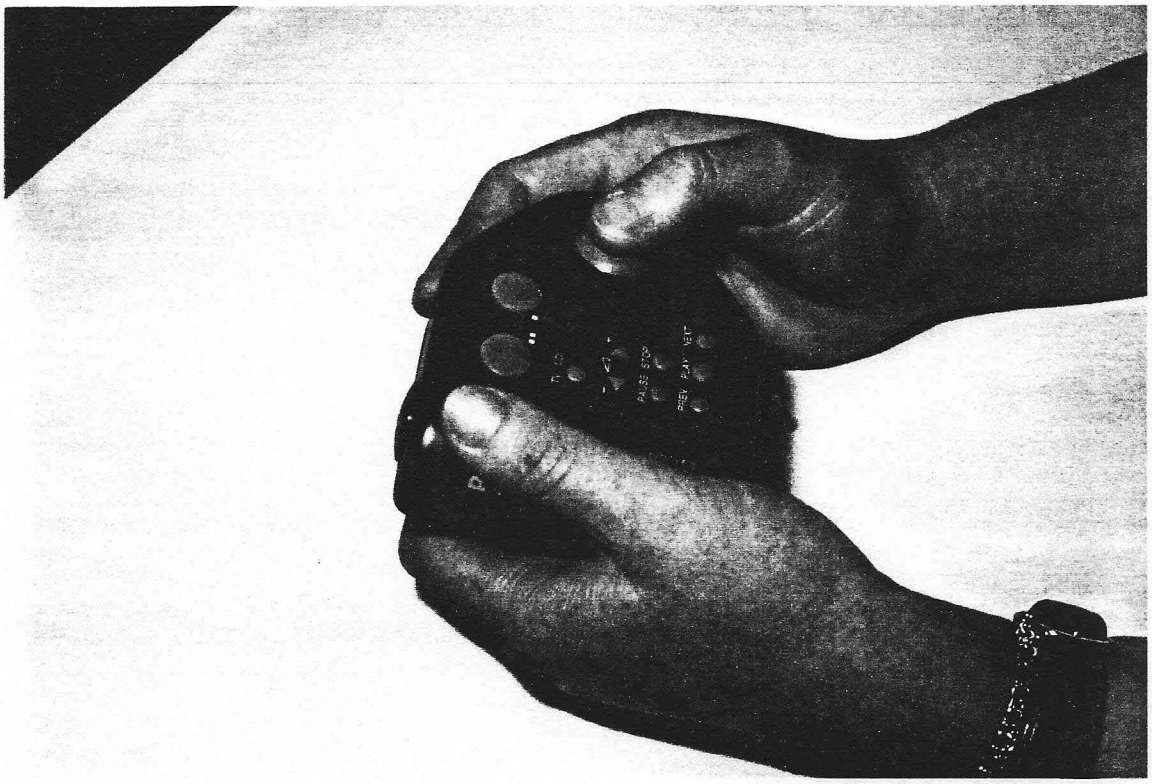
Most subjects found this model comfortable and easy to hold. But several subjects also found it unattractive. The balance of the device also received comment. Most subjects found the main body of the device heavy. To hold it securely required a firm grip, otherwise it tended to slip out of the users hand. This would be a significant disadvantage if the device were used over extended periods. Holding the device at the bottom of the handle was more comfortable for some, but this put the action buttons out of easy reach. In general subjects found that the action buttons were in the right place. For some who tended to hold the device low down on the handle the buttons were too far towards the top. Some subjects therefore suggested putting the buttons below the primary control rather than above it. Others found the CD player buttons at the bottom not easily reachedand had to use two hands to operate them. Two subjects were concerned that they could easily break the 'glass' cover for the infra red emitter. Subjects commented that the device falls over when laid on a table which was impractical.

5.4.2 Model FG

Positive votes	11
Negative votes	34
Balance	23-
Rank order	7

The shape was not generally liked. It was compared with a shaver, a small vacuum cleaner, and a gun. As to comfort subjects were divided. Most subjects found it comfortable to hold, and a good fit to the hand. Some found it uncomfortable. The shape made it unlikely that it would slip out of the hand. The shape also made it easy to aim. The balance was found to be poor with the weight biased to the head Use of the action buttons and the ball was found difficult to use with one hand. When the control is used with one hand the hand covers the CD player buttons. The position of the action buttons on the far side were found to be poorly positioned (right side for left handed use and vice versa). The action buttons were not easily reached with the thumb controlling the ball. The CD player buttons on the handle created an uncomfortable grip for the hand. Most subjects were inclined not to use the holder. One subject suggested throwing it away. The holder was not regarded as practical.

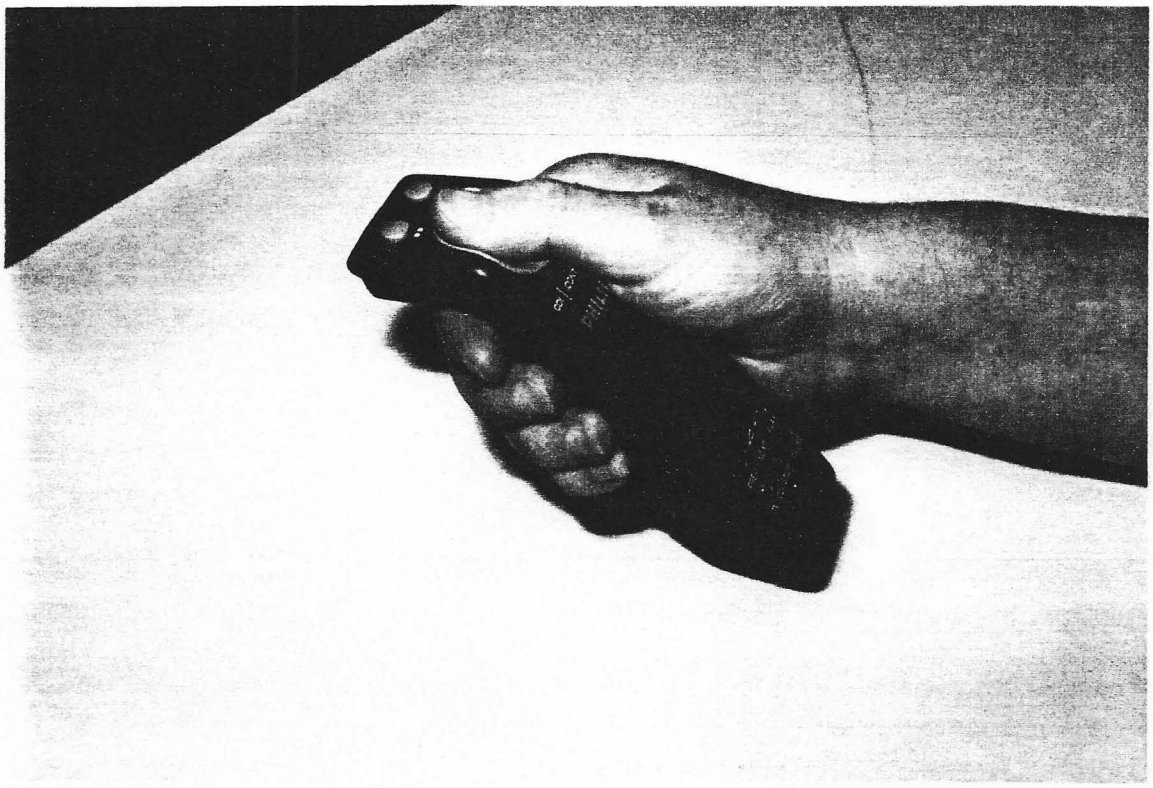
5.4.3 Model FT



Positive votes	5
Negative votes	31
Balance	26-
Rank order	8

The need to use two hands was not liked. For most subjects it was not regarded as comfortable, or providing a good fit to the hand. One subject pointed out that for educational titles the user may well want to write while using the device, another that the user may want to hold a cup of coffee. The action buttons were in a good position, but generally found to be too small. The ring round the ball was found to be uncomfortable. The buttons should be aligned more closely to the arc of thumb movement when moving the thumb off the ball. Felt solid and well balanced, but for some it was regarded as too heavy.

5.4.4 Model FP



Positive votes	14
Negative votes	13
Balance	1+
Rank order	5

This model was found to be comfortable to hold by some, but not easy to control with one hand. Several subjects commented on the need to alter their grip to hold it securely and operate the action buttons. Other subjects found it unstable, clumsy, and awkward to handle. The centre of gravity was too far down for some. Subjects hands tended to slip to the front. This model had an angular bottom; one can place it on a table but it was not so comfortable. The action buttons were postioned on the upper surface. This was liked less than to the side. The action buttons were not so easy to reach. The buttons were too far to the front to hold comfortably. Some subjects found it difficult to reach over the cursor pad to reach the buttons. The buttons, for some subjects, should be larger and flush with the surface. The main cursor control element was found to suit the thumb but not suitable for handling the cursor. Subjects preferred the dynamic aspect of moving the thumb to move the cursor, rather than only applying pressure on the same spot.

5.4.5 Model RCB

Positive votes	24
Negative votes	10
Balance	14+
Rank order	2

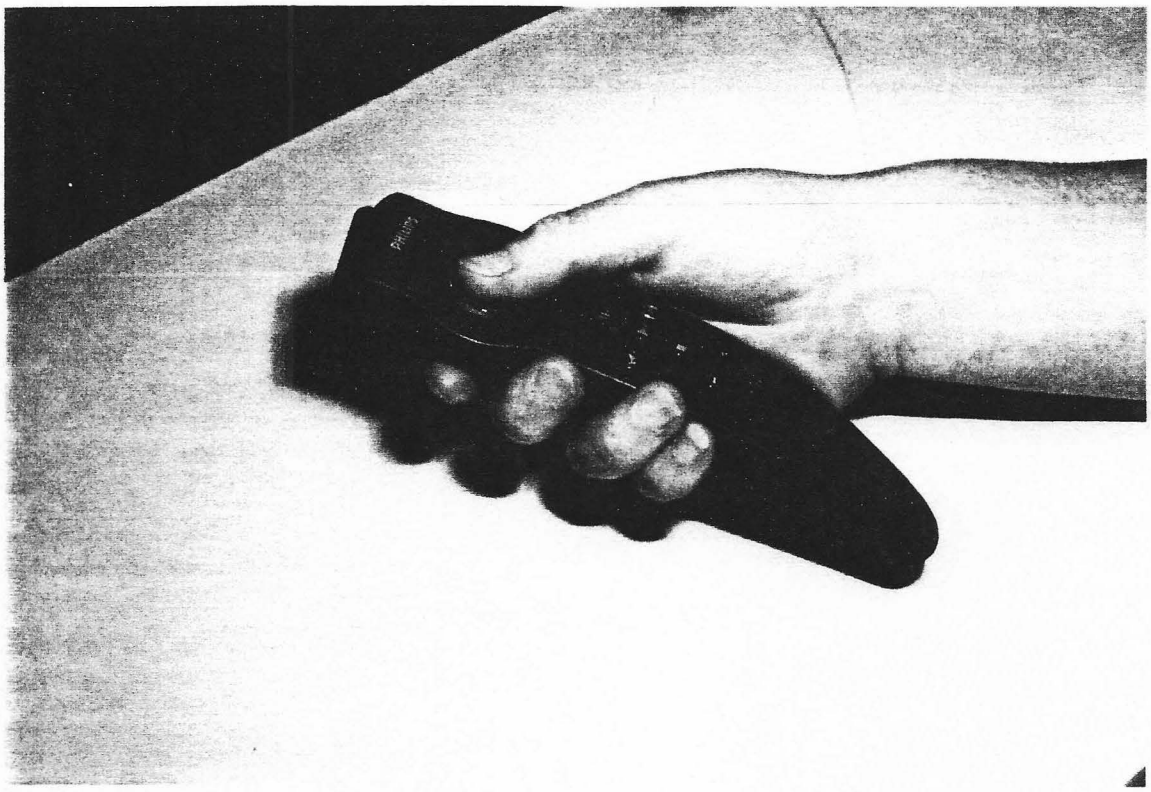
Generally received a positive reaction. Subjects liked the curvature of the body. The style appeared modern and fluent. Most subjects found it easy to handle. The curvature made it comfortable to hold. For most subjects it fitted the hand well although some commented that it might not be so comfortable for people with small hands, particularly in the area around the ball. Most subjects thought it would be easy to control with one hand without a change of grip. One negative comment was that the end of the model was too sharp. The power and play buttons were positioned directly under the thumb which made them easy to reach for most subjects. The holder received several negative comments. The model sat too tightly in the holder. The holder was bigger than the model itself. Several subjects thought that the holder would probably not be used.

5.4.6 Model RFB

Positive votes	13
Negative votes	6
Balance	7+
Rank order	4

Most subjects found the shape easy and comfortable to handle with a good fit to the hand. This model was criticised by some for being too angular and too long. For some subjects this made the shape uncomfortable. One subject found the point on the tail of the model particularly uncomfortable. Most subjects thought the configuration of the track ball and the action buttons would be easy to control with one hand. One subject criticised the height of the ball. The user would have to reach over the ball to use the action buttons. This might be awkward for people with small hands. The action buttons were easily identified. One subject would have preferred the action buttons to be above the ball. Several subjects did not like the CD player buttons on the rear of the device. One subject thought the CD player buttons should be closer to the top of the model. One subject preferred the CD Player buttons on the back so as to avoid accidently operating them. One subject preferred not to be able to lay down the model with the ball face down. One subject commented that he would have to stretch out his arm to ensure a reliable connection with the Infra Red detector.

5.4.7 Model RCP



Positive votes	31
Negative votes	2
Balance	29+
Rank order	1

For most subjects the overall impression was a good, well finished, modern design. Some subjects criticised it for being not functional and having little appeal. One subject did not like the curve of the body. All subjects found the model easy to handle and most regarded the fit to the hand as good. Some regarded this model as the most comfortable as reflected in the ranking data. Several subjects commented on the good balance. Most subjects liked the position of the action buttons in relation to the pressure pad. The action buttons were easily reached with the thumb without a change of grip. The inner angle of the action button surface was criticised for being too sharp by one subject. One subject noted that the use of the pressure pad technology reduced the height of the surface relative to the action buttons. For this subject reaching the action buttons was made somewhat easier than the similar model using track ball technology. One subject would have preferred the model to be heavier and a little shorter. One subject liked the position of the CD player buttons on the top side of the model. They liked the possibility of controlling the pressure pad, action buttons and the CD player buttons with one hand. One subject expected to have to stretch to direct the Infra Red beam towards the TV to obtain a reliable connection.

5.4.8 Model RFP

Positive votes	15
Negative votes	5
Balance	10+
Rank order	3

Most subjects found this model a good, well finished design. Some subjects criticised the model for having a 'straight look' and without a curved body. One subject preferred the track ball technology rather than the pressure pad. Most subjects found the body easy to hold and very comfortable. For most subjects the fit to the hand was good. One subject found the body uncomfortable because the body was straight. Several subjects compared this model with the previous one, RCP. Many subjects saw this model as more or less the same. Most subjects found the action buttons well positioned. The action buttons were easily reached. Many of the comments were the same as for the RCP model. It was clear from the comments that some users noticed the difference in the surface angle between RCP and this model. Three subjects clearly preferred the shape and angle of the action buttons in the RCP model. The CD player buttons could be reached using one hand. At least one subject regarded this as a positive point.

6.0 RECOMMENDATIONS

Design directions for the next generation CD-I general purpose input devices

6.1 Handling the Device

Fit to the hand

- anthropometric requirements need to be defined in order to encompass 95% of expected user population. Special consideration needs to be given to hand sizes of children. Minimum age groups need to be specified. Specific data requirements will depend on the particular device configurations proposed.

Physical relation between the primary control element and the action buttons

- position both on the top surface.
- place buttons adjacent to primary control element to enable easy coordinated operation,

Weight

- guidelines on weight and location of centre of gravity for device relative to grip position need to be determined.

Tactile quality

- smooth non-slip mat finishes especially for surfaces of control elements.

Grip and control

- user must be able to operate the primary control element and operate action buttons without significant change in grip,
- user must be able to operate the device with one hand, left or right
- do not compel the user to use two hands

Convenience

- user must be able to hold and carry the device in one hand,
- the device should be wireless

6.2 Controlling the cursor

Ease of control

- user must be able to position the cursor quickly and accurately
- enable a smooth control over change of cursor speed from zero to maximum speed at users discretion without perceptable thresholds
- the user should not perceive an absolute limit to maximum cursor speed that is determined by the system, eg Control Display ratio of the Yellow TBall, absolute speed limit imposed by Current RC.

Responsiveness of primary control element

- ensure direct response to user actions,

- ensure no perceptable delays in response to user actions,
- a balance is required between sensitivity (ratio of pixels to physical movement of device) and movement resistance in device mechanism. Specification method and recommendations need development.

6.3 General aesthetic quality of design

Style
- complement existing AV consumer electronics products
- colours, materials, and forms should complement each other.

Integrated design
- action buttons and CD player controls should not be 'add ons'

Shape
- provide rounded, smooth, slim shape.

Robustness and quality of fit
- provide 'solid feel' to device.
- housing parts should fit tightly together, no creaks or squeaks, no visible splits.
- primary control element and button housings should not feel loose in operation,

6.4 Usability measures

Subjective performance

The precise dimensions and the form of presentation used for collecting subjective data require review.

Objective performance

The objective data now available as a result of this test and the previous test require review in order to improve the accuracy of the data collected and enhance the long term benefit of usability testing of input devices.

Development of Usability test procedure

The tasks designed for usability tests must be reviewed. The cursor control tasks should reflect the range of tasks users of CD-I are likely to encounter when using CD-I with the general purpose device. The most notable cursor control tasks not covered in either this test or the previous test were system paced tracking tasks which require the user to respond rapidly.

The test procedure so far developed has proved to be successful to a point. However to improve the validity and value of the test procedure consideration must be given, amongst other aspects, to:

- how much experience should subjects have in the use of the input devices under test before reliable data can be measured?
- how many subjects are required in order to obtain a reasonable level of statistical reliability?
- what range of cursor control tasks are the most effective at discriminating between devices?
- can a set of 'Title independant' tasks be designed that provide a more objective basis for evaluating devices in the future?
- how should movement types and target sizes be best classified?