

22
7/24/95 J50

UCRL-ID-119816

A Human Engineering and Ergonomic Evaluation of the Security Access Panel Interface

Chris Hartney
William W. Banks

Human Factors Engineering and Engineering Geosciences Group
Applied Research and Engineering Division
Mechanical Engineering Department

February 1995


Lawrence
Livermore
National
Laboratory



FESSP

Fission Energy and Systems Safety Program

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

This is an informal report intended primarily for internal or limited external distribution. The opinions and conclusions stated are those of the author and may or may not be those of the Laboratory.

This report has been reproduced
directly from the best available copy.

Available to LOE and DOE contractors from the
Office of Scientific and Technical Information
P.O. Box 62, Oak Ridge, TN 37831
Prices available from (615) 576-8401, FTS 626-8401

Available to the public from the
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Rd.,
Springfield, VA 22161

UCRL-ID-119816

A Human Engineering and Ergonomic Evaluation of the Security Access Panel Interface

Chris Hartney
William W. Banks

February 1995

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER

ru

Summary

The purpose of this study was to empirically determine which of several security hardware interface designs produced the highest levels of end-user performance and acceptance. The FESSP Security Alarms and Monitoring Systems program area commissioned our study as decision support for upgrading the Argus security system's primary user interface so that Argus equipment will support the new DOE and DoD security access badges.

Twenty-two test subjects were repeatedly tested using six remote access panel (RAP) designs. Lawrence Livermore National Laboratory (LLNL) uses one of these interface designs (called RAP B in our study) in its security access booths. (See the six photographs in Appendix D.)

Along with the RAP B *insert-style* reader, we tested five prototype RAP variants, each with a different style of *swipe* badge reader, through which a badge is moved or swiped (see Table 2). We asked the untrained test subjects to use each RAP while they described to us how they thought they should respond so that the system would operate correctly in reading the magnetic strip on a security badge.

With each RAP variant, subjects were required to make four successful card reads (swipes) in which the card reader correctly read and logged the transaction. After each trial, a subject completed a 10-item interface acceptance evaluation before approaching the next RAP. After interacting with the RAP interfaces (for a total of the six RAP trials), each subject completed a 7-item overview evaluation that compared and ranked the five experimental RAPs, using the original (RAP B) insert style as a standard.

This six-trial process was completed once, using a badge with a radiation dosimeter attached, and again using a badge without a dosimeter. We took body measurements and asked demographic questions of each subject between the runs with and without the dosimeters.

We measured two sets of dependent variables during the test trials. Objective *performance* measures consisted of the frequency and type of errors made during card swiping operations and the time subjects took to complete four successful swipes. Subjective *preference* measures consisted of requesting subjects to rank their preference for each RAP operated during the test trials.

We collected additional subjective data in the form of recorded subjects' comments. We performed a within-subjects (repeated measures) multivariate analyses of variance (MANOVA), comparing linear combinations of performance and preference data across the six RAP configurations. Planned, post hoc comparisons of the five prototype RAP performance and preference means were contrasted, following the MANOVA.

Based upon the measures of performance and preference of the perceived ease of operation, we recommend that RAP D (with a protruding horizontal, 0° configuration of the badge reader) be targeted for implementation. This recommendation comes with the qualification that although we found several statistically significant differences between RAP means, these differences were small in magnitude. And the mean-squared error term was relatively large, indicating the limiting influences of small sample size and small signal, and a high degree of within-groups variability.

Contents

Introduction	1
Method	2
Test Subjects	2
Instruments	3
Performance	4
Preference	6
Procedure	7
Results	8
Preference	8
Performance	10
Discussion and Recommendations	13
RAP D	13
Other RAPs and Design Flaws Discussed	15
Appendix A. Subject Preference of RAPs	A-1
Appendix B. Errors by RAP and Type	A-2
Appendix C. Ease of Procedure: Tables C-1 and C-2	A-3
Appendix D. Photographs of RAP's A-F	A-4

Tables

Table 1. Test subject characteristics	2
Table 2. RAP badge reader configurations	3
Table 3. Dimensions of the remote access panel (RAP)	3
Table 4. Errors recorded by the RAP and the experimenter	5
Table 5. Mean preference scores with and without dosimeter	9
Table 6. Simple comparisons for preference mean group differences	10
Table 7. Mean errors with and without dosimeter	10
Table 8. Simple comparisons for error mean group differences	11
Table 9. Mean time for four successful swipes (with and without dosimeter)	12

Figures

Figure 1. Three views of the test badge with dosimeter attached	4
Figure 2. Mean preference scores with and without dosimeter	9
Figure 3. Mean errors with and without dosimeter	11
Figure 4. Mean time for four successful swipes (with and without dosimeter)	12

Introduction

The remote access panel (RAP) is a personnel access control device used for passage from one security area to another within the LLNL complex. The RAP allows an individual to transmit information (a personal identification number, or PIN) to the main processing system, which then determines that person's eligibility for clearance into a restricted area. A person's badge information and PIN must first be matched by the main processing system with a site-specific access list.

Security devices must restrict uncleared individuals, yet allow authorized individuals to pass relatively freely. It is also important that these devices be user-friendly.¹ In a previous study by Banks and Moore (1988),² human factors engineering design standards were applied to security monitoring stations in an attempt to design system interfaces that were "easy to use and understand" by a wide spectrum of individuals, including the physically challenged.

This study is in direct response to new LLNL-required changes to the physical dimensions of access cards (badges) and the design change from the current insert-style card reader to a side-swipe reader. These design changes were functionally required by the upgraded ID card technology that changed from a 27-character to a modern 40-character coded magnetic strip (typically located on the back of the card). Using the results of this study, we intended to establish the best operational configuration from a human factors engineering perspective, while incorporating the use of the side-swipe readers with minimal modification to existing RAPs. In this case, "best" was operationally defined as that card and reader configuration that minimized human and machine time and errors during use, and maximized user acceptance as measured by scaled reports of preference from users.

Under controlled laboratory conditions, we objectively assessed five proposed alternative human/machine interface configurations of RAPs along with the current RAP B used by LLNL in its daily security operations. (This B configuration has been used at LLNL for approximately 6 years.) We statistically analyzed the performance and preference data collected from these trials, using repeated-measures multivariate analyses of variance (MANOVA). We planned simple follow-up comparisons to determine the configuration that (1) produced the least errors; (2) produced the most successful badge reads in the least amount of time; (3) was most preferred by users, in terms of aesthetics and perceived comfort and performance; and (4) was perceived as most intuitive and "easy" to use.

A typical RAP provides a badge-reading slot, a telephone-styled keypad for entering the user's personal identification number (PIN), functional keys for performing various diagnostic and specialized tasks, and a liquid-crystal display (LCD).

Depending on the type of RAP badge reader, users begin the process of gaining entry into a restricted area by either inserting an ID card or badge into the reader and rapidly removing it (in the case of the currently used RAP B) or by smoothly swiping the card along the badge-reading slot, without pausing. With both types of reader, users must be sure that the magnetic strip is correctly oriented and positioned when inserted and withdrawn or when swiped through the reader.

The RAP's computer system identifies users through coded information on the magnetic strip on the badge. The RAP display then visually prompts users to enter their PINs, which are then entered on the telephone-styled keyboard. If the entered PIN is on the approved access list, the RAP prompts users to enter the area. If an error occurs, the RAP prompts users to retry the badge insertion or swipe, and then reenter the PIN; or the RAP informs users that they are not on the access list and cannot enter the security area. This study focuses narrowly on the initial step of this human/machine dialogue. Specifically, we were interested in how well (fast and accurately) the user can determine how to correctly hold and operate the card or badge to successfully complete a transaction that allows the badge to be read by the security monitoring system.

¹Banks, W.W., Jr., and Weimer, J. *Effective Computer Display Design*. Englewood Cliffs: Prentice Hall, 1992.

²Banks, W.W., Jr., and Moore, J.W. "Optimizing Man-Machine Performance of a Personnel Access Restriction Security System." Paper prepared for submittal to the International Nuclear Materials Management 29th Annual Meeting, Las Vegas, Nevada, 1988.

Method

Test Subjects

We nonrandomly selected ten female and twelve male volunteer subjects from the LLNL population of full-time summer employees and visiting students during August and September of 1994. Demographic and bioinformation was recorded for each subject, including gender, age, height, handedness, body dimensions (including height to shoulder, arm length, and hand size), and degree of prior exposure to the operation of card-reading devices (see Table 1).

Table 1. *Test subject characteristics; n = 22.*

	Mean	Low	High	
Height (inches)	68.46	62	77	
Age (yrs)	38.11	17	58	
Exposure to card readers (use during previous month)	54.636	2	165	
Years at LLNL	9.62	<1	21	
		No.		No.
Reading ability	9th to 12th grade	4	Collegiate	18
Job type	Researcher	2	Clerical	6
	Manager	1	Technical	8
	Professional	5		
Computer experience	None	0	Novice	1
	Intermediate	13	Expert	8
Years working with computers	0-1 yrs	1	1-3 yrs	5
	3-5 yrs	3	>5 yrs	13
Experience with technical devices	Low	7	Intermediate	4
	High	2	Extensive	9
Handedness	Right	20	Left	1
	Both	1		
Sight	20/20	7	20/20 corrected	13
	Bifocals	2		
Education level	No high school diploma	3	Some college	6
	Tech degree	3	Professional	1
	BA	8	MA	1

We explained to each test subject the reasons for our experiment only after the test trials were completed and all data had been collected. Requests for volunteer test subjects were made to five different LLNL work groups so that we might include a representative range of different backgrounds, especially in the areas of exposure to card-reading devices, years at LLNL, job type, and age.

Instruments

Remote Access Panels

The RAPs³ themselves are identical except for the configuration of the badge reader, as shown in Table 2.

Table 2. RAP badge reader configurations.

RAP	Position of badge reader
A	Vertical (flush-mounted RAP)
B*	Horizontal insert (LLNL)
C	Facing up 90° vertical
D	Protruding horizontally; 0°
E	Flush-mounted, horizontal; 0°
F	Angled up 45°

*All readers are swipe readers except for RAP B.

Characteristics the RAPs hold in common are shown in Table 3.

Table 3. Dimensions of the remote access panel (RAP).

Height	7.00 in. (17.78 cm)
Width	13.25 in. (33.66 cm)
Depth*	8.00 in. (20.32 cm)
Housing	injection-molded polycarbonate

*Includes the 1-in. protrusion from swipe.

Wall-mounted or outdoor RAP panels at LLNL are typically oriented with the bottom of the RAP 36.5 inches from the floor. Booth-housed RAPs are mounted at this same height—users must enter and exit through two opposing security doors in telephone booth-style structures during the process of traveling into and out of secured areas. At LLNL, this type of structure is called a CAIN booth. The interior of the CAIN booth measures 44 inches side to side and 46 inches front to back.

The location of our testing was the prototype-testing facility, Trailer 6951, Lawrence Livermore National Laboratory, Livermore, California. The six test RAPs were mounted on the inside walls of the facility. We placed masking tape on the floor to delineate the lateral and horizontal constraints that would be produced by CAIN booth walls. Luminance levels were measured with a photospot luminosity instrument and were found to range from 9.2 to 10.1 millilamberts, with an average over 6 readings of 9.8 millilamberts.

³The present study is solely interested in the badge reader, although the reader is inseparable from the RAP panel, and any preference judgment involves both the reader and the panel. For ease of presentation, RAP will be used to refer to the panel as a whole as well as the experimental swipe readers.

Badges

The badges used in the study were the standard issue items used by LLNL security as identification badges, having a plastic loop and metal clip attached the badge so that they can be attached to the user's clothes or cord around the neck. One set of experimental trials used badges that had a radiation dosimeter attached to the same loop, hanging behind the badge (Figure 1).

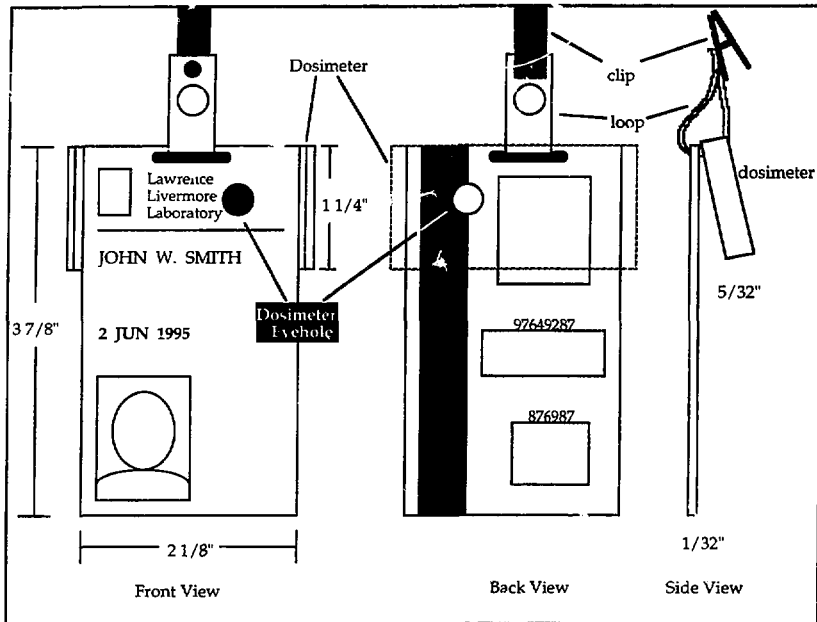


Figure 1. Three views of the test badge with dosimeter attached.

Additional instruments included a Stanley 60-inch (adjustable pull-out) measuring tape, a Pentax PV-C11A camcorder and Pentax tripod, SONY High Quality VHS-2 videocassettes, and a Hanhart model 341 stopwatch.

Performance

The performance measurements consisted of human error frequency and operation time, both measured by the internal clocks and the error-capturing capability of the RAPs. We also crosschecked the data obtained from the RAPs.

Errors

We recorded the number and type of errors occurring during the operational time of each test. The types of errors are listed in Table 4.

Table 4. Errors recorded by the RAP and the experimenter.

RAP-recorded error	Reason for error
Time-out	Swiping too slowly.
Alignment	Badge not swiped firmly and flatly along reader slot.
Parity	Code reading errors—unsuccessful read due to a variety of possible errors in the swipe, including too fast, too slow, crooked, and errors generated in the reader itself.
Too few	Incomplete read usually due to the dosimeter "eyehole" causing an interruption so that too few bytes of information are read. The reader reads the initial starting bits, but interprets the hole as indicating the read is over.
No read	Reader senses badge activity but makes no read whatsoever.
Experimenter-recorded error	
No read	Misaligning badge (magnetic strip).
No read	Swiping so quickly or slowly that the RAP does not sense activity.

A subject's error score was simply the difference of the total number of attempted swipes minus the number of successful swipes.

Errors Caused by the Dosimetry Badge Attachment. The current LLNL badges have a detachable dosimeter (a plastic, rectangular box that houses a radiation dosage detector; see Figure 1), which hangs from the same clip and lies directly behind the top section of the badge. To accommodate the "eyehole" through which radiation can be detected, a circular hole ($\frac{3}{8}$ in. diameter) is cut in the top right corner of the face (photo-side) of the badge. Since the magnetic strip runs the length of the left side on the back of the badge, this dosimeter eyehole cuts a $\frac{1}{8}$ in. half-circle into the magnetic strip. These eyeholes were the source of many "misreads" during the trials, because the badge reader interpreted the break in the continuity of the magnetic strip as the end of the strip, and then terminated the read. We discarded these misreads as programming errors and did not include them in the analysis.

Two points should be made about the dosimeter eyeholes. First, the tendency of the card reader to interpret the break as the end of the magnetic strip and to terminate the read was a programming, not a mechanical, flaw and was later corrected. All swipe readers were reprogrammed to read the magnetic strip regardless of the presence of an eyehole and regardless of the direction of swipe. (Note that the orientation of the magnetic strip, e.g., up or down, is still a factor.)

Second, we recognize that despite the exclusion of the dosimeter eyehole errors from the analysis, they may have contributed to other types of errors. We occasionally encountered a "chain reaction" of human errors, in which one machine-induced error led a subject to hastiness or confusion, thereby producing other errors. We believe, however, that this did not produce a significant or systematic increase in errors, because (1) we provided corrective instructions to the subject when observing this type of error, and we believe these instructions reassured the subject's confidence in performing the task, and (2) this type of error occurred only at the start of a trial. Therefore, the continuity or flow of the procedure from (re)start to four successful reads was intact, closely resembling (after inspection of the data) a trial with no eyehole errors.

Time

We measured time by viewing the videotapes after the trials. We measured the total time of the trial, starting at the moment the badge made first contact with the device and ending at the sound of the fourth successful read, as indicated by the RAP's beep message. With this total time, we then calculated the average time in seconds per successful swipe by dividing by the number of successful swipes (usually four).

Preference

Preference ratings were assessed through both questionnaires (described below) and our judgment of the "accuracy" of the user's intuitive sense of how to use the badge readers.

Questionnaires

We assessed the subjective preference ratings through questionnaires administered immediately after each trial. The questionnaires consisted of nine test items using a Likert seven-point scale (1="Strongly Disagree" and 7="Strongly Agree"):

1. I knew how to use the RAP console simply by looking at it.
2. It was easy to correctly swipe the badge four times.
3. The dosimeter attached to the badge made using this console more difficult than it would be without the dosimeter.
4. The badge fit smoothly into the slot.
5. The badge slid easily through the badge reader.
6. It was easy to reach (or gain access) to the badge-reader slot on the console.
7. My hand and arm were comfortable as I swiped the badge.
8. The overall appearance of this RAP console is visually pleasing.
9. Overall, looking at and using this RAP console was a pleasing and easy experience.

We used a discussion period with test subjects after the experiment to informally assess the impact of the questionnaire items and to ensure that the attitudes of the subjects were represented by the recorded responses they assigned to the questionnaire. We verbally asked each subject about each item, to ensure that what they indicated on the form was compatible with their intended response.

Subjects' preference scores were then computed as the average of their responses to the nine items, with the maximum possible score being a 7 and the minimum a 1. Item number 3 was only relevant when the dosimeter was attached to the badge, but this item seemed to cause confusion and was therefore removed from all computations.

Intuitive Ease of Understanding the Procedure, and RAP Acceptance Rankings

Two additional preference measurements were the subjects' intuitive sense of the proper procedure for a successful read and their ranking of RAP acceptability.

We measured the test subjects' intuitive sense of proper operation by observing them and asking them to attempt a swipe (badge read) while they described how they would do so without instruction from anyone else. We recorded the correctness (right, wrong) of the response along with the badge orientation employed in relation to the reader (magnetic strip facing up or down, left or right, toward or away) and the direction of the swipe (left to right, right to left, up or down). We also asked subjects to express their hypothesis for the correct reading procedure on the three RAPs determined to be the most popular.

We used a five-item preference ranking at the end of each set of trials (once with dosimeter and once without the dosimeter attached). We asked subjects to choose the RAP they felt (1) had performed most efficiently, (2) looked most efficient, (3) looked most aesthetically pleasing, and (4) was most recommendable for use at the Lab; finally, we asked subjects to (5) rank the five experimental RAPs by simply listing them in order of their personal preference.

Procedure

Subjects were scheduled to appear individually at various times throughout the day and on different days of the week for their twenty-five minute assessment. Upon arrival, they read a prepared statement describing the nature and purpose of the study; they then indicated their consent or refusal to participate. All those who read the statement gave their consent. We answered questions succinctly and truthfully throughout the experiment and during the briefing after the experiment. The same experimenter administered all aspects of every trial for the entire study.

Each of the six configurations of the RAP were randomly assigned a number, 1 through 6. Except for RAP B, which was always tested first because it is the current standard at LLNL, the order of testing was counterbalanced by using random-ordered sets of these same numbers, so that no systematic order of testing occurred.

Before the test trials began, we asked subjects whether they had used any sort of badge reader in the past—at the bank, the supermarket, or here at the Laboratory. They were told that the RAPs to be used in the experiment were very similar to these types of devices, and that for our trials they would be inserting into (in the case of RAP B) or sliding a badge across the readers.

Subjects then approached the first RAP and were shown its display, where swipe results were displayed and where they could read error messages. They were told that correct reads would be accompanied by a short beep, and unsuccessful reads by a longer beep. We then demonstrated this.

Finally, we instructed subjects to read aloud which errors, if any, were made, telling them that we would glance over their shoulder to check as well. This last step of reading aloud was incorporated to slightly slow down the subjects, since we found in a pilot study that subjects would often repeatedly swipe so quickly that they would clear an error message before the observer had a chance to record it.

At the start of each trial, we asked each subject to approach the device and attempt a successful swipe or to otherwise demonstrate what he or she felt was required to perform a successful badge reading. This initial try was observed and recorded to assess the intuitive preference factor. When necessary, we corrected the subject with something like "That's fine, except the magnetic strip has to face up instead of down." We then asked the subject to make successive swipes until the reader had signaled four successful swipes. Each trial was videotaped, with the performance measure of time recorded later via tape observations.

The testing process can be broken down into three sequential steps. In step 1, the user determines the proper method of holding the card for the correct badge reading. This initial step was unaided; it depended upon the user's previous experience or technical knowledge associated with RAPs and other devices using card readers, and also upon the user's intuition. This step was visually aided by the shape, placement, and appearance of the badge reader. In step 2, the user inserts or swipes the card. In step 3, the user receives audible feedback from the RAP in the form of a short beep for a successful read or a longer beep for an error—along with visual LCD feedback in the form of a tabulation of successful and unsuccessful reads. During the testing, we gave additional feedback in the form of verbal comments such as "Good" or "Try it again."

At the conclusion of the experiment, we reminded the subjects of the nature of the study, the anonymity of their responses, and the lack of connection to any work-related issues. We thanked the subjects for participating and told them that the results of the study would be made available in future months.

Results

The five experimental RAPs were compared in two multivariate analyses of variance (MANOVA) using variables of time, intuitive correctness of hypothesized operating procedure, and preference rankings.⁴ In each MANOVA, the five RAPs were treated as independent variables, with the same subjects participating in each of six trials and producing scores on two dependent variables: error and preference. The first MANOVA tested the preference scores and errors when the dosimeter was in place, and the second MANOVA tested the same scores when the dosimeter was not attached to the badge.

These statistical tests assessed whether relatively large, nonrandom differences existed between the mean scores of the dependent variables as a function of different RAP and badge design/configuration.⁵ A significant MANOVA indicates that the differences in the performance/preference variables for each of the five RAPs cannot be explained by chance. (Chance errors are those that will happen regardless of the design of the device. A significant difference indicates that more errors occurred on one RAP design than on another or that the dosimeter was attached to the badge.)

The dosimeter and nondosimeter groups were tested separately because preliminary paired t-tests showed that scores of the two groups were significantly different,⁶ which means analyzing them at the same time might hide, or, alternatively, falsely suggest real differences among the RAPs.

Both omnibus MANOVAs were significant (with dosimeter, $F(8, 14) = 4.279$, $p \leq 0.009$; without dosimeter, $F(8, 14) = 3.739$, $p = 0.015$), indicating at least one significant difference between RAP means per analysis. We then conducted planned simple comparisons of the mean error scores or the mean preference scores of two RAPs at a time.

Preference

The most accurate and reliable estimate of subject preference was deemed to be the posttrial preference questionnaires. These questionnaires were more detailed and were answered immediately following the trial, as opposed to the subject rankings of the RAPs. The rankings provided a comparison measure for supplementary information in the recommendation process.

While the dependent variable mean values associated with each independent variable condition indicate that significant differences existed as a function of the different experimental conditions (see Table 5 and Figure 2), a more refined statistical analysis based on each RAP mean value would be more precise in determining which specific RAP was relatively better (statistically significant) than another RAP design. We subsequently performed a limited number of paired comparisons as a follow-up on the significant effects resulting from the MANOVAs (see Table 6).

To facilitate making paired comparisons, we ranked the preference score groups and tested each of the top three scores against its next-highest ranking score. In other words, the top-ranked score was tested against the second-ranked score, the second against the third, and the third against the fourth (Table 6). The comparison of RAP B with the next most-highly rated RAP provided the only significant finding. (Because we used separate estimates of error for each comparison, we included RAP B in these tests.)

⁴We decided before the experiment that due to the scarcity of test subjects, all variables would not be included in one MANOVA, to reduce the ratio of variables to subjects and thereby increase the power of the design.

⁵Actually, the MANOVA combines the error and preference scores into a new, weighted score that takes into account the correlation between the two, and then compares the differences in the RAPs in the same fashion as would an ANOVA with just one dependent variable. In other words, MANOVA accounts for the fact that subjects may make substantially more errors on RAP X than they did on RAP Y, but still favor RAP X. MANOVA in effect balances the two variables and then compares all five RAPs, taking into account the variation in the scores in each category.

⁶For error scores, $t(21) = 3.61$ and $p = .001$; and for preference scores, $t(21) = -5.19$ and $p = .000$.

Table 5. Mean preference scores with and without dosimeter; n = 22.

	Mean	Std Dev	Min	Max	Rank
RAP A (with)	5.190	1.080	3.14	6.63	3.5
(without)	6.028	0.563	4.75	7.00	2.33
RAP B (with)	6.420	0.570	4.88	7.00	1.5
(without)	6.579	0.422	5.50	7.00	1.5
RAP C (with)	5.190	1.190	2.75	6.50	3.5
(without)	5.903	0.918	2.63	6.88	2.33
RAP D (with)	5.280	1.140	2.38	7.00	2.33
(without)	5.881	1.035	2.75	7.00	4.0
RAP E (with)	4.350	1.310	2.00	6.75	5.5
(without)	4.746	1.200	2.63	6.63	5.5
RAP F (with)	4.980	1.070	2.88	6.88	6.5
(without)	5.516	1.226	2.86	7.00	6.5

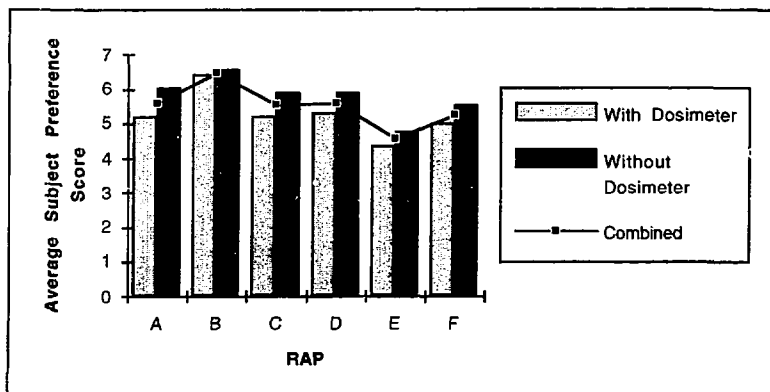
**Figure 2.** Mean preference scores with and without dosimeter; n = 22.

Table 6. Simple comparisons for preference mean group differences; $df = 21$.

	RAPs compared	F ratios
With dosimeter	RAP B by RAP D	5.13*
	RAP D by RAP C	0.20
	RAP D by RAP A	0.27
Without dosimeter	RAP B by RAP A	6.63*
	RAP A by RAP C	0.60
	RAP C by RAP D	0.12

* $p < 0.1$

Performance

Errors

Knowing of a dependency between time and errors, we decided to omit the variable of time in our statistical analysis, because it is typically the error that causes an increase in time.

For the follow-up comparisons of the error scores (see Table 7 and Figure 3), we focused on RAP D, which clearly ranked the best (notwithstanding RAP B). Without the dosimeter attached, only the difference between RAP B and RAP D was significant in the paired comparisons. With the dosimeter, RAP D had significantly fewer errors than RAP C (see Table 8).

Table 7. Mean errors with and without dosimeter; $n = 22$.

	Mean	Std dev	Rank
RAP A (with)	1.227	2.287	6
(without)	0.455	1.057	5
RAP B (with)	0.000	0.000	1
(without)	0.000	0.000	1
RAP C (with)	0.773	1.343	3
(without)	0.318	0.568	4
RAP D (with)	0.182	0.501	2
(without)	0.182	0.589	2
RAP E (with)	1.046	1.046	5
(without)	0.364	0.954	6
RAP F (with)	0.864	1.283	4
(without)	0.273	1.077	3

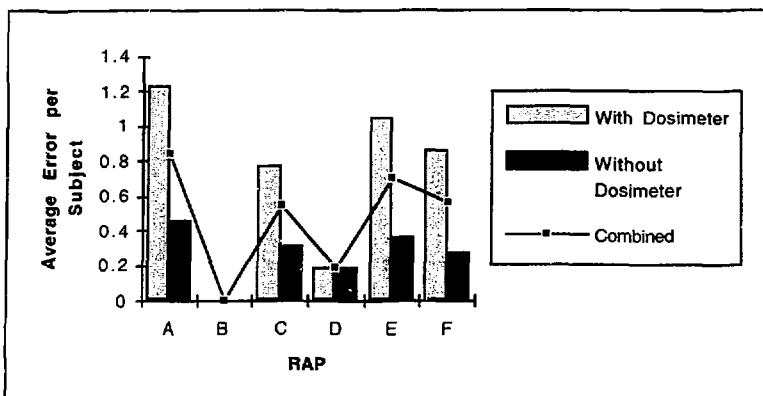


Figure 3. Mean errors with and without dosimeter; $n = 22$.

Table 8. Simple comparisons for error mean group differences; $df = 21$.

	RAPs compared	Absolute t value
With dosimeter	RAP B by RAP D	1.70*
	RAP D by RAP C	1.89*
	RAP C by RAP F	0.21
Without dosimeter	RAP B by RAP D	1.45*
	RAP D by RAP F	0.34
	RAP F by RAP C	0.17

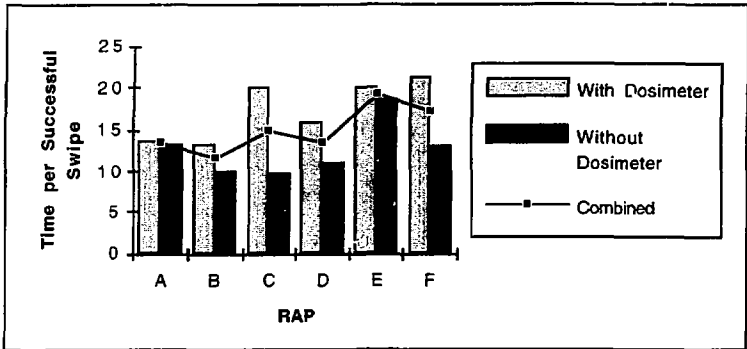
* $p < 0.1$

Time

Correlated time patterns, like error ranking, can be seen in the average time data (see Table 9 and Figure 4). Whenever we consider performance data, our focus is on errors rather than time. Time was not included in the main statistical analysis, due to its nonorthogonal nature and the time data's departure from the assumptions of a normal distribution (high degree of positive skew). We assumed a strong positive correlation between total number of errors and time, but the presence of an insignificant correlation ($r = .232$) suggests a degree of confounding rather than a true degree of independence. Note that the error tabulations do not include the dosimeter eyehole errors, while the timing tabulations do.

Table 9. Mean time for four successful swipes (with and without dosimeter); n = 22.

	Time (sec)	Rank		Time (sec)	Rank		
RAP A	(with)	13.73	2	RAP D	(with)	16.05	3
	(without)	13.20			(without)	11.00	3
RAP B	(with)	13.30	1	RAP E	(with)	20.06	4
	(without)	10.06			(without)	18.55	6
RAP C	(with)	20.15	5	RAP F	(with)	21.35	6
	(without)	9.73			(without)	13.13	4

**Figure 4.** Mean time for four successful swipes (with and without dosimeter); n = 22.

Discussion and Recommendations

On the basis of the objective performance and preference measurements taken under limited sample size, we recommend that RAP D be the candidate for implementation. We qualify our recommendation by pointing out that our finding of statistical significance may lack statistical power (1-beta), due to the relatively small sample sizes used.

The consequences of failing to heed our recommendation and choosing another design may be relatively low, because the mean performance differences between designs was numerically marginal, even though statistically significant. For example, when we used a matched t-test to compare RAP B with the experimental RAP's configuration of lowest frequency of errors and highest preference scores, RAP B was always statistically significant (fewer errors, more preferred). This result is quite probably due to user familiarity with the RAP B configuration. Only two errors were recorded for RAP B during the entire study: one parity error, and one no-read when the dosimeter blocked full insertion.

RAP D

RAP A, with the highest combined score, had a high error rate, which minimized it as a viable alternative. RAP B had the fewest errors overall (see Appendix B). With its horizontal or 0° orientation, RAP D had the second highest combined preference score (see Appendix A). RAP D was ranked in the top three most often (RAP D, 34 times; RAP A, 32 times; RAP C, 28 times).

Compared to trials without dosimeters, trials conducted with dosimeters attached produced more operation errors for all prototypes except RAP D—for which error rates were equal with or without a dosimeter. Therefore, RAP D will be a reliable performer regardless of the fate of the dosimeters. Without dosimeters, RAP F had one less error than RAP B (3 vs 4), but RAP F had the second highest number of errors with the dosimeter and was ranked fourth or fifth more often than first or second.

RAP D seemed to be the easiest to understand and the most successful to operate. It was most often subjectively rated the most quick and efficient to use. It also matched the subjects' intuitive sense of magnetic strip orientation 20 out of 22 times, topped only by low-ranked RAP E (see Appendix C). Throughout testing, RAP D was among those mentioned by subjects as most favored, and was in all instances rated as satisfactory.

Though the time data ranks RAP D behind RAP A when the dosimeter is attached, and behind RAP C when it is not, RAP D's consistent standing as third (behind RAP B and C, and C and D, alternatively) gives it the higher overall rating. RAP D is higher because when the dosimeter is switched, both RAP A and RAP C fall to the fifth rank.

Note that an added benefit to a horizontally mounted reader is the reduction in the amount of dust and debris entering the slot.

Other Recommendations

The following recommendations are based on our observations and the subjects' input during the trials.

Color

We recommend that the reader be designed with contrasting colors highlighting the location and line of the slot. The black-on-white look of the other RAPs, especially A and F, would solve two problems if also incorporated into RAP D. The positioning of the slot would then be more visually apparent, (1) increasing ease of operation, seen in the ability to use the device without training, and (2) thus decreasing errors due to insertion and alignment problems.

Black-on-white or black-on-beige color is adequate. Ideally, the slot should be color coded with the instructional display, so that a red arrow indicating direction of swipe would correspond to a red slot. However, practically and aesthetically speaking, this may not be necessary.

Tapered Channel

We recommend a short, tapered entry channel that visually indicates where the badge should be entered and that physically guides the badge. However, we do not recommend that the tapered channel require the tight and elongated swipe required in the E RAP. The channel should be a slight extension of the sort already existing on the swipe readers. Perhaps the channel should not begin with the same depth as the rest of the slot, but rather begin as a shallow track that gradually increases in depth. An important aspect of this channel should be the flat surface against which the badge rests. This acts as a sort of table upon which the swipe is begun, and then momentarily acts as a guide to assure the badge is swiped flatly and evenly.

Swipe Direction and Magnetic Strip Orientation

Subjects seemed to prefer having a choice in swiping their badge left to right or vice versa, although when asked to choose, in nearly every instance they preferred left to right, which afforded them a sort of "backhand" swipe. Note that all but one subject was right-handed.

Magnetic strip orientation should be down and away from the user.

Surface

The plastic swiping surfaces should always be hard and smooth. A "sandy" surface perhaps looks better, but it also increases the friction between badge and reader, as well as wearing down the badge over time.

Dosimeter Problem

It is not necessary that any additional modeling be done to provide a method to push the dosimeter aside as it is swiped. We say this not only because the dosimeter may be phased out in coming months—the structure of the swipe reader naturally blocks the dosimeter, so that when the card is swiped with the dosimeter trailing the badge, the dosimeter is naturally bumped out of the way.

Display

We recommend that a *simple* instructional display be placed on the wall directly over the RAP, depicting a hand that is holding the badge with the photo-side up. The display should also depict an arrow, indicating swiping from right or left. (For LLNL, the display should indicate swiping from left to right. When the dosimeter is no longer used, then an either-direction display can be posted, or a display can be omitted.)

Outdoor Use

When located outdoors, a RAP must be enclosed to protect the badge reading device from the weather. Gary Michalak's proposed housing—where the entire RAP housing is hinged and opens away from the wall, thus removing its sides as an impediment against swiping—is an excellent recommendation.

To facilitate either-direction swiping, having the housing split open—half moving left, half moving right—is another idea. The motion required to open the housing should be a one-handed process, with the movement of both sides of the housing mechanically linked so that the two move in tandem.

Other RAPs and Design Flaws Discussed

RAP F and RAP A had positive aspects that may have been somewhat confounded by shortcomings in the study.

RAP F

The reader's 45° angle used on RAP F was a plus to some and, to about the same number of people, a big minus. To use the 45° reader angle, holding the badge with fingers on top and thumb underneath allows a surer grip and seems to produce steadier swipes—resulting in fewer alignment and time-out errors; however, this was not an intuitive way to hold the card for most users. The alternative, with the thumb on top, forced an unnatural backward bend of the wrist.

An instructional display could be used to enhance the probability that more users would hold the badge correctly. However, we recommend that the most intuitive and natural positioning be presupposed in designing the reader, to ensure that most users are comfortable with the least information. Further research and testing would be required to determine whether an optimal angle would be helpful, but without that additional study, we recommend RAP D.

The design of RAP F (with the magnetic strip face up) presented a serious problem for test subjects to determine the correct side of the badge for swiping and was obviously the least intuitively understood by most test subjects. A downward-facing orientation of the strip would have been more intuitive.

RAP A

RAP A was aesthetically the most pleasing, and it was also ranked highly most often. However, its error rate was very high. The height of the reader, the need to face the magnetic strip to the right (which seemed unnatural to many subjects), and dosimeter problems seemed to produce the high error rate. The RAP badge reader's position directly on the flat backing seemed to worsen the dosimeter problem by not allowing the dosimeter room to be moved away from the badge (as in the E RAP).

Note that the RAP was fitted to the wall a bit loosely; this design is intended to be built-in, but in this case was simply bolted on. This may account for some number of the observed errors.

For flush-mounted RAPs, the most important alteration should be flattening the angle of the keypad to the horizontal position. While direct data collection of the PIN entry was not part of this study, several subjects mentioned an inability to see the keys, and discomfort in attempting to enter their code.

Appendix A. Subject preference of RAPs (in frequency of subject ratings).

Times rated	RAP ¹														
	A (vertical)			C (90°)			D (0°; horizontal)			E (0°; machined)			F (45°)		
	w/2	w/o ³	Total	w/	w/o	Total	w/	w/o	Total	w/	w/o	Total	w/	w/o	Total
Quick and efficient	4	5	9	5	5	10	5	7	12	2	0	2	4	5	9
Efficient in appearance	7	7	14	3	2	5	2	5	7	3	4	7	6	5	11
Aesthetically pleasing	7	8	15	1	1	2	3	5	8	5	4	9	5	3	8
Recommendable	3	5	8	5	6	11	5	6	11	1	1	2	4	5	9
Totals	21	25	46	14	14	28	15	23	38	11	9	20	19	18	37
Ranked 1st	5	7	12	4	5	9	6	6	12	1	0	1	5	4	9
Ranked 2nd	6	7	13	4	4	8	5	3	8	1	4	5	5	4	9
Ranked 3rd	4	3	7	4	7	11	7	7	14	5	2	7	1	3	4
Ranked 4th	3	1	4	4	3	7	3	3	6	6	7	13	5	7	12
Ranked 5th	2	2	4	3	3	6	1	2	3	5	6	11	5	3	8
Combined⁴	22	27	49	14	15	29	19	21	40	9	8	17	19	16	35
Rank (combined)	1	1	1	4	4	4	2*	2	2	5	5	5	2*	3	3

¹If tested, we believe that RAP B would rank third or fourth aesthetically and would rank among the top three overall. However, RAP B was not offered as an option in our evaluations, so that we might avoid prior experience/exposure as a contaminant of our testing.

²With dosimeter.

³Without dosimeter.

⁴Times chosen most efficient, most aesthetic, and preference-ranked 1 or 2.

*Tied for 2nd rank.

Appendix B. Errors by RAP and type.¹

Type of Error	RAP															Total Errors (%)		
	A (Vertical)			C (90°)			D (0°; Horizontal)			E (Machined)			F (45°)			w/	w/o	Total
	w/2	w/o ³	Total	w/	w/o	Total	w/	w/o	Total	w/	w/o	Total	w/	w/o	Total			
Time-out	2	1	3	4	1	5	1	2	3	8	0	8	3	0	3	18	4	22
Parity	16	9	25	11	3	14	3	0	3	10	7	17	12	2	14	52	21	73
Alignment	3	0	3	1	0	1	0	2	2	3	1	4	4	1	5	11	4	15
Too Few	3	0	3	1	0	1	0	0	0	0	0	0	0	0	0	4	0	4
No Read	0	0	0	0	2	2	0	0	0	0	0	0	1	0	1	1	2	3
Totals	24	10	34	17	6	23	4	4	8	21	8	29	20	3	23	86	31	117
Ratio (errors:success)	.273	.113	.193	.193	.068	.131	.045	.045	.045	.239	.091	.165	.227	.034	.131	.195	.070	.133
Rank	5	5	5	2	3	2*	1	2	1	4	4	4	3	1	2*			

¹These tabulations assume that the badge had been entered with the proper orientation, that is, with the magnetic strip in the proper position. Errors were counted individually; for example, three misaligned swipes were recorded as three alignment errors.

²With dosimeter.

³Without dosimeter.

*Tied for 2nd rank.

Appendix C-1. *Ease of procedure: subject's intuition.*

	RAP				
	A (Vertical)	C (90°)	D (Horizontal)	E (Machined)	F (45°)
Mag strip required to face:	Left	Away from user	Down	Down	Up and away
Subjects' "intuition"	left: 11	away: 15	up: 3	up: 2	up/away: 12
	right: 11	toward: 7	down: 19	down: 20	down/toward: 10
Direction of subjects' swipe	down: 21	left-right: 9	left-right: 16	left-right: 16	left-right: 10
	up: 1	right-left: 13	right-left: 6	right-left: 6	right-left: 12

Appendix C-2 was compiled at the end of several assessments when subjects were asked to reconsider which direction of swipe and orientation of the badge (magnetic strip) was most intuitive and most comfortable.

The purpose of this data is twofold. First, the configuration chosen for implementation will have accompanying procedural displays, but the process should be intuitive nonetheless. Second, this data is relevant to badge design issues, such as the dosimeter eyehole that partially cuts through the magnetic strip on the badge, and the possibility that the RAPs will be placed in boxes or in other confined spaces that may limit accessibility from one or the other side.

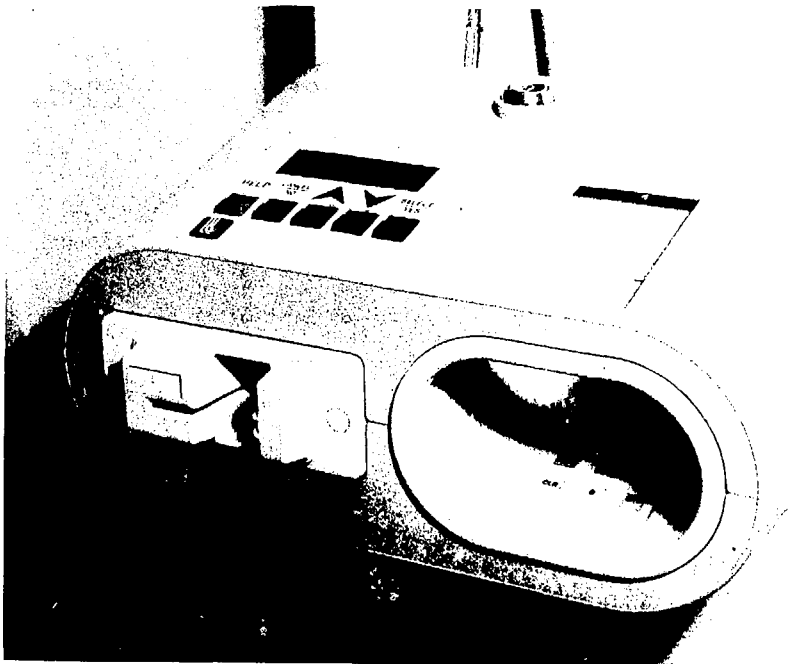
Appendix C-2. *Ease of procedure: subject's preferred procedure.*

	RAP		
	A (Vertical)	D (Horizontal)	F (45°)
Mag strip required to face:	Left	Down	Up and away
Subjects' "intuition"	left: 6	up: 1	up/away: 3
	right: 3	down: 6	down/toward: 7
Direction of subjects' swipe	down: 3	left-right: 7	left-right: 6
	up: 0	right-left: 7	right-left: 12
			no preference: 2

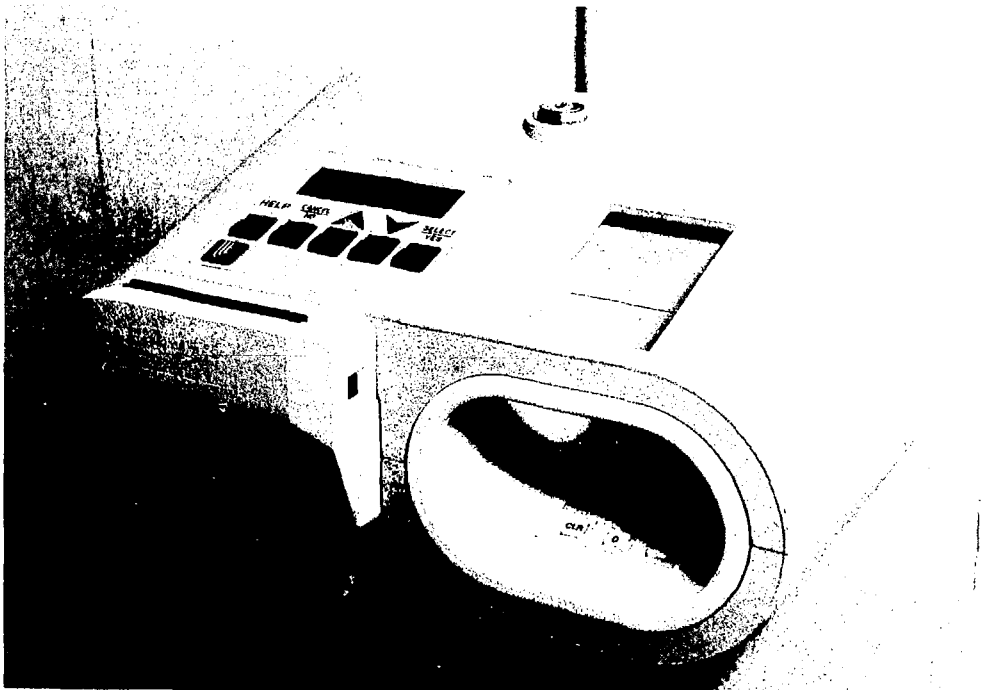


RAP A. Vertical (flush-mounted).

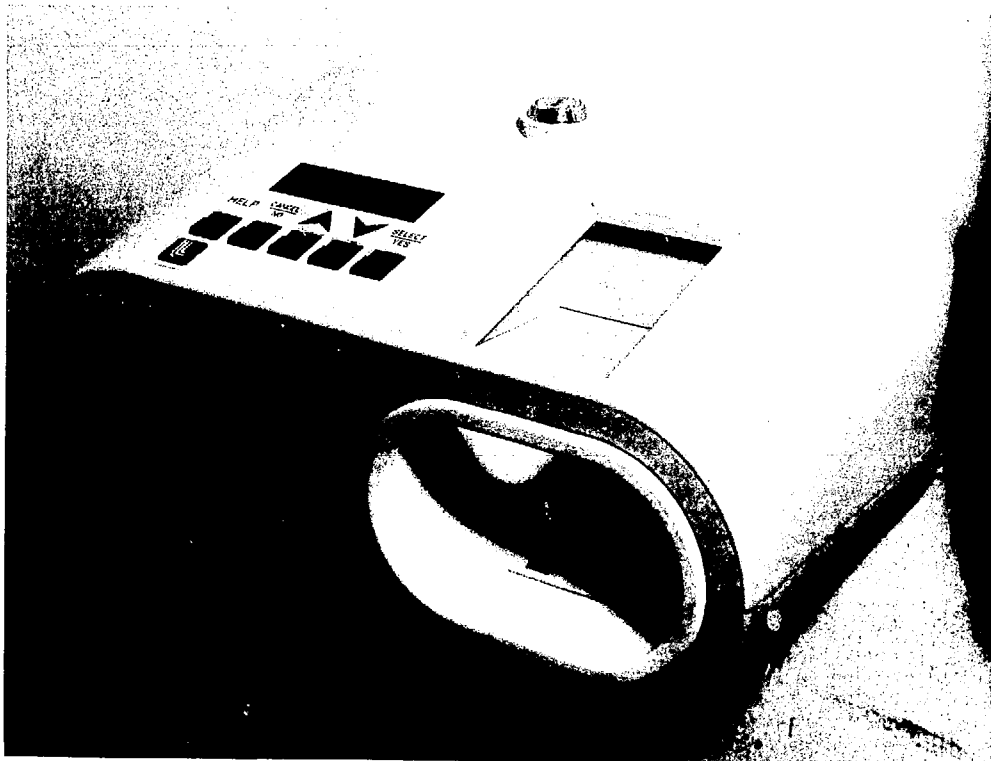
(This RAP is modified with a swipe reader. The user instructions on the RAP still show an insert-style reader.)



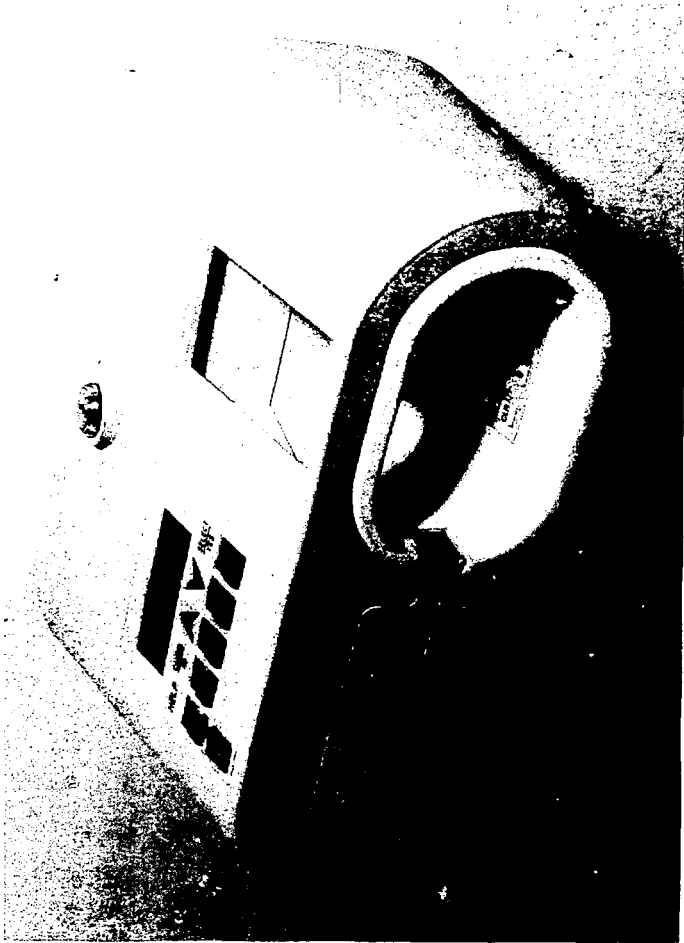
RAP B. *Horizontal insert (LLNL).*



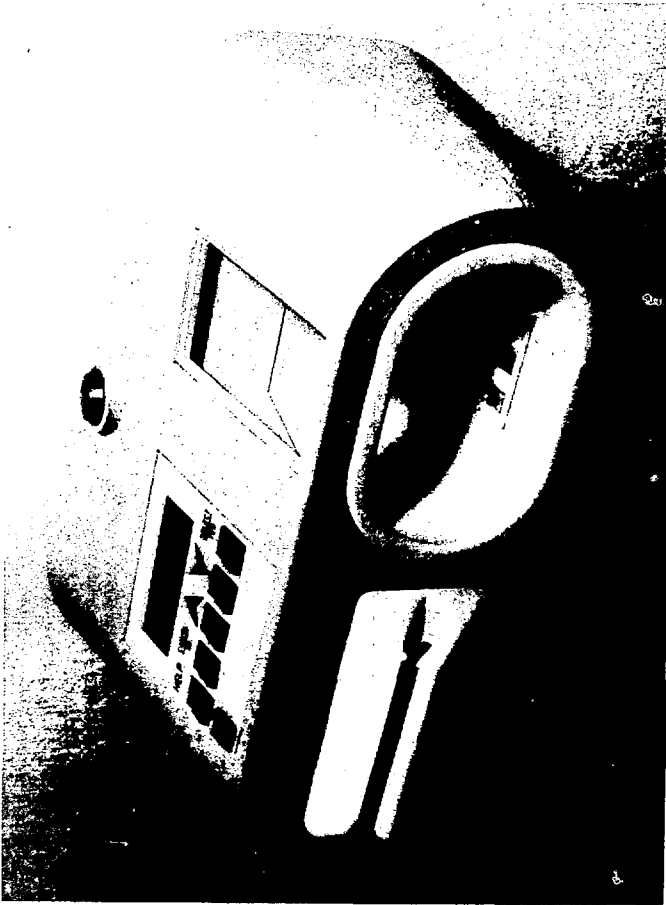
RAP C. Facing up 90° vertical.



RAP D. *Protruding horizontally; 0°.*



RAP E. Flush-mounted, horizontal; 0°.



RAP F. Angled up 45°.

**Technical Information Department • Lawrence Livermore National Laboratory
University of California • Livermore, California 94551**