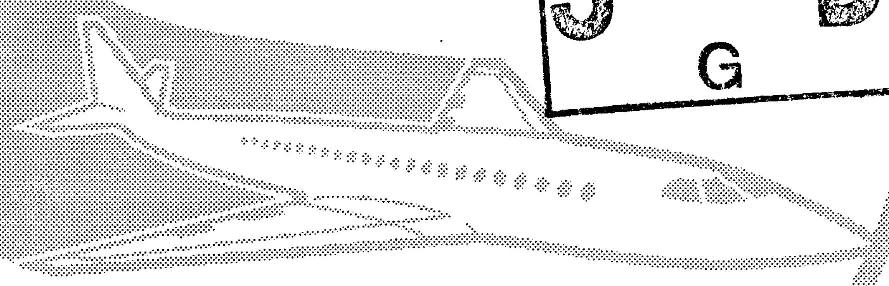
DOT/FAA/CT-95/41

FAA Technical Center
Atlantic City International Airport,
N.J. 08405

Initial Development of Selection Instruments for Evaluating Applicants As Potential Airline Passenger Baggage Screeners For Conventional X-ray Technology

J. L. Fobes, Ph.D.
D. Michael McAnulty, Ph.D.



July 1995

Final Report

This document is available to the U.S. public through the National Technical Information Service, Springfield, Virginia 22161.

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/CT-95/41		

4. Title and Subtitle	5. Report Date
INITIAL DEVELOPMENT OF SELECTION INSTRUMENTS FOR EVALUATING APPLICANTS AS POTENTIAL AIRLINE PASSENGER BAGGAGE SCREENERS FOR CONVENTIONAL X-RAY TECHNOLOGY	July 1995
	6. Performing Organization Code AAR-510
7. Author(s) J. L. Fobes, Ph.D., and D. Michael McAnulty, Ph.D.	8. Performing Organization Report No.
9. Performing Organization Name and Address Federal Aviation Administration Technical Center Aviation Security Research and Development Division Atlantic City International Airport, NJ 08405	10. Work Unit No. (TRAIS)
	11. Contract or Grant No. 93-G-037
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Aviation Administration Office of Civil Aviation Security 800 Independence Ave., S.W. Washington, DC 20591	13. Type of Report and Period Covered Final
	14. Sponsoring Agency Code ACS-1

15. Supplementary Notes

Report prepared by: Gerald D. Gibb, James S. Baker, Animesh Banarjee, Don Bennett, and Sam C. Kelly III, of Embry-Riddle Aeronautical University

16. Abstract

Currently there is a paucity of research that has been conducted to improve the selection process for airline passenger baggage screener personnel. Careful selection of these personnel is important to obtain the best qualified personnel and to mitigate the characteristically high employee turnover rate that is associated with the training losses in this industry.

This report describes the development and evaluation of two computer-based cognitive instruments that were examined as possible selection tools. The results from a concurrent validity pilot study demonstrated a strong relationship between measures from the predictor instruments and selected performance criteria.

17. Key Words	18. Distribution Statement
Personnel Selection, Aviation Security, Hidden Figures Test, Test Development	This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
UNCLASSIFIED	UNCLASSIFIED	23	

Form DOT F1700.7 (8-72) Reproduction of completed page authorized

Introduction

This report summarizes the initial development of two instruments that were proposed as possible selection tests for airline passenger baggage screener personnel. A significant portion of this work describes the procedures employed and results found to assess the psychometric properties and concurrent validity of the two instruments. The work was conducted as a precursor to performing a predictive validation study to mitigate the costs involved in such an effort and to ensure the best possible chances for screener success with conventional X-ray technology.

Considerable emphasis in this phase of the research program was invested in ensuring the instruments (a) demonstrated an acceptable range of individual differences, (b) displayed comparatively brief learning curves, (c) were internally consistent, and (d) demonstrated some degree of concurrent validity. An additional concern was to evaluate whether alternative strategies could be employed by examinees to effectively circumvent the purpose of the test instruments. Substantial effort was devoted to developing automated instruments that had easily understood, self-contained instructions to eliminate the need for an administrator.

The research basis for the development of these instruments can be found in two previous Federal Aviation Administration technical reports (Lofaro et. al., 1994a; 1994b). These reports address findings from a review of the related literature and a job task analysis, respectively. Furthermore, additional research to support the development of these instruments, and the underlying cognitive abilities, is based on work by Cantor (1994). In a series of studies with airport security personnel and cargo ship inspectors, Cantor found a relationship between target detection performance and assessments of field dependence/independence on the Embedded Figures Test (EFT). The results of his work indicated the potential for predicting target detection performance made using scores from this assessment tool. The EFT is a paper-and-pencil instrument (Witkin, Oltman, Raskin, & Karp, 1971) that requires respondents to disembed geometric figures from complex backgrounds.

The test battery was intentionally kept short in consideration of the operational requirement to reduce the time and cost involved in selecting airline passenger baggage screeners. In addition, previous results from the literature and work conducted in this research program did not reveal many abilities and traits required for success in this occupational field.

Method

Subjects

In the initial evaluation of the instruments, 67 undergraduate college students participated as volunteers. Thirty subjects were used in evaluating the Hidden Figures Test (HFT) and 37 subjects participated in the effort for the Hidden Patterns Test (HPT). No student was administered both tests.

In the concurrent validation phase of the research, 25 certified airport screeners employed by ITS were administered both of the tests. There were 11 male and 14 female screener subjects.

All screeners were employed in aviation security for a minimum of 8 months and had previously completed specialized training in weapons detection in July, 1994. The specialized training consisted of about 90 minutes of computer-based training on improvised explosive device (IED) detection conducted on the EG&G Linescan TnT system. This system depicts X-ray images from an IED library and trains the screeners in IED recognition. Complete screener IED detection performance data were available for 20 screeners and partial performance data were available for 4 more screeners. The data for two subjects was lost from computer storage on one test instrument.

Instruments

Hidden Figures Test. This version of the Hidden Figures Test (HFT) is a two-part instrument with 32 items equally divided between the parts. The instrument is a computer-based administration of the published paper-and-pencil test (Ekstrom, French, Harmen, & Dermen, 1976). Computer-based administration was selected to permit the assessment of reaction time measures in addition to the standard accuracy measures.

The instrument is introduced to subjects as "a test of your ability to find which one of five simple figures can be found in a more complex pattern." The administration is completely automated and requires no experimenter intervention once initialized. Unlike the paper-and-pencil instrument, only one test item is presented at a time. Subsequent test items are not presented until a response has been entered for the previous test item, which is permanently removed from the computer monitor. Each item presentation maintains the response set of five target patterns horizontally across the top of the computer monitor. Each of the five response targets are numerically labeled from 1 through 5 directly beneath the corresponding target patterns. The response set remains on the monitor any time the test administration is in progress and a test item is present. The complex background patterns are presented beneath the response set at approximately the center of the screen. The complex background patterns vary in size, shape, and complexity. Subjects are directed to respond by entering their choice on the computer numeric keypad.

The current version also differs from the original version in that the example items are dynamically displayed for subjects by first presenting the complex background pattern and then highlighting the target figure in color. The subjects perform two example items. Feedback is provided after each example response. For both correct and incorrect responses, the correct response (target pattern) is highlighted in color on the background pattern. Correct responses are provided with the textual feedback, "That's correct," on the monitor. Incorrect responses are provided feedback on the monitor with the text, "Sorry, only the (correct response number) figure can be found in the pattern."

Several other directions appear interspersed in the instruction set and are provided to the subjects before beginning the test. The following directions are all provided textually via the computer software and are subject paced.

1. Subjects are informed there is only one of the target figures in each complex pattern.

2. Each target figure will always be right side up and exactly the same size as the one in the complex pattern.

3. No additional lines can be added to the complex pattern. The target figure must be traceable on the existing complex pattern. This is further demonstrated graphically in conjunction with the textual instruction.

4. A score is calculated by subtracting the number of incorrect responses from the number of correct responses. Subjects are informed that it is not to their advantage to guess unless they are able to eliminate at least some of the possible responses that they know are wrong.

5. Information is provided that advises there is no time limit, but to work quickly because response times are assessed and are therefore important.

6. The number of test items is clearly identified.

Because reaction times are recorded, each part of the test begins with a 10 s countdown that is prominently displayed in the center of the monitor with the preface, "Test will begin in __ seconds." The textual command "Begin" appears 1 s after the countdown is completed, followed 500 ms later by the first test item. A 10 s rest period occurs between the first and second parts of the test.

Hidden Patterns Test. The Hidden Patterns Test (HPT) was developed as a two-part instrument with 100 items equally divided between the parts. Test items do not increase in difficulty level throughout the test. This assessment tool is a computer-based administration of the paper and pencil version developed by the Educational Testing Service (Ekstrom et al., 1976). The computerized version of the instrument permits assessment of both accuracy and response time measures.

The HPT is introduced as "a test of the ability to recognize a figure that is hidden among other lines." Instructions inform subjects that speed of recognition and response is important, but not to sacrifice accuracy.

Similar to the HFT, the HPT is fully automated. Test items are displayed one at a time on the computer monitor and are removed once a response is made. Each trial displays the target figure prominently in the center of the monitor when the test administration is in progress. The target figure (to which the test item is compared) appears 3 cm below the test item. Both test items and target figures are identical to those used in the original paper-and-pencil instrument.

The design of the instrument is similar to the automated HFT in that (a) examples are dynamically presented, (b) examples use a color trace of the test item for illustration, (c) a red 'x' is displayed during example items when the test item cannot be found in the target figure, and (d) all responses are made using a numeric keypad. Seven examples of the test item appearing within the target figure are provided as are four examples when a match is not present. All examples are self-paced and are presented in multiple orientations.

The instructions further inform subjects that the test item may be found in the target figure in an upside down, rightside up, or rotated configuration. Similar to the HFT, an individual's performance is evaluated as the number of correct responses minus the number of incorrect responses. Subjects are advised that it is not to their advantage to guess unless they can eliminate one of the choices because the distribution of matched versus unmatched items is disproportionate.

Each part of the test begins with a 10 s countdown identical to that used with the HFT. A brief, subject-controlled, rest period also occurs between the two parts of the test. Only response keys 1 and 2 are activated after the instruction set terminates. Subjects are instructed to depress the number 1 key for those patterns in which the model appears and the number 2 key when the model does not appear in the pattern.

Procedures

All testing was conducted in small, comfortable, quiet office spaces with a test administrator nearby. The instruments were administered using IBM-486 microprocessors with standard computer keyboards and 17-inch diagonal color monitors.

Psychometric Evaluation. Undergraduate college students were briefed that the study involved evaluating applicant selection instruments for airport passenger screener positions. The subjects were informed that the data were being analyzed only to assess the psychometric properties of the instruments. Following the briefing, the test administrator initiated the automated testing paradigm. Because the instruments had self-contained instructions and demonstrations, no further communication between subjects and the test administrator was necessary.

After completing the tests, the test administrator debriefed the participants. The debrief focused on the cognitive processes used by each subject to perform the HFT. The research team was particularly interested in the cues and techniques employed by subjects to solve the test problems. The purpose of the debrief was to assess if alternate strategies could be used that circumvent the cognitive dimensions of interest.

Concurrent Validation. The preadministration briefing to the screeners differed substantially from that provided to the college students. Because the screeners were currently employed in aviation security positions, great care was taken to assure participants that the results would not affect their jobs and the data would not be available to their employer. The participants were further assured that their performance was not evaluated and entered into their training or performance records. The participants were told they were selected only because of their participation in an earlier FAA training study in July 1994, and that this effort was conducted to evaluate the relationship between their performance on the current test instruments and their performance during the earlier training program. All screeners agreed to participate in this effort.

The screener subjects were not debriefed after completing the test battery. Performance feedback was not provided to any of the screener subjects.

During the test administration procedures at San Francisco International Airport, both experimenters observed that multiple test items would be presented in rapid succession as a result of a single keypress. It was clear that some subjects were maintaining pressure on the response keys longer than necessary when entering a response. The software was written such that unintentional responses would be recorded under these circumstances and thereby advance the presentation of items without subjects intentionally making appropriate decisions. These observations were later confirmed during the data analysis. A review of all individual responses indicated reaction times well below the expected threshold (i.e., 160 - 300 ms.). Data collected using the two test instruments in another study revealed a similar problem.

The software was modified after data collection to ignore responses with reaction times below 550 ms. This was accomplished by writing a timing subroutine that was activated after each response was entered. Installing this code eliminated the problem of inadvertent responding. In addition to this software modification on each instrument, the researchers developed software code to "lockout" all responses that were not part of the response set.

Results

Several analyses were conducted on both instruments to assess the psychometric properties of each one. These analyses were used to determine if the change from a paper-and-pencil format to a computer-based presentation adversely affected the test characteristics. Of key interest were the following: (a) the capability to measure individual differences, (b) the presence of speed-accuracy trade-offs, and (c) the internal consistency of the instruments. Posttest subject debriefing data were also examined to determine if alternate strategies were employed by the subjects to circumvent the assessment of the intended abilities. These results are presented in the subsections entitled HFT and HPT.

Further analyses were conducted to determine the concurrent validity of the instruments. Performance data on detecting IED's from 24 screeners at San Francisco International Airport were compared with performance on each of the two instruments. These results are presented in the last subsection.

HFT

The means and standard deviations for accuracy score for the student group ($\underline{n}$= 30) were M =7.00, SD = 9.03, and M = 5.13, SD = 8.70 for parts one and two, respectively. The HFT has standard deviations that are larger than the means because of the range of scores possible. A score is calculated by subtracting the number of incorrect responses from the number of correct responses, so the scale range is -32 to +32. The large standard deviations in relation to the means indicate this instrument has considerable individual variability. Correct and incorrect reaction times for the two parts were M = 47.69 s, SD = 23.43 s, and M = 72.84 s, SD = 48.47 s, respectively. Overall, the HFT had a mean correct reaction time of 49.02 s, SD = 15.25 s.

An examination of individual differences was further pursued by comparing mean accuracy scores of both parts of the instrument by dividing the sample into upper and lower quartile groups using the accuracy score data. The mean accuracy scores for both parts of the instrument for the upper quartile group ($\underline{n} = 7$) were M = 16.00, SD = 0, and M = 13.14, SD = 1.07. Respective mean accuracy scores for the lower quartile group ($\underline{n} = 7$) were M = -6.29, SD = 2.93, and M = -8.86, SD = 3.02. T-tests for differences yielded $\underline{t}$ (12) = 18.15, ($\underline{p}$<0.001) for part one and $\underline{t}$ (12) = 20.14, ($\underline{p}$< 0.001) for part two of the instrument in comparing quartile groups. The computer-based version of the HFT is effective in discriminating performance among examinees.

Item analysis between the quartile groups indicated that all 32 items were predictive of performance differences, albeit of different magnitudes. All but two items resulted in differences at a minimum of the 50% level. That is, at least half of the lower quartile group scored incorrectly on all items that the upper quartile group responded to correctly. Based on these results, all 32 items were retained for the final instrument.

Although no attempt was made to vary the difficulty level of test items, it was clear that the items were not of equal difficulty. As expected, the more complex patterns in the HFT were comparatively more difficult, as can be seen from the reaction time data (see Appendix A). Reaction time data provided a direct measure of item difficulty level.

A split-half reliability analysis was also conducted to assess the internal consistency of the instrument. The first and second parts of the instrument were significantly correlated, $\underline{r}$ = 0.92 ($\underline{p}$<0.01).

The presence of a speed-accuracy trade-off strategy in performing the test was assessed by examining the relationship between accuracy score for both parts of the test combined and mean correct reaction time across all 32 items. The Pearson correlation coefficient was not sigificant. This result was expected because the test instructions directed subjects to maintain a balance between the two performance goals.

Finally, no evidence was found to suggest that subjects were using alternative strategies to perform the HFT. The data from the debriefings (see Appendix B) indicated that subjects performed the HFT by matching geometric shapes, lines, or angles to solve the problems.

<u>HPT</u>

The means and standard deviations of the undergraduate college group for the accuracy score ($\underline{n} = 37$) were M = 38.27, SD = 7.10, and M = 41.13, SD = 4.93 for the first and second parts, respectively. These data indicate that the instrument is sensitive to individual differences as it demonstrates moderate variability in the scores. Correct and incorrect reaction times were M = 2.82 s, SD = 1.23 s, and M = 4.05 s, SD = 2.36 s, respectively for both test sections collapsed.

The sample was divided into upper and lower quartile groups using the mean accuracy measures of both parts of the test to examine individual differences further. The mean accuracy measure for both sections of the instrument for the upper quartile group ($\underline{n} = 9$) were M = 45.33, SD = 2.82, and M = 47.11, SD = 2.77. Respective mean accuracy scores for the lower quartile group ($\underline{n} = 9$) were M = 28.00, SD = 2.20, and M = 34.66, SD = 3.34. T-tests for differences between quartile groups yielded $\underline{t}$ (16) = 11.34, ($\underline{p} < 0.001$) for section one, and $\underline{t}$ (16) = 16.51, ($\underline{p} < 0.001$) for section two of the test.

The split-half reliability between the first and second sections of the HPT using mean accuracy scores was $\underline{r} = 0.44$ ($\underline{p} < 0.01$). Although significant, there is only moderate internal reliability across the instrument. The Pearson Product-Moment correlation coefficient between the total accuracy measure for both sections of the test collapsed against mean correct reaction time was $\underline{r} = 0.35$ ($\underline{p} < 0.05$). These data indicate there is a moderate speed-accuracy trade-off for this instrument.

Concurrent Validity Analyses

Complete test battery and performance evaluation data sets were available for 24 certified airport screeners that had completed training. Because of the problem found with inadvertent responding due to the excessive pressure maintained on the response keys by some of the subjects in this sample, the data were first reviewed to eliminate invalid responses. All responses below 500 ms were discarded from the data set. Because this resulted in an unequal number of test items presented across subjects, the data were converted to a percent correct (PC) measure by dividing the number correct by the number of valid responses entered. Means and standard deviations for PC were M = 41.18, SD = 30.64, and M = 78.52, and SD = 20.28, for the HFT and HPT instruments, respectively. Means and standard deviations for the mean correct reaction time measures (MCRT) were M = 41.02 s, SD = 34.41 s, and M = 4.76 s, SD = 1.93 s, for the respective instruments.

The primary measure of screener job performance was the Probability of Detection (Pd). Pd is a measure of the number of targets detected divided by the number of targets. The mean Pd before training was M = 26.6, SD = 22.3; the mean Pd after training was M = 43.1, SD = 29.6. A second measure of performance utilized was operator sensitivity (d'), which is calculated using Pd and the Probability of False Alarm (Pfa). This measure is derived from Signal Detection Theory (SDT). Pretraining mean d' was M = 0.97, SD = 0.70. After training, mean d' increased to M = 1.44, SD = 0.87.

Pearson correlation coefficients were calculated to examine the relationship between the test battery measures and the performance criteria of Pd and d'. The HFT measure demonstrated moderate-to-strong relationships to each of the criteria, both before and after the training. The correlation between the HFT PC and Pd was $\underline{r} = 0.75$ ($\underline{p} < 0.005$) and $\underline{r} = 0.59$ ($\underline{p} < 0.005$) for pretraining and posttraining conditions, respectively. Correlations between the MCRT measure and Pd were $r = 0.34$ ($p < 0.05$) and $r = 0.24$, NS, for the two training conditions.

In examining the relationship between the HFT measures and d', correlations of $r = 0.77$ ($p < 0.005$) and $r = 0.46$ ($p < 0.02$) were found for PC under the pretraining and posttraining conditions. The relationships between the MCRT measure of the HFT and d' were substantially weaker: $r = 0.43$ ($p < 0.02$) for the pretraining condition and $r = 0.34$ ($p < 0.05$) under the posttraining condition.

The HPT demonstrated a comparatively weaker relationship to the performance criteria measures. The correlations between the PC measure and Pd for the pretraining and posttraining condition were $r = 0.09$, NS, and $r = 0.44$ ($p < 0.02$). Correlations for PC and pretraining and posttraining d' were $r = 0.38$ ($p < 0.03$) and $r = 0.54$ ($p < 0.005$). No significant relationships were found between the MCRT measure and any of the performance criteria.

Before performing multiple regression analyses to assess the effect of combining the test battery measures against the various performance criteria, the researchers examined the relationship between the two test instruments. The correlation between the PC measures was $r = 0.29$; between the MCRT measures it was $r = 0.13$. Neither coefficient is significant, indicating that combining measures from both instruments may increase the observed relationship with the performance criteria.

A multiple regression model for both pretraining and posttraining Pd that included accuracy scores and correct reaction times from both instruments yielded R^2s of 0.58, $F(4, 16) = 5.54$, ($p < 0.005$), and 0.458, $F(4, 17) = 3.60$, ($p < 0.03$), respectively. The multiple regression models for d', pretraining and posttraining, that included the same predictor variables yielded R^2s of 0.658, $F(4, 15) = 7.23$, ($p < 0.002$), and 0.431, $F(4, 14) = 2.65$, NS. The regression equations for these models are shown in Table 1. These data indicate that the predictor variables from the test instruments account for considerable variance in performance. However, all subjects were previously trained and experienced in the airline passenger baggage screening position before being participants in the training study or this current effort. The test instruments should yield similar or stronger results in a predictive validation effort that utilizes newly hired employees as subjects.

Finally, the researchers examined the relationship between the performance measures from the two training conditions. The correlation between the pretraining and posttraining conditions for Pd and d' were $r = 0.61$ ($p < 0.001$) and $r = 0.47$ ($p < 0.01$), respectively. The correlation between the pretraining d' and posttraining Pd was $r = 0.52$ ($p < 0.005$). The relationship between pretraining Pd and posttraining d' was $r = 0.49$ ($p < 0.01$). These data indicate a reasonable level of stability across the training conditions with respect to performance assessment.

Table 1
Regression equations for Pretraining and Post Training Performance Measures

Pd (pretraining)	= 0.15 + .00637 HFT Accuracy - .00051 HFT MCRT - .00207 HPT Accuracy + .0097 HPT MCRT
Pd (posttraining)	= .0148 + .00439 HFT Accuracy - .00042 HFT MCRT + .00418 HPT Accuracy - .0372 HPT MCRT
d' (pretraining)	= - .516 + .0183 HFT Accuracy -.00075 HFT MCRT + .00432 HPT Accuracy + .0880 HPT MCRT
d' (posttraining)	= .455 + .00319 HFT Accuracy + .00473 HFT MCRT + .0183 HPT Accuracy - .133 HPT MCRT

Discussion

The goals of this research were to evaluate the psychometric properties of two instruments developed for possible use in selecting airline passenger baggage screener personnel and to determine if further work is warranted in conducting a predictive validation study in the operational environment. The initial phase of this work, evaluating the psychometric properties of the instruments, indicated that both instruments were sensitive to individual differences as shown by the significant differences found between upper and lower quartile groups when the sample was divided based on test performance data.

Further analyses demonstrate that both instruments have acceptable levels of internal consistency, although the HPT is less consistent. Little evidence was found to indicate problems with subjects employing alternate strategies to perform the tests, or the presence of speed-accuracy trade-offs. The debrief data indicate that all subjects used a strategy consistent with the intent of the instrument. In completing the HFT subjects would match lines, simple patterns, or angles in choosing their reponses.

The results from the concurrent validity analysis indicated that a predictive validation effort is warranted. The correlations between the test measures and nearly all the performance criteria demonstrated that a moderate-to-strong relationship existed between the predictors and criteria. The PC measure of the HFT in particular was found to account for 35 - 56% of the variance for the Pd criteria.

The data from the multiple regression analyses indicated a substantial improvement in predicting the four performance criteria over the single-order correlations. The amount of variance accounted for in predicting the performance criteria ranged from 43 - 66%. This indicates that the two instruments account for unique proportions of the variance. These results justify the conduct of a predictive validation study using both instruments.

REFERENCES

Cantor, M.B. (1994). Aviation Security: The Human Factor. Security Technology & Design, pp 60-65.

Ekstrom, R., French, J., Harmen, H., and Dermen, D.. (1976). Kit of Factor-Referenced Cognitive Tests. Princeton, NJ: Educational Testing Service.

Lofaro, R.J., Gibb, G.D., Garland, D.J., Guide, P.C., Baker, J.S., & St. Laurent, M. (1994a). Review of the Literature Related to Screening Airline Passenger Baggage. (DOT/FAA/CT-94/108). Atlantic City International Airport, NJ: NTIS # AD-AZ88 651/3/XAB.

Lofaro, R.J., Gibb, G.D., Garland, D.J., Guide, P.C., Baker, J.S., & St. Laurent, M. (1994b). A Protocol For Selecting Airline Passenger Baggage Screeners (DOT/FAA/CT-94/107). Atlantic City International Airport, NJ: NTIS # AD-AZ88 649/7/XAB.

Witkin, H.A., Oltman, P.K., Raskin, E., & Karp, S.A. (1971). A Manual for the Embedded Figures Test. Palo Alto, CA: Consulting Psychological Press.

APPENDIX A

HPT and HFT Reaction Times by Items Correct and Incorrect Response

Hidden Figures Test Reaction Time Measures by Item

Test Item	Correct Reaction Time (s)	Incorrect Reaction Time (s)
1	43.26	52.78
2	44.53	35.04
3	59.01	82.39
4	25.90	37.32
5	59.04	94.34
6	44.61	54.11
7	51.62	39.28
8	66.09	76.76
9	47.83	53.73
10	62.30	31.38
11	32.53	32.67
12	43.51	25.76
13	58.94	27.81
14	47.02	48.88
15	75.75	84.28
16	65.95	63.27
	---------- Rest Interval -----------	
17	40.63	31.07
18	42.97	40.12
19	44.50	45.79
20	44.86	45.51
21	40.22	43.38
22	38.23	32.55
23	36.39	29.80
24	38.61	49.80
25	53.75	36.09
26	42.97	31.07
27	55.01	56.28
28	44.80	33.55
29	53.41	96.50
30	64.00	62.31
31	81.99	83.39
32	57.98	26.86

Hidden Patterns Test Reaction Time Measures by Item

Test Item	Correct Reaction Time (s)	Incorrect Reaction Time (s)
1	2.66	2.54
2	4.01	2.82
3	1.78	2.85
4	2.95	4.34
5	2.26	2.04
6	2.74	4.74
7	2.48	1.96
8	2.74	3.34
9	3.57	3.58
10	3.70	2.55
11	4.29	1.87
12	1.63	0.88
13	2.97	1.64
14	3.42	2.64
15	1.90	3.08
16	2.47	2.41
17	3.55	9.89
18	2.43	2.08
19	1.92	2.69
20	2.10	1.75
21	2.95	3.27
22	2.20	3.27
23	2.95	2.31
24	2.04	1.80
25	2.46	0.96
26	1.73	2.37
27	1.49	1.29
28	3.39	1.71
29	2.47	3.64
30	2.31	3.86
31	2.48	1.79
32	2.56	2.28
33	2.72	1.92
34	1.45	2.86
35	2.21	1.48
36	1.68	1.84
37	2.02	1.84
38	2.79	1.00
39	1.59	3.84
40	3.16	3.58

41	2.71	1.04
42	5.60	2.20
43	1.96	1.70
44	2.10	1.26
45	1.52	1.98
46	1.18	1.21
47	3.12	2.15
48	5.02	3.28
49	3.66	2.89
50	4.16	5.18
	------------ Rest Interval -------------	
51	2.51	3.80
52	3.16	3.68
53	4.16	2.79
54	2.70	3.40
55	3.61	4.39
56	2.02	4.52
57	2.51	5.18
58	2.07	3.72
59	4.74	4.04
60	3.14	5.27
61	3.17	3.25
62	1.65	1.21
63	2.95	7.95
64	5.02	3.00
65	1.67	3.98
66	2.47	2.83
67	2.90	1.37
68	3.31	2.31
69	1.96	2.01
70	2.26	2.51
71	2.77	2.67
72	1.78	2.68
73	2.16	2.09
74	3.13	2.58
75	3.36	3.95
76	2.61	1.97
77	2.40	2.69
78	4.65	1.41
79	4.18	4.73
80	3.60	3.00
81	2.54	4.89
82	2.95	2.80
83	2.59	2.49
84	2.50	2.27

85	2.74	3.17
86	4.10	6.72
87	4.12	2.03
88	3.35	0.82
89	2.54	2.39
90	3.40	2.53
91	2.43	1.32
92	2.21	1.32
93	2.40	2.87
94	3.03	2.22
95	2.15	2.89
96	3.55	0.99
97	2.85	1.92
98	4.68	2.32
99	3.23	2.92
100	3.03	3.42

APPENDIX B

Abbreviated Subject Responses During HFT Debriefing of HFT

RESPONSES

They just jumped out after awhile, pointed at the screen.

Looked for the longest lines and matched shapes.

It was difficult at first, then guessed a lot.

Finger pointed at the screen, some were simply visible. Matched lines.

Used 3D. Divided the complex patterns. Formed images.

Visibly chose random shapes to fit, then chose each number in sequence. The lines helped in recognition.

Demo helped. Certain aspects of each shape helped, the corners mostly. Always looked for the same shapes first.

Took one at a time and matched pieces of the shapes to the puzzle. Some just came easily.

Most were easily recognizable. Selected parts of the shape and by trial-and-error, solved them.

Matched the shapes easily. (Subject finished fast)

Shape recognition. Matched similar parts.

Recognized the shapes by combination of clues and seeing them as they occured, that is easy to see.

Matched similar lines. Traced some and some just appeared.

After a while, they just became easy to recognize.

Followed the picture scope and finger drawing. (Whatever!)

At first they were difficult, but became easy as I was able to match similar lines.

Systematically tried each shape to find each solution.

Traced each shape on the screen. (Took a long time.)

Systematically placed each shape in the pattern.

Looked for the longest lines. Looked hard.

Matched the angles. Looked for the lines.

Finger pointed, tilted head, and looked from a distance.

Using fingers to find the shapes.

Virtually placed each shape in the patterns. Matched lines (He was the fastest).

Associated parallel lines, and basic shapes. Line lengths helped.

Estimated least likely, and matched lines.

Traced some, while some just popped out.

Matched lines. Process of elimination. Distinguishing figures in certain shapes just popped out.