TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FAA-AM-72-23		

4. Title and Subtitle	5. Report Date
A STUDY OF THE STATE-TRAIT ANXIETY INVENTORY AND THE ASSESSMENT OF STRESS UNDER SIMULATED CONDITIONS	May 1972
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Roger C. Smith, Ph.D.	

9. Performing Organization Name and Address	10. Work Unit No.
FAA Civil Aeromedical Institute P. O. Box 25082 Oklahoma City, Oklahoma 73125	11. Contract or Grant No.
	13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address Office of Aviation Medicine Federal Aviation Administration 800 Independence Avenue, S.W. Washington, D. C. 20591	OAM Report
	14. Sponsoring Agency Code

15. Supplementary Notes

This research was completed under AM-A-72-PSY-34.

16. Abstract

One approach to the measurement of stress in ATC and other aviation personnel is through assessment of anxiety. One means of accomplishing this is to use the State-Trait Anxiety Inventory (STAI), a 40-item self-report technique. However, the STAI may be susceptible to effects of biasing response sets. In order to determine if stress can be simulated on the STAI, the responses of 283 subjects obtained under normal instructions were compared to their responses obtained when asked to make it appear as if they were experiencing a high degree of stress. It was found that while both the A-State (current anxiety level) and A-Trait (anxiety proneness) scales were influenced by this response set, the effect was greater for the A-State scale. Using the A-State raw score alone as an index of bias resulted in 95.7% accuracy in protocol classification. A two-step procedure, using eight items and the A-State raw score, was subsequently devised which reduced false-positive classifications of legitimate psychiatric patient profiles from 49% to 38% with no reduction in discriminatory power for "normal" records. The application of the index to clinical and research settings in aviation was discussed.

17. Key Words	18. Distribution Statement
Stress Personnel Assessment Psychological Evaluations Anxiety	Availability is unlimited. Document may be released to the National Technical Information Service, Springfield, Virginia 22151, for sale to the public.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	5	$3.00

Form DOT F 1700.7 (8-69)

A STUDY OF THE STATE-TRAIT ANXIETY INVENTORY AND THE ASSESSMENT OF STRESS UNDER SIMULATED CONDITIONS

I. Introduction.

One approach to the measurement of stress in air traffic control and other aviation personnel is through the assessment of anxiety. There are a number of physiological and psychological techniques for accomplishing such measurements; however, many involve considerable inconvenience or discomfort, time, or are not suitable for repeated administration. One recently developed technique which may circumvent many of these problems is the State-Trait Anxiety Inventory (STAI).[5] This device produces scores indicating the individual's current level of anxiety, i.e., A-State, and how prone the individual is to experience anxiety, i.e., A-Trait. It is easily administered and scored and, since it contains only 40 items which are to be rated on four-point scales, it takes only a few minutes for the respondent to complete. It is also supported by a substantial normative data-base from a variety of different populations. For these reasons, it has already been employed in the study of stress in air traffic controllers.[4]

Although the STAI is a promising inventory for the measurement of anxiety, it may be susceptible, as are most self-report techniques, to the effects of biasing response sets. In particular, there are situations in which respondents might tend to elevate the level of anxiety which they report on the STAI scales. For example, records obtained from some individuals engaged in professions such as law enforcement, air traffic control work, piloting aircraft, medical surgery, and the like might be biased by over-emphasis of the impact of job-related stress upon their affective states. The purpose of this investigation was twofold. The first purpose was to determine to what extent STAI scores are influenced by attempts of respondents to emphasize the appearance of stress. The second purpose of the study was to then establish procedures for identifying STAI protocols biased by such tendencies.

The assistance of Gary Hutto and Charles Abbott in the conduct of this study is gratefully acknowledged. The author is also indebted to Charles D. Spielberger who supplied the psychiatric patient protocols used in some of the analyses.

II. Method.

The STAI was administered twice to each of 283 paid student volunteers (200 males, 83 females) recruited primarily from the University of Oklahoma. The first administration of the STAI was made with the standard instructions for both the A-State and A-Trait forms. These instructions request the respondent to rate each of the 20 A-State items according to their current feelings and each of the 20 A-Trait items according to how they generally feel. Each item is answered by indicating which of four response categories best describes to what degree items such as "I am tense" apply to the respondent. For A-State items, the categories range from "not at all" to "very much," while for A-Trait items the range is from "almost never" to "almost always." The *S*s were not told that a second administration was planned. After completion of the first STAI, there was a two to three minute break, then *S*s were asked to complete the STAI a second time. The first 131 *S*s were each asked to respond as would an employee who wanted to make it appear that he was in an extremely stressful job. The remaining 152 *S*s were given similar instructions with the addition of a request to "try to prevent the examiner from knowing that you are exaggerating your responses." The *S*s were without access to their first STAI records during the second STAI administration. Five *S*s from the first group, and six *S*s from the second group, did not understand the instructions, or were unable to complete the task as directed and were subsequently eliminated from the sample.

TABLE 1.—Mean A-State and A-Trait raw scores for authentic STAI records (first administration) and STAI records simulated for stress under bias and concealed-bias instructions (second administration).

Group	A-State			A-Trait		
	Authentic	Simulated	t	Authentic	Simulated	t
Stress Bias	34.02	69.78	33.13***	34.90	63.73	20.46***
Concealed Stress Bias	34.88	60.01	31.80***	36.52	56.08	22.01***
t	0.95	12.62***		1.60	5.29***	

***$p < .001$

III. Results.

As expected, there were large differences between STAI scores obtained under normal conditions and those obtained following the response-bias instructions (Table 1). Both of the stress instruction groups obtained higher mean A-State and A-Trait scores on the stress-simulated STAI administration than on the form taken with standard instructions. The stress-simulated scores were somewhat higher for the group given the simpler instructions than for the group given explicit instructions to conceal their attempts to bias the records. It was also apparent in comparing the scores of the biased A-State and A-Trait scales that the A-State measure increased more than the A-Trait measure when respondents attempted to exaggerate stress, t (270) = 10.02, $p < .001$.

There were no significant differences between males and females on the A-State or A-Trait scales under either the standard or simulated conditions. Therefore, all subsequent analyses were conducted without respect to the sex of the respondent.

For each item, the percentage of *S*s who indicated higher anxiety on the second than on the first rating was determined for the entire group of *S*s. For items on the A-State scale, the mean percentage of *S*s showing increased anxiety per item on the second administration was 80.17 (S.D. = 6.25), while for items on the A-Trait scale, the mean percentage was 73.06 (S.D. = 8.99). The percentages for specific items on the A-State scale ranged from 60.0 to 87.9; the range for items on the A-Trait scale was 59.3 to 85.0.

Since the effects of bias were rather uniform across items, it was felt that the total raw scores for the STAI scales might serve as well as any sub-group of items in discriminating between authentic and stress-simulated records. It was found that using scores from the A-State scale alone provided the best index of response bias after considering the scores from the two scales separately and summed together. The distribution of the scores indicated that the most efficient classification of records resulted when raw scores of 48 or more were considered "stress-simulated" (Table 2). The overall hit rate using this score was 515 correct profile classifications out of 544 records (94.7%). This score included only 12 of the 272 (4.4%) authentic records (false positives) while 255 of the 272 (93.8%) stress-simulated scores were correctly identified. Of the biased records, 120 of the 126 (95.2%) records from the group simply instructed to exaggerate stress and 135 of the 146 (92.5%) records from the *S*s who tried to conceal the stress-simulation were correctly classified as simulations. The difference between the proportions of correct classifications for the two groups given different stress-simulation instructions was not significant, and substantial differences were not apparent except toward the upper limit of the distribution of scores. (This was also true of subsequent analyses; therefore, all remaining findings are presented for the pooled instruction groups.) To identify 95% of the stress-simulated records required a cutting score of 45 which resulted in a false-positive rate of 7.7%.

Use of the raw scores for discriminating between stress-simulated and authentic A-Trait records was not as successful as was the case with the A-State scale; however, it was possible to achieve a reasonable hit rate using 47 as the cutting score (Table 2). This resulted in the correct classification of 483 of 544 profiles (88.8%). Of the 272 stress-simulated records, 235 (86.4%) were included by this score, while 24 of 272 (8.8%) authentic records were misclassified. To identify 95% of the stress-simu-

lated A-Trait records required a cutting score of 35, but this resulted in the misclassification of 144 (52.9%) of the authentic records.

While the procedure using raw scores may serve as an effective screen in a "normal" population, a perusal of the norms for the STAI[5] indicates that the false-positive rate is likely to be quite high if this approach is applied to other populations. For example, use of 48 as the cutting score for the A-State scale would result in false-positive rates of 52% for neuropsychiatric patients, 44% for prison inmates, and 36% for medical/surgical patients. Therefore, the records of the experimental groups, as well as 125 records from Spielberger's neuropsychiatric patient normative sample were subjected to further analysis. The neuropsychiatric sample was selected for analysis since this was the normative group with the highest mean A-State and A-Trait scores. The analysis consisted of the determination of the number of *S*s in each condition using each of the four possible ratings for each item. It was found that ratings indicating the lowest level of anxiety (responses which added only one point to *S*'s score) were relatively common among authentic and patient records, but were infrequent in stress-simulated protocols. An index was then developed using those specific items that were given "one" (lowest anxiety) ratings on no more than 10 stress-simulated records. Each set of STAI records (authentic, stress-simulated, and patient) was then divided in half on an odd-even basis. The number of "one" ratings on the 12 A-State and 5 A-Trait items selected for analysis in each record was determined. It was found that among records with high scores (48 or more on A-State), a combination of six A-State items (numbers 1, 2, 3, 10, 11, 12) and two A-Trait items (numbers 21 and 30) provided the best differentiation between authentic and patient records on the one hand, and stress-simulated records on the other. If any

TABLE 2.—A-State and A-Trait cumulative raw score distributions for authentic and two types of stress-simulated records.

Score	A-State				A-Trait			
	Authentic (N = 272)	Simulated (N = 126)	Concealed Simulation (N = 146)	Sum Correct[a] (N = 544)	Authentic (N = 272)	Simulated (N = 126)	Concealed Simulation (N = 146)	Sum Correct[a] (N = 544)
65 or more	0	93	48	413	0	70	25	367
64	0	97	52	421	0	77	35	384
63	0	101	62	435	0	80	39	391
62	0	103	72	447	1	84	42	397
61	1	104	78	453	4	85	55	408
60	2	108	89	467	4	87	64	419
59	2	109	96	475	6	90	71	427
58	2	110	104	484	6	93	78	437
57	2	110	106	486	6	95	85	446
56	4	112	112	492	7	97	92	454
55	5	114	113	494	9	97	93	453
54	6	116	117	499	10	98	101	461
53	7	118	119	502	12	99	105	464
52	7	119	123	507	15	100	111	468
51	8	119	126	509	17	100	115	470
50	10	119	131	512	18	103	118	475
49	10	119	133	514	20	104	120	476
48	12	120	135	515	21	106	124	481
47	16	120	136	512	24	110	125	483
46	19	120	137	510	30	112	126	480
45	21	120	139	510	38	115	127	476
44	26	120	139	506	44	115	129	472
43	32	121	139	500	52	117	130	467
42	39	121	140	494	57	117	134	466
41	43	122	141	492	68	117	135	457
40	55	125	141	483	78	118	135	447

[a]Number of records correctly classified if all scores at or above that level are designated as biased.

record contained two or more "one" ratings among these eight items, it was then classified as authentic, and the remaining protocols were then screened for stress-simulation by the raw score procedure previously described. Using this two-step procedure resulted in a reduction of six in the number of patient records misclassified as biased and a reduction of two in the number of misclassified authentic records. None of the classifications of simulated records in this half of the sample was changed by this procedure. When applied to the other half of each set of STAI records, it was found that seven patient, two authentic, and two stress-simulated records were removed from that group of records classified as biased by use of the cutting score of 48. Thus, considering the entire sample together, the number of authentic records misclassified was reduced from 12 to 8 (2.9%), the number of patient profiles classified as biased was reduced from 61 (48.8%) to 48 (38.4%), while the number of stress-simulated records correctly identified decreased from 255 to 253 (93.0%).

If only A-State scores are combined, the six A-State items serve most effectively for the first step in the screening process. Applied to all A-State records, the two-step procedure, again employing a cutting score of 48, resulted in a reduction of 11 false-positive classifications in patient records at the cost of excluding two correct classifications of stress-simulated protocols. There was no effect on authentic records in this case.

When considering A-Trait scores alone, the procedure which resulted in the most successful classification rate for the half-sample split on an odd-even basis was to classify as "valid" any A-Trait record which contained a "one" rating on any of the four items 21, 26, 27, and 30, then classify the remaining records by the cutting score of 47. This resulted in a reduction of five patient and three authentic false-positive classifications at the cost of incorrectly excluding a single stress-simulated record. For the second half of the sample, this procedure reduced patient false-positives by seven, and authentic false-positives by one, while increasing misclassified stress-simulated records by one. Considering the entire sample, the two-step procedure reduced the number of misclassifications of patient protocols from 62 (59.6%) to 50 (40.0%), and of authentic records from 24 to 20 (7.4%). The reduction in correct identifications of stress-simulated records was from 235 to 233 (85.7%).

IV. Discussion.

While the findings indicate that STAI scores are quite susceptible to intentions to exaggerate stress effects, it is also clear that the STAI user can screen records for this type of bias with relatively little effort and with considerable effectiveness. While the application of this screening approach to other "normal" populations may diminish the effectiveness somewhat from the 95% hit rate attained with the experimental population, the normative data for "normal" student groups suggest that at worst the false-positive rate would probably not exceed 14% using the raw score alone, and would be less using the two-step procedure. The overall effectiveness is also dependent upon the cutting score chosen by the user; however, it is possible to vary cutting scores across a considerable range to maximize correct identification of either authentic or stress-simulated protocols. For example, if the user can tolerate the loss of legitimate records required by reduction of the cutting scores, it is possible to achieve 95% assurance that stress-simulated records will not be accepted as legitimate.

The question as to the use of the two-step procedure in preference to the use of the raw score alone is an open one. The evidence suggests that the impact of using the item screening prior to application of a cutting score is probably minimal. Therefore, it seems reasonable to suggest that the raw-score-only approach be used unless the user has a particular need for legitimate high-scoring profiles. It should be emphasized that as with all such indices, STAI records classified as stress-simulated can only be considered as "suspect" since there is relatively little basis for discriminating between records reflecting legitimate intense anxiety and records biased by stress-simulation. Thus, with patients, both neuropsychiatric and surgical, prison inmates, and other samples where high anxiety may be common, the false-positive rate will be quite high when using this screening procedure. Under circumstances where the false-positive rate is high, or where loss of authentic records cannot be tolerated, the screening procedure might serve as a "flag" to indicate the need for further evaluation of an individual's affect status.

In comparison to the hit rate for stress-simulation attained with the STAI, it was found that the most efficient index for the same stress-simulation bias in the Composite Mood Adjective Checklist[2] achieved 90% accuracy in classification of simulated and authentic records.[3] The widely used *F* minus *K* index developed by Gough[1] for detecting simulation of psychopathology on the MMPI has a maximum overall hit rate of approximately 86%, with a limit of 88% accuracy in identifying biased records. Thus, it appears that the efficiency of the screening procedures described in this study compares favorably with the efficiency of indices generally considered useful for identifying response biases.

In sum, it appears that the STAI can be used for measuring the anxiety of air traffic controllers, or workers in other high-demand occupations, with considerable confidence that the validity of findings will not be reduced by the unsuspected effects of biasing response sets. It should therefore prove a most useful technique for both long- and short-term studies of stress in air traffic controllers where a brief, easily administered measure of anxiety is desired.

REFERENCES

1. Gough, H. G.: The F Minus K Dissimulation Index for the MMPI, JOURNAL OF CONSULTING PSYCHOLOGY, 14:408–413, 1940.
2. Malmstrom, E. J.: Composite Mood Adjective Check List, Unpublished Manuscript, University of California, Los Angeles, 1968.
3. Smith, R. C.: Assessment of a "Stress" Response-Set in the Composite Mood Adjective Checklist. FAA Office of Aviation Medicine Report No. AM–71–14, 1971.
4. Smith, R. C., C. E. Melton, Jr., and J. M. McKenzie: Anxiety in Air Traffic Controllers. Paper presented to the 43rd Annual Scientific Meeting of the Aerospace Medical Association, Miami, Florida, May 1972.
5. Spielberger, C. D., R. L. Gorsuch, and R. E. Lushene: *Manual for the State-Trait Anxiety Inventory.* Palo Alto, California, Consulting Psychologists Press, Inc., 1970.

24058