NAMRL- 1299

PERFORMANCE CONSISTENCY ON A PERCEPTUAL-MOTOR TASK AS A CORRELATE OF ACHIEVEMENT MOTIVATION

Gerald B. Thomas and Robert W. Clipper

June 1983

NAVAL AEROSPACE MEDICAL RESEARCH LABORATORY
PENSACOLA, FLORIDA

Approved for public release; distribution unlimited

PERFORMANCE CONSISTENCY ON A PERCEPTUAL-MOTOR TASK AS A CORRELATE OF ACHIEVEMENT MOTIVATION

Gerald B. Thomas and Robert W. Clipper

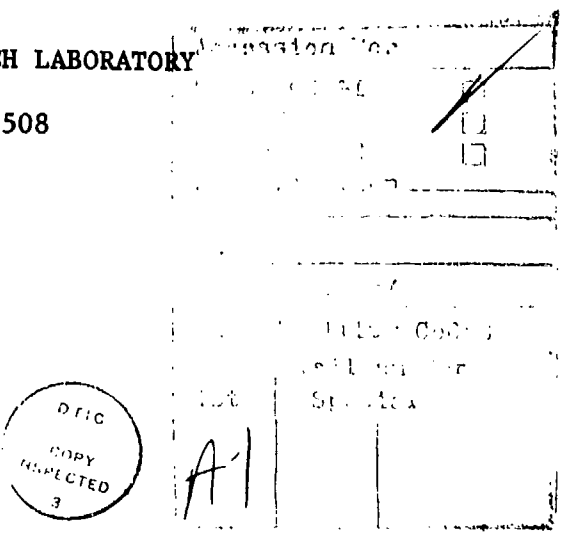
Naval Medical Research and Development Command
MF5852801A-0002

Reviewed by
Ashton Graybiel, M.D.
Chief Scientific Advisor

Approved and Released by
Captain W. M. Houk, MC, USN
Commanding Officer

17 June 1983

NAVAL AEROSPACE MEDICAL RESEARCH LABORATORY
NAVAL AIR STATION
PENSACOLA, FLORIDA 32508



SUMMARY PAGE

THE PROBLEM

Because of its potential usefulness in such areas as personnel selection, performance readiness, etc., the development of a reliable and valid objective measure of achievement motivation has been the goal of researchers for decades. To date, however, the most widely accepted measures of this construct have been instruments based on pencil-and-paper or projective techniques. In the course of our research to develop auditory fitness standards for Naval aviators, we investigated the possibility that consistency of performance could reflect quality of effort, and it is from these exploratory investigations that the present report results.

FINDINGS

The findings from a series of studies in which student Naval aviators and flight officers were administered traditional measures of achievement motivation (i.e., the Thematic Apperception Test and the Mehrabian Achieving Tendency Scale) as well as tests of reaction time consistency indicated that performance on a particular reaction time paradigm repeatedly correlated significantly ($p < .05$) with the Thematic Apperception Test measure of achievement motivation. Furthermore, the test-retest reliability of the reaction time consistency index averaged .70, an improvement over the typical .30 to .40 reliability coefficients associated with the traditional, less objective measures. Finally, in a very limited validation study, the consistency values of student Naval aviators who had passed primary flight training tended to differ (though not significantly so; $p < .10$) from those who had attrited.

RECOMMENDATIONS

It is recommended that additional research utilizing larger populations be conducted to optimize the paradigm's design and provide additional validation information.

ACKNOWLEDGMENTS

The authors would like to thank all student aviators who volunteered to participate in this project and ENS Bruce Ortel, MSC, USNR who coordinated their scheduling. Gratitude is also extended to LT E. A. Trautman, MSC, USNR for providing selection test scores for the population and to LT Frank C. Petho, MSC, USN for supplying the subjects' flight training results.

Achieving a goal depends on the presence of at least two factors--requisite ability and sufficient desire. Tests of general and specific aptitudes have been in use since the time of Binet and are well established in personnel selection regimens as measures of task-required abilities or propensities. The second factor, goal oriented motivation, has been more difficult to objectively assess in the human model with the result being the current absence of a universally accepted, valid, reliable, and easily administered measure.

Our interest in this area of human assessment stemmed from the need to ascertain the purposefulness with which subjects of different age groups performed on a prospective test battery designed to determine the auditory fitness of Naval aviators. Traditional reliance on absolute performance was discarded as an option because many sensory and higher level auditory abilities were known to be age-dependent. Therefore, with no suitable standard test of achievement motivation available, we undertook a series of studies to investigate an alternative approach.

The purpose of the present research is to investigate the possibility that consistency of performance (vice absolute performance) could reflect a person's desire to do well in an experimental situation. More specifically, in our design, the question becomes, "Do aviation officer candidates who score well on traditional measures of achievement-related motivation also show a greater degree of consistency in their performance on a perceptual-motor task?"

The emphasis on consistency of performance instead of absolute measures stems from the thinking that traditional indices of task performance are based on measures of central tendency which, in fact, reflect both the ability and the desire of the participant. Indices so derived, therefore, tell an ambiguous story. A closer look at consistency of performance, on the other hand, might provide a more valid and potentially less confounded indication of a subject's interest in succeeding. Only to the extent that the mean and variance distributions were statistically correlated would inherent abilities possibly be contributing to the scores.

Although consistency of performance as a measure of achievement motivation has not been explicitly tested in the past, there have been a number of authors who have investigated persistence of effort (e.g., 1, 5, 18, 21). Research in this area has indicated that the relationship between task persistence and achievement motivation is a function of the probability of success in the task, success or failure feedback, relevance of the task to desired goals, intelligence, etc. Unfortunately, the methods employed in the preceding studies have been of the type where metric sensitivity is somewhat coarse (e.g., the number of arithmetic problems attempted). Studies in vigilance could also be viewed as investigations of persistence (of attention), but the time required for the administration of the vigilance paradigm could be prohibitive.

In view of the preceding considerations, it was decided to employ a reaction time paradigm to assess persistence of effort (be it attentional, acquisitional, etc.). The reaction time procedure has a number of useful characteristics including good potential sensitivity (measurement in milliseconds), minimal confounding motor skills effects (single finger key press), ease of administration (computer controlled), and versatility (numerous possible stimulus-response manipulations). On the other hand, as the designs progress from simple configurations (one stimulus/one response) to complex, choice arrays (multiple stimuli/multiple responses), increased involvement of perceptual and cognitive factors occurs along with a departure from a potentially "pure" measure of achievement motivation. Nevertheless, a reaction time paradigm seemed best suited for investigating the performance consistency notion.

Regarding the use of reaction time variability as the principal measure of interest, research has been conducted utilizing this index but most of these studies have dealt with clinical issues. For example, several studies have shown significant differences between normal and retarded persons (3, 17) and a relationship between reaction time variability and degrees of brain damage (6, 10) and the incidence of Parkinsonism (9). Reaction time variability has been less successful in differentiating among individuals with similar levels of intelligence, however (23).

The latter finding is of potential significance in the present studies because the populations utilized were aviation officer candidates who could be assumed to be relatively homogeneous with respect to intelligence, career striving, etc. Answering the question of whether a reaction time paradigm could be devised which provided sensitivity sufficient to differentiate among members of this select group with its potentially restricted range was one purpose of Experiment I.

A second, related concern was whether any of the available tests of achievement motivation which were to be correlated with the reaction time variability measure were similarly sufficiently sensitive.

Two tests of achievement motivation were selected for use in the present series of experiments---the Thematic Apperception Test (TAT) and the Mehrabian Achieving Tendency Scale (MATS).

The TAT was originated by Murray (16) as a clinical tool and revised by McClelland, Atkinson, Clark, and Lowell (14) for use in the measurement of "need for achievement" (nAch). It is a projective technique in which individuals are asked to view a "neutral" scene and then write a story about the scene using their "creative imagination." The story is then analyzed by trained judges according to a standardized scoring scheme for its achievement related content. The TAT is perhaps the best known and most widely used measure of achievement motivation and, during its three decades of existence, has been refined and its relationship to other achievement measures replicated a number of times (e.g., 2, 4, 13). As Klinger (11) notes, however, there have also been a significant number of researchers who have failed to find relationships.

One refinement of the TAT has been the inclusion of a qualifying variable--fear of failure. Typically, this has been determined through the use of the Test Anxiety Questionnaire (TAQ) developed by Sarason and Mandler (20). The TAQ is a pencil-and-paper instrument which, when combined with the TAT, provides a "resultant measure" useful in the explanation of the sometimes reported non-linear relationship between nAch and achievement.

Despite the fact that the TAT-derived measure of need for achievement is the oldest and most popular such technique, it is not without its drawbacks and detractors. From a practical standpoint, the time and personnel required for its administration and scoring can, in some instances, be prohibitive. Secondly, Entwistle (7) cites the TAT's relatively poor test-retest reliability of about .3 to .4 as a major problem and questions its validity in light of the fact that, at least in one study, the verbosity of the scored stories correlated more highly with achievement than did their achievement-related content. Others (e.g., 11) question the very "thought sample" method upon which the TAT is based.

The MATS is a more recent attempt (15) at the quantification of achievement-related motivation and is in the form of a pencil-and-paper, semantic differential questionnaire which is objectively scored. It, too, is typically combined with a "fear of failure" score (e.g., the TAQ) and becomes the RAM (resultant achievement motivation).

To emphasize the difficulty of developing a valid and reliable measure of the somewhat ethereal achievement motivation construct, it should be recognized that the typical correlation between the TAT-derived and the MATS-based measures seldom exceeds .3 and is usually closer to .2 (22).

Clearly then, as stated earlier, measuring motivation to achieve has been a challenging undertaking with no presently available device enjoying universal acceptance.

Experiment I

The principal purpose of this initial study was to determine whether the consistency of performance/achievement motivation notion had any basis whatsoever in fact. To effect this determination, a design manipulating several reaction time variables and including a number of achievement-related indices was utilized.

Besides the aforementioned TAT and MATS values (including their resultant measures), the performance of our subject population on aviator selection instruments was also included. These latter values were comprised of scores on the following instruments:

1. Academic Qualifications Test (AQT) -- a 60-minute test of general intelligence.

2. Flight Aptitude Rating (FAR)

   a) Mechanical Comprehension Test (MCT) -- a 40-minute test dealing

with the ability to perceive physical relationships.

b) Spatial Apperception Test (SAT) -- a 10-minute test of the ability to orient in space.

c) Biographical Inventory (BI) -- an untimed questionnaire regarding personal history, interests, attitudes, etc.

Although none of the preceding selection tests necessarily reflects the present state of achievement motivation, the possibility exists that some of them (e.g., the AQT) mirror past levels of achievement orientation or an ongoing trait of desire to succeed.

Regarding the reaction time paradigm, an approach catholic in nature was adopted whereby several potentially critical variables were manipulated. The intent with this approach was to provide a paradigm which would allow any indication of a relationship between reaction time parameters and achievement motivation to be observed. Subsequent research would be directed toward refining the technique and identifying critical variables. The variables manipulated and their rationale for inclusion are described below.

1. Task complexity. Two types of reaction time paradigms were employed. The first was a classic "simple" design with a single stimulus requiring a single response throughout the trial blocks. The second paradigm consisted of four possible stimuli occurring on a given trial and calling for one of two response choices. In addition, the latter paradigm's stimuli were not equiprobable. The purpose of this feature was to ensure a challenging scenario for the subjects as well as to test the finding by Atkinson and Litwin (1) that individuals differing in levels of achievement motivation sometimes make different risk/payoff choices.

2. Feedback. In both the simple and choice reaction time paradigms, feedback (in milliseconds) was either provided or not provided after each trial. It is generally acknowledged that an improvement in reaction time performance (generally mean reaction time) occurs when a subject is informed of his response latencies. Although individual differences in this improvement have not been reported, it is conceivable that persons differing in internal vs. external motivation or control would show differential performance as a function of feedback. Also, it can be inferred from McClelland's work that the presence of feedback is necessary for differentiating motivational levels (12, 14).

3. Stimulus modality. Both visual and auditory stimuli were employed in the present study primarily in deference to the fact that our ultimate populations of interest are aviators who, in the course of operations, must process signals of both modalities.

In addition to the aforementioned measures, the number of errors each subject produced was also recorded for use as an index of performance quality.

METHOD

Subjects. A total of 28 male Aviation Officer Candidates (AOCs) undergoing preliminary testing at the Naval Air Station, Pensacola, Florida served as volunteer subjects in Experiment I.

Apparatus. The reaction time tests were conducted in a sound-proof Industrial Acoustics Chamber (Model 400) with stimuli being manually presented and controlled by a Gerbrands G1360 reaction time apparatus. Auditory stimuli were produced by a Maico MH-20 audiometer and presented binaurally over a set of Telephonics TDH-39 headphones at a sound pressure level (SPL) of 78 dB. Visual stimuli were in the form of illuminated discs which measured 32 mm in diameter and were located on the response panel of the subject at a viewing distance of approximately 50 cm. Response times were measured to the nearest millisecond with a Gerbrands G1271 digital clock.

The four slides which made up the TAT portion of the study included: 7BM "father-son"; 8BM "boy and operation"; "inventors" picture; and "a student" (14). The slides were projected in a darkened room by a Kodak Carousel AF-2 projector onto a 1.5 m square Duralite projection screen.

Procedure. Each session began with the simple reaction time task (SRT1) in which each of four stimuli (a red light, a green light, a high (3000 Hz) tone, and a low (250 Hz) tone) were presented individually in discrete blocks of 10 trials each. Subjects were provided feedback (response time in milliseconds) after each trial in the form of an oral report by the experimenter.

The second phase of the session introduced the choice reaction time task in which any of the four stimuli could occur on a given trial. Subjects were instructed to respond as "quickly and accurately as possible" by depressing a key with the index finger of the dominant hand if either the high tone or the green light occurred and with a key press by the index finger of the non-dominant hand if either the low tone or the red light occurred. Furthermore, each subject was informed beforehand of the probabilities of occurrence of the four stimuli: high tone --- .55; green light --- .05; low tone --- .15; and red light --- .25 (see Figure 1). A card describing the preceding stimulus-response and probability relationships was placed before the subject to serve as reference throughout the session. After completing 50 trials of this RT task without feedback (CRT1), another 50 choice reaction time trials with feedback (CRT2) were presented. A fourth block of 50 reaction time trials (choice) without feedback (CRT3) was then completed and the test session concluded with a re-run of the simple reaction time task (SRT2) without feedback (see Table I). This order of presentation was the same for all 28 subjects.

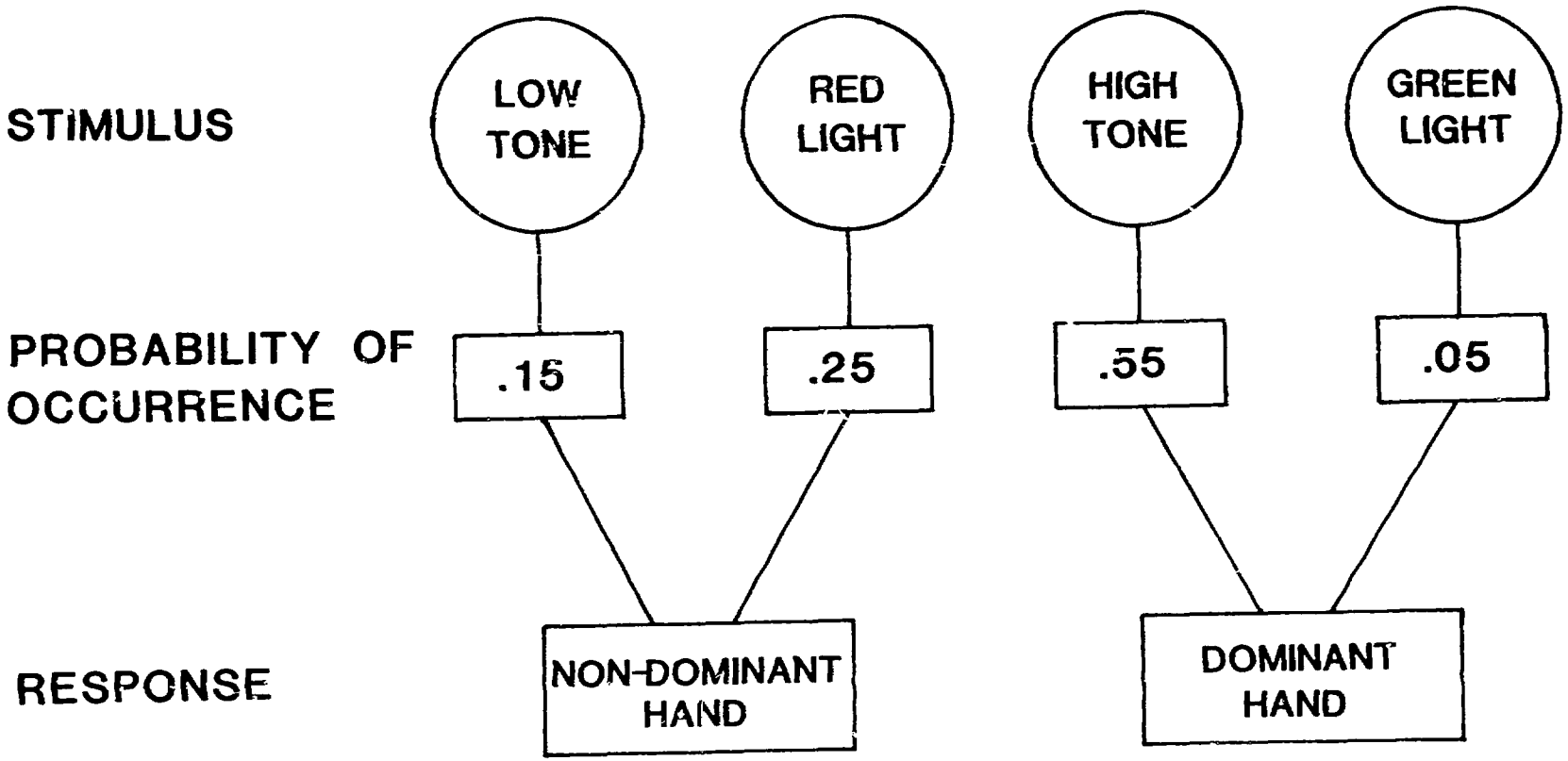


FIGURE 1. STIMULUS-RESPONSE PAIRINGS AND ASSOCIATED PROBABILITIES OF OCCURRENCE FOR THE CHOICE REACTION TIME PARADIGM.

Table I

Description and Order of Reaction Time Tasks for Experiment I.

ORDER	1	2	3	4	5
REACTION TIME TASK	Simple	Choice	Choice	Choice	Simple
FEEDBACK	Yes	No	Yes	No	No
NUMBER OF TRIALS	40	50	50	50	40
DESIGNATIONS	SRT1	CRT1	CRT2	CRT3	SRT2

Throughout the session each trial was initiated by the appearance of a "ready" light on the response panel. The time between the appearance of the ready light and the onset of the stimulus varied randomly between one and five seconds during the simple reaction time tasks and was a fixed three-second period during the choice reaction time trials. The intertrial interval (ITI) averaged five and eight seconds for the simple and choice reaction time tasks, respectively.

After completing the reaction time trials, each subject was administered the TAT, MATS, and TAQ.

The TAT measures were gathered following procedures outlined by McClelland, et al. (14). Following a 20-second viewing of each of the four neutral slides, the subjects were requested to create a four-minute written story about each slide such that the following questions were answered: What is happening? What led up to the situation? What is wanted and by whom? What will happen? The stories were then scored by two practiced judges whose inter-judgment reliability was .95.

The MATS and TAQ were then completed by the subjects at their own pace.

RESULTS

With regard to the reaction time tasks, more than one performance scoring scheme is possible with the hypothesis under test and the paradigms employed.

First, with reference to the choice reaction time tasks, it could be argued that a subject desirous of doing well on the task would develop a perceptual-motor strategy which would optimize response latencies to all four

possible stimuli. That is, despite the fact that the stimuli vary in their probabilities of occurrence and, therefore, in their potential mean latencies (i.e., probabilities and latencies tend to be inversely related (e.g., 19)), a subject could adopt a strategy which "equalizes" his responses to the four stimuli. In light of this possibility, response latencies and standard deviations were summed over all stimuli for each condition and analyzed.

Table II is a matrix of selected correlations relating reaction time, achievement motivation, and aviator selection test variables. As can be seen, three variable pairs were significantly related ($p < .05$):

* TAT and CRT1 standard deviation with an r of -.43

* TAT and CRT2 standard deviation with an r of -.38

* MATS and TOTAL ERRORS with an r of -.38

Choice reaction time means approached significance with the TAT but failed to reach the .05 level. Not shown in the matrix are the correlations between reaction time means and corresponding standard deviations (these averaged r's of .47 ($p < .01$)).

A second method of scoring performance on the reaction time measures is based on the possibility that a subject would develop a strategy whereby he would emphasize one stimulus at the expense of the others. For example, Atkinson and Litwin (1) found that individuals high in nAch and low in fear of failure tended to choose alternatives which had medium risks and medium payoffs. An individual such as this might choose to perform as well as possible on the red visual stimulus with its probability of occurrence of .25 and attend less actively to the other stimuli. Subjects with other nAch/fear of failure configurations would, in theory, tend to concentrate on other stimuli. To check this possiblity, the 28 subjects were divided into quartiles ranging from those with high nAch and low fear of failure scores (relative to the group as a whole) to those with low nAch and high fear of failure scores. Employing a Chi-squared technique, no differential preference for the four stimuli occurred for the different quartiles.

After mathematically equating the reaction times for the four stimuli in the choice reaction time conditions (based on simple reaction time values), individual correlations between each stimulus and the various achievement-related measures were calculated. No significant relationships were found between the types of stimuli and scores of achievement motivation.

Table III contains the correlations between the aviator selection test scores and the reaction time and achievement motivation values. No significant relationships were noted although an inverse relationship between test anxiety and the AQT approached significance. Other tentative trends included an inverse relationship between reaction time means and standard deviations and the AQT; the reaction time standard deviations and the SAT; the MCT and reaction time means; and weak but generally positive correlations between all selection test scores and the MATS, RAM, TAT and TAT-TAQ.

Table II

Correlation Matrix of Reaction Time and Achievement Motivation Scores

	TAT	MATS	TAQ	TAT-TAQ	KAM
MEANS					
SRT1	-.10	-.15	.14	-.13	-.19
CRT1	-.34	.06	.12	-.27	-.04
CRT2	-.35	.04	.07	-.25	-.03
CRT3	-.25	.06	.08	-.20	-.02
SRT2	-.32	.06	-.03	-.21	.06
STANDARD DEVIATIONS					
SRT1	.04	-.26	.17	-.05	-.27
CRT1	-.43*	.03	.06	-.32	-.02
CRT2	-.38*	.08	-.22	-.10	.19
CRT3	-.10	.12	.05	-.13	.04
SRT2	-.16	-.03	-.17	-.03	.09
ERRORS	-.14	-.38*	.08	-.14	-.28

*p <.05 (r = .37)

Table III

Correlations of Reaction Time and Achievement Motivation Scores with Aviator Selection Test Variables

	AQT	FAR	MCT	SAT	BI
MEANS					
SRT1	.06	.17	-.07	.11	.26
CRT1	-.21	.15	-.19	.18	.16
CRT2	-.29	.09	-.21	.08	.14
CRT3	-.28	.13	-.16	.10	.17
SRT2	.28	.17	.11	-.02	.21
STANDARD DEVIATIONS					
SRT1	.12	-.04	-.18	-.03	.04
CRT1	-.14	-.29	-.14	-.30	-.26
CRT2	-.21	.03	.00	-.26	.08
CRT3	-.23	.09	.07	-.20	-.13
SRT2	.09	-.22	-.33	.01	-.19
TAT	-.02	.02	.22	.17	.15
MATS	.14	.21	.35	.11	.26
TAQ	-.35	-.06	-.08	.08	-.05
TAT-TAQ	.24	.06	.21	.07	-.01
RAM	.31	.17	.27	.02	.19

DISCUSSION

The finding that achievement motivation as measured by the TAT was significantly related to two of the three choice reaction time variability measures suggests that the idea that performance consistency reflects a need for achievement could have merit. This notion is further buttressed by the fact that all of the choice reaction time variability scores were properly (i.e., inversely) related to TAT scores eventhough not all reached statistically significant levels. The failure of the perceptually and cognitively less demanding SRT condition to follow a similar general trend indicates that the process is more complex than was initially hoped and that higher level perceptual and cognitive functions will have to be introduced and investigated to fully understand the relationship. Illustrative of this is the finding that while the first choice reaction time condition without feedback (CRT1) was significantly related to TAT scores, the temporally removed, but functionally identical, CRT3 provided an insignificant correlation of -.10. Between these two conditions, the choice reaction time task with feedback, CRT2, also provided a statistically significant correlation with the TAT. From these results it is conceivable that a task acquisition process is at work in this situation in addition to, and perhaps superior to, relevant motivational variables. Subsequent experiments were conducted to examine this possibility.

The resultant TAT-TAQ scores were negatively correlated with all reaction time conditions although none reached statistical significance. Similarly, nine of the ten reaction time mean measures were negatively correlated with TAT-derived measures but, these too, failed to reach levels of statistical significance. Despite the dearth of statistically significant relationships in the preceding analyses, the number of correlations in the predicted direction implies that quality of reaction time task performance and TAT-derived need for achievement scores share some common ground.

Reaction time value correlations with the MATS and RAM were mixed and minimal in degree with the exception of the correlation between total errors and the MATS ($r = -.38$; $p < .05$). Of note is the contrast between the MATS-based measures of achievement motivation and those provided by the TAT, and their respective relationships with reaction time means and standard deviations. This lack of agreement, as noted earlier, is not uncommon and the obtained correlation between the TAT and MATS of .22 (not shown in the tables) is typical (22). It is, nonetheless, troubling when two instruments purportedly measuring the same construct are in such lack of agreement. In this regard, it should be recalled that the ultimate purpose of the TAT is not revealed to the subject; he is led to believe that his stories are tests of "creative imagination." The MATS' intent, on the other hand, is forthright and apparent. In a like fashion, the subjects in the reaction time tasks were told to "respond as quickly and as accurately as possible" with no mention of consistency of performance. Perhaps the commission of errors (an obvious index of performance) parallels the overt nature of the MATS, whereas consistency is similar to the quasi-deceptive character of the TAT. In other words, reaction time errors and answers on the MATS might be more susceptible to the demand characteristics of the experimental situation than

are TAT stories and reaction time consistency.

The lack of an obviously strong relationship between the aviation selection test scores and measures of reaction time or achievement motivation implies an absence of "real world relevance" for the latter instruments. However, it should be noted that the aviation selection test measures employed in this study account for only about 36% of the variance associated with the success of a pilot in flight training (8) thereby leaving an ample amount of variance unexplained.

Finally, the results also provide no support for the predictions regarding stimulus preference and motivation type that were inferred from the findings of Atkinson and Litwin (1) relating to risk/pay-off choices, although this study, of course, is not a critical test of those findings.

The principal purpose of Experiment I was to determine if there was any demonstrable relationship between reaction time task performance and traditional measures of achievement motivation. Subsidiary, but inextricably related, to this question were the queries regarding the sufficient sensitivities of the instruments measuring all relevant variables. It appears from the results of Experiment I that there is some justification in responding to the preceding questions in the affirmative. Nevertheless, a number of other questions require resolution before unequivocal conclusions can be drawn. The following experiments address some of these uncertainties.

Experiment II

The second experiment in the series had several aims. First, this study sought to replicate the major finding of Experiment I, namely, the relationship between reaction time consistency and TAT scores. Secondly, to determine whether the order of presentation of the choice reaction time conditions was important, feedback to the subjects regarding each trial's reaction time (the differentiating variable of the three choice reaction time tasks in Experiment I) was counterbalanced. Finally, to reduce experimenter effects, the reaction time paradigms were computer-presented and controlled.

METHOD

Subjects. A total of 32 male student Naval aviators (SNAs) and Naval flight officers (NFOs) awaiting school assigment at the Naval Air Station, Pensacola served as volunteer subjects.

Apparatus. As noted earlier, in a departure from the technique used in Experiment I, all instructions and stimuli were presented, and responses recorded and tabulated, by a Hewlett-Packard HP-85 micro-computer. The computer was programmed to present visual and auditory stimuli using its display screen and internal tone generator and to control all timing events, including response latencies, using its internal clock. All data were stored on tape for later analyses. Feedback, when it was provided, occurred as a numerical readout on the HP-85's screen.

The introduction of the preceding instrumentation automation also required the substitution of the word "READY" for the warning light used in Experiment I and the replacement of the previously used red and green visual stimuli with the numerals "8" and "3," respectively. The digits measured approximately 2mm x 1.5mm on the HP-85's monochromatic screen and were viewed from an average distance of approximately 30 cm. Similarly, the "high" and "low" tones of Experiment I acquired center frequencies of 3000 Hz and 250 Hz, respectively, and a free field SPL of 65 dB at the ears.

Responses were made by pressing one of the two outermost "Special Function" keys located directly below the screen. A plexiglas platform was constructed to shield the keyboard and to allow the subject to rest his hands and forearms while making the left and right index finger key presses.

Measures relating to achievement motivation were determined using the materials described in Experiment I.

Procedure. All procedures were identical to those of Experiment I, with two exceptions. First, only two choice reaction time conditions, CRT1 and CRT2, were employed along with the introductory SRT1. Furthermore, the feedback associated with the CRT conditions was counterbalanced so that half the subjects received feedback on CRT1 and no feedback on CRT2 (Group AB), whereas the other half received no feedback on CRT1, but were provided their response latencies on CRT2 (Group BA). Group BA's order is comparable to that used in Experiment I.

The second major procedural change involved the shortening of the inter-trial interval (ITI) from the eight seconds employed in Experiment I to three seconds.

RESULTS

The data from Experiment II were analyzed using the overall means and standard deviations of each reaction time condition for each group of subjects.

These analyses failed to show any significant correlations between the reaction time measures and scores on the achievement motivation instruments.

Additional analyses were postponed pending the identification of the cause of this failure to replicate the findings of Experiment I.

DISCUSSION

The failure to replicate the results of Experiment I was cause for some concern and prompted a close examination of the procedural changes incorporated in Experiment II. It was surmised that the drastic shortening of the ITI from eight to three seconds could have interfered with the attentional processes required to perform the task optimally. Experiment III tested this hypothesis.

Experiment III

Experiment III was a second attempt to replicate the findings of Experiment I. Emphasis was placed on the correlations of CRT1 and CRT2 with the TAT measure of achievement motivation.

METHOD

Subjects. A total of 20 male SNAs and NFOs from the same pool as those in Experiment II served as volunteer subjects. No subject who had participated in Experiments I or II took part in Experiment III.

Apparatus. All materials and equipment were identical to those employed in Experiment II.

Procedure. The procedure was the same as that used in Experiment II with the exception that the ITI was lengthened to a value of eight seconds (identical to that which was used in Experiment I). Two groups of 10 subjects each were administered the reaction time paradigm according to the order described in Table IV.

RESULTS

Table V contains the correlations of selected reaction time variables with the TAT for both groups of subjects. As can be seen, the only condition to significantly correlate with the TAT was the CRT1 of Group BA, although CRT2 approached statistical significance.

No other analyses performed resulted in the identification of statistically significant relationships.

DISCUSSION

Two important findings resulted from Experiment III. The first deals with the fact that apparently the results of Experiment I are replicable, as Group BA's correlation of -.70 with the TAT indicates. Secondly, although not unequivocal, a comparison of the results of Experiments II and III suggests the importance of the ITI in this design.

It also appears that the order in which the reaction time conditions occur bears some significance. If feedback were an unimportant variable in this paradigm and order were, one would expect similar degrees of relationship between the CRT's of both groups and the TAT. Instead, only in that group where feedback was initially absent (i.e., Group BA) did a significant correlation with achievement motivation, as measured by the TAT, result. Should this finding reliably recur, an explanation of its mechanism of operation would require additional research.

Because Experiment I failed to show any particular theoretical or utilitarian value of employing stimuli of differing probabilities, and because we desired to make the choice reaction time tasks somewhat less

Table IV

Description and Order of Reaction Time Tasks for Experiment III

Order	1	2	3
Reaction Time Task	Simple	Choice	Choice
Feedback			
Group BA	Yes	No	Yes
Group AB	Yes	Yes	No
Number of Trials	40	50	50
Designations	SRT1	CRT1	CRT2

Table V

Correlations between the TAT and Reaction Time Values for Two Orders of Feedback Presentation

	TAT
Group BA	
CRT1	
Standard deviation	-.70*
Mean	-.31
Errors	.05
CRT2	
Standard deviation	-.48
Mean	-.13
Errors	-.35
Group AB	
CRT1	
Standard deviation	-.12
Mean	-.03
Errors	-.39
CRT2	
Standard deviation	.01
Mean	.01
Errors	.35

*p <.05 (r = .63)

complicated, it was decided to run additional subjects in a paradigm where all stimuli had equal probabilities of occurring.

Experiment IV

The purpose of the fourth experiment in the series was to investigate the effects of equating the probabilities of occurrence of the four stimuli. It was also a further test of the findings of Experiments I and III.

METHOD

Subjects. Twenty male SNAs and NFOs from the same source as the preceding experiments served as volunteer subjects.

Apparatus. All materials and equipment were identical to those used in Experiments II and III.

Procedure. The procedure employed was the same as that used in Experiment III with the exception that all four stimuli had probabilities of occurrence of .25. The group size and order of presentation of the reaction time conditions were as in Experiment III (described in Table IV).

RESULTS

Table VI contains the correlations of the reaction time measures with the TAT for the two groups of ten subjects. As shown in the table, three conditions were significantly related to the TAT. As in Experiments I and III, the correlation between the variability of CRT1 for Group BA was significant at the .05 level with a correlation coefficient of -.68. The mean of CRT1 for Group BA was also significantly correlated ($r = -.67$; $p < .05$) with the TAT measure. In addition, the variability of CRT2 for Group BA correlated at the .05 level ($r = -.71$) with the TAT.

None of the reaction time means or standard deviations was significantly correlated with the TAT for Group AB but the errors associated with CRT1 did correlate significantly ($r = -.68$; $p < .05$).

No other analyses were attempted.

DISCUSSION

Once again, the major finding of Experiment I was replicated; some choice reaction time conditions without external feedback tend to correlate significantly with the TAT-derived measure of need for achievement. In this experiment, however, CRT2, the reaction time condition which followed the initial "no feedback" CRT1, also reached significance. This is a replication of Experiment I in which both CRT1 and CRT2 were statistically significant.

Table VI

Correlations between the TAT and Reaction Time Values for Two Orders of Feedback Presentation Under Conditions of Equal Stimulus Probability

	TAT
Group BA	
CRT1	
Standard deviation	-.68*
Mean	-.67*
Errors	-.08
CRT2	
Standard deviation	-.71*
Mean	-.58
Errors	-.33
Group AB	
CRT1	
Standard deviation	-.07
Mean	.16*
Errors	-.68*
CRT2	
Standard deviation	-.06
Mean	-.17
Errors	-.18

*$p < .05$ ($r = .63$)

Also in this experiment, the mean of CRT1 for Group BA was significantly correlated with the TAT for the first time (although means for CRT1 and CRT2 in Experiment I approached significance).

Finally, the errors for CRT1 of Group AB were also correlated with the TAT but this finding will await replication before being discussed.

Regarding the stimulus probability question, it doesn't appear, at least in our design, that stimulus probabilities exert a critical effect. As a result, our attempt to alter the cognitive load by eliminating disparate stimulus probabilities was not measurably successful in Experiment IV.

Another question which demanded immediate attention related to the role of continued performance on the tasks. In other words, what is the test-retest reliability of the reaction time variability measure?

Experiment V attempted to answer this question.

Experiment V

The occurrence of a significant relationship between choice reaction time variability and the TAT in Experiments I, III, and IV illustrate the replicability of the effect but provides no information regarding the task's reliability within a given group of subjects. In this experiment, most of the subjects who served in Experiment IV were retested on the choice reaction time task to determine its repeatability.

METHOD

Subjects. A total of 18 of the 20 subjects who served in Experiment IV also participated in this experiment.

Apparatus. The same apparatus used in Experiment IV was employed in Experiment V. No achievement motivation scores were gathered.

Procedure. All subjects had participated in Experiment IV, therefore, new stimulus orders were generated for each session, but stimulus probabilities remained at .25 throughout the testing. Because of time constraints and the fact that the CRT1 condition of Group BA was of primary interest, all subjects were run on the SRT1 (with feedback, 40 trials), the CRT1 (without feedback, 50 trials), then the CRT2 (with feedback, 50 trials). A total of two retests per subject was completed with the intact population. Due to the temporal instability of the student aviator population, the time between retests varied from one to seven days depending upon the individual aviator.

RESULTS

Table VII contains the correlations between the original test and each of the two retests. As can be seen, the correlations between the three tests are highly significant.

No other analyses were performed.

Table VII

Test-retest Correlations of Standard Deviations for CRT1

		(Test) 1	(Retest) 2	(Retest) 3
(Test)	1	1.0		
(Retest)	2	.74***	1.0	
(Retest)	3	.69**	.73***	1.0

** p <.01 (r = .59)
*** p <.001 (r = .71)

DISCUSSION

The average correlation between the three administrations of the reaction time tasks of about .7, while not in the area of excellent reliability, is still an improvement over the .3 to .4 of such tests as the TAT (7).

The relatively good correlations obtained is not to say that there was no change in absolute performance over the three sessions. In fact, most subjects tended to improve their speed and consistency but not significantly so and not differentially.

Experiment V represented a relatively short-term retest of reliability where positive transfer from one session to the next could have occurred. It would be interesting to gather reliability data on tests administered weeks or months apart instead of days. Unfortunately, the nature of our target population makes that difficult.

The results of Experiment V represent a cursory treatment of the reliability question associated with the reaction time variability task. Whether this task bears any relation to real world events remains unknown, despite the fact that its correlation with nAch as measured by the TAT is significant. It should be remembered that the TAT is far from a perfect predictor of achievement in real life, so validations against such an instrument are of questionable worth. With this in mind, we examined the flight school progress of those aviators who had served as subjects in Experiment I. Unfortunately, only 20 of them had finished primary flight training or had withdrawn from the program for one reason or another at the time of this writing. Therefore, the results which follow should be viewed as being very tentative.

Flight School Success of Subjects Participating in Experiment I

Information available for analysis at the time of this report indicated that nine subjects had completed primary flight training successfully, seven had dropped out of the program (two after completing flight training), three had incomplete data due to medical problems or service transfer, and one was found to be not physically qualified for flying.

For the purposes of this analysis, those nine who had successfully completed training and had advanced to a "post primary" assignment constituted the "Pass" group. Those seven who had dropped out of the program for whatever reason made up the "Fail" group. The remaining four for whom data were incomplete were not considered in the analysis.

Because of the very small numbers of subjects available, the non-parametric Kruskal-Wallis Analysis of Variance technique was employed to compare the Pass and Fail groups along selected dimensions of interest. Table VIII contains the H and p values for the reaction time variability measure (CRT1), the TAT, and the standard aviator selection tests currently in use by the U. S. Navy.

As Table VIII reveals, no measure attained traditional levels of significance in differentiating the aviators who had passed primary flight training from those who had attrited. It is interesting to note, however, that the reaction time variability measure occurred in the two groups in the expected direction and came closest to meeting the statistical significance criterion. The SAT also approached, but did not reach, statistical significance. Another point worthy of mention is the fact that, although the CRT1 variability measure and the TAT had a correlation coefficient of -.43 in the original population from Experiment I, the TAT scores showed little tendency to occur differentially in the Pass/Fail groups. Conclusions, however, must be postponed until data from this and other populations are complete.

General Discussion

The series of studies described in this report sought to investigate the validity of the idea that consistency of performance on a perceptual-motor task might bear some relationship to achievement-related motivation as measured by traditional pencil-and-paper methods.

Four of the five experiments showed a significant correlation between a particular reaction time condition and the need for achievement as measured by the TAT. Although the degrees of correlation were at the .05 probability level of statistical significance, the repeatability of the correlations, which ranged from -.43 to -.70 with different groups of student aviators, points to the general nature of the relationship.

Table VIII

Kruskal-Wallis Analysis of Variance Values of the Pass versus Fail Groups for Several Assessment Techniques

Measure	H	p
CRT variability	2.90	.10
SAT	2.01	.17
AQT	.80	.40
FAR	.70	.43
TAT	.11	.76
MCT	.02	.90
BI	.00	.99

Having answered the primary question regarding consistency of performance and need for achievement, light was shed on several other aspects of the problem. For example, the importance of a sufficiently long intertrial interval in the reaction time procedure was identified, as was the apparently minimal influence of differing stimulus and response probabilities in the paradigm.

Of relevance to the ultimate utility of the reaction time technique was the finding that the reliability coefficients of about .7 are better than the traditional, less objective metrics. Also promising was the tendency for the consistency measure to occur differentially, though not significantly so, in the two dichotomous, Pass/Fail groups of student aviators. This latter relationship to success in flight school requires additional investigation, however.

Along with the preceding findings, a number of additional questions arose concerning the exact nature of the phenomenon being investigated. For example, the clear tendency for the strongest relationship between reaction time consistency and nAch to occur when the initial presentation of the choice reaction time paradigm was administered without feedback is perplexing. The underlying mechanism of the phenomenon does not appear to be related to the acquisition of the task because it recurred during the retesting series in Experiment V. Perhaps a complex interaction of the effects of meaningful feedback with different levels (e.g., high vs. low) or types (e.g., internal vs. external) of achievement motivation is of some

importance in this situation.

Other questions requiring answers involve such practical matters as the optimal paradigm and scoring system for maximizing the effectiveness of the reaction time technique and the minimal number of trials required for the appearance of the relationship.

Of course, additional replications as well as further research into the reliability and validity questions are required before perceptual-motor consistency measures can be viewed as viable objective indices of achievement motivation.

In conclusion, although much work remains to be completed, this initial series of exploratory studies indicates that quality of performance need not be defined solely in terms of measures of central tendency and that consistency of performance may reflect a need for achievement or similar construct previously obtainable only by less objective means.

REFERENCES

1. Atkinson, J. W. & Litwin, G. H. Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure. Journal of Abnormal and Social Psychology, 60:52-63, 1960.

2. Atkinson, J. W. & Raphelson, A. C. Individual differences in motivation and behavior in particular situations. Journal of Personality, 24:349-363, 1956.

3. Baumeister, A. A. Behavioral inadequacy and variability of performance. American Journal of Mental Deficiency, 73:477-483, 1968.

4. Birney, R. C. The achievement motive and task performance: A replication. Journal of Abnormal and Social Psychology, 56:133-135, 1958.

5. Brown, M. Determinants of persistence and initiation of achievement related activities. In: J. W. Atkinson & J. O. Raynor (Eds.), Motivation and Achievement. Washington, D.C.: Hemisphere, 1974.

6. Bruhn, P. & Parson, O. A. Continuous reaction time in brain damage. Cortex, 7:278-291, 1971.

7. Entwisle, D. R. To dispel fantasies about fantasy-based measures of achievement motivation. Psychological Bulletin, 77:377-391, 1972.

8. Griffin, G. R. & Mosko, J. D. Naval aviation attrition 1950-1976: Implications for the development of future research and evaluation. NAMRL Report 1237, Naval Air Station Pensacola: Naval Aerospace Medical Research Laboratory, 1977.

9. Hicks, L. H. & Birren, J. E. Aging, brain damage, and psychomotor slowing. Psychological Bulletin, 74:377-396, 1970.

10. King, H. E. Psychomotor changes with age, psychopathology, and brain damage. In: A. T. Welford & J. E. Birren (Eds.), Behavior, Aging and the Nervous System, Springfield, IL: Thomas, 1965.

11. Klinger, E. Fantasy need achievement. Psychological Bulletin, 66:291-306, 1966.

12. Korman, A. K. The Psychology of Motivation. Englewood Cliffs, N. J.: Prentice-Hall, 1974.

13. Lowell, E. L. The effect of need for achievement on learning and speed of performance. Journal of Psychology, 33:31-40, 1952.

14. McClelland, D. C., Atkinson, J. W., Clark, R. A., and Lowell, E. L. The Achievement Motive. New York: Appleton-Century-Crofts, 1953.

15. Mehrabian, A. Measures of achieving tendency. Educational and Psychological Measurement, 29:445-451, 1969.

16. Murray, H. A. Techniques for a systematic investigation of fantasy. Journal of Psychology, 3:115-143, 1936.

17. Nettleback, T. and Brewer, N. Effects of stimulus-response variables on the choice reaction time of mildly retarded adults. American Journal of Mental Deficiency, 81:85-92, 1976.

18. Raynor, J. O. and Entin, E. E. Achievement motivation as a determinant of persistence in contingent and noncontingent paths. In: J. O. Raynor and E. E. Entin (Eds.), Motivation, Career Striving, and Aging, Washington, D.C.: Hemisphere, 1982.

19. Remington, R. J. Analysis of sequential effects in choice reaction times. Journal of Experimental Psychology, 82:250-257, 1969.

20. Sarason, S. B. and Mandler, G. Some correlates of test anxiety. Journal of Abnormal and Social Psychology, 47:810-817, 1952.

21. Weiner, B. The effects of unsatisfied achievement motivation on persistence and subsequent performance. Journal of Personality, 33:428-442, 1965.

22. Weiner, B., Johnson, P., and Mehrabian, A. Achievement motivation and the recall of completed and incompleted exam questions. Journal of Educational Psychology, 59(3):181-185, 1968.

23. Welford, A. T. Reaction Times. New York: Academic Press, 1980.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: NAMRL-1299	2. GOVT ACCESSION NO. AD-A135 933
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)*: Performance Consistency on a Perceptual-Motor Task as a Correlate of Achievement Motivation	5. TYPE OF REPORT & PERIOD COVERED
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Gerald B. Thomas and Robert W. Clipper	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Aerospace Medical Research Laboratory, Naval Air Station, Pensacola, Florida 32508	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62758N F58524 MF5852801A 0002
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Medical Research and Development Command, Naval Medical Command, National Capital Region, Bethesda, Maryland 20814	12. REPORT DATE: 17 June 1983
	13. NUMBER OF PAGES: 26
14. MONITORING AGENCY NAME & ADDRESS *(If different from Controlling Office)*	15. SECURITY CLASS. *(of this report)*: UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

Mr. Robert W. Clipper is currently LTjg R. W. Clipper assigned to the Naval Aerospace Medical Institute, Pensacola, Florida.

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Achievement motivation
Auditory reaction time
Performance consistency
Personnel selection

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

A series of studies involving Student Naval Aviators, Naval Flight Officers and Aviation Officer Candidates was conducted to determine whether consistency of performance on a perceptual-motor task was related in any way to scores on traditional pencil-and-paper tests of achievement motivation. Several parameters of a multi-modality reaction time (RT) paradigm were manipulated (including the presence or absence of feedback; stimulus and response probabilities; and paradigm type (simple RT vs. choice RT)) and correlated

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601 |

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

with measures of achievement motivation derived from the Mehrabian Achieving Tendency Scale and the Thematic Apperception Test (TAT). Additional indices investigated included the Test Anxiety Questionnaire, scores on selection tests for student aviators, and RT means and error rates. The results indicated that a particular choice RT paradigm repeatedly resulted in correlations of .43 to .69 ($p < .05$) between performance consistency and TAT-based measures of achievement motivation and that the test-retest reliability of the consistency measure was on the order of .70. No other relationships appeared repeatedly throughout the five studies. A very limited validation study relating relevant measures with the ultimate performance of Experiment I's subjects in flight school indicated that a performance consistency measure might be a useful supplement to existing selection tests in predicting success in flight training. It is recommended that additional research be conducted to replicate and refine the RT consistency measure with the ultimate goal being an objective measure of achievement motivation.