BIBLIOGRAPHY OF WOMBAT ARTICLES

Publications by Aero Personnel

Roscoe, S.N., & Carol, L. (1987). Wonderful original method for basic airmanship testing. *Proceedings of the Fourth International Symposium on Aviation Psychology*, 493-499. Columbus, OH: Department of Aviation, The Ohio State University.

The WOMBAT was developed to aid in the prediction and assessment of pilot performance. It was designed to be minimally impacted by external experiences, such as playing video games and using desktop flight simulators. Hardware for the WOMBAT consists of a left- and a right-hand joystick and a keyboard. The left-hand joystick moves forward and backward and controls two parallel lines on the screen. It is used to track the vertical height of an expanding and contracting hexagon. The right-hand joystick controls a small cross on the screen that is used to track a circle. The right-hand joystick also has a trigger that controls an automatic tracking feature. This feature is available only when the testee is tracking well. The automatic tracking feature, however, is designed to fail periodically. When it fails, it continues to track the input function but does so inaccurately.

The WOMBAT includes three bonus tasks that a testee should perform as frequently as a possible. In the digit-canceling task a series of numbers is presented at short intervals. After three numbers have been presented, the testee must press the number on the keyboard matching the first number presented in the series. In the quadrant location task, the screen is divided into four quadrants with eight numbers in each quadrant. The numbers range from 1 to 32. The testee must identify the numbers in ascending order and press the number on the keyboard corresponding to its quadrant. The three-dimensional figure rotation task consists of two three-dimensional figures that are presented beside each other. The testee must quickly determine if the figures are mirror images, the same image presented in the identical orientation, or the same image presented in a different orientation. A testee must also indicate how certain he/she is of the accuracy of his/her response. If a testee indicates that he/she is certain about a response and the response is correct, the testee will score more points than if the response is correct, but the testee indicates less confidence in the answer. If a testee indicates uncertainty about the answer, and if the response is in fact incorrect, the testee is allowed additional opportunities to provide a correct response.

The tracking tasks can be operated in either velocity mode or acceleration mode. In the velocity control mode, the parallel lines and cross move only when the joysticks are moved. Furthermore, the speed at which the lines and cross move is proportional to the displacement of the joysticks, i.e. the further the joysticks are displaced from neutral, the faster the cursor travels (the greater the velocity). In

the acceleration mode, the greater the displacement of the joysticks from neutral, the greater the acceleration of the cursors. When the joysticks are in the neutral position, the lines and cross move at a constant velocity (zero acceleration).

The WOMBAT program generates a number of performance measures: a total score, a total tracking score, a total bonus score, and a score on each of the three bonus tasks. Other dependent measures are also available for each of the bonus tasks. To score well on the WOMBAT, the testee must determine what isimportant and set priorities accordingly. The testee must remember to do a specific task at a specified future time. Finally, the testee must not allow frustration to impede performance.

Roscoe, S.N., Corl, L., & LaRoche, J. (1997). *Predicting human performance*. Pierrefonds Quebec: Helio Press.

This book reviews many of the issues surrounding operational aptitude testing. The authors note that for aptitude tests to be valid, the complexity of the test should match the complexity of the construct being tested. The authors further state that for complex operations an individual must be able to, "evaluate and integrate information about all relevant events, conditions, and resources, quickly assess changes in situational priorities, and allocate attention accordingly." Tests that assess the skills and abilities required by complex operations will be lengthy of necessity. Such tests, regardless of their administration time, will be cost effective because of the benefits of hiring highly capable applicants.

Roscoe, S.N., Corl, L., & LaRoche, J. (1999). *Keeping the picture. The measurement of flow control*. Pierrefonds Quebec: Helio Press.

This book is concerned with the WOMBAT-FC (Flow Control), which is designed to assess people for jobs involving the dynamic flow of information. Examples of such jobs include air traffic controllers, train and port traffic supervisors, control room operators, and vehicle dispatchers. The book is divided into three major sections. The first section presents a brief history of the development of human engineering as a discipline and its contributions to the aerospace industry. The second section describes the primary and secondary tasks that comprise the WOMBAT-FC and discusses the calculation of concurrent feedback measures. The third and final section deals with software installation, test administration, and the structure of the data files.

Publications in Refereed Journals

Caponecchia, C, Zheng, W.Y., Regan, M. A (2018). Selecting trainee pilots: Predictive validity of the WOMBAT situational awareness pilot selection test. Applied Ergonomics, 73 (2018), 100-107.

This study was designed to test the ability of the WOMBAT test to predict performance outcomes in a sample of ab-initio pilots. The data collected in the study were flight time to first solo, flight time to license, and instructor ratings of student performance at key phases of the flight training curriculum. Information on whether participants discontinued the flight training program was also collected. The authors report that scores on the WOMBAT test were associated with reductions in the time required to reach milestones throughout the training program (roughly 3 minutes to first solo per 1 point WOMBAT increase, and 8 minutes to license. Significant correlations are reported between WOMBAT scores and instructor ratings.

O'Brian, K., S. & O'Hare, D. (2007). Situational awareness ability and cognitive skills training in a complex real-world task. *Ergonomics*, *50* (7), 1064-1091.

The authors conducted three experiments. In all three studies, scores on the WOMBAT were assumed to reflect a situational awareness ability. The goal of the first study was to examine if cognitive skills training could improve situational awareness ability as reflected in the WOMBAT scores. In the second study they assessed the link between underlying situational awareness ability as measured by WOMBAT and performance on TRACON (a PC-based ATC simulation) and SAGAT (a recall test). In the last study they determined if encouraging participants to plan and consider the interrelations of elements in the environment could bolster situational awareness.

Participants in the first experiment were 28 students from the University of Otago who ranged in age from 18 to 36 years old. Participants received TRACON training, and 1 week later were administered the WOMBAT 4.0. Participants were randomly assigned to either the Cognitive Management or Procedural Training group. Following training, participants were required to complete a TRACON simulation test.

Results indicated that type of training was a significant predictor of TRACON test scores [$t(22)= -2.07$, $p=.05$]. However, WOMBAT scores did not significantly predict performance on TRACON [$t(22)=-1.20$, $p>.24$]. Participants who scored lower on the WOMBAT and received procedural training performed more poorly on the TRACON task than did either higher scorers or lower scorers who received cognitive management training. Additionally, compared with those with high WOMBAT scores, participants with low scores delayed more aircraft and made fewer commands.

Participants in the second study were 11 female and 9 male students from the University of Otago, The participants ranged in age from 20 to 39 years old. Participants were trained on TRACON. After initial training, participants engaged in four sessions of approximately 1 hour each. SAGAT measures were taken during the third session, and WOMBAT testing took place in the fourth session.

Scores on SAGAT, TRACON, and WOMBAT were all statistically significantly correlated with each other. WOMBAT, however, had a higher association with TRACON performance [$r(16)=0.691$, $p=0.001$] than with SAGAT performance [$r(16)=0.519$, $p=0.031$].

O'Hare, D. (1997). Cognitive ability determinants of elite pilot performance. *Human Factors, 39*(4), 540-552.

Two studies are reported in this paper. The goal of the first study was to test the theory that the WOMBAT measures an individual's ability to maintain situational awareness. The goal of the second study was to determine if elite pilots have higher levels of situational awareness than other pilots.

In Study 1 participants (24 men ages 26-62) were given both the WOMBAT and the Walter Reed Performance Assessment Battery. The Walter Reed Performance Assessment Battery measures the individual abilities that were hypothesized to underlie performance on the WOMBAT. Participants spent 60 min completing the WOMBAT, and performance scores were recorded every 10 min.

The WOMBAT scores included a total score, a tracking score, a total bonus score, and a score on each of the three bonus tasks: a three-dimensional figure rotation and matching task, a quadrant location task, and a two-back serial digit canceling task. Bonus scores significantly improved over time [$F(5, 90) = 1l.03$, $p<0.0001$], but tracking scores did not. Age only correlated with the percent correct scores on the Manikin Test (part of the Walter Reed Performance Battery). The WOMBAT scores showed significant diminishing correlations with computer game experience across the six time blocks. After the first 10-min block, the correlation between the WOMBAT scores and computer game experience was higher [$r(24)=.64$, $p<0.001$] than at the end of the last 10-min block [$r(24)= .47$, $p<0.05$]. Lastly, scores on the Pattern Recognition Test (part of the Walter Reed Battery) significantly correlated with WOMBAT performance on the last 10-min block [$r(24)=.57$, $p<.01$]. Ultimately, while computer-game experience, pattern recognition ability, and mental rotation ability were found to contribute significantly to initial WOMBAT performance, only pattern recognition ability remained a significant contributor to final WOMBAT performance.

For Study 2 the author selected a group of male pilots attending the Omarama Cup Soaring competition in Omarama, New Zeland. Of the pilots selected, eight were considered elite and six were considered highly experienced. A group of 12 non-pilots matched on age and occupational status were also selected for comparison. The WOMBAT was administered to participants to determine if those who scored higher on the test were also more highly ranked in competitive flying.

Results indicated no significant differences among age or occupational status groups. Elite pilots achieved higher WOMBAT scores than experienced pilots, and experienced pilots had higher scores than non pilots. The WOMBAT was not affected by individual abilities other than pattern recognition.

Government Publications

Bourne, L. E., & Yaroush, R. A. (2003). *Stress and cognition: A cognitive psychological perspective.* (NAG2-1561). National Aeronautics and Space Administration.

This paper is a literature review of cognitive psychology as it relates to issues of stress. The WOMBAT is mentioned as a test of situational awareness and stress tolerance, but no data related to the WOMBAT were presented in this paper.

Darlington, K., Palacio, L. V., Dowler, T., & LeDuc, P. (2006). *Situational awareness, crew resource management and operational performance in fatigued two-man crews using three stimulant countermeasures*. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.

The authors used the Duo WOMBAT with various other tasks to assess the effectiveness of stimulants (dextraomphetamine, caffeine, and modafinil) on the performance of fatigued Army pilots in the USAARL UH-60 flight simulator. Participants were 32 rated, current army aviators who were tested in pairs over a period of 7 days.

Participants were kept awake for 87 hours. The WOMBAT was administered to participants during 3 fixed-interval sessions throughout the day and night. Participants were also required to perform a variety of maneuvers in the simulator for 1 hour. The WOMBAT scores consisted of a total score, a tracking score, a total bonus score and a score on each of the three bonus tasks: a three-dimensional figure rotation task, a two- back serial digit canceling task, and a quadrant location task. Results indicated significant Drug by Session interactions for the digit cancellation [$F(15, 105) = 2.120$, $p<0.001$] and tracking subtasks [$F(15, 105) = 2.717$, $p<0.001$]. For the various flight maneuver tasks, there were no significant session main effects or drug by session interactions. Thus, the WOMBAT was found to be a more sensitive measure of the effectiveness of the stimulants in reducing fatigue than performance in the simulator.

Hyland, D. T., Kay, E. J., & Deimler, J. D. (1994). *Age 60 study, part IV: Experimental evaluation of pilot performance.* (DOT/FAA/AM-94/23). Oklahoma City, OK: Federal Aviation Administration.

In response to the requirement in FAA Part 121 that pilots retire at age 60, the authors examined the relationship between pilot aging and performance. Participants were 40 B727-rated male pilots between the ages of 41 and 71. Participants were administered Cogscreen (a battery of cognitive abilities test), WOMBAT, and Flitescript (a test of pilots' representations of situational awareness in long-term memory). Following the completion of these tests, participants were required to accomplish a series of tasks using a flight simulator. Performance on the simulator tasks was evaluated by trained raters. After completing the tests and simulator tasks, participants were given a questionnaire to assess how

effectively they believed each measure assessed their performance or aptitude to perform in a real-world flying situation.

The WOMBAT scores included a total score, a tracking score, a total bonus score, and a score on each of the three bonus tasks: a three-dimensional figure rotation and matching task, a quadrant location task, and a two-back serial digit canceling task. Cogscreen scores included an accuracy score, a speed score and a total score. Flightscript scores consisted of a response latency score and a response accuracy score.

Only Cogscreen correlated with simulator performance. WOMBAT bonus scores were correlated with all three Cogscreen scores. Additionally, pilots who performed better on the WOMBAT bonus tasks also performed more accurately on the Cogscreen subtests. WOMBAT total scores were also correlated with Cogscreen accuracy and total scores. Prior video game experience was found to correlate with some WOMBAT scores. Pilots who had greater experience with video games having similar qualities to WOMBAT performed better on the WOMBAT tracking (r=.32) and bonus (r=.32) tasks. Flightscript scores were not significantly correlated with either WOMBAT or Cogscreen scores. Lastly, older participants were rated lower than younger participants on simulator maneuvers, and older participants performed more poorly on all three predictor tests.

LeDuc, P., Riley, D., Hoffman, S. M., Brock, M. E., Norman, D., Johnson, P., et al. (1999). ***The effects of sleep deprivation on spatial disorientation.*** **(USAARL 2000-09). Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.**

The purpose of this study was to investigate the effects of fatigue on aviator response to in-flight disorientating events. Participants were eight UH-60 rated pilots who ranged in age from 27 to 48 years old. Participants were required to complete a 1-hour UH-60 simulator flight, objective and subjective measures of alertness, as well as cognitive measures that included WOMBAT, MATB, and SYNWORK. MATB requires subjects to monitor simulated fuel levels, pump status, engine performance, and other aspects of aircraft status while performing a tracking task and occasionally changing radio frequencies. SYNWORK is a battery of tests that includes a Sternberg memory task, an arithmetic task, a visual monitoring task, and an auditory task. Half of the participants completed the testing process first when they were sleep deprived; the other half first, when fully rested. Sleep deprivation lasted for 40 hours.

Participants experienced decreased performance on both the MATB and SYNWORK when fatigued. WOMBAT scores included a total score, a tracking score, a total bonus score, and a score on each of the three bonus tasks: a three-dimensional figure rotation and matching task, a quadrant location task, and a two-back serial digit canceling task. The total bonus scores as well as performance on the three-dimensional figure rotation task and the quadrant location task decreased significantly for fatigued participants as compared to rested participants.

LeDuc, P., Rowe, T., Martin, C., Curry, I., Wildzunas, R., Schmeisser, E., et al. (2009*). Performance sustainment of two man crews during 87 hours of extended wakefulness with stimulants and napping.* (USAARL 2009-04). Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.

This study was designed to assess the effect of countermeasures (dextroamphetamine, caffeine, modafinil, and naps) on simulated sustained operations. Participants were 30 male and 2 female UH-60 rated pilots. The participants ranged in age from 19 to 55 years. Flight performance was assessed in a UH-60 flight simulator. The effects of both the countermeasures and sleep deprivation on cognitive processes were assessed through the use of the WOMBAT, the Duo WOMBAT, the Psychomotor Vigilance Task, and four subtests of the Cambridge Neuropsychological Test Battery. These subtests measured simple and choice reaction time, speed of matching to sample, speed of visual information processing, and spatial planning.

Pairs of participants were assigned to one of three stimulant groups or to a placebo group. The participants completed a familiarization/training period, which was followed by a baseline collection period. The participants then completed an 87-hour sleep deprivation period, which was followed by a 2-hour nap and four more testing sessions.

Of the five tests of cognitive processes, only the test of speed of visual information demonstrated a main effect of drug on performance. Similarly, scores on both the WOMBAT and the Duo WOMBAT showed no significant main effect of drug. All of the tests of cognitive processes showed significant declines in performance with sleep deprivation. In contrast, neither the WOMBAT nor the Duo WOMBAT was sensitive to sleep deprivation.

Tirre, W. C., & Gugerty, L. J. (1999). *A cognitive correlates analysis of situational awareness.* (AFRL-HE-AZ-TR-1998-0086). Brooks Air Force Base, TX: U.S. Air Force Research Laboratory.

As part of the literature review of this study, the WOMBAT was mentioned as a measure of situational awareness. The authors referred to the O'Hare (1997) publication regarding the WOMBAT and suggested that the .63 correlation between the first and final 10 min of WOMBAT performance could be used as a reliability estimate.

Other

Emery, C.D. and Holding, D.H. (1993). Practice effects on the WOMBAT device. *Proceedings of the International Symposium on Aviation Psychology*, 394-397. Columbus, OH: Department of Aviation, The Ohio State University.

This study compared the performance of 20 commercial airline pilots and 20 college students on the WOMBAT. The commercial pilots were all management pilots at a large international air carrier. The students performed the WOMBAT twice, approximately 1 week apart; the airline pilots, once.

The 90-min WOMBAT testing session was divided into 10-min intervals and the performance of the pilots was compared to that of the first testing session of the students. The pilots performed better than the students across all nine time intervals. The pilots also continued to improve throughout the testing session, whereas the students' scores declined in the last 30 min of the testing session. No test-retest reliability information is presented for the students.

Additionally, the authors obtained supervisory ratings of the commercial pilots' stick-and-rudder skills, cockpit resource management skills, and scores on an unspecified number of flight training check rides. The stick-and-rudder and cockpit resource management skills were rated on a scale of 1 to 5. The correlations between the WOMBAT scores and the supervisors' ratings and the check ride scores ranged from -.12 to .27.

Frey, B. F., Thomas , M., Walton, A. J., & Wheeler, A. (2001). WOMBAT as an example of situational awareness testing in pilot selection: An argument for the alignment of selection training, and performance. *Proceedings of the 11th International Symposium on Aviation Psychology*, 1-6. Columbus, OH: Department of Aviation, The Ohio State University.

The authors conducted two studies. The first study was designed to determine the predictive validity of the WOMBAT for training performance in ab initio pilot students. The second study tested the hypothesis that the WOMBAT scores were negatively related to situational awareness error scores in a flight simulator.

In Study 1 participants were 30 male and 2 female ab initio students with ages ranging from 18 to 31 years. Each participant completed the WOMBAT test as a preliminary requirement for acceptance into the course. During the initial states of the course a flight exam was given and approximately 8 months into the program participants were given another flight exam. WOMBAT was used to predict three measures of training performance: 1) flying hours at the time of the flight exam, 2) the first flight exam score, and as a measure of academic performance, their 3) grade average.

WOMBAT scores included a total score, a tracking score, a total bonus score, and a score on each of the three bonus tasks: a three-dimensional figure rotation and matching task, a quadrant location task, and a two-back serial digit canceling task. The only statistically significant relation was a positive correlation between grade average and the WOMBAT total bonus score [$r(30) = .364$, $p<.05$].

For Study 2, participants were 20 male and 1 female ab initio students with ages ranging from 18 to 31 years. Participants prepared a flight plan and began a cross country simulated flight. The simulation

was periodically frozen and students were asked to complete questionnaires related to their situational awareness at each time period.

There was a negative correlation between the total WOMBAT bonus scores and situational awareness errors made during the flight simulation [$r(19)= -.423$, $p<.05$] (the better the WOMBAT score, the fewer the errors). The multiple correlation between the total WOMBAT scores, the total bonus scores, and flight simulator performance was not significantly higher than the correlation between the total bonus scores and flight simulator performance.

Grugle, N. L. (2005). ***Understanding the effects of sleep deprivation on executive function, complex task performance and situation awareness*****. Unpublished Dissertation, Blacksburg, VA.**

This dissertation explored the impact of sleep deprivation on cognitive processes, complex task performance, and situational awareness. Five tasks were used to assess the effect of sleep deprivation on cognitive processes. One of these, the Psychomotor Vigilance Task, requires sustained attention and is known to be sensitive to sleep deprivation. Three tasks were from the Kit of Factor Reference Tests and assessed inductive reasoning, deductive reasoning, and planning. The fifth task assessed decision making. The WOMBAT-FC was used to measure the effect of sleep deprivation on complex task performance and situational awareness.

The participants were 26 men and 22 women ranging in age from 18 to 39 years old. Baseline data were obtained prior to the beginning of the sleep deprivation period. Each WOMBAT testing session required 60 min.

The WOMBAT total score actually improved from the baseline (mean score 94.3, SD=50.1) to the sleep-deprived condition (mean score 98.9, SD=55.5), as did the total bonus score. Grugle also examined performance on the tracking task and each of the three bonus tasks. The three-dimensional rotation task and the quadrant location task showed significant improvements in performance from the baseline to the sleep deprivation condition. The two-back memory task and the tracking task showed no significant effects. Similarly, of the five tests assessing cognitive processes, only the Psychomotor Vigilance Task showed an effect of sleep deprivation.

Grugle divided the 60-min WOMBAT testing period into 5-min increments and examined changes in performance with time on task. Performance in both the baseline and sleep deprived conditions improved with time on task. The slopes of the best-fitting lines did not differ between the two conditions, indicating that the rate of improvement was statistically indistinguishable under the baseline and sleep deprived conditions. The intercept for the sleep deprivation condition was slightly higher than the intercept for the baseline condition, suggesting a small carry-over effect from the baseline to the sleep deprivation condition. This difference, however, was not significant.

Because tests of cognitive processes were included in this study, the relation between the total score on the WOMBAT and scores on the tests of cognitive processes was examined. Grugle regressed scores from the four tests of cognitive processes on the total WOMBAT score. Scores from the working memory, inductive reasoning, and planning tasks significantly predicted total score in the baseline condition. The model accounted for 32.8% of the variance in the total WOMBAT score.

Prew, S., J. (1997). *Pilot Selection and Recruitment*. *Civil Aviation Training*, 8(5), 12-16.

The author reviewed the WOMBAT and provided a description of both the standard and the Duo WOMBAT tests.

Telfer, R. (1986). Pilot judgment training: The Australian study. *Proceedings of the Third International Symposium on Aviation Psychology*, 265-273. Columbus, OH: Department of Aviation, The Ohio State University.

This study was designed to assess the validity of pilot judgment training. The author asserted that airmanship is best tested through novel situations to avoid contamination from prior experiences. The author believes the WOMBAT is such a test.

For more information, please contact

Aero Innovation inc.

info@aero.ca

www.aero.ca