Female and Male Aviators Are Not Affected Differently By Sleep Deprivation and Continuous Task Demands

By
John A. Caldwell, Jr., Ph.D.
J. Lynn Caldwell, Ph.D.

Aircrew Health and Performance Division

June 1996

Approved for public release; distribution unlimited.

U.S. Army Aeromedical Research Laboratory
Fort Rucker, Alabama 36362-0577
Notice

Qualified requesters

Qualified requesters may obtain copies from the Defense Technical Information Center (DTIC), Cameron Station, Alexandria, Virginia 22314. Orders will be expedited if placed through the librarian or other person designated to request documents from DTIC.

Change of address

Organizations receiving reports from the U.S. Army Aeromedical Research Laboratory on automatic mailing lists should confirm correct address when corresponding about laboratory reports.

Disposition

Destroy this document when it is no longer needed. Do not return it to the originator.

Disclaimer

The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation. Citation of trade names in this report does not constitute an official Department of the Army endorsement or approval of the use of such commercial items.

Human use

Human subjects participated in these studies after giving their free and informed voluntary consent. Investigators adhered to AR 70-25 and USAMRDC Reg 70-25 on Use of Volunteers in Research.

Reviewed:

Richard R. Levine
LTC, MS
Director, Aircrew Health and Performance Division

Released for publication:

John A. Caldwell, Ph.D.
Chairman, Scientific Review Committee

Dennis F. Shanahan
Colonel, MC, MFS
Commanding
Recent changes in U.S. military regulations have brought about the assignment of female aviators to combat roles. This has been the topic of serious debate, but there are few facts concerning differences between males' and females' abilities to withstand combat-relevant stressors. This study was conducted to determine whether there are gender differences in responses to a common operational stressor, sleep deprivation. Six male and six female UH-60 helicopter pilots were exposed to a 40-hour period of continuous wakefulness and tested on flight performance and mood. The flight performance results indicated that gender produced virtually no operationally significant differences in the effects of sleep loss. Furthermore, although mood evaluations showed that women felt less tense and more energetic overall than their male counterparts, there were no interactions between sleep deprivation and gender. Thus, male and female aviators appear equally capable of performing flight-related tasks despite moderate sleep loss.
SUMMARY

Recent changes in U.S. military regulations have brought about the assignment of female aviators to combat roles. This has been the topic of serious debate, but there are few facts concerning differences between males' and females' abilities to withstand combat-relevant stressors. This study was conducted to determine whether there are gender differences in responses to a common operational stressor, sleep deprivation. Six male and six female UH-60 helicopter pilots were exposed to a 40-hour period of continuous wakefulness and tested on flight performance and mood. The flight-performance results indicated that gender produced virtually no operationally-significant differences in the effects of sleep loss. Furthermore, although mood evaluations showed that women felt less tense and more energetic overall than their male counterparts, there was no indication of interactions between sleep deprivation and gender. Thus, male and female aviators appear equally capable of performing flight-related tasks despite moderate sleep loss.
Table of contents

Introduction ......................................................................................... 3
Methods ................................................................................................. 4
   Apparatus ......................................................................................... 4
   Procedure ......................................................................................... 5
Results .................................................................................................. 6
   Flight performance ........................................................................... 6
   Subjective mood data ....................................................................... 13
Discussion ............................................................................................. 17
References ............................................................................................. 19
Appendix A - Manufacturer’s list ......................................................... 21

List of figures

1. Effects of gender and iteration on performance of the low-level navigation .... 7
2. Effects of gender and time of day on the straight-and-levels ....................... 8
3. Effects of gender and time of day on performance of the climbs .................. 9
4. Effects of gender and time of day on performance of the left standard-rate turns . 10
5. Effects of gender and iteration on performance of the descents .................... 11
6. Effects of gender and time of day on performance of the descents ............... 11
7. Effects of time of day and iteration on performance of the right standard-rate turns . 12
List of figures (Continued)

8. Effects of gender and time of day on performance of the left descending turn ........... 13

9. Effects of gender and time of day on tension-anxiety scores .......................... 14

10. Effects of gender and time of day on depression-dejection scores .................... 15

11. Effects of time of day on vigor-activity scores ........................................... 16

12. Effects of gender and time of day on fatigue-inertia scores ............................ 16

13. Effects of gender and time of day on confusion-bewilderment scores ............... 17
Introduction

In April of 1993, the United States' military restrictions on women in combat were relaxed, and shortly thereafter, the first female pilots began training to fly attack missions. The lifting of this 45-year-old law was debated throughout the nation, but little empirical data was presented to support positions on either side (Nelan, 1993; Johnson, 1991; Donnelly, 1991). Opinions about gender equality, the suitability of women for strenuous duty, sending young mothers into combat, and the possibility of women war casualties were considered by the U.S. Congress and the Department of Defense (Hurrell and Lukens, 1994).

Presently, women are authorized to serve in combat aviation roles and many other military occupations with the exception of "combat arms" positions in Army and Marine ground forces. The impact this will have on future military conflicts remains to be determined. However, at present there is little evidence that there are differences in how well males and females can cope with militarily-relevant stressors in aviation.

Flying an aircraft is a complex cognitive task which requires a high level of intelligence and excellent sensory capacities (hearing and eyesight), but not necessarily physical strength. Since gender differences in intelligence, hearing, and vision are negligible (Tyler, 1965), there is little reason to believe that females are less competent than males under normal flight circumstances. However, it is conceivable that gender might interact with operational stressors in ways that could create performance differences between men and women in some situations. One possibility is that female and male pilots may be affected differently by sleep deprivation, although presently there is no reason to believe this is the case. Such a possibility should be explored because of the likelihood that sleep loss and fatigue will be encountered in operational aviation scenarios.

Generally, fatigue produces decrements in performance and skill after continuous mental and/or physical work (Porcu et al., in press). Fatigue also results from the sleep deprivation commonly encountered during sustained operations. Inadequate sleep degrades the central nervous system as evidenced by increases in EEG theta activity and reductions in EEG alpha power (Lorenzo et al., 1995). The operationally-significant impact of sleep loss is the impairment in soldier performance due to slower reactions, reductions in vigilance, decrements in cognitive abilities, and changes in affect (Krueger, 1991).

The failure to attain adequate sleep can be dangerous and costly. According to one report, sleep deprived drivers contributed to 1,225 traffic fatalities, 45,000 disabling injuries, and almost 2 billion dollars worth of accident-related costs during 1988 in the United States alone (Webb, 1995). Sleep-deprived pilots may account for a significant proportion of aircraft accidents as well, but this is difficult to substantiate since mishap reports are more likely to ascribe causation to global factors such as "human error." However, it is noteworthy that "human error" has been identified by Ramond and Mozer (1995) as a contributing factor in
50-80 percent of aviation mishaps, and there is evidence that fatigue has been the cause of deficient resource allocation, impaired attention, and inadequate control in the cockpits of ill-fated flights (Billings and Reynard, 1984). Whether females and males differ in their susceptibility to these problems is unknown.

The present study explored the effects of sleepiness and fatigue on the flight performance and psychological states of helicopter pilots. In addition, since there have been no published studies examining gender differences in the effects of fatigue on pilots, this study compared the responses of males and females.

Methods

Six female and six male UH-60 pilots were tested at the U.S. Army Aeromedical Research Laboratory (USAARL). Subjects were not permitted to consume caffeine during the protocol. Only one of the subjects was a cigarette smoker. The average amount of flight experience for the females was 748 hours, and the average amount for the males was 723 hours. The average weights of the women and men were 133.8 and 166.0 pounds respectively. The data reported here represents a reanalysis of a portion of data from an investigation of the effects of a stimulant versus a placebo. However, all of the data in this report are from the placebo condition only.

Apparatus

UH-60 flight simulator

Flights were conducted in a UH-60 helicopter simulator with a 6 degrees of freedom motion base and a full-fidelity visual cockpit. Flight data (heading, airspeed, altitude, etc.) were acquired with a DEC VAX 11/780.* The acquired data were converted to composite flight scores (Jones & Higdon, 1991).

Profile of mood states

Subjective mood evaluations were made with the Profile of Mood States (POMS). The POMS is a 65-item paper and pencil test which measures affect or mood on six scales: tension-

*See manufacturer’s list
anxiety, anger-hostility, depression-dejection, vigor-activity, fatigue-inertia, and confusion-bewilderment (McNair, Lorr, and Droppleman, 1981).

**Procedure**

Each subject completed several simulator flights and POMS questionnaires throughout a baseline day (prior to sleep deprivation) and during the actual sleep deprivation period. In addition, there were other tests which will not be reported here.

**Flight performance**

During each simulator flight, subjects performed a profile of standard flight maneuvers. These consisted of low-level navigation and nontactical, upper-airwork maneuvers which required subjects to fly the simulator "on instruments" using no external visual references. There were 16 standardized maneuvers which the subjects flew each time. The first group of upper-airwork maneuvers was flown with the automatic flight control system (AFCS) trim engaged (the normal mode when flying the UH-60), and the second group was flown with the AFCS trim turned off. The AFCS trim system enhances the stability and handling qualities of the aircraft/simulator, and when the AFCS is turned off, accurate flight control becomes much more difficult. Following the low-level navigation, there were four straight-and-levels (1 with AFCS off), two left standard-rate turns (1 with AFCS off), three right standard-rate turns (1 with AFCS off), two standard-rate climbs (both with AFCS on), three standard-rate descents (all with AFCS off), and one left descending turn (with AFCS off).

During each maneuver, subjects were required to maintain precise control over the parameters important for that specific type of maneuver (i.e., heading, altitude, airspeed, etc.). For instance, heading control was evaluated during straight-and-level flight, but not during turns. Scores indicative of how well the subject flew each maneuver were calculated in two steps. First, the control scores for the parameters relevant to each maneuver were determined using limits which were sufficiently stringent to prevent ceiling effects (e.g., subjects rarely were able to attain a perfect score). For each parameter, as long as subjects maintained aircraft control within the most precise specified limits, a perfect score of 100 would result. For example, if a subject never deviated from the assigned heading by more than 1 degree, he/she would earn a heading control score of 100. Larger deviations produced lower scores. Second, the scores from each parameter were averaged into a single composite score. Thus, if a subject scored 100 on heading, 85 on altitude, and 90 on airspeed, he/she earned a composite score of 91.7 for that particular maneuver.
Profile of mood states

During POMS questionnaires, subjects indicated how they felt in terms of 65 "mood states." Answer sheets were hand-scored to yield results on six dimensions (tension-anxiety, anger-hostility, depression-dejection, vigor-activity, fatigue-inertia, and confusion-bewilderment).

Test schedule

Upon arrival at USAARL, subjects were given a medical evaluation after signing an informed-consent agreement. Subjects with past psychiatric or cardiac disorder, a history of sleep disturbances, or any current significant illness would have been rejected, but none of these problems were found. After the first day, subjects received three training sessions on the UH-60 simulator flights and the POMS. After training, subjects went to sleep at 2300. The continuous wakefulness period began after a full night of sleep (the subject was awakened at 0700). He/she then completed three baseline sessions (at 0900, 1300, and 1700), each of which included the simulator flight and the POMS. In addition, there was a POMS given at 2340 on the baseline day. The aviator was not allowed to sleep after baseline testing. Instead, he/she began testing under sleep-deprivation at 0100. On deprivation days, there were five equally-spaced sessions (at 0100, 0500, 0900, 1300, and 1700) as well as a final POMS administration (at 2225) prior to recovery sleep.

Results

Flight performance

General

Composite flight scores from each maneuver were analyzed in repeated measures analyses of variance (ANOVA) in which the factors were gender (males, females), session (0100, 0500, 0900, 1300, 1700) and, where appropriate, iteration (i.e., turn 1, turn 2, turn 3). Significant effects were followed up with analysis of simple effects and/or contrasts. Corrections for violations of the compound symmetry assumption were made by using the Huynh-Feldt adjusted degrees of freedom.

Navigation

The composite flight scores (based on heading, altitude, slip, and roll) for the four parts of the navigation segment of the flight profile (leg 1-leg 4) indicated there was an interaction between gender and navigation leg (F(3,30)=4.83,p=.0074). Analysis of simple effects showed
that, while there were differences among the legs in both males and females (p<.05), the pattern of differences was affected only slightly by gender as can be seen in Figure 1. In the males, performance on the second leg was better than performance on all the others, and performance on the third leg was worse (p<.05). In the females, performance on the second leg was not different from performance on any of the others with the exception of the third leg, which again, was worse than all the others (p<.05). There were also overall differences among the navigation legs regardless of gender (F(3,30)=43.08, p<.0001). This was because (with gender collapsed) the second leg was best and the third leg was worst (p<.05). There were no other effects on the navigation portion of the profile, but the session effect approached significance (p=.07). There were no overall gender differences.

![Composite Score vs Navigation Legs](image)

**Figure 1.** Effects of gender and iteration on performance of the low-level navigation.

**Straight-and-levels**

The ANOVA on the composite scores (based on heading, airspeed, altitude, slip, and roll) during the straight and levels (SLs) revealed two main effects. The first was a session effect (F(4,40)=6.16, p=.0006) which was due to better performance at 0100 than at any other time of the day and poorer performance at 0900 than at 1300 or 1700 (p<.05). This effect is depicted in Figure 2. The second was an iteration effect (F(2,06,20.58)=26.19, p<.0001) because of performance being best on the first SL and poorest on the fourth (p<.05). That the fourth SL
produced the lowest scores is not surprising given that it was flown without the AFCS trim system engaged. There were no gender-related differences on this maneuver.

Figure 2. Effects of gender and time of day on the straight-and-levels.

**Climbs**

The composite scores (based on heading, airspeed, slip, roll, and climb rate) on the two climbs (both flown with the AFCS engaged) also indicated statistically significant effects on the iteration factor ($F(1,10)=13.24, p=.0045$) and the session factor ($F(4,40)=3.99, p=.0081$). The iteration effect was because performance on the first climb was better than performance on the second. The session effect was due to declining performance between 0100 and 0900 and between 0500 and 0900 ($p<.05$). However, performance at 1300 recovered to the level seen earlier at 0100 so that it was better than what was observed at 0500 and 0900 ($p<.05$) as can be seen in Figure 3. There were no gender-related effects on this maneuver.
Figure 3. Effects of gender and time of day on performance of the climbs.

Left standard-rate turns

The composite scores (based on turn rate, altitude, airspeed, slip, and roll control) from the two left standard-rate turns revealed iteration (F(1,10)=88.57, p<.0001) and session (F(4,40)=3.69, p=.0119) differences. The iteration effect was due to the first turn (with AFCS engaged) being better than the second (without the AFCS). The session effect was due to better performance at 0100 than at 0500 (p=.057) and 0900 (p<.05), and poorer performance at 0900 than at 1300 when there was an afternoon recovery in flight abilities (p<.05). This time-of-day effect is shown in Figure 4. Whether the subjects were male or female did not affect performance in any respect.
Figure 4. Effects of gender and time of day on performance of the left standard-rate turns.

Descents

The composite scores (based on heading, airspeed, slip, roll, and descent rate) during the three descents (all flown without the AFCS) revealed several effects. Most important was the interaction between gender and iteration (F(1.96,19.57)=4.99, p=.0183) which occurred because of a substantial drop in performance between the first and third descents in the males (p<.05) and the absence of a similar decline in the females (see Figure 5). In addition to this effect, there was an overall difference among the descents with gender collapsed (F(1.96,19.57)=3.62, p=.0467). This supports what was seen in the males in that there was a significant decline from the first to the third descent (p<.05) while none of the other comparisons showed differences. Finally, there was a main effect on the session factor (F(4,40)=9.93, p<.0001) due to the fact that performance was better at 0100 than at any other time of day, and performance at 0500 was better than performance at 0900; however, performance recovered in the afternoon so that it was better at 1300 than it was at 0500 or 0900, and it was better at 1700 than it was at 0900 (p<.05), as can be seen in Figure 6.
Figure 5. Effects of gender and iteration on performance of the descents.

Figure 6. Effects of gender and time of day on performance of the descents.
Right standard-rate turns

The composite scores (based on turn rate, altitude, airspeed, slip, and roll control) on the three right turns indicated an interaction between iteration and session ($F(8,80)=2.17$, $p=.0388$) and a main effect on the iteration factor ($F(2,20)=3.99$, $p=.0081$). The interaction, shown in Figure 7, was due to significant differences among the sessions only during the last turn (flown with the AFCS turned off) where performance was better at 0100 than at 0500, 0900, or 1300; better at 0500 than at 0900; and worse at 0900 than at 1700 ($p<.05$). The iteration main effect was attributable to the fact that performance on the third turn was lower than performance on the first or the second, and that performance on the first turn was lower than performance on the second ($p<.05$). There were no gender-related differences on this maneuver.

![Figure 7](image)

Figure 7. Effects of time of day and iteration on performance of the right standard-rate turns.

Left descending turn

The scores on the left descending turn (based on turn rate, airspeed, slip, roll, and descent rate) showed there was an overall session effect ($F(4,40)=2.94$, $p=.0322$) which was partially consistent with what was observed in the other maneuvers. In this case, performance at 0100 was clearly better than performance at 0500 and 0900 ($p<.05$), as can be seen in Figure 8. However, the typical afternoon recovery in performance which was seen in several other maneuvers did not attain statistical significance here.
Figure 8. Effects of gender and time of day on performance of the left descending turn.

Subjective mood data

General

Data from each of the six scales of the POMS were analyzed with repeated measures ANOVAs in which the grouping factor was gender (male, female) and the within factor was session. Unlike the flight data, which was not collected late in the evenings, there were POMS administrations at 2340 on the baseline day (preceding sleep deprivation) and at 0340, 0740, 1140, 1540, 1940, and 2225 on the deprivation day. Thus, there were seven levels of the session factor in this analysis. Significant effects were followed up with posthoc contrasts (there were no interactions necessitating analysis of simple effects).

Tension-anxiety

The analysis of the tension-anxiety scores revealed an overall difference between males and females on this scale ($F(1,10)=8.79$, $p=.0142$) due to the males expressing greater tension.
than the females (the mean scores were 8.5 and 2.8 respectively). There also was a session main effect (F(6,60)=4.65, p=.0006) due to lower tension scores at 2340 (before sleep deprivation) than during every session except 1940 and 2235 on the deprivation day; higher tension at 0740 than at 1140, 1940, and 2225; and higher tension at 1540 than at 1940 or 2235 (p<.05). These differences are depicted in Figure 9. There was no gender-by-session interaction.

Figure 9. Effects of gender and time of day on tension-anxiety scores.

Depression-dejection

The depression-dejection scores showed differences across the testing sessions as well, but there were no other significant effects. The session effect (F(6,60)=2.53, p=.0298) was due to lower depression scores at the end of the baseline day than at 0340, 0740, and 1140 on the deprivation day; and higher scores at 1140 on the deprivation day than at 1940 and 2225 (p<.05), as can be seen in Figure 10.
Figure 10. Effects of gender and time of day on depression-dejection scores.

Vigor-activity

Scores from the vigor-activity scale revealed a main effect attributable to gender (F(1,10)=5.68, p=.0384), and a main effect attributable to session (F(6,60)=5.39, p=.0002), but there was no gender-by-session interaction. The overall difference between males and females was due to the fact that vigor scores were higher among the women than they were among the men (the mean scores were 16.0 versus 9.3 respectively). The session effect was due to higher vigor at the end of the baseline day (at 2340) than at any of the sessions on the deprivation day (p<.05). In addition, vigor scores were lower at 0740 on the deprivation day than they were at 1540 or 1940 (p<.05). These session effects are depicted in Figure 11.

Fatigue-inertia

The analysis of the fatigue-inertia scores showed no differences between males and females; however, there was a main effect on the session factor (F(6,60)=12.29, p<.0001). This was due to lower fatigue scores at 2340 on the baseline day than at any of the sessions on the deprivation day (p<.05). Also, fatigue was lower at 0340 than at 0740, 1140, and 2225, while fatigue was higher at 0740 than at 1540 or 1940 (p<.05) (See figure 12).
Figure 11. Effects of time of day on vigor-activity scores.

Figure 12. Effects of gender and time of day on fatigue-inertia scores.
Confusion-bewilderment

The analysis of confusion-bewilderment scores indicated no gender differences, but as was the case with fatigue scores, there was a session main effect ($F(6,60)=7.70$, $p<.0001$). Contrasts for this effect showed that there were lower scores at 2340 on the baseline day than there were at any of the sessions on the deprivation day ($p<.05$). Also, the scores were lower at 0340 than at 0740, while they were higher at 0740 than at 1140, 1540, 1940, or 2225 ($p<.05$) as can be seen in Figure 13.

![Figure 13. Effects of gender and time of day on confusion-bewilderment scores.](image)

Discussion

Both the flight data and the subjective mood reports showed that there were no practically-significant differences between the males and females tested in this study. In fact, of the seven sets of flight maneuvers examined here, there were no overall gender effects and no interactions between gender and sleep deprivation on any of them. On two of the seven sets (low-level navigation and straight descents), there were interactions between gender and the maneuver iterations which, in both cases, were due to larger differences across the iterations in males than in females; however, the operational significance of this effect is small given there
were no significant differences between males and females within any of the individual iterations.

Besides these performance effects, the POMS data did show that the males were more tense and anxious than the females and that the females felt more vigorous than the males throughout the deprivation period. Reasons for this discrepancy are not readily apparent, but it is interesting to note that these overall differences did not interact with the degree of sleep deprivation. In other words, there were basic differences between the self-reported feelings of the men and women, but they did not become differentially more or less pronounced as a function of sleep deprivation. Also, it is noteworthy that while the women reported feeling more vigorous than the men throughout the deprivation period, this did not translate into superior performance.

Instead, the performance of the men and women was remarkably similar. As was revealed in the session main effects, performance always started out at the highest level at 0100 (only 2 hours past normal bedtime) and then declined significantly to its lowest at 0500 or 0900. Performance then frequently increased from the morning to the afternoon (usually at 1300) and often remained high at 1700. However, afternoon performance did not recover to the level seen at the first test session of the day. This pattern of morning degradation followed by an afternoon circadian recovery is consistent with data reported elsewhere in the literature (Collins, 1977; Penatar et al., 1993).

The effects of sleep deprivation on flight performance were consistent with the effects of sleep loss on subjects' self-reported mood states. For instance, vigor ratings declined steadily from baseline to 0740 and then gradually improved from 0740 to 1940; whereas fatigue increased from baseline to 0740 and then declined from 0740 to 1940. Thus, the subjects (both males and females) had fairly accurate perceptions of their own alertness difficulties as the sleep deprivation period progressed.

In conclusion, the data from the present study show that while both male and female aviators are compromised by inadequate sleep, these effects are not differentially accentuated or diminished as a function of gender. Thus it appears that men and women are equally capable of enduring a major battlefield stressor (i.e., inadequate sleep). Future work should be conducted to clarify whether soldier gender may have an impact on resistance to other types of operational strains.
References


References (Continued)


Appendix A.

Manufacturer's list

Digital Equipment Corporation
P.O. Box C52008
Nashua, NH 03061-2008