Original Article

Effect of hypobaric hypoxia on subjective perception of mental workload measured with NASA Task Load Index

Dahiya YS [*], Tripathi KK [+]

ABSTRACT

NASA Task Load Index (NASA TLX), employed for prototype and systems evaluation, is a multidimensional subjective workload index which provides an overall workload score based on a weighted average of ratings in six dimensions. Out of its six dimensions, three relate to demands imposed and the other three relate to interaction of the subject with the task. Hypoxia may affect many of these constituent dimensions. Present study examined the effect of hypobaric hypoxia on the subjective perception of mental workload measured using NASA TLX. The study examined the effect of short term exposure (15 min) to mild hypobaric hypoxia (10,000 ft) and moderate hypobaric hypoxia (15,000 ft) simulated in a hypobaric chamber on NASA TLX measured during engagement of 16 healthy male volunteers in a working memory task with two difficulty levels. Order of administration of the two task variants and three altitude conditions was counterbalanced. Statistical procedure involved a two way analysis of variance. Results showed that the ratings in all dimensions of NASA TLX (except frustration) exhibited significant effect of task difficulty. However, no significant effect of task difficulty was observed on weights. On the other hand, no significant effect of hypoxia was observed in ratings, weights and weighted workload in any of the dimensions of NASA TLX as well as overall weighted workload despite a fairly high internal consistency amongst the constituent dimensions of NASA TLX as discerned through Cronbach's Alpha. In conclusion, NASA TLX scores did conform to objectively defined task difficulty and demands. Also, the scores provided some diagnostic information regarding the source of load in the task. Results support the contention that subjective verbal reports do constitute legitimate psychological data. Workload derived from subjective measures (such as NASA TLX) is not sensitive enough to be affected by small deficits in the working memory during short term exposure to mild to moderate hypoxia.

IJASM 2009; 53(2): 34-43

Keywords: Mental workload, NASA TLX, Hypoxia, Working memory

Mental workload is a concept employed for prototype and system evaluation [1]. Despite lack of a universally accepted definition and imperfection in assessment techniques, researchers have found measuring mental workload to be useful. Federal Aviation Administration (FAA) and European Joint Airworthiness Regulations (JAR) have specified that a formal mental workload assessment must be conducted as part of the certification procedure for all new aircraft and subsequent to any modifications [2]. However, the authorities remain 'delightfully vague' regarding its methodology [3].

The measures of mental workload assessment should be unobtrusive, reliable, sensitive to a wide range of variations and diagnostic with respect to change in demand. A number of psychological and physiological tools have been tried for the evaluation of mental workload. However, no tool has, so far, been made available to meet all these qualitative requirements [4, 5] and researchers mostly use subjective indices for the purpose. This is due to their high face validity, simplicity, sensitivity to a variety of work conditions, user acceptance and ease of administration. The subjective indices can be uni-dimensional or multi-dimensional. NASA

* Chief Instructor, No 1 AMTC, IAF, Hindon

+ DPMO & Senior Advisor (Av Med), Central Air Command, IAF

Date of submission: 25 Sep 09

Date of approval : 30 Nov 09

Task Load Index (NASA TLX) is one such measure which evaluates workload on six dimensions viz. mental demands, physical demands, temporal demands, own performance, effort and frustration.

There is evidence in the literature to conceive that hypoxia affects certain higher mental functions [6-21] which in turn may affect some of the constituent dimensions of NASA TLX. Hence, NASA TLX may be affected in hypoxia which is one of the stressors most commonly encountered in aviation. Despite technical advancements, exposure to mild/moderate hypoxia is commonplace.

Federal Aviation Regulation (FAR) 91.211 prescribes requirement of supplemental oxygen for the pilots in general aviation only above 12,500 ft Mean Sea Level (MSL). Directorate General of Civil Aviation (DGCA) regulations also permits unrestricted flying up to 10,000 ft without supplemental oxygen and flying for 30 minutes without oxygen supplementation between 10,000 ft and 13,000 ft [22]. Hypoxic exposures can also occur due to an inboard leak (i.e. a leak of cabin air into the mask cavity).

Exposure to mild hypoxic insult is known to affect mood state [7, 17], cognitive attributes [8,10,11,12,16,17] and working memory [9,17]. It is also conceived to have an effect on time estimation due to an indirect modulation through an increase in heart rate [20,21].

It would be reasonable to infer that mental workload and handling qualities of an aircraft (workload indices are often used to characterise the handling qualities of an aircraft) as assessed using subjective workload indices are liable to variation on exposure to hypoxia which occurs commonly during flying.

The present study examined the effect of mild to moderate hypoxia on performance and on one or more of the six constituent dimensions of NASA TLX in a PC based task involving visuo-spatial working memory with two levels of difficulty on the ground level (Bangalore: 3,159 ft AMSL) and at 10,000 ft and 15,000 ft simulated in an altitude chamber.

Materials and Methods

Experiment Design: It was a 'within subject (repeated measure)' design wherein each subject was examined six times viz. at ground level, at 10,000 ft and 15,000 ft simulated altitudes breathing air with two task variants depending upon the difficulty levels i.e. 'Difficult' and 'Easy' ones. Both easy and difficult tasks were administered three times (under normoxia, hypoxia 10,000 ft and hypoxia 15,000 ft conditions).The subjects were randomly allocated to one of the four groups which were defined according to the sequences of hypoxia exposure and the difficulty of task administered. Four subjects were allocated, at random, to a particular sequence. In each sequence, task administration was counterbalanced for the two difficulty levels as well as two altitudes.

Subjects: 16 male subjects (Age 29±5 yr; Height 174±9 cm; Weight 77±14 Kg) participated in the study. The subjects were essentially low landers and were ascertained to be healthy through history and clinical examination. The experimental protocol and risks involved were explained to them and a written consent was obtained. Ability to ventilate middle ears was ensured through the examination of ears for the mobility of tympanic membranes with the Valsalva manoeuvre on ground and a standard 'ear clearance' run.

Experimentation: Experiment was conducted in the decompression chamber installed at the Department of High Altitude Physiology and Hyperbaric Medicine at Institute of Aerospace Medicine, Bangalore, at the same time of the day, (between 1300 h and 1600 h) to avoid circadian variations in the performance. Every subject was asked to practice the working memory task at least twice for 5 minutes (at ground level). This stabilized the performance of the subject and made him familiar with the test procedure. Difficulty in the tasks was varied through manipulations in task attributes. Each subject performed the working memory task (easy or difficult) which was immediately followed by estimation of mental workload in performing that task using NASA TLX at ground level and then at simulated altitude. Once the desired altitude (10,000 ft or 15,000 ft) was attained, the task was performed after an interval of 15 minutes. At least 72 hr gap was ensured between the two recordings in the same subject.

Working Memory Task: A PC based 'Visuo-spatial Working Memory Task' was administered. In the task, the subject was asked to identify a target stimulus in an array of similar stimuli. The stimuli were one of the 90 geometrical figures of different shapes. Size of these geometrical figures was approximately 3.75 cm x 2.5 cm. The array comprised of 4 such stimuli (2 on the right and 2 to the left) arranged horizontally with inter-stimulus separation of approximately 1.25 cm. One of these stimuli might/might not be the target stimulus. The subject was asked to respond, by different key presses, if the target appeared on the right or on the left of the array or it did not appear in the array at all. There were two unique features of the task; first, the subject had to wait till the disappearance of both target and array before executing a response (otherwise the response was not accepted by the programme) and secondly, the next target appeared only when the subject had responded. All occurrences were kept equi-probable. The difficulty of the task was varied by manipulating the intrinsic task attributes. Characteristics of the two task variants are given as below in Table 1.

Table 1: Characteristics of the task

Task Attributes	Easy Task	Difficult Task
Pre-target Delay	500 ms	500 ms
Exposure Time of Target	500 ms	200 ms
Pre Array Delay	500 ms	500 ms
Exposure Time Array	1000 ms	400 ms
Inter-stimulus Delay	500 ms	500 ms

Estimation of Mental Workload [23]. The NASA TLX is a two-part evaluation procedure consisting of both weights and ratings. Three separate computer programs are used: 'WEIGHTS' is used to collect weights; 'RATINGS' is used to collect ratings; and 'COMBINE' is used to combine them into an overall weighted workload score. The first requirement is to obtain numerical ratings for each scale that reflect the magnitude of that factor in the given Working Memory Task. The 'RATINGS' program presents the six scales on the monitor. Subjects respond by marking each scale at the desired location, using the mouse of the computer. Each scale is presented as a line divided into 20 equal intervals anchored by bipolar descriptors (e.g. High/Low). Ratings were obtained following the entire task.

The subjects then evaluated the contribution of each factor (its weight) to the workload of the task. These weights account for two potential sources of between-rater variability: differences in workload definition between raters within tasks and differences in the sources of workload between tasks. In addition the weights themselves provide diagnostic information about the nature of the workload imposed by the task. There are 15 possible pair-wise comparisons of the six scales. The 'WEIGHTS' program presents each pair to the subject on the monitor. Subjects select the member from each pair that contributed more to the workload of that task. The computer tallies the number of times that each factor was selected. The tallies can range from 0 (not relevant) to 5 (more important than any other factor). The 'COMBINE' program then computes the overall workload score for each subject by multiplying each rating by the weight given to that factor by that subject. The sum of the weighted ratings for each task is divided by 15 (the sum of the weights).

Statistical Analysis: Normality of the distribution of the data was examined using Shapiro Wilk's Test. A two way analysis of variance (ANOVA) was employed. The two factors were difficulty (with 2 levels viz. easy and difficult) and altitude (with 3 levels viz. ground level, 10,000 ft and 15,000 ft). Level of significance was set at $p < 0.05$. Additionally, reliability analysis was also done and Chronbach's á was calculated to examine the internal consistency amongst the 'ratings' of the constituent dimensions of NASA TLX. For the scales that are used for research tools (like NASA TLX), á value of 0.7 to 0.8 was regarded satisfactory [24].

Results

The reaction time and accuracy in the working memory task exhibited significant effect of task difficulty ($F = 17.417$; $p = 0.001$ & $F = 177.245$; $p = 3.14E-09$ respectively). Significant effect of hypoxia was observed in reaction time ($F = 3.883$; $p = 0.033$) but not in accuracy ($F = 0.341$; $p = 0.714$). An insignificant interaction effect was seen between task difficulty and altitude on accuracy and reaction time (Fig. 1).

The ratings on all dimensions of NASA TLX (except Frustration) exhibited significant effect of task difficulty (Fig. 2). No significant effect of task difficulty was observed on weights. A significant effect of task difficulty was observed in weighted workload in only mental demands ($F = 12.203$; $p = 0.003$), temporal demands ($F = 3.547$; $p = 0.079$), own performance ($F = 6.065$; $p = 0.026$) and effort ($F = 6.295$; $p = 0.024$) (Fig. 3).

On the other hand, no significant effect of hypoxia was observed in ratings, weights and

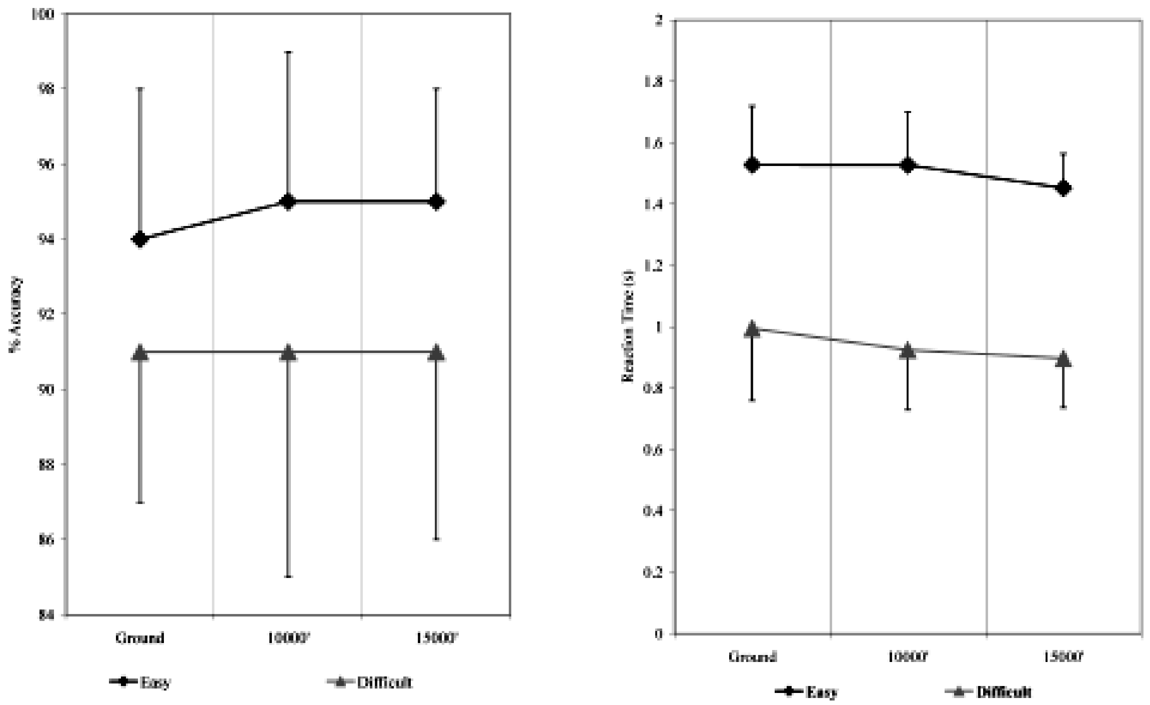

Fig. 1: Percentage Accuracy and Reaction Time in easy and difficult task variants as a function of altitude

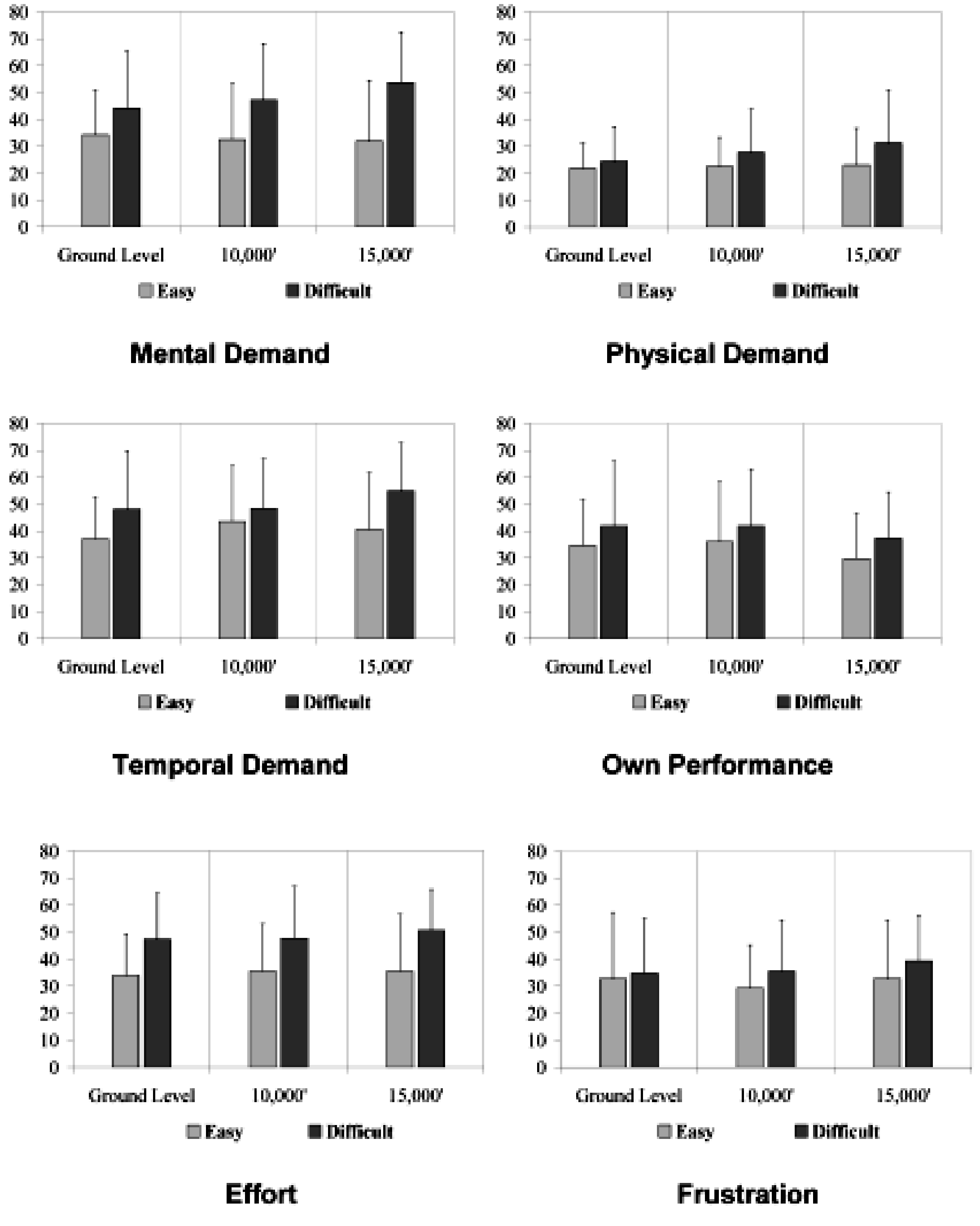

Fig. 2: 'Ratings' in the Six Constituent Dimensions of NASA TLX
(Note: Bars are the mean values and error bars represent + 1 SD)

weighted workload in any of the dimensions of NASA TLX as well as overall weighted workload (Fig. 4).

Discussion

The present study was conducted with an aim to examine the effect of hypobaric hypoxia on the subjective perception of mental workload. To accomplish this, mental workload was assessed using a multi-dimensional subjective index (NASA Task Load Index) during engagement of 16 subjects in a PC based task at ground level as well as at two

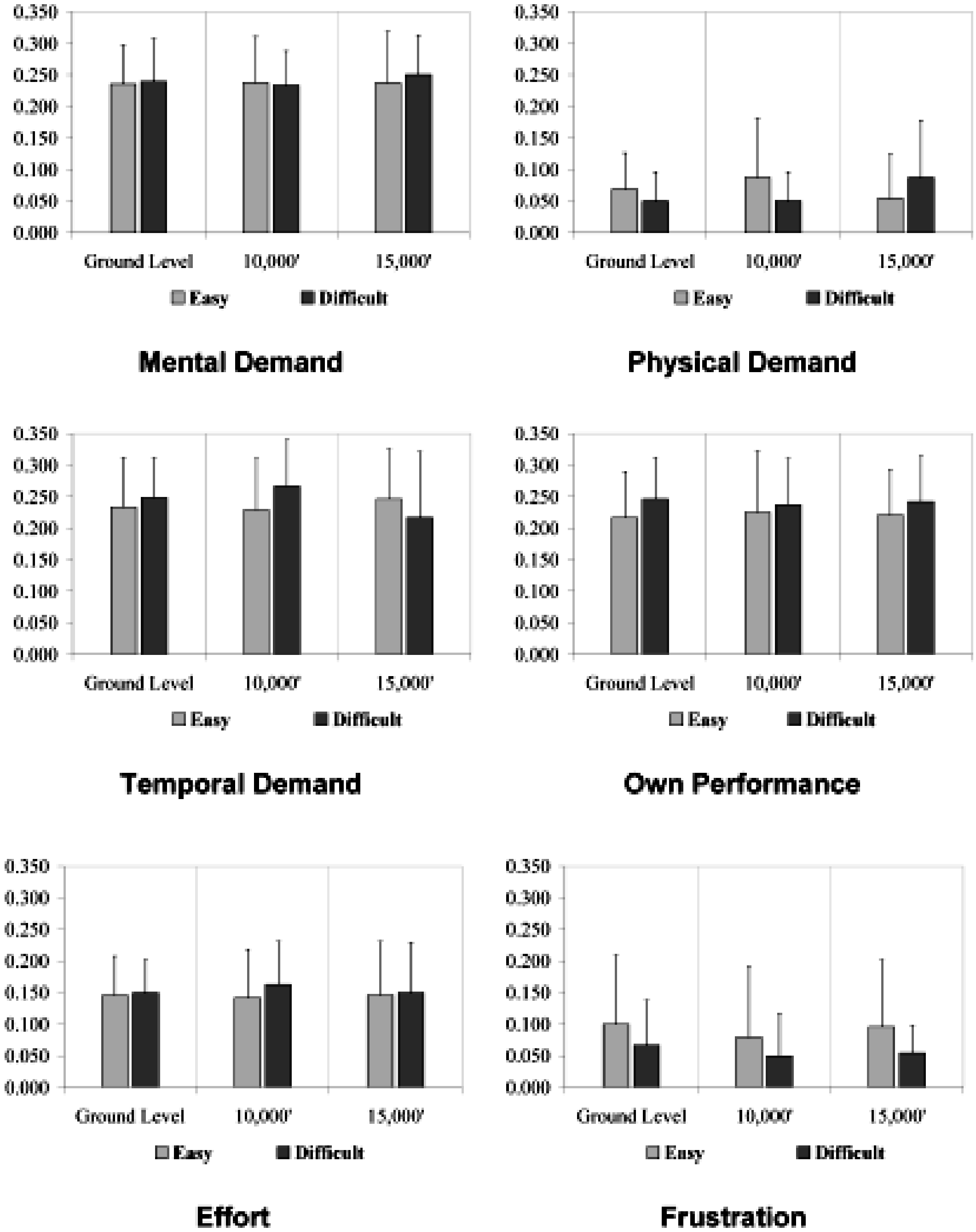

Fig. 3: Weights in the Six Constituent Dimensions of NASA TLX (Note: Bars are the mean values and error bars represent + 1 SD)

altitudes (10,000 ft and 15,000 ft) simulated in a Explosive Decompression Chamber (EDC). The above task is established to have a locus of load on visuo-spatial working memory which is recognizable by NASA TLX [25]. Two task variants with varying difficulty levels were chosen for the study so that the data could be submitted to an 'Additive Factor Analysis' paradigm of Sternberg [26]. This was done with the aim to examine whether the effects of hypoxia and task difficulty were additive or interactive.

In view of the established effects of hypoxia

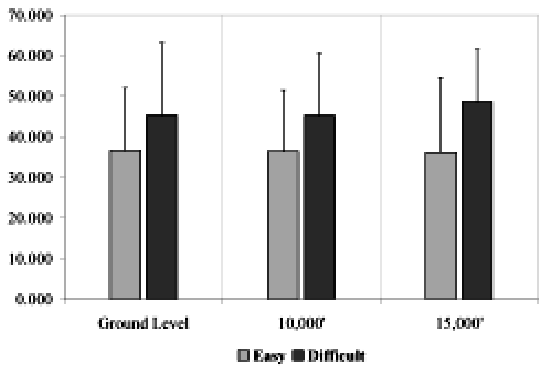

Fig. 4: Overall Weighted Workload using NASA TLX (Note: Bars are the mean values and error bars represent + 1 SD)

on many of these dimensions, either directly or indirectly, it was hypothesized that the workload assessment using NASA TLX would be affected under hypobaric hypoxia.

On analysis of intrinsic task attributes and task performance, it was found that difficulty/memory load in the difficult task variant was increased by decreasing the exposure time of target and array to 40% of the corresponding values in the easy task. Pre-target delay and inter-stimulus delay was not changed. This manipulation amounts to an effective degradation of stimulus quality and absence of preview – the two variables reported to increase the task difficulty [27]. Target and array in these tasks inflict concurrent load on working memory and therefore can be considered analogous to classical 'concurrent working memory load methods'. Recall of target is influenced by the viability of rehearsal in the early retention interval. A much reduced exposure time in the difficult task permitted less time for such rehearsals and resulted in a significant main effect of task difficulty on accuracy. The accuracy, however, did not exhibit a significant main effect of hypoxia. Interaction effect between hypoxia and task difficulty was also insignificant. The last two effects indicate that the effect of task difficulty was not modified by hypoxia.

On the other hand, reaction time exhibited significant effect of both task difficulty as well as hypoxia. However, there was no interaction effect. Thus, reaction time was significantly reduced in difficult task as compared to the easy one. It also reduced significantly with altitude hypoxia. The reduction in reaction time with task difficulty and hypoxia may appear to be strange at the first instance. In a classical experimental paradigm, reaction time should have increased with hypoxia and task difficulty. The above paradox is however, intelligible in the light of certain task characteristics which are inherent to the task used in this study. As explained earlier, target and array in these tasks inflicted concurrent load on working memory analogous to classical 'Concurrent Working Memory Load Paradigm'. Any kind of degradation in the stimulus quality (exposure time in the present study) and/or a deficit in the working memory (under the effect of hypoxia in the present study) amounted to an effective increase in the concurrent load on the working memory affecting the viability of rehearsal in the early retention interval. In such a 'concurrent memory load' type of situation, accuracy is adversely affected with an increase in the time for which the above stimuli (target and the array competing for the identical cognitive resources) are retained [28, 29, 30]. Therefore, in order to maintain accuracy, subjects tended to respond more promptly with an increase in the task difficulty and under the conditions of hypoxia which is known to cause a deficit in the working memory [9].

It is noteworthy that accuracy exhibited a significant reduction with task difficulty. However, it was preserved under hypoxia with a change in the performance strategy which necessitated a significant change in the reaction time. Therefore, preservation of accuracy under the influence of hypoxia in the present study can not be viewed as a lack of effect of hypoxia on the task performance. Results do suggest a deficit occurring in the working memory under the condition of hypobaric hypoxia and it was manifest at 15,000 ft but not at 10,000 ft. This is in consonance with many studies which have attempted to examine the effect of mild hypoxia on working memory. Both storage and retrieval of information in the working memory are reported to be affected under hypobaric hypoxia [9]. However, the experiment design in the present study does not permit to infer as to whether storage or retrieval was affected under hypoxia. In any case, it was not the primary objective of the study.

Nonetheless, an insignificant interaction effect between task difficulty and altitude on reaction time, seen in the light of 'Additive Factors Method' of Sternberg [26] suggests that the influence of hypoxia and task difficulty is additive and affect two different stages in the human information processing.

Scores in NASA TLX: NASA TLX scores did conform to objectively defined task difficulty and demands. Further, the scores provided some diagnostic information regarding the source of load in the task.

However, there was no significant effect of mild to moderate hypoxia on the ratings, weights or weighted load in any of the six dimensions of NASA TLX. It may seem paradoxical at the first instance because there was a definite deficit in working memory under moderate hypoxia as evident from the results of performance in the Working Memory Task (vide supra). This apparent ambiguity can be understood in view of the controversy which exists even on date, regarding the value of such verbal reports like NASA TLX, to be considered as legitimate psychological data. This approach raises the issue of the extent to which humans are consciously aware of the information processing. Nisbett and Wilson opined that there was almost no conscious awareness of perceptual or memory processes [31]. On the contrary, Ericsson and Simon strongly argued that verbal reports constitute legitimate data. They suggested that subjects were able to report on the information heeded in working memory and that the inaccurate reports resulted from requesting information which was not directly heeded [32]. Also, researchers on subjective mental workload indicate that humans have some underlying quantifiable representation of mental workload [33]. Term 'Primary Memory' has been used to describe a form of Short Term Memory (STM) equivalent to the contents of consciousness. Thus, it is expected that workload created by tasks involving STM may be accessible to subjective report even if other aspects of mental workload are not [34]. However, because of limitations of verbal report, subjective measurements of mental workload do not, necessarily, include all the information relevant to mental workload [35].

Even though, it was not the primary objective of the study, its results do throw some light on the validity of NASA TLX. Verbal reports, in the form of NASA TLX, constitute legitimate psychological data as apparent from the values of Cronbach's Alpha suggesting an excellent internal consistency [24]. Nonetheless, it does not *pari passu*, establish the sensitivity of the scale. Deficit in the working memory during exposure to moderate hypobaric hypoxia seems to be too subtle to be detected through NASA TLX. Or else, the subjective perception could have been clouded by certain amount of 'euphoria' which occurs during such hypoxic exposure. The time frame in which the effects of hypoxia on NASA TLX were studied and other evidence available in the literature is more in favour of the former.

Conclusion

The present study examined the effect of short term exposure (15 minutes) of mild and moderate hypobaric hypoxia (10,000 ft and 15,000 ft simulated altitudes, respectively) on the subjective perception of mental workload measured with NASA TLX during engagement of 16 healthy male volunteers in a working memory task with two difficulty levels. Moderate hypoxia had a significant effect on task performance. However, subjective perception of mental workload remained unaffected. It was despite a fairly high internal consistency amongst the constituent dimensions of NASA TLX as discerned through Cronbach's Alpha. Results of the study lend support to the contention that subjective verbal reports do constitute legitimate psychological data. However, workload indices derived from such reports are not sensitive enough to be affected by small deficits in the working memory during short term exposure to mild to moderate hypoxia.

Conflict of interest: None

References

1. Moray N. Final report of the experimental psychology group. In N Moray, editor. Mental workload - Its theory & measurement. New York: Plenum; 1979 p. 101-14.

2. Kantowitz BH, Casper PA. Human workload in aviation. In: Weiner EL, Nagel DC, editors. Human factors in aviation. San Diego: Academic Press; 1988: 157-87.

3. Cooper CE, Harper RP. The use pilot rating in the evaluation of aircraft handling qualities, NASA Tech. Report No. TN-D-5153. Moffett field, CA: NASA Ames Research Centre; 1969.

4. Johannesen G. Workload and workload measurement. In N Moray, editor. Mental workload - Its theory & measurement. New York: Plenum; 1979: 3-12.

5. Hartman BO, Mc Keneie RE, editors. Survey of methods to assess mental workload, AGARD AG 246. France: Neuilly-sur-Seine; August 1979.

6. Pelamatti G, Pascotto M, Semenza C. Verbal free recall in high altitude: proper names vs. common names. Cortex. 2003; 39: 97-103.

7. Barbara SH, Louis EB & Harris RL. Elevation dependant symptom, mood and performance changes produced by exposure to hypobaric hypoxia. The International Journal of Aviation Psychology. 1988; 8(4):319–34.

8. Bartholomew et al. Effect of moderate levels of simulated altitude on sustained cognitive performance. Journal of Aviation Psychology. 1999; 9: 351-9.

9. Fowler M, Pralic H, Brabant M. Acute hypoxia fails to influence two aspects of short term memory: Implications for the source of cognitive deficits Aviat Space Environ Med. 1994; 65: 641-5.

10. Green RG, Morgan DR. The effect of mild hypoxia on a logical reasoning task. Aviat Space and Environ Med. 1985; 56:1004-8.

11. Denison DM, Poulton EC, Ledwith F. Complex reaction times at simulated cabin altitudes of 5,000 ft and 8,000 ft. Aerospace Medicine. Oct 1966; 1010–3.

12. Nesthus T, Rush LL, Wreggit SS. Effect of mild hypoxia on pilot performance at general aviation altitudes. Final report. US department of transport, Federal Aviation Administration. National technical information service, Sprinfield, Virginia. April 1997.

13. Mc Carthy D, Corban R, Legg S, Faris J. Effect of mild hypoxia on perceptual-motor performance: a signal-detection approach. Ergonomics. 1995; 38(10):1979-91.

14. Blogg LS, Gennser M. Cerebral blood flow velocity and psychomotor performance during acute hypoxia. Aviat Space Environ Med. 2006; 77: 107-13.

15. Kelman GR, Crow TJ, Bursill AE. Effect of mild hypoxia on mental performance by a test of selective attention. Aerospace Medicine. 1969; 40(3):301-3.

16. Kelman GR, Crow TJ. Impairment of mental performance at simulated altitude of 8,000 ft. Aerospace Medicine. 1969; 40(9):981-2.

17. Harding RM, Gradwell DP. Hypoxia and hyperventilation. In: Ernsting J, Nicholson AN, Rainford DJ, editors. Aviation Medicine. 3rd Ed.

London: Butterworths; 1999: 43-58.

18. Paul MA, Fraser WD. Performance during mild acute hypoxia. Aviat Space Environ Med. 1994; 65: 891-9.

19. Gold RE, Kulak LL. Effect of hypoxia on aircraft pilot performance. Aerospace Medicine. 1972; 43(2):180-3.

20. Jamin T.Apnea induced changes in time estimation and its relation to bradycardia. Aviat Space Environ Med. 2004; 75(10):876-80.

21. Toman N. Subjective time estimation and autonomic nervous system. PhD Thesis from:Fachbereich Humanmedizin, Freie Universität Berlin. 2004:[1 screen]. Available from: URL: http://www.diss.fuberlin.de/2004/270/indexe.html. Accessed July 15, 2007.

22. Directorate General Civil Aviation. Civil aviation requirements, Section: 2 – Airworthiness series 'I' Part II, 24 April 1992. Aircraft equipment and instruments Rev. 1, 17 September 1999. Available from: URL: http://www.dgca.nic.in/rules/car-ind.htm. Accessed October 01, 2005.

23. Users' Manual. NASA task load index (TLX) v.1.0. Human Performance Research Group Moffit Field, CA: NASA Ames Research Centre. Available from: URL: http://www.cc.gatech.edu/classes/AY2005/cs7470_fall/papers/manual.pdf. Assessed on October 01, 2005.

24. Martin JB, Douglas GA. 28th Statistical notes: Cronbach's Alpha. In: Series of Occasional Notes on Medical Statistics. BMJ. 1997; 314: 572.

25. Tripathi KK. Psycho physiologic evaluation of human mental workload. PhD Thesis from National Institute of Mental Health and Neurosciences (Deemed University), Department of clinical psychology, Bangalore; 2001.

26. Sternberg S. The discovery of processing stages: Extensions of Donders' method. Acta Psychologica, International Journal of Psychonomics. 1969;30:276-315. Available from: URL:http://www.sciencedirect.com. Assessed on August 20, 2007.

27. Wickens CD. Attention. In: PA Hancock, editor. Human factor psychology. Amsterdam: Elsevier Science Publishers; 1987; 47: 28-80.

28. Yeh YY, Wickens CD. The effect of varying task difficulty on subjective workload. In: Proceedings of the 29th Annual Human Factors Society Meeting; 1985; Santa Monica. CA: Human Factors Society; 1985: 765-9.

29. Wickens CD, Kramer AF, Vanasse L, Donchin E. The performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information processing resources. Science. 1983; 221:1080-2.

30. Wickens CD, Sandry DL, Vidulich M. Compatibility and resource competition between modalities of input, output and central processing. Human Factors. 1983;25(2):227-48.

31. Nisbett RE, Wilson TD. Telling more than what we know: Verbal reports on mental processes. Psychological review. 1977;84(3):231-59.

32. Ericsson KA, Simon HA. Verbal Reports on Data. Psychological Review. 1980;87(3):215-51.

33. Revesman ME. Assessment of SWAT accuracy. In: Proceedings of the Human factor Society, 29th Annual Meeting. Santa Monica, CA; 1985. 183-7.

34. Klapp ST. Short term memory limits in human performance. In: PA Hancock, editor. Advances in psychology. Amsterdam: Elsevier science publishers; 1987. (Human factor psychology; Vol 47).

35. Yeh YY, Wickens CD. Dissociation of performance and subjective measures of mental workload. Human Factors. 1988; 30:111-20.