REPORT
ISSN 1174 - 1406
Volume 23 No. 17 1998

Mental Workload of Mechanised Processing with a Single Grip Harvester

Mark Sullman and Patrick Kirk

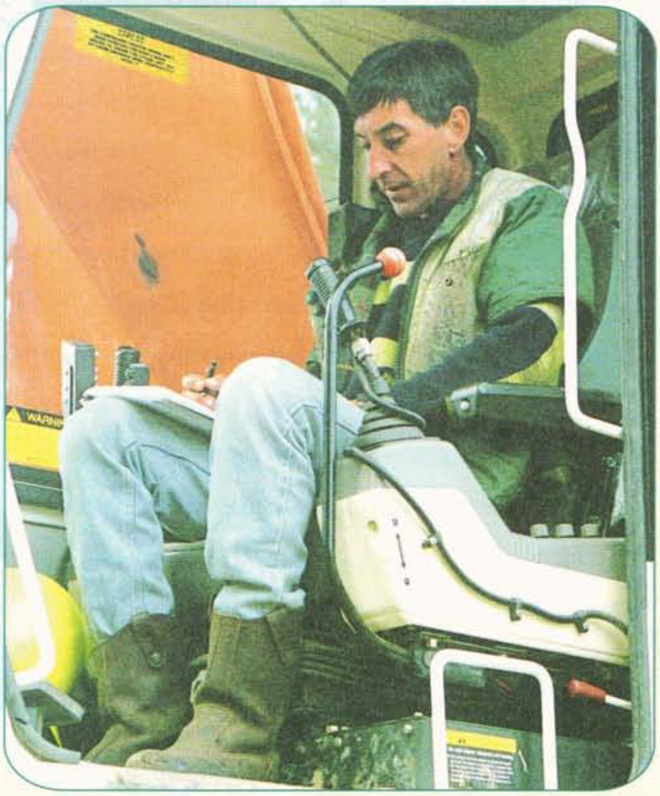


Figure 1 - Operator filling in TLX

Summary

A study was conducted to measure mental workload in single grip harvester operators processing logs, and to trial the acceptability of measuring mental workload in the forest industry using the NASA - TLX method.

- The NASA - TLX proved to be an effective tool for measuring the mental workloads of the operators of forest machinery.
- The operators experienced workloads that were similar to simulated air traffic control work during very busy periods and the simulated flying of an F-16, but were considerably higher than those experienced by airline pilots and simulated flying of a helicopter.
- Based on these findings, mechanised processing with a single grip harvester can be ranked as a high workload task.

Introduction

With the New Zealand forest industry's increasing reliance upon mechanisation comes an increase in the mental workloads associated with forest work (Sullman and Gellerstedt, 1997). This requires knowledge on the level of mental workload required to efficiently complete these tasks. This will enable industry to: assess whether mental workloads are too high; develop and test methods for reducing mental workload (such as equipment redesign, improving the human-computer interface); assess whether additional tasks can be added without any decrease in the worker's performance; assess whether new tools impose too high a workload; and decide whether mental workload is leading to undesirable consequences such as high accident rates, high turnover and absenteeism. Unfortunately, although there has been extensive research undertaken on the physical workloads of New Zealand forest work (Ford, 1995; Hartsough and Parker, 1997; Kirk and Sullman, 1995; Kirk, Sullman and Parker, 1996; Sullman and Byers, 1996), only Sullman and Gellerstedt (1997) have measured mental workload to date.

Liro Forestry Solutions
Private Bag 3020, Rotorua, New Zealand
Telephone: +64 7 348 7168 Facsimile: +64 7 346 2886
Email: pat.kirk@fri.cri.nz

There are basically four different categories of workload measurement. These are: **primary** task measures, **secondary** task measures, **physiological** measures and a **subjective** measures and limitations.

1. Primary Task Measures

Primary task measures are ideal, in that they provide an indication of operator and system performance. However, primary measures do not provide any indication of the spare capacity an operator has (Sirevaag et al. 1993). For example, two forwarder operators can be working at the same level, but one may be using their full mental capacity, while the other may have enough spare mental capacity to perform an additional activity. Another disadvantage is that primary task measures are not sensitive to changes in workload, and are not easily transferred from one task to another (Charlton, 1996).

2. Secondary Task Measures

Secondary tasks can be used to assess residual mental capacity, which is not consumed by the performance of the primary task (Liu and Wickens, 1994). The theoretical assumption behind the use of secondary task measurement is that the processing resources of the operator are limited, and that the operator's performance will deteriorate when several activities compete for the same resources (Wickens, 1992).

Changes in performance on the secondary task are assumed to reflect changes in the mental demands of the primary task. However, in some cases, the secondary task reduces performance on the primary task (Sirevaag et al. 1993). From a safety point of view, this intrusion is unacceptable. The ideal secondary task does not interfere functionally with the performance of the primary task, but does require some attention.

Time estimation has been shown to be a secondary task that possesses this advantage (Michon, 1966). Time estimation does not use the same sensory-motor paths as a large range of operators' tasks and can be learned, implemented and scored easily (Hart, 1975). Also, their performance is relatively stable with respects to learning effects (Hart, 1975). In the case of single grip harvester operators, the safety issue is not an excessive one. A much bigger issue is the interruption of production. Another problem with the use of secondary tasks is that the addition of a secondary task changes the nature of the primary task and therefore contaminates any workload measures obtained (O'Donnell and Eggemeier, 1986). Secondary tasks also lack global sensitivity and transferability to other tasks (Charlton, 1996).

3. Physiological Measures

There are a large number of physiological measures being used to measure mental workload, including; ECG (Electrocardiogram), EOG (Electrooculogram), EDA (Electrodermal activity), and EEG (Electroencephalogram). Measures of heart rate and heart rate variability have often been used as measures of mental workload (Sirevaag, et al. 1993). It is thought that mental work is reflected in a reduction in the person's heart rate variability (HRV). Unfortunately, there has been some inconsistency in the success of HRV in measuring mental workload (Meshkati, 1988). Also, dealing with the data, once it has been collected, is fairly complex due to the large quantities of data. In order to reduce the data into a more manageable size, it must first be analysed by Fast Fourier Transformation (FFT), before a calculation of the actual variability can be made. This requires special software which is very expensive. A second alternative is to look at the pattern graphically to see if there is a visual difference.

4. Subjective Measures

Subjective measures still remain the most popular method of measuring mental workload, and are the method against which other methods are evaluated. In fact some researchers believe that subjective ratings of mental workload come closest to tapping the essence of mental workload and generally provide the most valid, sensitive and practically useful indicators (Hart and Staveland, 1988). Subjective ratings also have high face validity, as they measure the individual's experience of mental workload. They are also low cost, easy to implement, low intrusiveness, high correspondence with operator performance measures and high operator acceptance (Sirevaag et al. 1993). However, in some cases subjective measures have not been found to correlate well with performance measures. Subjective measures are also: more sensitive to global than specific processing demands; more sensitive to intermediate, than high levels of workload; and are most reliable when collected soon after task performance.

NASA Task Load Index (TLX)

One of the most frequently used and most validated subjective measure is the NASA Task Load Index (TLX) (Charlton, 1996). The TLX is a multidimensional rating procedure that provides an overall workload score based on a weighted average of ratings on six sub-scales: mental demand, physical demand, temporal demand, performance, effort and frustration (Figure 1 for definitions). The first three sub-scales reflect the task demands placed upon the operator. The remaining three sub-scales characterise the interaction between the operator and the task (Desmond and Hoyes, 1996).

This study used one of the most validated methods of subjectively measuring mental workload - the NASA Task Load Index (TLX) (Charlton, 1996; Desmond and Hoyes, 1996; Hancock, 1989; Hancock, 1996; Liu and Wickens, 1994; Sirevaag et al., 1993) to measure the mental workload associated with processing stems using a single grip harvester.

Method

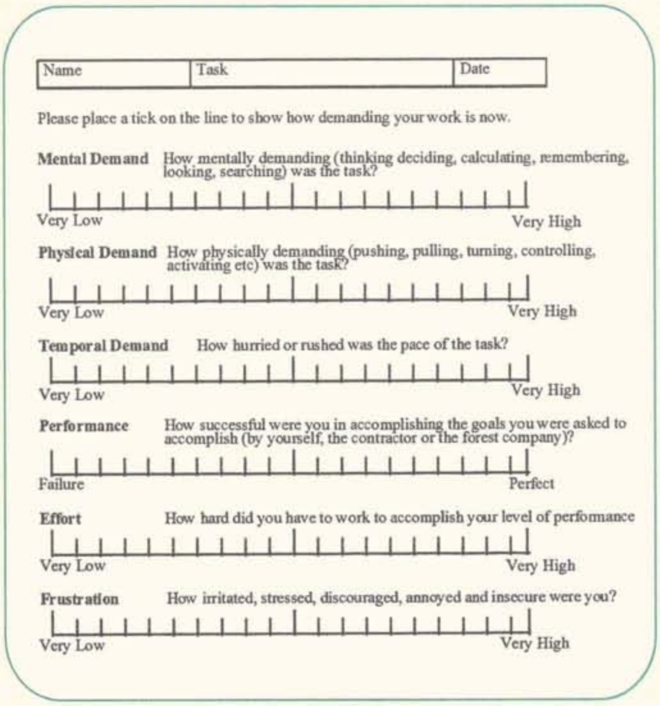
Subjects

The subjects were three harvester operators who were processing stems into logs using a single grip harvester. The average processing experience was 1.8 years (range 6 months to 3 years). Each operator was monitored for four days, with data being averaged over all four days to calculate the overall averages for each operator.

Procedure

The NASA-TLX was used according to the protocols established by previous researchers (Hart and Staveland, 1988). There are basically two steps to the protocol.

Step 1: Participants make pairwise comparisons between the six sources of mental workload (mental demand, physical demand, temporal demand, frustration, effort and performance). The subjects were presented with every possible pair of scale titles in a forced choice paradigm (making 15 comparisons). The subjects were required to identify which of the pair contributes the most to their personal definition of mental workload. Each pair was presented twice, in reverse order, to control for the position effect. The number of times each source was chosen was tallied, and the average of the two comparisons of the same pairs was used as the source of workload weight. This pairwise comparison generated a weighting for each source of mental workload, where the highest possible rating is 5 and the lowest 0. The total number of weights add up to 15. Step 1 was completed prior to beginning Step 2.



Name	Task	Date

Please place a tick on the line to show how demanding your work is now.

Mental Demand How mentally demanding (thinking deciding, calculating, remembering, looking, searching) was the task?

Very Low — Very High

Physical Demand How physically demanding (pushing, pulling, turning, controlling, activating etc) was the task?

Very Low — Very High

Temporal Demand How hurried or rushed was the pace of the task?

Very Low — Very High

Performance How successful were you in accomplishing the goals you were asked to accomplish (by yourself, the contractor or the forest company)?

Failure — Perfect

Effort How hard did you have to work to accomplish your level of performance

Very Low — Very High

Frustration How irritated, stressed, discouraged, annoyed and insecure were you?

Very Low — Very High

Figure 2 - NASA TLX rating scale

Step 2: Each hour the operators were asked to rate their perceived workload on the six individual sources of mental workload on a scale which ranged from 0 (very low mental workload) to 100 (very high mental workload) (Figure 2). These scores represent the raw ratings. The overall workload rating was found by multiplying the raw ratings by the source weights (derived in Step 1) and dividing by 15.

Results

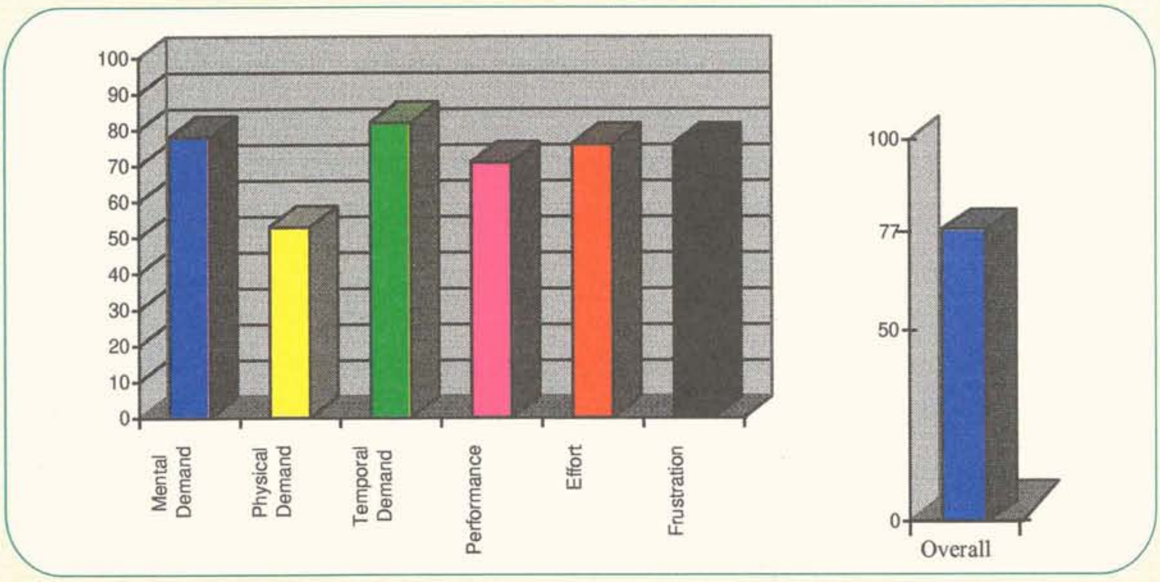


Figure 3 - Components of mental workload for Operator 1

Figure 3 shows that Operator 1 rated temporal demand as the highest followed by mental demand (100 = very high mental workload and 0 = very low mental workload). Figure 3 also shows that Operator 1 perceived physical workload to be considerably lower than the other components. In fact this operator did not believe that physical demand contributed to the overall mental workload at all (as shown in the weighting process). Using the six components and the weightings derived from the paired comparisons, the overall workload was 77.

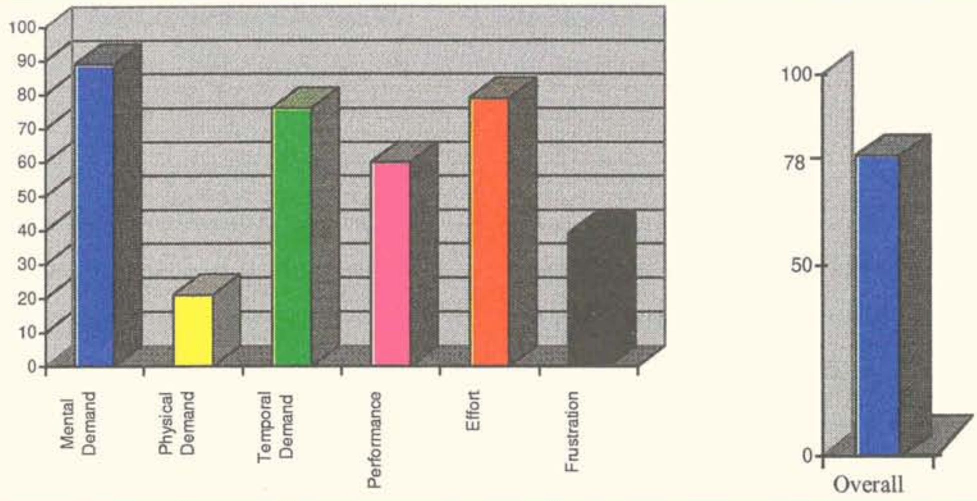


Figure 4 - Components of mental workload for Operator 2

Operator 2 rated the mental demand to be the highest, followed by the amount of effort required and then the temporal demand (Figure 4). The physical demand was again given the lowest rating. As with Operator 1, this operator did not believe that physical demands contributed to the overall mental workload (as seen in the weighting process). Using the six components and the weightings derived from the paired comparisons, the overall workload was 78.

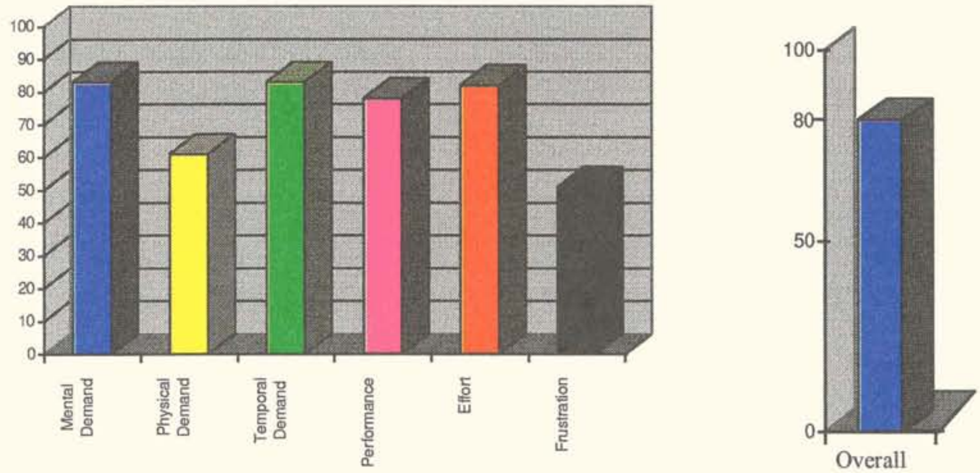


Figure 5 - Components of mental workload for Operator 3

Operator 3 rated the mental demand the highest, followed by temporal demand and effort (Figure 5). Unlike the other two operators, Operator 3 rated frustration to be the lowest, with physical demand being second lowest. This operator also thought that the physical demand did contribute to mental workload (as shown by the weightings generated from the paired comparisons). Using the six components and the weightings derived from the paired comparisons, the overall workload was 80.

Discussion

The overall rating across all three operators was 78. This figure is almost the same as: simulated air traffic control in high workload situations, 77 (Desmond and Hoyes, 1996); simulated take offs and landings in an F-16, 75 (Hancock, 1989). The mental workload rating found in this study was higher than: operating a helicopter simulator 64 (Sirevaag et al., 1993); taking off at Los Angeles airport in a Boeing 767-300, 58 (Samel, Wegmann and Vejvoda, 1996); simulated supermarket work 55 (Lui and Wickens, 1994); air traffic control under easy conditions 52 (Desmond and

Hoyes, 1996); landing a Boeing 767-300 at Hamburg airport, 47 (Samel et al., 1996); simulated driving of a private car 35 (Lansdown, 1997); and flying a Boeing 767-300, 33-17 (Samel et al., 1996).

Subjective measures of mental workload (such as used here) measure the individual's experience of mental workload. Therefore, the three operators in this study experienced workloads that were similar to simulated air traffic control work during very busy periods and simulated flying of an F-16, but were considerably higher than those experienced by airline pilots, simulated driving a car, simulated flying a helicopter and simulated supermarket work.

Based on these findings, mechanised processing with a single grip harvester can be ranked as a high workload task. This finding is backed up by other researchers using different measures of mental workload (Gellerstedt, 1993; Sullman and Gellerstedt, 1997).

Conclusions

- This study shows that the NASA - TLX can be used by the forest industry to effectively determine the mental workload of machine operators.
- The three operators perceived very high mental workloads.
- The overall mental workload ratings were remarkably similar across the three operators, ranging from 77 to 80.
- Across the three operators, mental demands of the task were rated as the largest source of mental workload, followed by the temporal demands. The physical component was rated as he lowest source of mental workload.

References

Charlton, S. G. (1996). Mental workload test and evaluation. In T.G. O'Brien and S.G. Charlton (Eds.), Handbook of Human Factors Testing and Evaluation. Lawrence Erlbaum: Mahwah, New Jersey.

Desmond, P.A. and Hoyes, T.W. (1996). Workload variation, intrinsic risk and utility in a simulated air traffic control task: Evidence for compensatory effects. Safety Science 22(1-3), 87-101.

Ford, D. (1995). An Ergonomic Analysis of Manual Versus Chainsaw High Ladder Pruning of Pinus radiata in New Zealand. A thesis submitted in partial fulfilment of the requirements of the degree of Master of Philosophy at Massey University.

Gellerstedt, S. (1993). Thinning with a forestry machine: The mental and physical work. Garpenberg, Sweden: Swedish University of Agricultural Science.

Hart, S.G. (1975). Time estimation as a secondary task to measure workload. In Procedings of the 11th Annual Conference on Manual Control. NASA.

Hart, S.G. and Staveland, L.E. (1988). Development of the NASA task load index (TLX): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.), Human mental workload (pp. 139-183). Amsterdam: North-Holland.

Hancock, P.A. (1989). The effect of performance failure and task demand on the perception of mental workload. Applied Ergonomics 20 (3), 197-205.

Hancock, P.A. (1996). Effect of control order, augmented feedback, input device and practice on tracking performance and perceived workload. Ergonomics 39(9), 1146-1162.

Hartsough, B. and Parker, R. (1997) Manual pruning of Douglas-Fir. New Zealand Journal of Forestry Science, 26, 449-459.

Kirk, P., Sullman, M. and Parker, R. (1996). The impact of fatigue on the safety, comfort and productivity of clearfell tree fallers. LIRO Report 21, 18.

Lansdown, T.C. (1997). Visual allocation and the availability of driver information. In T. Rothengatter and E. Carbonell Vaya (Eds.). Traffic and Transport Psychology. Amsterdam: Pergamon.

Liu and Wickens, C.D. (1994). Mental workload and cognitive task automaticity: an evaluation of subjective and time estimation metrics. Ergonomics 37(11), 1843-1854.

Meshkati, N. (1988). Heart rate variability and mental workload assessment. In P.A. Hancock & N. Meshkati (Eds.). Human Mental Workload. Amsterdam: North-Holland.

Michon, J.A. (1966). Tapping regularity as a measure of perceptual motor load. Ergonomics 9, 401-412.

O'Donnell, R.D. and Eggmeier, F.T. (1986). Workload assessment methodology. In K.R. Boff, L. Kaufman, & J.Thomas (Eds.), Handbook of perception and human performance, Vol 2: Cognitive processes and performance. New York: Wiley.

Samel, A., Wegmann, H.M. and Vejvoda, M. (1996). Aircrew fatigue in long-haul operations. In L. Hartley (Ed.). Proceedings of the Second International Conference on Fatigue and Transportation: Engineering, Enforcement, and Education Solutions. Fremantle, Australia.

Sirevaag, E.J., Kramer, A.F., Wilkens, C.D., Reisweber, M., Strayer, D.L. and Grenell, J.F. (1993). Assessment of pilot performance and mental workload in rotary wing aircraft. Ergonomics 36(9), 1121-1140.

Sullman, M. and Byers, J. (1996). Physical workload of planting under three different site conditions. LIRO Report 21, 14.

Sullman, M. and Gellerstedt, S. (1997). The mental workloads of mechanised harvesting. New Zealand Forestry, 42 (3), 48.

Wickens, C.D. (1984). Engineering psychology and human performance. New York: Harper Collins.