

3.3.4.4 Hart and Bortolussi Rating Scale

General description. Hart and Bortolussi (1984) used a single rating scale to estimate workload. The scale units were 1 to 100, with 1 being low workload and 100 being high workload.

Strengths and limitations. The workload ratings significantly varied across flight segments, with takeoff and landing having higher workload than climb or cruise. The workload ratings were significantly correlated to ratings of stress (+0.75) and effort (+0.68). These results were based on data from 12 instrument-rated pilots reviewing a list of 163 events.

Moray, Dessouky, Kijowski, and Adapathya (1991) used the same rating scale but numbered the scale from 1 to 10 rather than from 1 to 100. This measure was significantly related to time pressure but not to knowledge or their interaction.

Data requirements. The subjects need only the end points of the scale.

Thresholds. Low workload equals 1; high workload equals 100.

SOURCES

Hart, S.A. and Bortolussi, M.R. Pilot errors as a source of workload. *Human Factors*. 25(5), 545–556, 1984.

Moray, N., Dessouky, M.I., Kijowski, B.A., and Adapathya, R.S. Strategic behavior, workload, and performance in task scheduling. *Human Factors*. 33(6), 607–629, 1991.

3.3.4.5 Instantaneous Self Assessment (ISA)

General description. The Instantaneous Self Assessment (ISA) is a five-point rating scale (see table 27) that was originally developed in the United Kingdom to evaluate workload of air traffic controllers. ISA has since been applied to evaluating workload of Joint Strike Fighter pilots. On-line access to workload ratings was added to ISA and the resultant system renamed Eurocontrol Recording and Graphical display On-line (ERGO) (Hering and Coatleven, 1994).

TABLE 27

Instantaneous Self Assessment

ISA Button Number	Color	Legend	Definition
5	Red	VERY HIGH	Workload level is too demanding and unsustainable, even for a short period of time.
4	Yellow	HIGH	Workload level is uncomfortably high, although it can be sustained for a short period of time.
3	White	FAIR	Workload level is sustainable and comfortable.
2	Green	LOW	Workload level is low, with occasional periods of inactivity. Operator has considerable spare capacity and is relaxed.
1	Blue	VERY LOW	Workload level is too low. Operator is resting or not contributing to crew tasks.

Source From Hering, H. and Coatleven, G. ERGO (version 1) for Instantaneous Self Assessment of Workload (EEC Note No. 24/94. Brussels, Belgium: EUROCONTROL Agency, April 1994. (With kind permission of The European Organization for the Safety of Air Navigation (EUROCONTROL)[®] 1994 EUROCONTROL. All rights reserved.)

Strengths and limitations. Hering and Coatleven (1996) stated that the ISA has been used in ATC simulations since 1993. Castle and Leggatt (2002) performed a laboratory experiment to compare the workload estimates from three rating scales: ISA, NASA TLX, and the Bedford Workload Scale. They asked 16 pilots and 16 nonpilots to rate their workload using each of these three workload scales while performing the Multiple Attribute Task Battery. As a control, subjects also performed the task battery without rating their workload. Finally, subjects were asked to complete a face validity questionnaire. Average ratings for the 11 scales on the questionnaire were between 4 and 6 on a scale of 1 to 7 (7 being the highest positive rating). This was comparable to the other two workload measures.

There were, however, significant differences between the two groups. The nonpilots rated the ISA to be significantly more professional in appearance, and the pilots rated the ISA to be significantly more reliable. ISA was not sensitive to differences between pilots and nonpilots in the performance of a task battery designed to simulate flying a fixed-wing aircraft. The correlation between ISA and the Bedford Workload Scale was +0.49 and the NASA TLX was +0.55. The correlation with ratings of observers with ISA ratings was +0.80. The correlation with task loading on the Multiple Attribute Task Battery was highest for the ISA (+0.82), and lower for the NASA TLX (+0.57) and the Bedford Workload Scale (+0.53). There were no significant correlations between ISA rating and performance. Nor were there significant effects on performance whether or not the ISA rating was given. That was also true for the NASA TLX and Bedford Workload Scale scales. Internal consistency as measured by Cronbach's alpha varied between 0.43 and 0.78 for subjects and 0.64 to 0.81 for observers. Retest reliability for the same task performed 2 weeks later was +0.84. Instant ratings were reported to be more consistent than ratings made 2 min after the task.

Tattersall and Foord (1996) in a laboratory study using a tracking task reported that tracking task performance decreased when ISA responses were made, and therefore warned of its intrusiveness on primary task performance.

Lamoureux (1999) compared 81 categories of aircraft relationships in ATC and predicted versus ISA subjective workload ratings. The predictions were 73% accurate.

Data requirements. Use of the standard rating scale.

Thresholds. These vary from 1 to 5.

SOURCES

- Castle, H. and Leggatt, H. *Instantaneous Self Assessment (ISA) – Validity & Reliability (JS 14865 Issue 1)*. Bristol, United Kingdom: BAE Systems, November 2002.
- Hering, H. and Coatleven, G. *ERGO (version 1) for Instantaneous Self Assessment of Workload (EEC Note No. 24/94)*. Brussels, Belgium: EROCONTROL Agency, April 1994.
- Hering, H. and Coatleven, G. *ERGO (version 2) for Instantaneous Self Assessment of Workload in a real-time ATC simulation environment (EEC Report No. 10/96)*. Bruxelles, Belgium: EROCONTROL Agency, April 1996.
- Lamoureux, T. The influence of aircraft proximity data on the subjective mental workload of controllers in the air traffic control task. *Ergonomics*. 42(11), 1482–1491, 1999.
- Tattersall, A.J. and Foord, P.S. An experimental evaluation of instantaneous self-assessment as a measure of workload. *Ergonomics*. 39(5), 740–748, 1996.