

Young, M. S. & Stanton, N. A. (2002). It's all relative: defining mental workload in the light of Annett's paper. *Ergonomics*, 45(14), 1018-1020.

It's all relative: Defining mental workload in the light of Annett (2001)

Mark S. Young and Neville A. Stanton

Department of Design, Brunel University

Runnymede Campus, Englefield Green

Egham, Surrey TW20 0JZ

England

Tel: +44 1784 431341

Fax: +44 1784 439515

email: mark.young@brunel.ac.uk

Young, M. S. & Stanton, N. A. (2002). It's all relative: defining mental workload in the light of Annett's paper. *Ergonomics*, 45(14), 1018-1020.

It's all relative: Defining mental workload in the light of Annett (2001)

Abstract

Annett (2001) has made an in-depth critique of subjectivity in science, concentrating not only on those rating scales which are extensively used in ergonomics, but also the very philosophy of subjectivism itself. The present authors take this critique and use it to reappraise their own interpretation of mental workload, a key concept popularly assessed by subjective measurement. Having recently proposed a definition of mental workload (Young & Stanton, 2001), the present commentary considers deeper problems of subjectivity in the interpretation, perception, and measurement of mental workload. It is realised that there are fundamental problems inherent with such a subjective concept, particularly concerned with the dependence on context and lack of an absolute metric. However, it is also noted that the very point of ergonomics is to address contextual influences on human performance. Therefore, in the absence of a complete paradigm shift in ergonomics, the discipline will have to either accept or find ways of coping with the problems of subjectivity.

Young, M. S. & Stanton, N. A. (2002). It's all relative: defining mental workload in the light of Annett's paper. *Ergonomics*, 45(14), 1018-1020.

It's all relative: Defining mental workload in the light of Annett (2001)

On the basis of a comprehensive literature review, Young & Stanton (2001) arrived at the following definition of mental workload (MWL):

“The mental workload of a task represents the level of attentional resources required to meet both objective and subjective performance criteria, which may be mediated by task demands, external support, and past experience.” (p. 507)

The definition was offered for a construct which had previously evaded having its meaning pinned down. However, it was just an offering, and by no means argued as the ultimate formalisation of MWL. Indeed, it was Young & Stanton's (2001) own understanding of the available literature which led them to suggest the definition. Now, in the light of Annett's (2001) article, it seems there are three cans of subjective worms associated with it.

The first is the problem of interpretation. The authors had their own backgrounds and knowledge bases serving as (albeit possibly implicit) frames of reference and agendas for the review. Being researchers in cognitive ergonomics, it is unsurprising that their definition was couched in terms of attention and performance. This does not necessarily mean such an approach was right or wrong – any expert in the arena would undoubtedly have done a similar thing. A stress researcher would have reviewed the stress literature and defined MWL in terms of stress reactions. Whether or not Young & Stanton's

Young, M. S. & Stanton, N. A. (2002). It's all relative: defining mental workload in the light of Annett's paper. *Ergonomics*, 45(14), 1018-1020.

(2001) definition becomes accepted depends on the extent to which other researchers in the area agree with it, which in turn depends on the degree of overlap between their interests. This cuts to the heart of the intersubjectivity issue to which Annett (2001) refers, only at a more macro level: "Progress in science depends on the degree of shared meaning between individuals concerning their observations and experiences." (this issue). Annett argues that intersubjectivity is crucial for laboratory measurements to be valid. However, it is just as vital amongst those scientists making the measurements to agree on the concepts they are measuring. Science (with a capital "S") is, after all, a belief system, itself governed by the vagaries of subjective interpretation. Just like any other belief system, there has to be a certain leap of faith to accept some of the concepts and theories.

Secondly, Young & Stanton (2001) introduce subjectivity into the definition by allowing MWL to be influenced by the internal goals of the individual. This is where the definition gets quite knotty, since these goals will vary between individuals and across situations. With such a context-specific construct, one might argue whether a definition is of any practical use at all. To illustrate this, take the perception of time. Our definitions of time – seconds, minutes, hours, and even months – are largely arbitrary, and would be of little use if we were spending our lives on another planet. However, for all intents and purposes, the 24-hour day is sufficient, since there are few people concerned with measuring time relative to different planetary orbits or rotations. So, specific to the context of Earth, our definition of time is useful. But outside that context it begins to lose its practicality and even its 'absoluteness' – time can even be affected by

Young, M. S. & Stanton, N. A. (2002). It's all relative: defining mental workload in the light of Annett's paper. *Ergonomics*, 45(14), 1018-1020.

altitude, speed, or relative direction (cesium clocks have been shown to lose time when travelling eastwards, and gain time westwards). The same goes for MWL, but the context scale is much smaller – rather than planet to planet, it changes from person to person. If the meaning of MWL for a given person in a particular situation is different from that for another person in a different situation (or even the same person in a different situation etc.), defining it in these terms seemingly has as much use as an atomic clock on Jupiter. The Science of Ergonomics, though, is all about context. As Tom Stewart observed at this year's Ergonomics Society Lecture, every good ergonomist will tell you that human performance depends on its context. Predictions, therefore, can only be made with reference to that context, so by implication, all of our theories must be context-specific.

Finally, there is the complex issue of MWL measurement. Many authors claim that the use of subjective ratings may well be the only index of 'true' MWL (e.g., Hart & Staveland, 1988). Subjective MWL scores have been shown to be sensitive to perceived difficulty (Liu & Wickens, 1994) and demand for multiple resources (Hockey, Briner, Tattersall & Wiethoff, 1989). One particular advantage with subjective measures is that they are sensitive to changes in effort, when such effort maintains primary task performance at stable levels (Hockey et al., 1989).

One of the hurdles in deriving a subjective measure of MWL is validating it by correlation with other measures (such as physiological evidence) or against objective demand. Annett (2001) suggests the method of cross-modal matching to quantify

Young, M. S. & Stanton, N. A. (2002). It's all relative: defining mental workload in the light of Annett's paper. *Ergonomics*, 45(14), 1018-1020.

psychometric scales, and perhaps descriptors of weight (e.g., this task is 'heavy', 'light' etc.) would provide an adequate analogy for MWL. But this raises the question of whether there is such a thing as 'objective demand', or whether – especially in a field trying to improve design – it is simply what people perceive that is important. If the latter is the case, it will mean there is no need for the use of an anchor task to calibrate subjective scales.

Criticisms of subjective measurement techniques are primarily concerned with the metacognitive abilities of the operator (Petrucci & Cloutier, 1992; Praetorius & Duncan, 1988). That is, given the fact that the measures are necessarily administered post-task, one might question the reliability of self-reports, particularly for long task durations. A recent attempt to develop an instantaneous (i.e., 'on-line') measure of subjective MWL met with limited success (Tattersall & Foord, 1996). Unfortunately, whilst it correlated with other measures of workload, it was disruptive to the primary task. These effects are further complicated by a Heisenberg-esque uncertainty principle – that the act of measurement itself is also demanding, so may influence MWL ratings.

In exploring a philosophical view of MWL, then, it would seem that the many problems of definition and measurement are down to the subjective nature of the beast. Whether it is the measurement, the experience, or even the very Science of MWL, there really is no absolute truth, only perceptions of truth. Hardly anyone would argue against the existence of such a thing as MWL, though, since we all experience it frequently. But the fact that it means different things to different people at different times causes problems

Young, M. S. & Stanton, N. A. (2002). It's all relative: defining mental workload in the light of Annett's paper. Ergonomics, 45(14), 1018-1020.

for applied research. How are we supposed to design a system for "optimal MWL" when such a state will probably only exist for an infinitesimally brief period of time? As every good ergonomist will tell you, "It's all relative", as context is definitely the current zeitgeist in Ergonomics. Unless a better paradigm emerges, we will have to find ways of working within it. Perhaps the key will be to draw up a list of "mindspace" tables to design for cognitive ergonomics, just like the Bodyspace data for physical ergonomics (Pheasant, 1986). With such data, the logical positivist approach to MWL research (which we in particular have taken in our laboratory) could be justified as more science than art.

References

Annett, J. (2001). Subjective rating scales: Science or art? Ergonomics, this issue.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), Human mental workload (pp. 138-183). Amsterdam: North-Holland.

Hockey, G. R. J., Briner, R. B., Tattersall, A. J., & Wiethoff, M. (1989). Assessing the impact of computer workload on operator stress: the role of system controllability. Ergonomics, 32(11), 1401-1418.

- Young, M. S. & Stanton, N. A. (2002). It's all relative: defining mental workload in the light of Annett's paper. Ergonomics, 45(14), 1018-1020.
- Liu, Y., & Wickens, C. D. (1994). Mental workload and cognitive task automaticity: an evaluation of subjective and time estimation metrics. Ergonomics, 37(11), 1843-1854.
- Petrusic, W. M., & Cloutier, P. (1992). Metacognition in psychophysical judgment: an unfolding view of comparative judgments of mental workload. Perception and Psychophysics, 51, 485-499.
- Pheasant, S. T. (1986). Bodyspace. London: Taylor & Francis.
- Praetorius, N., & Duncan, K. D. (1988). Verbal reports: a problem in research design. In L. P. Goodstein, H. B. Anderson, & S. E. Olsen (Eds.), Tasks, errors and mental models: a festschrift to celebrate the 60th birthday of Professor Jens Rasmussen (pp. 293-314). London: Taylor & Francis.
- Tattersall, A. J., & Foord, P. S. (1996). An experimental evaluation of instantaneous self-assessment as a measure of workload. Ergonomics, 39(5), 740-748.
- Young, M. S. & Stanton, N. A. (2001). Mental workload: theory, measurement, and application. In W. Karwowski (Ed.), International encyclopedia of ergonomics and human factors: Volume 1 (pp. 507-509). London: Taylor & Francis.