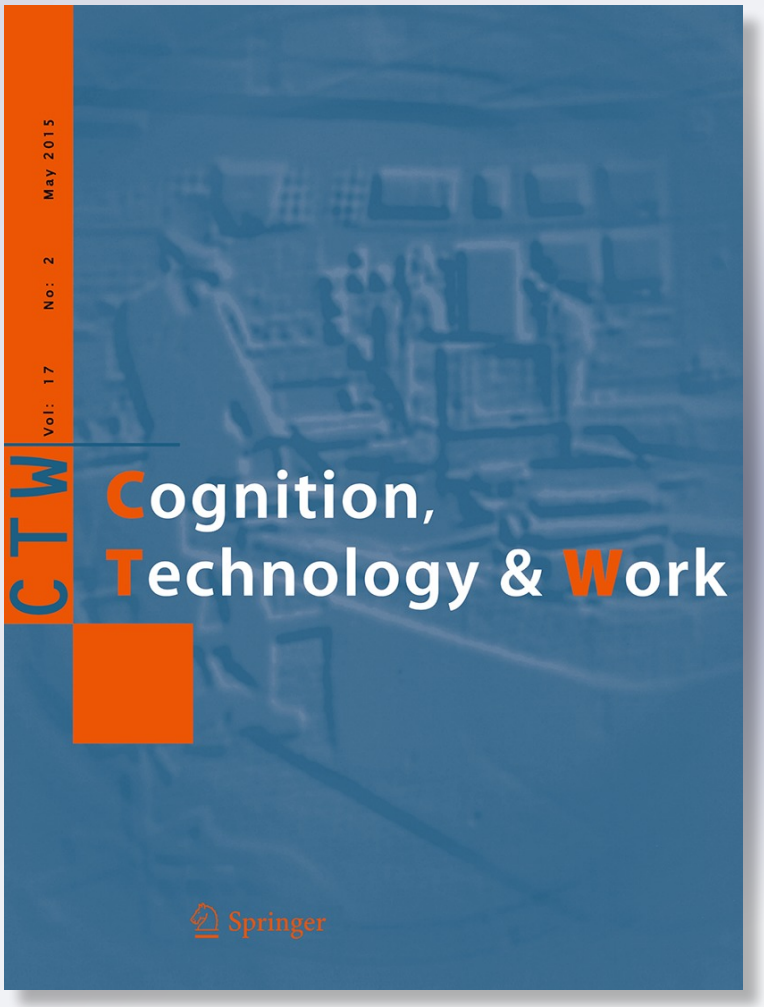
From figments to figures: ontological alchemy in human factors research

Sidney W. A. Dekker & James M. Nyce

Cognition, Technology & Work

ISSN 1435-5558
Volume 17
Number 2

Cogn Tech Work (2015) 17:185-187
DOI 10.1007/s10111-015-0321-7



Springer

Cogn Tech Work (2015) 17:185–187
DOI 10.1007/s10111-015-0321-7

ORIGINAL ARTICLE

From figments to figures: ontological alchemy in human factors research

Sidney W. A. Dekker · James M. Nyce

Received: 9 September 2013 / Accepted: 8 January 2015 / Published online: 24 January 2015

Abstract To bridge the gap between psychology and engineering, human factors research routinely engages in a kind of ontological alchemy which turns putative mental constructs into measurable numbers. We should not overestimate the ontological status of our constructs, nor underestimate the extent to which our own research that makes them "real."

Keywords Workload · NASA TLX · Ontology · Epistemology · Human factors research · Alchemy

Living for so long on the threshold between psychology and engineering has meant that human factors can have a hard time remaining faithful, or credible, to both. But it tries. We have long realized that not all statements by operators or practitioners about what they experience (or what goes on "in their minds") can be accommodated pragmatically or analytically within the confines of a typical human factors study. Or, for that matter, we might not even want to accept that those statements themselves (like those gathered in a talk-aloud or verbal process tracing study) are authentic to or even congruent with the experience. Lived experience, after all, can escape almost all theoretical and rhetorical attempts to fence it in, ours included. Even to assume that experience can be trapped in any explicit way reflects nostalgia for a time when both science and its objects of study were seen as more certain, more objectively attainable. But human factors is a science of pragmatics and seems unwilling to acknowledge (let alone be hindered by) such finer points of phenomenological concern. Those of us involved with the design of systems, for instance, may need to deliver findings in a form that is useable for engineers. This means of course figures, or numbers. In the history of the discipline, and its projects, the measurement of workload is a great example of the epistemological reduction deployed by human factors. This reduction is often achieved in oblique ways: Those involved in the research are either unaware or regard them as part of the natural course of doing the business of research.

One example is how NASA TLX turns workload into a series of operationalisms (mental demand, physical demand, temporal demand, performance, effort, and frustration) and provides respondents with ordinal scales to mark the experienced extent of each. It is interesting to see how this echoes traditions and ideas in psychology that were foresworn by human factors long ago. Wilhelm Wundt, in his Leipzig laboratory, once set out to develop a chronometry of mind, but had to admit that it was too bold a research goal. The measurement of workload by operationalisms such as those of the NASA TLX (e.g., temporal demand on a scale) reflects a pale version of his ambition. Moreover, it uses a version of a Wundtian experimental research technique—introspection (or rather, retrospection)—which asks subjects to "look inside" as it were and reflect on their experience. Introspection became so mistrusted as a psychological research method that it

S. W. A. Dekker (✉)
Safety Science Innovation Lab, School of Communication, Humanities and Social Sciences, Griffith University, N16 Macrossan Building, Nathan Campus, Brisbane, QLD 4111, Australia
e-mail: s.dekker@griffith.edu.au

S. W. A. Dekker
School of Psychology, The University of Queensland, Brisbane, QLD, Australia

J. M. Nyce
Department of Anthropology, Ball State University, Muncie, IN, USA

gave rise to the mental Ice Age of behaviorism in the early part of the twentieth century. Its oblique reintroduction in workload measurement, however, is possible because it fits the empiricist and quantitativist credentials that human factors believes it needs to burnish.

This is what we mean by ontological alchemy: As if by alchemy, mental events or figments of an event—empirically accessed through individual acts of introspection or retrospection—become transmuted into both numbers and placeholders in a numbered set. The result is an invention, a creation, a fabrication, indeed a construct-like workload. Alchemy, a forerunner of chemistry, was concerned with the transformation of matter: turning the worthless into the useful and the valuable. Ontological alchemy is the transformation of a judgment about a mental experience into a number: turning it into something useful (and valuable) too. Once workload has become a figure, our analytic machinery stops. It has done its job. We do not need to concern ourselves further with whatever high order claims, like those about epistemology or ontology, our informants, our subjects, our participants, might want to put before us. Whatever "reality testing" is necessary has already occurred. It has been taken care of by the measurement, the tool, the index.

The critical issue here is that it enables us to apply to informant statements almost any kind of cryptanalysis or exegesis the Western mind is capable of. In the name of achieving credibility, plausibility, or rationality, this can go quite far (Dekker and Nyce 2004). A study into the workload of air traffic controllers with flight strips used ordinal scales like those of the NASA TLX. Once the numbers were in, researchers made a leap of faith, conflating analysis and interpretation. Ordinal scales were transmuted into ratio ones (or ordinal scales were mistaken for ratio scales) so that a statistical analysis became possible (Albright et al. 1996). This is ontological alchemy at its best:

> The form listed all factors with a 9.6 cm horizontal line next to each. The line was marked low on the left end and high on the right end. In addition, a vertical mark in the center of the line signified the halfway mark. The controllers were instructed to place an X on the line adjacent to the factor to indicate a response… scales were scored by measuring distance from the right anchor to the mark placed by the controller on a horizontal line (in centimeters)… Individual repeated measures ANOVAs [were then conducted].

Based on the cryptanalysis, these various alchemical steps facilitated (construct to ordinal scale, ordinal scale to ratio numbers, ratio numbers to statistics) the study concluded that flight strips are superfluous. Of course, such a conclusion, and the results on which it is based, is only as stable and legitimate as the cryptanalysis itself. However, such an approach to human activity has built a set of assumptions that are often unacknowledged but are powerful in ruling in and out what is "believable" and "plausible" about these forms of actions. Humans, however, are self-constituting and reflecting actors, not stable objects in the natural world. Giddens referred to this as the double hermeneutic: The study of human activity must necessarily be based on people's self-interpretation (Giddens 1984). First there are self-interpretations among those people who are studied in human factors research. This is the interpretive work necessary to convert their lived experiences into a tick mark on a scale. The second hermeneutic applies to the human factors researchers themselves, who not only determine workload to be about temporal, mental and physical demands, about frustration levels and effort, but then take those results and help give them meaning in a context that sets their value. It is in these operations Carrier would argue that many of the problems of our analytic enterprise occur because we are largely unaware that we tend to use "folk categories" derived in various ways from our own culture as though they are legitimate, "scientific" categories (Carrier 1992).

The end result is almost always the same: Statements of others are assessed by and reduced to categories and meanings derived from the literature and our own understandings of what is possible and probable and plausible and necessary in the worlds that might need (and fund) our findings. It is this kind of strategic retreat from the other (the operator, the practitioner) that underlies and characterizes the discipline's analytic categories and models. The kind of repudiation of participant testimony can be seen as an analytic "fall from grace." Only some kind of decipherment—no matter how partial: a set of figures for example—is considered the natural and logical endpoint of the human factors research project. Referral to a single, simple gold standard of what is real or relevant (a workload index) seems legitimate because, after all, everybody in the field does it. More than anything, statements derived from these figures become statements about us; about our own understandings of how the world we work in lives. We might not even be aware how much this influences the accounting we do of the thoughts and practices of others, nor how much we miss of the experiences of others as a result.

De Winter (2014) offers what Flyvbjerg would call a pre-paradigmatic escape. What phrenology once did for psychology, neuroscience will do for human factors: It will come to the field's ontological rescue. The argument goes like this (Dekker et al. 2010): It has not yet been demonstrated, after all, that it is impossible for human factors to achieve the rigor and objectivity of the "hard" sciences. It

may take more time, and more evidence, particularly more biological or physical evidence that correlates with mental states before a science of this kind can be achieved. But scientific progress can be helped along if we do not give up on measuring "real" things; "objectively" real things, visible things. The end result is that that which cannot be seen must be turned into something that can. Because then it is real (Wilkin 2009). But, of course, even the measurable objects of hard science have been relativized in what has been called the "universality of hermeneutics" (Feyerabend 1993). Just like the measurable bumps on crania that, according to phrenology, once "correlated" with criminality, the cranial blood flows of neuroimaging also become objective facts because they have been historically conditioned and constructed as such (Wallerstein 1996). De Winter's escape is not just spurious. It is a rearguard action intended to insure that science (of a particular sort) is seen both as legitimate and possible. But whatever the science, it must never overestimate the ontological status of its own constructs and be aware that the relationship between these constructs and facts often says more about us than about the world we study.

References

Albright CA, Truitt TR et al (1996) How controllers compensate for the lack of flight progress strips. National Technical Information Service, Arlington

Carrier JG (1992) Occidentalism: the world turned upside-down. Am Ethnol 19(2):195–212

De Winter JCF (2014) Controversy in human factors constructs and the explosive use of the NASA TLX: a measurement perspective. Cogn Technol Work 16(3):289–297

Dekker SWA, Nyce JM (2004) How can ergonomics influence design? Moving from research findings to future systems. Ergonomics 47(15):1624–1639

Dekker SWA, Nyce JM et al (2010) Epistemological self-confidence in human factors research. J Cogn Eng Decis Mak 4(1):27–38

Feyerabend P (1993) Against method. Verso, London

Giddens A (1984) The constitution of society: outline of the theory of structuration. Polity Press, Cambridge

Wallerstein I (1996) Open the social sciences: report of the Gulbenkian Commission on the restructuring of the social sciences. Stanford University Press, Stanford

Wilkin P (2009) The ideology of ergonomics. Theor Issues Ergon Sci 11(3):230–244