

# ÜBER EINIGE AKTUELLE PROBLEME DER ERFASSUNG, MESSUNG UND BEURTEILUNG DER PSYCHISCHEN BELASTUNG UND BEANSPRUCHUNG

· *psychische Belastung* · *psychische Beanspruchung* · *Messverfahren* · *Gütekriterien*

## ZUSAMMENFASSUNG

Die zunehmende Bedeutung der psychischen Belastung und Beanspruchung als Komponente der Gesamtbelastung erfordert Messverfahren, die eine messtheoretisch befriedigende Erfassung dieser Konstrukte erlauben, und zwar unter praktischen wie wissenschaftlichen Gesichtspunkten. Die Diskussion einiger der mit der Erfassung dieser Konstrukte verbundenen Probleme zeigt, dass dabei sowohl konzeptuelle - was soll gemessen werden - wie methodische Probleme - wie kann das gemessen werden - sowie Probleme der Bestimmung der Messgüte - wie lässt sich eine angemessene Messgüte bestimmen und sicherstellen - auftreten. Unter Bezug auf generalisierungstheoretische Ansätze wird gezeigt, wie Verfahren zur Erfassung dieser Konzepte überprüft und beurteilt werden können.

## PRAKTISCHE RELEVANZ

Die Diskussion der bestehenden Probleme begrifflicher Unschärfe und ihre Bedeutung für die Erfassung und Beurteilung psychischer Belastung und Beanspruchung soll zu einer kritischen Beurteilung der Gebrauchstauglichkeit verschiedener Messverfahren anregen und ihrem unreflektierten Einsatz vorbeugen.

## PROBLÈMES D'ESTIMATION, DE MESURE ET D'ÉVALUATION DE LA CHARGE MENTALE (ASTREINTE)

· *Contrainte* · *charge mental (astreinte)* · *mesure* · *critères psychométriques*

## RESUME

La valeur de la contrainte et de l'astreinte psychique comme composante de la charge globale, exige d'être mesurer avec une technique de mesure satisfaisante, qui permet d'appréhender cette approche du point de vue de références scientifiques pratiques. La discussion à propos des problèmes engendrés par ce concept montre, qu'aussi bien les problèmes conceptuels (ce qui doit être mesuré) que les problèmes de méthode (comment peut on effectuer les mesures), et que les problèmes de l'exactitude des mesures (comment définit-on l'exactitude d'une mesure et quelle est sa fiabilité) sont actuels. Nous référant à la théorie de la Généralisation, nous montrons comment les mesures effectuées pour appréhender ce concept peuvent être évaluées et validées.

## IMPORTANCE PRATIQUE

Cette discussion sur les problèmes de mesure de la charge mentale a pour but d'encourager les utilisateurs potentiels d'instrument de mesure de la charge mentale à évaluer de façon critique la pertinence et l'utilisabilité de ces instruments et de prévenir toute utilisation inappropriée.



# ON SOME PROBLEMS OF ASSESSING, MEASURING AND EVALUATING MENTAL WORKLOAD

· *mental stress* · *mental strain* · *measurement* · *psychometric criteria*

## SUMMARY

Based on the increasing importance of mental workload, both as a component of total workload and legal or quasi legal requirements in the EU, instruments for assessing, measuring or evaluating mental workload are required which allow for an objective, reliable, valid, sensitive and diagnostic assessment of the intended object of measurement, i.e. mental stress, mental strain or the effects of mental strain in the operator, e.g. mental fatigue, monotony or satiation (as defined in ISO 10075). It is argued that it is most important to make clear distinctions between these intended objects of measurement, because blurring the distinction will lead to unreliable and invalid results. This is demonstrated by reference to studies using generalizability theory (in this issue), where it could be shown that some of the proposed indicators of mental workload, e.g. heart rate variability, are not able to differentiate different kinds (diagnosticity) and levels (sensitivity) of cognitive workload, but instead allow for a differentiation among operators according to their (habitual) level of arousal. It should thus be made clear whether the intention is to measure and differentiate between people (e.g. according to their strain) or working conditions (with regard to the workload they impose upon the operator). These intentions require different approaches of evaluating the psychometric properties of measurement instruments, as described in ISO/CD 10075-3, and it is argued, that a generalizability-theoretical approach can be used to solve these problems and for estimating the psychometric properties of an instrument.

It is further argued that no general measure of mental stress or mental strain can be developed, since both are not unidimensional concepts, varying on a dimension from low to high, but that only for components, differentiating specific aspects or components of mental stress or strain adequate measurement instruments can be developed, as has been done for some of the effects of mental strain in the operator. It does not make sense, to allow for a compensation of cognitive complexity by time restrictions, since these seem to put demands on different resources, which of course may interact. And it does not make sense to compensate or simply not to discriminate between fatigue and monotony.

As a consequence measurement of mental workload would first of all require some more elaborated theoretical analyses of mental workload which should allow for the development of adequate measurement instruments, if accompanied by more sophisticated measurement approaches.

## PRACTICAL RELEVANCE

This discussion of problems relating to the measurement of mental workload is intended to encourage potential users of instruments for measuring mental workload to critically evaluate the suitability and usability of these instruments and to prevent an unreflected use of such instruments.

## 1 EINLEITUNG

Man kann den Inhalt dieses Beitrages entgegen der klassischen Gliederung nach Einleitung, Hauptteil und Schluss vielleicht besser nach einer Menge klassischer W-Fragen (warum, weshalb, wieso) gliedern, und das soll daher hier auch so erfolgen. Fragen wir uns also *warum* und *wozu* man eigentlich psychische Belastung und Beanspruchung erfassen oder sogar messen muss, soll oder kann, *was* man da eigentlich erfassen will, *wer* das tun sollte, *womit*, *wie*, etc., und *welche* Probleme sich dabei ergeben und *wohin* sich die Dinge unserer Ansicht nach entwickeln sollten.

## 2 WARUM UND WOZU?

Die Frage nach dem warum und wozu lässt sich, wie ein Blick durch die Literatur schnell zeigt, modisch ein- aber auch substantiell vielfältig beantworten, wobei wir letzteres vorziehen. Dabei bemühen wir zur Begründung, dass es sich dabei nicht um eine besonders neue Fragestellung handelt, gerne Max Rubner, der bereits 1911 in seiner Denkschrift zur Errichtung eines Kaiser Wilhelm Instituts für Arbeitsphysiologie darauf hinwies, dass die Analyse von Problemen der psychischen Belastung, auch wenn er das damals etwas anders nannte, eine dringend zu lösende Aufgabe darstellt:

„Unter Arbeitsleistung versteht man gemeinhin nur die Leistung des Menschen als Arbeitsmaschine. Unsere Kulturentwicklung lässt aber diese Form menschlicher Tätigkeiten mehr und mehr zurücktreten, sie hat dafür andere Arbeitsformen geschaffen, bei deren Ausführung es mehr auf den Intellekt als auf die mechanische Kraft ankommt, jene tausendfältigen Arbeitsformen, bei denen durch die Überwachung von Maschinen, bei der Ausführung von Kleinarbeit infolge von Arbeitsteilung die Aufmerksamkeit, Geschicklichkeit, Gewissenhaftigkeit und Dauer der Leistung das Entscheidende ist. Bei der ersten Arbeitsweise liegt also der Schwerpunkt auf mechanischem Gebiet, bei der zweiten mehr auf dem nervösen Gebiet und den Leistungen des Gehirns, d.h. die Arbeit beruht auf dem richtigen Funktionieren der Sinnes-tätigkeit, der Wahrnehmungsfähigkeit und Willensschulung. Aufgabe des Instituts für Arbeitsphysiologie soll es daher sein, in großen Zügen die menschlichen Leistungen körperlicher und geistiger Natur unter den verschiedenen Lebensaufgaben, aber auch hinsichtlich der Eigenart der Kinderzeit, des

jugendlichen Alters, des Greisenalters, bei Mann und Frau, bei verschiedenen Konstitutionen und Rassen, nach Maß und Zahl zu erforschen. ... Die Ergebnisse der Forschungen sollen praktischen wie wissenschaftlichen Zielen dienen.“ (Rubner 1911, zitiert nach Rutenfranz et al. 1981)

Wichtig erscheint insbesondere auch der letzte Satz dieses Zitates, der darauf verweist, dass die Ergebnisse dieser Forschungen „praktischen wie wissenschaftlichen Zielen“ (und zwar in dieser Reihenfolge!) dienen sollen. Damit wird bereits 1911 die praktische Relevanz der Lösung von Fragen der psychischen Belastung hervorgehoben. Die Neuentdeckung und Betonung der Bedeutung dieser Fragen in unserer Zeit mutet daher manchmal ausgesprochen überraschend an; noch überraschender allerdings die gelegentlich anzutreffende Argumentation, dass wir das heute alles schon im Griff hätten. Dass dies nicht so ist, lässt sich leicht belegen, u.a. auch durch die Beiträge in diesem Themenheft.

Trotz der schon damals betonten Bedeutung der psychischen Belastung und der tatsächlich erfolgten Einrichtung des Kaiser-Wilhelm- und späteren Max-Planck-Institutes für Arbeitsphysiologie mit einer eigenen Abteilung für Arbeitspsychologie war der Umfang der Bearbeitung - und insbesondere der Lösung - von Fragestellungen aus dem Bereich der psychischen Belastung, Beanspruchung und ihrer Folgen im Rahmen der Arbeitswissenschaft und Arbeitspsychologie - aus den verschiedensten Gründen - eher bescheiden, so dass Rutenfranz (1969) in seiner Denkschrift zur Errichtung eines Nachfolge-Institutes für das von der Max-Planck-Gesellschaft inzwischen in ein Institut für Systemphysiologie umgewidmete Institut zusammenfassend feststellen musste: „... dass viele Probleme arbeitsphysiologischer Forschung, die Anlass zur Gründung des Max-Planck-Institutes für Arbeitsphysiologie gaben, nicht gelöst werden konnten. Die Dringlichkeit solcher Lösungen wächst jedoch im Zuge der industriellen Entwicklung ständig, wobei sich Schwerpunktverlagerungen arbeitsphysiologischer Forschungen abzeichnen“ (Rutenfranz 1969), wobei er insbesondere auch auf Fragen der psychischen Belastung hinwies, u.a. auch auf die bereits von Rubner hervorgehobene Zunahme von Überwachungstätigkeiten und den dadurch bewirkten Funktions- und Belastungswandel des arbeitenden Menschen.

Heute, rund 30 Jahre später, erscheint uns diese Feststellung immer noch aktuell - und

zwar aktueller denn je zuvor, weil die Probleme durch die Dynamik der Entwicklung und des Einsatzes neuer Technologien noch einmal verschärft wurden und wohl auch weiter werden, wie der Vergleich von Leitwarten der chemischen Industrie und der darin anfallenden Tätigkeiten aus den siebziger Jahren (Singer et al. 1970) mit modernen Prozessleitsystemen (Meyer et al. 2001; Schomann et al. 2001) zeigt.

Mit diesen Veränderungen hat die Forschung in diesem Bereich offensichtlich nicht Schritt gehalten, teils weil andere Schwerpunkte gesetzt wurden, teils weil für die Bearbeitung dieser Fragen auf einer eher grundlagenorientierten Ebene kaum Forschungsunterstützung zu gewinnen war und ist, teils sicher auch wegen der Komplexität und Komplikationen der Fragestellungen in diesem Bereich.

Die Entwicklung zu Überwachungs-, Kontroll- und Steuerungstätigkeiten (Schmidtke 1966, 1993) und der damit verbundene Funktionswandel des Menschen im Arbeitssystem (Singer et al. 1970; Nachreiner 1988) im Bereich der Industrie, ebenso wie die offensichtliche Entwicklung der Arbeitstätigkeiten im Büro- und Dienstleistungsbereich zu in größerem Umfang psychisch, und zwar kognitiv wie emotional belastenden Funktionen, machen damit klar, warum man sich nun langsam verstärkt und intensiver als bisher mit den seit Jahren ungelösten Fragen der psychischen Belastung und Beanspruchung beschäftigen sollte. Dies erscheint um so wichtiger, wenn man - etwa im Sinne einer Vorlaufforschung - gestaltend in diese Entwicklung eingreifen will, um eine „Anpassung der Arbeit und der Maschine an den Menschen an Stelle des bisher geübten Hineinpressens des Menschen in die von Technik und Wirtschaft einseitig festgelegten Arbeitsbedingungen“ (Lehmann 1938, zitiert nach Rutenfranz et al. 1981) zu erreichen. Dabei erscheint es erforderlich, dass diese Forschung im Bereich der psychischen Belastung - im Kontrast zu aktionistischen Ansätzen zur ‚Bekämpfung von Fehlbelastungen‘ oder schnell hingehauenen Fragebogen nach dem Motto ‚Wie haben Sie’s denn nun empfunden oder wo drückt Sie psychisch der Schuh?‘ - wissenschaftlich seriös durchgeführt werden kann und wird, praxis- und grundlagenorientiert, um Rubner’s Reihenfolge noch einmal aufzugreifen, auch wenn das für die betroffenen Forscher intensiver und extensiver psychisch belastend sein dürfte.

Neben dieser aus der Sache selbst sich ergebenden Verpflichtung zur Beschäftigung mit



psychischer Belastung und Beanspruchung besteht bereits seit einigen Jahren auch eine formale Verpflichtung, sich mit diesen Fragen zu beschäftigen. So schreiben - auf der europäischen Rechtsentwicklung basierende - Gesetze (z.B. Arbeitsschutzgesetz) oder Verordnungen (z.B. Bildschirmarbeitsverordnung) den Arbeitgebern zwingend vor, die psychische Belastung am Arbeitsplatz zu beurteilen - ohne jedoch gleichzeitig zu spezifizieren *was da wie* beurteilt werden soll - um ggf. durch geeignete Arbeitsgestaltungsmaßnahmen vermeidbare Gefährdungen zu vermeiden; obwohl dies eigentlich unnötig ist, weil zumindest in der BRD in dieser Hinsicht alles in Ordnung sein soll, wenn man beispielsweise den Stellungnahmen zum Entwurf der DIN EN ISO 10075-2 Glauben schenkt. Dass die Praxis zumindest gelegentlich völlig anders aus-

sieht, wird einem rasch klar, wenn man sich mit halbwegs offenen Augen umsieht und die Realität mit den Vorgaben dieser Norm vergleicht (vgl. Nachreiner et al. 1998). Die Vehemenz der Einsprüche (mehr als 60, allerdings überwiegend von Arbeitgeber- oder Industrieseite und gleichlautend) macht deutlich, welche (sozial-)politische Brisanz damit diesem Thema ‚Psychische Belastung und Beanspruchung‘ zukommt.

Dabei dürfte dafür nicht zuletzt verantwortlich sein, dass das Thema selbst - auch von Fachleuten - schwer zu handhaben ist und die wissenschaftliche und insbesondere natürlich die vorwissenschaftliche Diskussion ausgesprochen unscharf oder auch konträr ist. Dies ist jedoch bei dem erreichten Stand der Forschung und der bisherigen (un-)systematischen Beschäftigung mit diesem The-

ma nicht verwunderlich. Der Gegenstand selbst ist also konzeptuell und daher zwangsläufig auch methodisch und messmethodisch schwer fassbar.

Nicht zuletzt aus diesem Grunde wurde daher im internationalen Normenausschuss Ergonomie die Forderung nach internationalen Normen formuliert, die sich mit den Begrifflichkeiten, den Konzepten und insbesondere der Messung und Erfassung beschäftigen sollten. Dies hat zur Entwicklung der Serie ISO 10075 geführt, deren Teil 1 sich mit der Terminologie, Teil 2 mit Gestaltungshinweisen und der zur Zeit in der Entwicklung befindliche Teil 3 mit Anforderungen an Messverfahren beschäftigt (vgl. Nachreiner 1999).

### 3 WAS SOLL ERFASST WERDEN?

Damit sollte terminologisch eigentlich klar sein, was mit psychischer Belastung, psychischer Beanspruchung und deren Folgen, z.B. psychische Ermüdung, Monotonie oder Sättigung, eigentlich gemeint ist; obwohl Stress - als Beanspruchungs*folge* verstanden - in der damaligen Diskussion noch nicht als normungsfähig durchsetzbar war und daher nicht in der Norm enthalten ist. In der englischsprachigen Version der Norm bezeichnet ‚stress‘ die Belastung, wohingegen die Beanspruchungs*folge* ‚Stress‘ im englischen als ‚stress response‘ zu übersetzen wäre.

Wenn in den einschlägigen legalen Vorgaben daher von der psychischen Belastung die Rede ist - zumindest wird das in der BRD so interpretiert -, die der Arbeitgeber zu beurteilen hat, dann müsste eigentlich klar sein - und das ist sicher ein Vorteil des auch in der Norm umgesetzten Belastungs-Beanspruchungs-Modells (Nachreiner 2001) -, dass die (Gesamtheit der) auf den Arbeitenden psychisch einwirkenden Bedingungen zu erfassen und in ihrer potentiellen Wirkung zu beurteilen sind, nicht jedoch deren psychische Repräsentation oder deren Auswirkungen im Arbeitenden gemeint sind, die stark von individuellen Verarbeitungsmechanismen geprägt sind.

Schaut man sich jedoch an, was derzeit in Wissenschaft und Praxis als psychische Belastung erfasst und diskutiert wird, erhält man einen kaleidoskopartigen Eindruck. So werden als die zur Zeit besonders dominanten Belastungsfaktoren die in Bild 1 dargestellten genannt, u.a. also hoher Verantwortung, Zeitdruck, Überforderung durch die

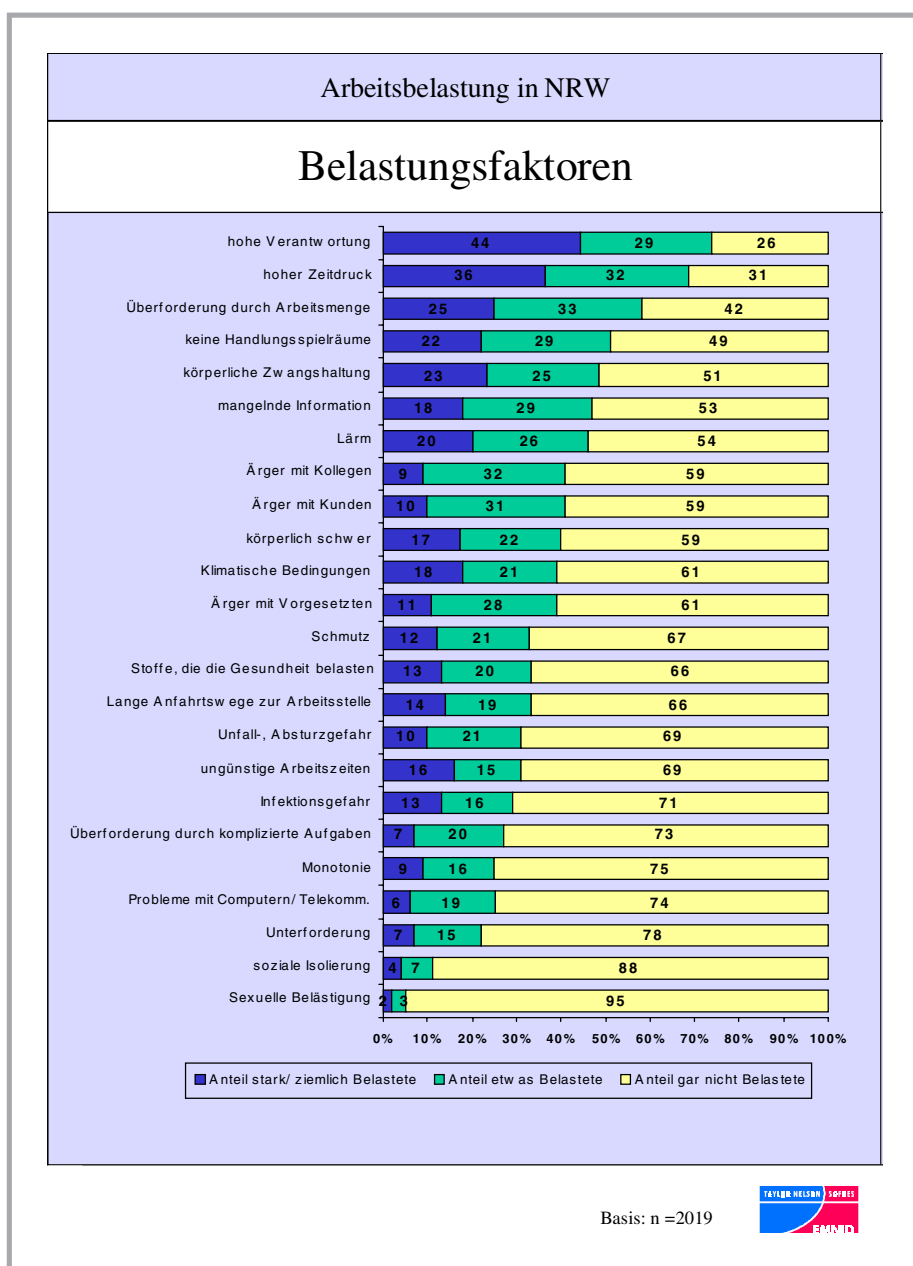


Bild 1: Belastungsfaktoren in NRW (nach: MASQT 2001)



Arbeitsmenge, kleine Handlungsspielräume - und damit dominant also psychisch wirksame Faktoren auf den ersten Plätzen. In einem Konzeptentwurf für eine Schwerpunktaktion ‚Psychische Fehlbelastungen und Stress‘ des BMA (2001) werden Belastungskonstellationen mit

- hohem Zeit- und Termindruck
- hohem Anforderungsdruck
- Zunahme der Arbeitsintensität
- Informationsmangel bzw. Überflutung
- Kommunikations- und Kooperationsbarrieren
- mangelnde Handlungs- und Entscheidungsspielräume
- fehlende soziale Unterstützung
- mangelnde Qualifikation und Erfahrungen

als in den Vordergrund tretend und für die Beschäftigten als Gesundheitsrisiken immer bedeutsamer werdend genannt. Bei genauerer Betrachtung fallen dabei allerdings einige Merkwürdigkeiten auf. So findet man *Zeit- und Termindruck*, *Anforderungsdruck*, sowie mangelnde Qualifikation und Erfahrung, wobei man sich fragt, ob es sich hier nicht eher um Reaktionen bzw. individuelle Merkmale der Arbeitenden handelt. Druck dürfte wohl eher einem Erleben entsprechen, und Erfahrung ist nach unserer Erfahrung an das Individuum oder die Gruppe gebunden, also ein Merkmal des oder der Arbeitenden. Auch die Emnid Befragung (Bild 1) wirft (außerhalb der Person existierende) Bedingungen ( $\approx$  Belastung) und innerhalb der Person ablaufende Reaktionen ( $\approx$  Beanspruchung, z.B. Desaktivierung, und Beanspruchungsfolgen, z.B. Monotonie) munter durcheinander und bezeichnet diese dann undifferenziert als Belastungen bzw. Belastungsfaktoren. Belastung hat unter dieser Perspektive, im Gegensatz zur arbeitswissenschaftlichen Konzeption, eine eindeutig negative Konnotation, Belastung ist damit immer auch etwas schlechtes, belästigendes. Vielleicht macht ja auch dieser, von der ergonomischen Terminologie abweichende Gebrauch des Begriffs der psychischen Belastung die Diskussion so chaotisch und aufgeladen (vgl. Nachreiner 2001; Oesterreich 2001).

Auffallend ist, dass in den genannten Aufzählungen z.B. die Belastung durch die Art

und Komplexität der zu verarbeitenden Information, deren zeitliche Dynamik, Darstellung und Codierung - also klassische Faktoren der psychischen Belastung - nicht auftreten. Das mag damit zusammenhängen, dass für deren Erfassung ergonomische Basiskompetenzen erforderlich sind, die in den Umfragen nicht unbedingt erwartet werden und selbst bei Fachkräften für Arbeitssicherheit nicht ohne weiteres vorausgesetzt werden können (vgl. Nachreiner et al. 1998).

Darüber hinaus fällt auf, dass die in der Emnid-Befragung genannten ‚Belastungen‘, wie auch die vom BMA genannten eigentlich relativ leicht behebbar wären und kaum der psychologischen Fachkompetenz bedürfen. Wenn die zeitliche Bindung (als äußeres, zu gestaltendes Merkmal zur Vermeidung von erlebtem Zeitdruck) zu eng ist, sind weitere Bindungen zu wählen oder die zeitlichen Bindungen aufzuheben, bei zu hohem ‚Verantwortungsdruck‘ oder zu hoher Verantwortung lässt sich die Belastung durch Reduzierung der Verantwortung erledigen, bei Überforderung durch die Arbeitsmenge wäre die Arbeitsmenge zu reduzieren, etc.. Damit scheint uns hier kein für die Arbeitspsychologie besonders interessantes Problem vorzuliegen. Das lässt sich auch mit allgemeinem (Management-)Wissen oder sogar common sense erledigen: man fragt die Arbeitenden, was sie (subjektiv) belastet und stellt das dann einfach ab.

Andererseits erscheint eine normkonforme Erfassung der psychischen Belastung als ausgesprochen problematisch. „Die Gesamtheit aller erfassbaren Einflüsse, die von außen auf den Menschen zukommen und psychisch auf ihn einwirken“ - denn das ist die psychische Belastung nach DIN EN ISO 10075-1:2000 - erfassen zu wollen, erscheint wenig erfolgversprechend (Schmidtke 1993); oder besser gesagt unmöglich, denn für diese *Gesamtheit* gibt es und kann es kein vernünftiges Bezugssystem geben, auf dem man das Phänomen abbilden könnte; oder wie sollte man sich eine Integration von Informationsdichte, zeitlicher Dynamik der Informationsverarbeitung, Verantwortung für Sachwerte, Zeitgrenzen der Aufgabebearbeitung, physikalischen und sozialen Umgebungsbedingungen die psychisch auf den Arbeitenden einwirken, vorstellen? und zwar am besten noch in einem einzigen Messwert? Auffälligerweise wird das allerdings von der körperlichen Belastung, für die ja ebenfalls kein gemeinsames Bezugssystem vorliegt, nicht gefordert. Hier sind durchaus Einzelmesswerte (unterschiedliche Parameter) für einzelne *Kompo-*

*nenten* der Belastung üblich, und zwar auf unterschiedlichen Bezugssystemen, z.B. für schwere dynamische Arbeit und einseitig dynamische Muskelarbeit, und niemandem würde wohl einfallen, diese als Maß der Gesamtbelastung zu integrieren. Vielleicht wäre das ja auch für einzelne *Komponenten* der psychischen Belastung akzeptabel, wobei dann immer noch die Frage wäre, auf welchen Dimensionen und mit welchen Bezugssystemen diese Abbildung von Teilkomponenten der psychischen Belastung erfolgen könnte.

Das gilt sicher auch für den Bereich der psychischen Beanspruchung. Auch hier würde sicher gerne ein einheitliches Maß gesehen, mit dem sich die psychische Beanspruchung, und das ist nach der DIN EN ISO 10075-1:2000 „Die unmittelbare (nicht die langfristige) Auswirkung der psychischen Belastung im Individuum in Abhängigkeit von seinen jeweiligen überdauernden und augenblicklichen Voraussetzungen, einschließlich der individuellen Bewältigungsstrategien“ erfassen ließe. Der Vorteil eines einheitlichen Maßes auf einem einheitlichen Bezugssystem bestünde darin, von der Beanspruchung im Sinne einer Umkehrfunktion auf die zu bewertende und zu gestaltende psychische Belastung zurück zu schließen. Dies sollte möglich sein, wenn die Beanspruchung eine Funktion der Belastung ist und individuelle Effekte per Mittelung eliminiert werden können (vgl. Schmidtke & Bubb 1993), um der Belastung dann eine einheitliche Größe zuzuordnen zu können. Aber auch dieses Vorhaben muss wohl zu den Akten gelegt werden, wie einige der Beiträge in diesem Heft belegen: auch für die psychische Beanspruchung dürfte es wohl keine einheitliche Dimension / Größe / Referenzsystem oder was auch immer geben (vgl. auch Schmidtke 1993). Auch hier stellt sich selbstverständlich die Frage nach dem Bezugssystem, und zwar selbst dann, wenn man Versuchspersonen im Labor oder Arbeitende an ihren Arbeitsplätzen bittet, ihre psychische Beanspruchung (was verstehen die eigentlich darunter?) insgesamt (also wohl integrativ) auf einer wie immer gearteten Skala einzuschätzen. Vielleicht wäre auch hier eine differenzierte Erfassung erfolgversprechender und ehrlicher - und theoretisch sinnvoller (Nachreiner 1981, 1982).

Richter und Mitarbeiter haben immer wieder, so auch in diesem Heft (Richter et al. 2002), belegt und betont, dass auch die Erfassung der kurzfristigen (beeinträchtigenden) Beanspruchungsfolgen nur differenziert Sinn macht, weil sich diese Beanspruchungs-



folgen (z.B. Ermüdung, Monotonie oder Sättigung) erstens deutlich unterscheiden, und zwar auf verschiedenen Indikationsbereichen, und zweitens von unterschiedlichen Belastungsbedingungen verursacht werden. Gestaltungs- und Interventionsmaßnahmen zur Vermeidung dieser Folgen müssen daher an unterschiedlichen Arbeitsbedingungen, d.h. an unterschiedlichen Komponenten der Belastung, ansetzen. Demerouti et al. (eingereicht) haben jüngst Hinweise gefunden, dass dies wahrscheinlich auch für längerfristige Beanspruchungsfolgen wie Burnout oder seine Teilkomponenten (Erschöpfung und Distanzierung) gilt.

Insofern sind allen Versuchen, psychische Belastung, psychische Beanspruchung oder deren Folgen undifferenziert oder integriert zu erfassen, keine Erfolgsaussichten beschieden. Offensichtlich ist damit eine differenzierte Erfassung aller Stufen im Belastungs-Beanspruchungs-Prozess erforderlich, und zwar sowohl aus praktischen, z.B. zum Zwecke der Prävention beeinträchtigender Beanspruchungsfolgen, wie aus theoretischen Gründen, z.B. um etwas über die Wirkmechanismen und ihre Dynamik zu erfahren und diese besser modellieren zu können. Dabei können und werden komplexere Modellannahmen als die des Belastungs-Beanspruchungs-Modells erforderlich sein, um zu sinnvollen und validen Modellen im Bereich der psychischen Belastung zu kommen (Nachreiner 1981).

Wichtig ist und bleibt aber zunächst einmal klar zu differenzieren, was da eigentlich erfasst wird oder werden soll: Phänomene *außerhalb* der Person, also personenunabhängige Bedingungen, oder aber *innerhalb* der Person sich abspielende Reaktionen einer Person auf diese Bedingungen, und dazu zählen auch deren Wahrnehmungen, subjektive Repräsentationen, Redefinitionen oder verbale Beschreibungen dieser situativen Bedingungen, die für eine valide Modellbildung, Prognose und Ableitung von Gestaltungsempfehlungen jedoch zunächst wieder objektiv, d.h. unabhängig vom Erfassenden, sowie reliabel, valide, sensitiv und diagnostisch aussagefähig zu erfassen sind (vgl. ISO/CD 10075-3). Es lässt sich nämlich leicht zeigen, dass zwischen den subjektiven Repräsentationen von Arbeitsbedingungen bei den Betroffenen, zumindest wie sie in Fragebögen wiedergegeben werden, und den von unabhängigen und unbeteiligten Beobachtern eingeschätzten Ausprägungen derselben Bedingungen keine allzu starken Zusammenhänge bestehen,

zumindest nicht in einer Größe, die den einen Messwert durch den anderen ersetzbar machen würde, und schon gar nicht, wenn man gängige Anforderungen an die Reproduzierbarkeit von Messergebnissen zu Grunde legt (vgl. Demerouti 1999). Das dies zu theoretischen wie praktischen Fallstricken führt, sollte eigentlich offensichtlich sein.

#### 4 WIE KANN MAN DAS ERFASSEN?

Wenn man also praktische wie theoretische Ziele verfolgt, ist eine saubere Erfassung oder Messung der jeweils interessierenden Phänomene unabdingbar. Dabei ergibt sich aber sofort die Frage: geht das eigentlich in diesem Bereich? Was sollten wir denn da messen? und wie? und hat das überhaupt etwas mit messen zu tun oder ist das alles nur Hokuspokus und Vortäuschung falscher Tatsachen?

Schmidtke (2002) sieht das bekanntermaßen und basierend auf eigenen Erfahrungen ausgesprochen kritisch. Wir tendieren hier zu einer etwas optimistischeren Sichtweise, nicht weil wir der Meinung wären, dass alles, was da auf dem (nicht zuletzt wegen der - ungenauen - gesetzlichen Vorgaben boomenden) Psychomarkt als Messung verkauft wird, auch eine Messung darstellt. Hier stimmen wir mit Schmidtke überein: was da manchmal als Messung ausgegeben wird, spottet jeder Beschreibung und wird den einschlägigen Gütekriterien in keiner Weise gerecht.

Andererseits bedeutet Messen nicht zwingend apparatives Messen oder pointer measurement im System der Standardmaßeinheiten, sondern nach gängiger messtheoretischer Auffassung (Orth 1974; Stevens 1951) die Abbildung eines empirischen Relativs auf einem numerischen Relativ nach bestimmten Regeln, wobei für unterschiedliche Gegenstandsbereiche unterschiedliche Skalenqualitäten und Messgüten erreichbar sind (Stevens 1946). Dabei ist die Frage der erforderlichen Messgüte allerdings nicht absolut zu beantworten, sondern richtet sich nach dem jeweiligen Zweck der Messung. Für die Auswahl einer von mehreren Alternativen reicht beispielsweise eine sicher reproduzierbare Rangreihe dieser Alternativen auf dem interessierenden Merkmal (z.B. Signalunterscheidbarkeit oder Monotonie) aus. Dieses Niveau ist oft auch hinreichend für die Überprüfung theoretischer Modellannahmen oder von Arbeits-hypothesen. Andererseits sind für bestimm-

te andere Entscheidungen, z.B. über die Schädigungslosigkeit oder die Beeinträchtigungsfreiheit bestimmter Arbeitsbedingungen, Punktschätzungen - mit den üblichen Konfidenzintervallen - notwendig. Hier reicht eine Rangreihe offensichtlich nicht mehr aus. Neuere Entwicklungen auf dem Gebiet der psychologischen Messtheorie, wie die von Cronbach et al. (1972) vorgestellte und in mehreren Beiträgen dieses Themenheftes (Nickel et al. 2002; Pfendler & Schütte 2002; Schütte 2002; Schütte & Nickel 2002) angewandte Generalisierbarkeitstheorie (G-Theorie) tragen dem Rechnung und unterscheiden daher zwischen Messgenauigkeiten für relative und für absolute Entscheidungen.

Wichtig ist allerdings, dass für jede Art von Messung die psychometrischen Gütekriterien bekannt sind, denn nur dann ist beurteilbar, wie objektiv, zuverlässig, valide, sensitiv und diagnostisch aussagefähig die jeweilige Messung ist, wobei die Kennwerte alleine noch nichts aussagen, wenn nicht klar ist, wie sie ermittelt wurden. Das gilt übrigens auch für subjektive Tatbestände (als Messgegenstand), wie die erlebte Schwierigkeit oder Anstrengung, die möglichst objektiv (d.h. unbeeinflusst durch den Messenden, als Merkmal des Messverfahrens) zu erfassen sind.

ISO/CD 10075-3: 2001 spezifiziert daher Anforderungen an Messverfahren - nicht jedoch Messverfahren selbst - die für Messungen im Bereich der psychischen Belastung und Beanspruchung auf 3 verschiedenen Gütestufen gelten (vergleichbar den Güteklassen in der Lärmmessung). Dabei wird zwischen *bedingungs-* und *personenbezogenen* Messungen differenziert und gefordert, dass die Gütekriterien für die Einsatzart bestimmt werden, für die ein Verfahren vorgesehen ist. Es ist also zu unterscheiden, ob die Belastungsbedingungen, als Merkmale der Situation, oder die individuellen Beanspruchungsreaktionen, als Merkmale der Person, den Messgegenstand darstellen. Soll etwa die Lesbarkeit von Text (als Belastungsgröße) auf verschiedenen Bildschirmgeräten über die Leistung bzw. die dabei aufzuwendende Mühe oder Beeinträchtigung erfasst werden (vgl. DIN EN ISO 9241-3), so geht es darum, die Bildschirmgeräte in eine reproduzierbare Rangreihe zu bekommen oder ihnen auf den Dimensionen Leistung und Beeinträchtigungsfreiheit jeweils einen reproduzierbaren Punktwert zuzuordnen zu können. Damit ist die Präzision dieser Zuordnung zu bestimmen, mit den Bildschirmgeräten als Merkmalsträgern,

auch wenn die Daten mit Hilfe der Reaktionen von Personen ( $\approx$  Messgeräte) erhoben werden. In diesem Fall ist eine Reliabilitätsprüfung per Konsistenzschätzung (z.B. gerade gegen ungerade Versuchsdurchläufe) völlig unerheblich, da sie lediglich die Zuverlässigkeit, mit der die Leistung der Versuchsperson in diesem Test ermittelt werden kann, prüft. Die Diskrimination der Leistung unterschiedlicher Versuchspersonen ist jedoch hier überhaupt nicht relevant.

Wichtig ist also zunächst jeweils festzustellen, was der intendierte Messgegenstand ist. Sind die Belastungsbedingung die Merkmalsträger, so kann der Messgegenstand psychische Belastung, bzw. eine Teilkomponente davon, auch über (den Rückschluss von) Beanspruchungsreaktionen der Arbeitenden erhoben werden, z.B. über die subjektive Einschätzung der Beanspruchung durch diese Bedingungen oder über psychophysiologische Beanspruchungskennwerte, wenn diese Messwerte objektiv erhoben werden und zwischen unterschiedlichen Ausprägungen der Belastung zuverlässig, trennscharf, valide und diagnostisch aussagefähig differenzieren (und personenspezifische Varianzanteile eliminiert werden können). Aber genau dies bedarf des validen Nachweises. Diese Trennung in bedingungs- und personenbezogene Messung und der Nachweis der diesbezüglichen Gütekriterien ist bedeutsam, weil unter einer ergonomischen Perspektive das Messinteresse in der Regel auf der Beurteilung der Belastung und weniger auf der individuellen Beanspruchungsreaktion (als Personenmerkmal) liegen dürfte und eine hohe personenbezogene Messgüte noch keine Gewähr für eine hohe bedingungsbezogene Messgüte darstellt.

Nickel et al. (2002) belegen anhand einer Analyse der psychometrischen Gütekriterien der Herzfrequenzvariabilität, dass diese Unterscheidung in der Tat von Bedeutung ist. So konnte Eilers (1999; Nickel et al. 2002) zeigen, dass die Herzfrequenzvariabilität als personenbezogenes Merkmal ausgesprochen (und für psychologische Messungen fast unglaublich) hohe Reliabilitäten aufweist, als bedingungsbezogenes Merkmal bei Daueraufmerksamkeitsaufgaben unterschiedlicher Art und Schwierigkeit jedoch praktisch unbrauchbar, weil unzuverlässig ist. Während man also einzelne Personen ausgesprochen gut differenzieren kann - aber das ist unter einer ergonomischen Perspektive in der Regel nicht interessant - lassen sich unterschiedliche Belastungsbedingungen anhand der Herzfrequenz-

variabilität der Versuchspersonen nicht diskriminieren, obwohl dieses Maß bis in die jüngste Zeit hinein von einigen Autoren als valides, sensitives und diagnostisch aufschlussreiches Maß der psychischen Beanspruchung (mental strain) bezeichnet wird (Boucsein & Backs 2000; Luczak 1998; Mulder et al. 2000).

Um so wichtiger erscheinen daher sauber kontrollierte Studien zur Bestimmung der Messgüte von Verfahren, und zwar zunächst im Labor um dadurch zunächst die interne Validität der Evaluationsstudien (vgl. dazu Nachreiner et al. 1987) zu sichern. Auf die in der Sekundärliteratur (z.B. Boucsein & Backs 2000) berichteten Zusammenhänge kann man sich nämlich nicht unbedingt verlassen, weil eine genaue und kritische Durchsicht der vorliegenden Primärliteratur zeigt, dass es vielen der berichteten Studien gerade an dieser primär zu sichernden internen Validität mangelt (vgl. dazu Nickel 2001), von der externen Validität ganz zu schweigen.

Erst wenn sich ein Verfahren bei einer - am besten unter Laborbedingungen zu sichernden - intern validen Untersuchungsanlage bewährt, lohnt sich der aufwändigere und Arbeitende in ihren realen Kontexten betreffende Versuch der Erprobung im Feld. Diesen Versuch bestehen dann einige Verfahren, die sich im Labor als hoch reliabel erweisen, unter den im Feld gegebenen Bedingungen, z.B. der der Reaktivität bei reaktiven Messansätzen wie Fragebogenerhebungen - denn der Sinn solcher Messungen ist in der Praxis einfach zu durchschauen und die Ergebnisse in der Regel mit Konsequenzen verbunden - nicht mehr (vgl. Eilers et al. 1989, 1990; interessant ist in diesem Zusammenhang, dass die Studie, die die hohe Reliabilität belegt (Eilers et al. 1986), gerne zitiert wird, während die Studien, die das Verfahren als invalide ausweisen, insbesondere unter Feldbedingungen (Eilers et al. 1989, 1990), nur höchst selten zitiert werden!).

Die in diesem Themenheft zusammengetragenen Studien zur Überprüfung der Gütekriterien einiger gängiger Messverfahren zur Erfassung der psychischen Beanspruchung (als bedingungsbezogenes Merkmal) belegen, welcher erhebliche methodische Aufwand notwendig ist, wenn man in diesem Bereich akzeptabel messen oder die interessierenden Gegenstände doch wenigstens einigermaßen sauber messen oder erfassen will. Dabei stellt ein Teil der Studien zunächst lediglich die Prüfung der notwen-

digen - aber nicht hinreichenden - Differenzierung arbeitsanalytisch und theoretisch klar differenzierbarer Belastungsbedingungen dar; für die Herzfrequenzvariabilität mit dem klaren Ergebnis, dass damit mentale oder kognitive Beanspruchung nicht zu erfassen ist, weshalb eine weitere Erprobung im Feld unterbleiben kann. Inwieweit damit jedoch Aktivierungsphänomene - oder emotionale Beanspruchungsreaktionen bzw. Stress im vorwissenschaftlichen Verständnis - erfasst werden können, bleibt weiterhin zu untersuchen, zunächst vorzugsweise unter kontrollierten Bedingungen im Labor.

Ebenfalls zu untersuchen bleibt schließlich auch, inwieweit sich die hier im Labor als einigermaßen brauchbar bzw. erfolgversprechend bewerteten Verfahren im Feld bewähren, denn das ist es ja, was uns eigentlich interessiert, insbesondere bei der Lösung praktischer Fragestellungen. Daher ist die Überprüfung und Bewährung eines Verfahrens im Rahmen intern valider Untersuchungsanlagen (z.B. unter kontrollierten Laborbedingungen) eine notwendige, aber noch keine hinreichende Bedingung für den Einsatz eines Messinstrumentes in der Praxis. Hier ist der Nachweis der externen Validität unabdingbar, der hier noch erbracht werden müsste bzw. bereits hinfällig ist, weil diese Minimalforderung nicht erfüllt ist. Nicht zuletzt aus diesem Grund dürfte die Forderung nach gegenseitiger Kontrolle von Feld- und Laborstudien eine lange Tradition innerhalb der Arbeitspsychologie haben.

Die in den Studien in diesem Themenheft dargestellten Vorgehensweisen und Ergebnisse zeigen auch, dass und wie die in der jetzt zur Abstimmung anstehenden ISO/CD 10075-3 verlangten Kennwerte ermittelt werden können und dass Vorgehensweisen, die sich lediglich auf eine Rangkorrelation zwischen zwei Beobachtern über verschiedenen Items beschränken (und sei es auch als sog. vollständige Doppelanalyse, Oesterreich 1992) - oder ähnliche Standards verfolgen - nicht mehr ganz dem State of the Art entsprechen. Damit wird bei vielen Arbeitsanalyseverfahren der Anschein einer höheren Reliabilität erweckt als tatsächlich vorhanden. Das bei Oesterreich (1992) angeführte fiktive Beispiel macht nämlich deutlich, dass nicht nur eine (für die bedingungsbezogene Analyse irrelevante) personenbezogene Konsistenzanalyse bei erheblichen Differenzen auf sämtlichen Items hohe Werte erreicht, sondern auch, dass sich innerhalb der Arbeitstätigkeiten (die hier das Messobjekt darstellen) auch eine perfekte Be-



obachterübereinstimmung ergäbe, und zwar trotz erheblicher Unterschiede zwischen den Untersuchern. Die gelegentlich vorgeschlagene Berücksichtigung von Niveauunterschieden zwischen den Beurteilern kann dabei nur als wenig taugliches Hilfsmittel angesehen werden. Sinnvoller erscheinen dagegen Verfahren, die systematische Varianzquellen (und zwar nicht nur als Haupt- sondern auch als Interaktionseffekte, z.B. wenn bestimmte Beurteiler nur bestimmte Items auf anderem Niveau benutzen) in den Griff kriegen, und damit eine saubere Messung der eigentlich interessierenden Messgegenstände erlauben. Die G-theoretisch orientierten Beiträge in diesem Heft (Nickel et al. 2002; Pfendler und Schütte 2002; Schütte 2002, sowie Schütte und Nickel 2002) belegen, dass der Weg, Fehlervarianz zu systematisieren und damit kontrollierbar zu machen, offensichtlich erfolgversprechender ist, insbesondere im Gegensatz zu klassischen Ansätzen, die den Messfehler nur als Zufallsfehler behandeln und auf Ausgleich durch Mittelwertbildung hoffen. Gerade die Ausparialisierung systematischer Fehler- bzw. bedingungsbezogener Varianzkomponenten lässt uns die eigentlich interessierenden Messwerte damit um so präziser schätzen; ein Vorteil, der in einem bisher - und sicher teilweise auch zu Recht - als von ausgesprochener Unschärfe gekennzeichneten Feld von besonderer Bedeutung sein dürfte. Es erscheint von daher allerdings bemerkenswert, dass selbst in einem Handbuch psychologischer Arbeitsanalyseverfahren (Dunkel 1999), in dem es auch um Verfahren zur Ermittlung psychischer Belastung geht, G-theoretische Konzepte offensichtlich keine Rolle spielen, wie ein Blick in das Sach- und Namensverzeichnis schnell zeigt.

ISO/CD 10075-3 spezifiziert keine Strategie der Überprüfung der psychometrischen Kennwerte als die einzig zulässige. Das würde auch keinen Sinn machen, weil diese Überprüfung nur entsprechend dem jeweils gewählten Messansatz erfolgen kann, d.h. dass für probabilistische Modelle andere Überprüfungen erforderlich sind als für klassische Messansätze. ISO/CD 10075-3 empfiehlt jedoch generalisierungstheoretische Überprüfungen, weil damit mit *relativ* geringem Aufwand eine simultane Überprüfung von Objektivität, Reliabilität, Validität, Sensitivität und Diagnostizität der Messung erfolgen kann, ein für bedingungsbezogenes Messen nahezu unschätzbare Vorteil, denn insbesondere unter Feldbedingungen sind den Möglichkeiten der (strengen) Überprüfung der psychometri-

schen Gütekriterien enge sachliche und finanzielle Grenzen gesetzt. Insofern stellen die hier vorgestellten Studien gleichzeitig eine erfolgversprechende - Erprobung dieses Nachweises der Gütekriterien dar, denn wie die Ergebnisse von Eilers (1999; Nickel et al. 2002) belegen, lässt sich sehr gut damit überprüfen, *wie präzise* ein Messverfahren tatsächlich *was* misst; in diesem Falle offensichtlich - und wie nach einigen besser kontrollierten Studien in der Literatur zu erwarten (vgl. dazu Nickel 2001) - nicht mentale Beanspruchung.

Interessant wäre, wenn die dazu erforderlichen Arbeitsschritte nicht müßig wären, weil es dazu einfachere und präzisere Messgeräte (z.B. Uhren) gibt, mit Hilfe der G-Theorie zu bestimmen, wie genau sich mit der Herzfrequenzvariabilität die Zeit seit dem Beginn des Versuches abschätzen lässt: denn das scheint die Herzfrequenzvariabilität als Aktivierungsindikator in den Versuchen von Eilers (1999) tatsächlich abzubilden. Das interessante an diesem Ansatz ist daher nach unserer Auffassung, dass man bei geeigneter Anlage des Untersuchungsdesigns für die Überprüfung der Messgüte eines Verfahren sehr schnell auf dessen Probleme stößt und diese dann auch quantitativ abschätzen kann. Darüber hinaus erlaubt der Ansatz Prognosen, in welcher Weise die Messgenauigkeit verbessert werden kann, und auch hier kann sehr schnell deutlich werden, dass die ggf. erforderliche Erhöhung der Messversuche über Items oder Versuchspersonen, um zu einer akzeptablen Messgüte zu kommen, für die Praxis irrelevant, weil undurchführbar ist. So belegt etwa Schütte (2002), dass für die Anstrengungsskala eine Anzahl von Messungen notwendig ist, die in der Praxis illusorisch sein wird. Nicht nur weil es sich um zu viele zeitaufwendige Einzelmessungen handeln müsste, sondern weil die Versuchsperson sich und uns unter Realbedingungen fragen würden, ob wir noch alle Tassen im Schrank haben, wenn wir sie denselben Messgegenstand mehrfach einschätzen lassen wollten.

Uns scheinen diese Hinweise auf Möglichkeiten zur Verbesserung der Messgenauigkeit eines Instrumentes (oder dessen Einsatzes) ein bemerkenswerter Vorteil dieses Vorgehens zu sein. Darüber hinaus lässt sich so auch die Sensitivität prüfen, also die Möglichkeit des Messverfahrens nicht nur zwischen tot und lebendig - oder zwischen Arbeit und Ruhe wie bei der Herzfrequenzvariabilität (vgl. Nickel et al. 2002) oder geschlossenen oder geöffneten Augen mit Hilfe des EEG (Hecker et al. 1980), sondern

insbesondere in den für die Praxis relevanten (mittleren) Bereichen zuverlässig und hinreichend zu differenzieren, ein Problem an dem Verfahren zur Erfassung der psychischen Beanspruchung (und hier sei beispielhaft auf die AZA-Skala von Bartenwerfer (1969) und die Versuche mit Doppel-tätigkeiten von Bornemann 1942a,b hingewiesen) häufig leiden. So kann es beispielsweise im Rahmen der Beurteilung der psychischen Belastung bei Bildschirmarbeit (als bedingungsbezogene Erfassung über bestimmte Komponenten der psychischen Beanspruchung) nicht hingenommen werden, wenn dazu ein Verfahren eingesetzt wird, das nur extreme Fehlgestaltungszustände erkennen lässt; wiewohl es manchmal durchaus erwünscht zu sein scheint, dass mit Hilfe der angewandten Verfahren nichts entdeckt wird, damit auch alles so bleiben kann wie es ist - obwohl man mit hinreichend sensitiven Instrumenten durchaus bedeutsame Gestaltungsmängel entdecken würde. Anders kann man sich den Einsatz bestimmter Verfahren kaum erklären.

Auch die Überprüfbarkeit der Diagnostizität, also der Möglichkeit und der Genauigkeit der Differenzierung unterschiedlicher Arten oder Formen der Belastung oder Beanspruchung (im Gegensatz zur Sensitivität, bei der es um die graduelle Abstufung *einer* Belastungsart oder Beanspruchungskomponente geht) erscheint als ein Vorteil des G-theoretischen Ansatzes, entspricht dies doch genau den bekannten Forderungen von Wendt (1956) oder Schmidtke (1965), dass eine Validierung von Verfahren zur Erfassung der Beanspruchung oder ihrer Folgen, z.B. psychische Ermüdung oder Monotonie, lediglich an der Dauer der Belastung oder ihrer Intensität theoretisch (und zwar inhaltlich wie messtheoretisch) wie methodisch defizitär ist und daher komplexere Validierungsstrategien erforderlich seien. Beide, Wendt (1956) wie Schmidtke (1965), haben sehr früh darauf hingewiesen, dass etwa für die Abschätzung der Arbeitsdauer einfachere, zuverlässigere und validere Instrumente existieren, insbesondere wenn man diese mit psychophysiologischen Verfahren vergleicht. So wären beispielsweise in den Versuchen von Eilers (1999) theoretisch signifikante und in ihrem Ausmaß relevante Varianzkomponenten für die *Interaktionen* von Schwierigkeitsstufen, Arten der Diskriminationsanforderungen *und* der Dauer der Belastung für einen akzeptablen Validitätsnachweis zu fordern gewesen, und das u.U. noch in einem differenzierbaren, spezifischen Muster, das den bei der Versuchsanlage zu vermutenden und in den

Leistungsdaten sich widerspiegelnden Nicht-Additivitäten entspricht. Die Ergebnisse zeigen, dass genau das nicht zutrifft. Lediglich für die Dauer seit Versuchsbeginn und die Versuchspersonen - also aus ergonomischer Perspektive uninteressante Varianzquellen - ergaben sich bedeutsame Varianzbeiträge, die eine zuverlässige, aber nicht intendierte Differenzierung der Versuchspersonen und der Dauer erlauben. Das bedeutet nicht, dass Dauereffekte für einen Validitätsnachweis entbehrlich wären; im Gegenteil, wenn die Beanspruchung eine Funktion von Intensität, Dauer und zeitlichem Verlauf der Belastung ist (und Intensität wie Dauer ein hinreichendes Ausmaß annehmen), dann sollten sich durchaus Effekte der Arbeitsdauer zeigen, zumindest ab einem bestimmten Niveau der Intensität der Belastung (bei Eilers (1999) operationalisiert durch die Schwierigkeit der geforderten Diskrimination), was sich in Interaktionseffekten ausdrücken sollte. Da diese nicht gefunden werden konnten und die HFV-Kennwerte sich hypothesenkonträr zu einer Beanspruchungsinterpretation verhalten, lässt sich daraus diagnostisch schließen, dass die HFV offensichtlich nicht die mentale Beanspruchung abbildet sondern wahrscheinlich eher allgemeine Aktivierung. Aber die kann wohl kaum mit psychischer Beanspruchung gleichgesetzt werden, auch wenn sich psychische Beanspruchung in Aktivierung äußern kann - aber eben nicht muss (vgl. Bartenwerfer et al. 1963).

Auch wenn man nicht bis zu den G-Koeffizienten durchrechnet (weil es sich bereits nach Inspektion der Varianzaufklärung nicht mehr lohnt!), kann der Ansatz der Varianzzerlegung der Messwerte über experimentelle oder quasi-experimentelle Designs erfolgversprechend zur Überprüfung der Messgüte eines Verfahrens eingesetzt werden, wie die Arbeit von Nickel (2001; Nickel et al. 2002) zeigt. Durch eine systematische Variation von theoretisch als bedeutsam betrachteten Variablen ist es hier gelungen nachzuweisen, dass sich mit der HFV maximal emotionale, nicht aber kognitive Beanspruchungsreaktionen erfassen lassen, und insbesondere nicht in dem explizit dafür in der Literatur benannten Bereich (vgl. dazu im Detail Nickel 2001), der Unterscheidung zwischen automatisierter und kontrollierter Informationsverarbeitung, die sich nach Sanders (1983) in der Zuwendung von gerichteter Aufmerksamkeit oder Effort unterscheiden sollen, was sich in einer Unterdrückung der HFV bei kontrollierter Informationsverarbeitung äußern sollte (vgl. auch Mulder et al. 2000), dies offen-

sichtlich aber nicht tun, wie Nickel (2001; Nickel et al. 2002) durch eine geschickte Operationalisierung belegt. Offensichtlich stimmt da etwas nicht bei den Verfechtern der HFV als Indikator mentaler Beanspruchung; entweder ist die theoretische Argumentation nicht ganz sauber, die Anlage der Validitätsstudien nicht intern valide (weil Drittverursachung, z.B. durch emotionale Beanspruchungskomponenten nicht ausgeschlossen werden kann), oder beides. Gut angelegte G-theoretische Überprüfungen lassen das schnell erkennen und können damit einen bedeutsamen Beitrag zur Abschätzung der Diagnostizität eines Verfahrens leisten.

Damit scheint uns relativ gut belegbar, dass der in ISO/CD 10075-3 empfohlene Ansatz zur Überprüfung von Verfahren zur Erfassung der psychischen Belastung, Beanspruchung oder ihrer Folgen eine der Forderungen und Zielsetzungen dieser Norm, nämlich die Spreu vom Weizen (sofern vorhanden) auf dem boomenden Psychomarkt zu trennen, erfüllen kann. Allerdings ist auch dieser Ansatz kein narrensicheres Allheilmittel. G-theoretische Ansätze verlangen für die Anwendung von Typ II Modellen (nach der Klassifikation von Eisenhart (1947) = Random-Effects-Modelle; bzw. unter bestimmten Bedingungen auch von gemischten Modellen, vgl. Pfendler & Schütte 2002) für die Schätzung der Varianzkomponenten die Realisierung von für die für die Generalisierung angestrebte Universa von Bedingungen (z.B. die zu unterscheidenden Stufen der Komplexität und Arten der Informationsverarbeitung) repräsentative Bedingungen - oder aber die Beschränkung der Generalisierungsweite auf Universa, für die die untersuchten Bedingungen repräsentativ sind. Das wird in der Ergonomie in der Regel zutreffender sein. Das beinhaltet aber auch, sich dieser Generalisierungsweite wohl bewusst zu sein, und nicht vom (ggf. ausgesprochen künstlichen) Labor auf das Feld (mit in der Regel ziemlich unterschiedlichen Bedingungen) zu generalisieren. Bevor wir Aussagen über die Güte eines Messinstrumentes für die Untersuchung realer Arbeitstätigkeiten akzeptieren, sollten wir danach den Nachweis der Gütekriterien für solche Bedingungen fordern, die die zu untersuchenden beinhalten - oder zumindest nicht definitiv ausschließen.

ISO/CD 10075-3 fordert neben der Angabe der Gütekriterien auch die Angabe, für welchen Messgegenstand diese Gültigkeit beanspruchen - ob also Belastungsmerkmale, Beanspruchungsmerkmale oder Beanspruchungsfolgen damit erhoben werden können

sollen - sowie die nachvollziehbare Dokumentation ihrer Überprüfung für jeden beanspruchten Gültigkeitsbereich. Auch damit soll noch einmal der Intention Nachdruck verliehen werden, dass beispielsweise eine Überprüfung der Reliabilität eines Kurzfragebogens zur psychischen Belastung, auf dem die Mitarbeiter angeben wie belastet sie sich erleben, als Konsistenzreliabilität innerhalb von Personen nichts, aber auch gar nichts mit seiner Aussagefähigkeit als Messinstrument zur Erfassung von - personenunabhängigen - Belastungsaspekten zu tun hat. Vielleicht lässt sich damit allenfalls die erlebte Belastung als Personenparameter abschätzen, der für eine bedingungsbezogene Beurteilung der Belastung u.U. allenfalls die Qualität eines Zufallsgenerators hat. Und es soll hier auch noch einmal darauf hingewiesen werden, dass man nicht davon ausgehen kann, dass durch Mittelwertbildung über verschiedene Personen reliable Messwerte für Bedingungen erreicht werden können - solange dies nicht definitiv belegt ist -, weil eine solche Mittelwertbildung nur Zufallsfehler mit Erwartungswert Null ausgleichen kann, nicht aber systematische Fehlertendenzen (bias), die häufig - und insbesondere unter Feldbedingungen - nicht ganz unwahrscheinlich sind (wie dies z.B. bei einer Frage nach der - subjektiv beurteilen - Fahrtüchtigkeit nach Alkoholkonsum und einem daran gekoppelten Führerscheinenzug zu erwarten wäre; hier würde sich wohl niemand auf die Aussagen der Betroffenen als Experten für ihren Zustand verlassen wollen). Auch hier erscheinen Messansätze, die solche systematischen Tendenzen abschätzbar und kontrollierbar machen, deutliche Vorteile zu bieten, und vielleicht bringt der Zwang, die Überprüfung der Messgüte nun - nach Art und Umfang - genau angeben zu müssen einige Autoren dazu, noch einmal gründlich darüber nachzudenken, was sie da eigentlich wie genau messen, z.B. ob psychische Belastung oder Arbeitszufriedenheit, die ja durchaus etwas miteinander zu tun haben können, aber definitiv nicht dasselbe sind, auch wenn beide über dieselbe Methode, die (wertende) Beschreibung der Arbeitsbedingungen durch die Betroffenen erhoben werden sollen.

## 5 WOHN ?

Vielleicht hilft das präzisere inhaltliche und methodische Nachdenken ja, einige der immer noch bestehenden Probleme der Messung von Komponenten oder Merkmalen der psychischen Belastung, Beanspruchung oder



ihrer Folgen zu lösen, damit wir Rubner's Forderung zur Erforschung der Probleme der psychischen Beanspruchung in Zukunft besser und solider als bisher nachgehen können. Dabei können wir auf ein universelles Maß, das alles in sich vereint, durchaus verzichten. Nicht nur weil das nicht machbar ist, wie auch die Beiträge in diesem Themenheft wieder belegen, sondern weil es nicht einmal wünschenswert wäre, weil solche Maße zwangsläufig zu einer durch nichts zu rechtfertigenden Übereinfachung führen müssten, die den tatsächlichen Gegebenheiten in diesem Gebiet nicht gerecht werden kann.

Erfolgversprechend für die Zukunft erscheint daher vor allem eine auf sauberen theoretischen Konzepten und empirischen Analysen aufbauende Entwicklung spezifischer Messinstrumente, die die zu erfassenden Merkmale oder Dimensionen der intendierten Messgegenstände so präzise wie möglich erfassen. Das verlangt jedoch zunächst einmal eine Klärung was da erfasst werden soll, für praktische wie für theoretische Zwecke. Das muss sich nicht ausschließen, muss aber auch nicht identisch sein. Wichtig erscheint vielmehr die Angemessenheit an den Verwendungszweck der Messung, und der kann sich bedeutsam zwischen der Lösung theoretischer, z.B. nach der genauen Wirkungsdynamik von bestimmten Belastungen, und praktischer Fragestellungen unterscheiden. Während wir bei der Klassifikation der Beanspruchungsfolgen, zumindest der kurzfristigen, theoretisch relativ klar sehen (vgl. Richter et al. 2002), auch wenn dabei manchmal noch methodisch nicht unerhebliche Probleme auftauchen, sieht das im Bereich der Belastung und Beanspruchung inhaltlich immer noch ziemlich unklar aus. Das Belastungs-Beanspruchungs-Modell ist schließlich lediglich ein formales Modell, das inhaltlich offen ist (Nachreiner 2001), und DIN EN ISO 10075-2 stellt eine Sammlung von Einzelmerkmalen dar, die bekanntermaßen zur psychischen Belastung beitragen. Aber so etwas wie eine Theorie der psychischen Belastung und Beanspruchung und der durch sie ausgelösten Folgen, insbesondere auch der längerfristigen, ist zur Zeit nicht erkennbar, auch nicht in den sog. ‚arbeitspsychologischen‘ Modellen bei Oesterreich (2001). Es erscheint uns daher sinnvoll, sich damit intensiver zu beschäftigen als bisher. Das setzt allerdings auch die Bereitstellung der erforderlichen Ressourcen voraus, weil sich das nicht mal so eben nebenher erledigen lässt, wie der Stand der Dinge zeigt.

Unter praktischen Gesichtspunkten sollten wir uns daher nach bestmöglicher theoretischer Klärung des Untersuchungsgegenstandes daran machen, spezifische Messinstrumente zu entwickeln, die die angestrebten Messgegenstände mit hinreichender Präzision zu erfassen erlauben (und dies nicht nur vortäuschen). Dafür gibt es bei angemessener Berücksichtigung methodischer Prinzipien bei der Konstruktion und Überprüfung der Instrumente durchaus gute Erfolgsaussichten, während uns Omnibus-Instrumente, die nach dem Motto ‚da gibt es einen Unterschied‘ zwischen irgendwas auf irgendeinem theoretisch kaum begründbaren Indikator von vornherein zum Scheitern verurteilt scheinen.

Dass es sich lohnt, für praktische wie für wissenschaftliche Zwecke, den Beziehungen zwischen technisch-organisatorischen Arbeitsbedingungen, deren subjektiver Repräsentation, den dadurch ausgelösten Reaktionen und den darauf basierenden Folgen auf spezifischen Skalen nachzugehen, zeigen die Bemühungen von Richter und Mitarbeitern sowie von Demerouti (1999) sowie Demerouti et al. (eingereicht), wohingegen uns Analysen von Zusammenhängen zwischen lediglich subjektiven Belastungs-, Beanspruchungs- und Beanspruchungsfolgen-Kennwerten nur als weitere und überflüssige Übungen zur Theorie der kognitiven Dissonanz (Festinger 1957) erscheinen.

## LITERATUR

- Bartenwerfer, H.:** Einige praktische Konsequenzen aus der Aktivierungstheorie. Zeitschrift für experimentelle und angewandte Psychologie, 16, 195-222, 1969
- Bartenwerfer, H.; Kötter, L. & Sickel, W.:** Beiträge zum Problem der psychischen Beanspruchung - II. Teil: Verfahren zur graduellen Beurteilung der psychischen Beanspruchung in der Industrie (Forschungsberichte des Landes NRW, Band 141, Nr. 1131). Köln: Westdeutscher Verlag, 1963
- Bornemann, E.:** Untersuchungen über den Grad der geistigen Beanspruchung. I. Teil. Ausarbeitung der Methode. Arbeitsphysiologie, 12, 142-172, 1942a
- Bornemann, E.:** Untersuchungen über den Grad der geistigen Beanspruchung. II. Teil. Praktische Ergebnisse. Arbeitsphysiologie, 12, 173-191, 1942b
- Boucsein, W. & Backs, R.W.:** Engineering psychophysiology as a discipline: Historical and theoretical aspects. In: R.W. Backs & W. Boucsein

(eds.): Engineering psychophysiology. Issues and applications. pp. 3-30. Mahwah: LEA, 2000

**Bundesministerium für Arbeit und Sozialordnung (BMAS):** Konzeptentwurf Schwerpunktaktion "Psychische Fehlbelastungen und Stress" im Rahmen der "Initiative für eine neue Qualität der Arbeit". Berlin: BMAS, 2001

**Cronbach, L.J.; Gleser, G.C.; Nanda, H. & Rajaratnam, N.:** The dependability of behavioral measurements: Theory of generalizability for scores and profiles. New York: Wiley, 1972

**Demerouti, E.:** Burnout: eine Folge konkreter Arbeitsbedingungen bei Dienstleistungs- und Produktionstätigkeiten. Frankfurt/Main: Peter Lang, 1999

**Demerouti, E., Bakker, A.B., Ebbinghaus, M., Nachreiner, F.:** From mental strain to burnout. (zur Veröffentlichung eingereicht)

**DIN EN ISO 9241-3: 2000:** Ergonomische Anforderungen für Büro Tätigkeiten mit Bildschirmgeräten. Teil 3: Anforderungen an visuelle Anzeigen. Berlin: Beuth, 2000

**DIN EN ISO 10075-1: 2000:** Ergonomische Grundlagen bezüglich psychischer Arbeitsbelastung. Teil 1: Allgemeines und Begriffe. Berlin: Beuth, 2000

**DIN EN ISO 10075-2: 2000:** Ergonomische Grundlagen bezüglich psychischer Arbeitsbelastung. Teil 2: Gestaltungsgrundsätze. Berlin: Beuth, 2000

**Eilers, K.:** Herzfrequenz und Herzfrequenzvariabilität als Indikatoren psychischer Beanspruchung. Zur Reliabilität und Validität von Herzfrequenz- und Herzfrequenzvariabilitätsmaßen als Indikatoren psychischer Beanspruchung bei Daueraufmerksamkeitsbelastungen. Frankfurt/Main: Lang, 1999

**Eilers, K., Nachreiner, F. & Hänecke, K.:** Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung. Zeitschrift für Arbeitswissenschaft, 40, 215-224, 1986

**Eilers, K., Nachreiner, F. & Böning, E.:** Zur subjektiven Skalierung psychischer Beanspruchung. Teil 1: Experimentelle Überprüfung der Validität verankerter Relativurteile. Zeitschrift für Arbeitswissenschaft, 43, 217-223, 1989

**Eilers, K., Nachreiner, F. & Böning, E.:** Zur subjektiven Skalierung psychischer Beanspruchung. Teil 2: Überprüfung der Validität verankerter Relativurteile in einer Felduntersuchung. Zeitschrift für Arbeitswissenschaft, 44, 24-29, 1990

**Eisenhart, C.:** The assumptions underlying the analysis of variance. Biometrics, 3, 1-21, 1947

**Festinger, L.:** A theory of cognitive dissonance. Stanford: Stanford University Press, 1957

**Hecker, R.; Schmidtke, H. & Wegener, H.:** Reliabilität und Validität spektraler EEG-Parameter als Indikatoren der psychischen Beanspruchung. Meisenheim: Anton Hain, 1980

**ISO/CD 10075-3: 2001:** Ergonomic principles related to mental work-load - Measurement and assessment of mental workload. (Committee Draft) Geneva: ISO, 2001

**Lehmann, G.; Atzlert, E.:** Die Med. Welt, 12, 1549-1550, 1938 (zitiert nach Rutenfranz et al 1981)

**Luczak, H.:** Arbeitswissenschaft. Berlin: Springer, 1998

**MASQT (Ministerium für Arbeit und Soziales, Qualifikation und Technologie des Landes Nordrhein-Westfalen):** Arbeitswelt NRW 2000. Belastungsfaktoren - Bewältigungsformen - Arbeitszufriedenheit. Düsseldorf: MASQT, Referat Presse und Öffentlichkeitsarbeit, 2001

**Meyer, I.; Nickel, P.; Schomann, C. & Nachreiner, F.:** Ergonomische Analysen der Gestaltungsgüte bildschirmgestützter Prozeßleitsysteme - Teil 1: Analyse der Schnittstellengestaltung. In: Gesellschaft für Arbeitswissenschaft e. V. (GfA) (Hrsg.): Arbeitsgestaltung, Flexibilisierung, Kompetenzentwicklung (47. Arbeitswissenschaftlicher Kongreß der GfA, Universität, Gesamthochschule Kassel, 14.-16.03.2001), S. 55-57. Dortmund: GfA-Press, 2001

**Mulder, G.; Mulder, L.J.M.; Meijman, T.F.; Veldman; J.B.P. & Van Roon, A.M.:** A psychophysiological approach to working conditions. In: R.W. Backs & W. Boucsein (eds.): Engineering psychophysiology. Issues and applications. pp. 139-159. Mahwah: LEA, 2000

**Nachreiner, F.:** Lärm: Isolierte Belastungsgröße oder Komponente der Gesamtbelastung? - Einige eher grundsätzliche Überlegungen zum Problem der Belastung und Beanspruchung durch Lärm am Arbeitsplatz. In: A. Schick (Hrsg.): Akustik zwischen Physik und Psychologie. Ergebnisse des 2. Oldenburger Symposiums zur Psychologischen Akustik. S. 211-214. Stuttgart: Klett-Cotta, 1981

**Nachreiner, F.:** Eigenzustandskalierung zur Integration von Beanspruchungen. In: W. Rohmert (Hrsg.): Ergonomie der kombinierten Belastungen. Bericht über die Herbsttagung der Gesellschaft für Arbeitswissenschaft e.V., Darmstadt 1981. S. 47-57. Köln: Otto Schmidt, 1982

**Nachreiner, F.:** Zur Belastung und Beanspruchung bei Überwachungs-, Kontroll- und Steuerungstätigkeiten. In: F. Nachreiner (Hrsg.): Aktuelle Pro-

bleme der Belastungs- und Beanspruchungsforschung. Festschrift zum 60. Geburtstag von Joseph Rutenfranz. S. 111-130. Frankfurt/Main: Lang, 1988

**Nachreiner, F.:** International Standards on Mental Work-Load - The ISO 10 075 Series. Industrial Health, 37, 125-133, 1999

**Nachreiner, F.:** Replik zu Oesterreich. Zeitschrift für Arbeitswissenschaft, 55, 173-175, 2001

**Nachreiner, F.; Meyer, I.; Schomann, C. & Hillebrandt, M.:** Überprüfung der Umsetzbarkeit der Empfehlungen der ISO 10075-2 in ein Beurteilungsverfahren zur Erfassung der psychischen Belastung (Schriftenreihe der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin. Forschung Fb 799). Bremerhaven: Wirtschaftsverlag NW, 1998

**Nachreiner, F.; Müller, G.F. & Ernst, G.:** Methoden zur Planung und Bewertung arbeitspsychologischer Interventionsmaßnahmen. In: U. Kleinbeck & J. Rutenfranz (Hrsg.): Arbeitspsychologie. Enzyklopädie der Psychologie D, III, 1. S. 360-439. Göttingen: Hogrefe, 1987

**Nickel, P.:** Sensitivität und Diagnostizität der 0,1 Hz-Komponente der Herzfrequenzvariabilität zur Erfassung psychischer Beanspruchung, Dissertation. Oldenburg: Carl von Ossietzky Universität, 2001 (im Druck)

**Nickel, P.; Eilers, K.; Seehase, L. & Nachreiner, F.:** Zur Reliabilität, Validität, Sensitivität und Diagnostizität von Herzfrequenz- und Herzfrequenzvariabilitätsmaßen als Indikatoren psychischer Beanspruchung. Zeitschrift für Arbeitswissenschaft, 56, 2002 (in diesem Heft)

**Oesterreich, R.:** Das Belastungs-Beanspruchungskonzept im Vergleich mit arbeitspsychologischen Konzepten. Zeitschrift für Arbeitswissenschaft, 55, 162-170, 2001

**Oesterreich, R.:** Überprüfung von Gütekriterien bedingungsbezogener Arbeitsanalyseverfahren. Zeitschrift für Arbeitswissenschaft, 46, 139-144, 1992

**Orth, B.:** Einführung in die Theorie des Messens. Stuttgart: Kohlhammer, 1974

**Pfendler, C. & Schütte, M.:** Untersuchung der Sensitivität von NASA-TLX und NASA-TLX-ZEIS bei einer Flugführungsaufgabe. Zeitschrift für Arbeitswissenschaft, 56, 2002 (in diesem Heft)

**Richter, P.; Debitz, U.; Schulze, F.:** Diagnostik von Arbeitsanforderungen und kumulativen Beanspruchungsfolgen am Beispiel eines Call Centers. Zeitschrift für Arbeitswissenschaft, 56, 2002 (in diesem Heft)

**Rubner, M.:** Denkschrift an den Senat der Kaiser Wilhelm Gesellschaft zur Errichtung eines Kaiser-Wilhelm-Institutes für Arbeitsphysiologie, 1911 (zitiert nach Rutenfranz et al. 1981)

**Rutenfranz, J.:** Zur Frage einer Weiterführung des ehemaligen Max-Planck-Institutes für Arbeitsphysiologie Dortmund. Gießen, Frühjahr 1969

**Rutenfranz, J.; Wahle, A. & Klimmer, F.:** Ein Geheimrat entwickelt die Arbeitsphysiologie - Geheimrat Rubner und seine Mitarbeiter. In: Deutsche Gesellschaft für Arbeitsmedizin (Hrsg.): Epidemiologische Ansätze im Bereich der Arbeitsmedizin. Arbeitsmedizinisches Kolloquium. (Bericht über die 21. Jahrestagung der Deutschen Gesellschaft für Arbeitsmedizin e.V., gemeinsam mit dem Verband Deutscher Betriebs- und Werksärzte e.V. - Berufsverband Deutscher Arbeitsmediziner, dem Hauptverband der gewerblichen Berufsgenossenschaften e.V. - Bonn. Berlin, 13. bis 16. Mai 1981). S. 495-506. Stuttgart: Gentner, 1981

**Sanders, A. F.:** Towards a Model of Stress and Human Performance. Acta Psychologica, 53, 61-97, 1983

**Schmidtke, H.:** Die Ermüdung. Symptome, Theorien, Messversuche. Bern: Huber, 1965

**Schmidtke, H.:** Überwachungs-, Kontroll- und Steuerungstätigkeiten (RKW Reihe Arbeitsphysiologie - Arbeitspsychologie). Berlin: Beuth, 1966

**Schmidtke, H.:** Mentale Beanspruchung durch informatorische Belastung. In: H. Schmidtke (Hrsg.): Ergonomie. 3. Auflage. S. 143-160. München: Hanser, 1993

**Schmidtke, H.:** Vom Sinn und Unsinn der Messung psychischer Belastung und Beanspruchung. Zeitschrift für Arbeitswissenschaft, 56, 2002 (in diesem Heft)

**Schmidtke, H. & Bubb, H.:** Das Belastungs-Beanspruchungskonzept. In: H. Schmidtke (Hrsg.): Ergonomie. 3. Auflage. S. 116-120. München: Hanser, 1993

**Schomann, C.; Meyer, I.; Nickel, P. & Nachreiner, F.:** Ergonomische Analysen der Gestaltungsgüte bildschirmgestützter Prozeßleitsysteme - Teil 2: Psychische Beanspruchung bei der Arbeit an Prozeßleitsystemen. In: Gesellschaft für Arbeitswissenschaft e. V. (GfA) (Hrsg.): Arbeitsgestaltung, Flexibilisierung, Kompetenzentwicklung (47. Arbeitswissenschaftlicher Kongreß der GfA, Universität, Gesamthochschule Kassel, 14.-16.03.2001). S. 59-62. Dortmund: GfA-Press, 2001

**Schütte, M.:** Bestimmung der bedingungsbezogenen Messgenauigkeit der Anstrengungsskala.



Zeitschrift für Arbeitswissenschaft, 56, 2002 (in diesem Heft)

**Schütte, M. & Nickel, P.:** Multivariate Analyse der bedingungsbezogenen Meßgenauigkeit von Beanspruchungsindikatoren. Zeitschrift für Arbeitswissenschaft, 56, 2002 (in diesem Heft)

**Singer, R.; Rutenfranz, J. & Nachreiner, F.:** Zur Beanspruchung des Menschen bei Überwachungs-,

Kontroll- und Steuerungstätigkeiten in der Industrie. Arbeitsmedizin, Sozialmedizin, Arbeitshygiene (ASA), 12, 314-319, 1970

**Stevens, S.S.:** On the theory of scales of measurement. Science, 103, 677 - 680, 1946

**Stevens, S.S.:** Mathematics, measurement, and psychophysics. In: S.S. Stevens (ed.), Handbook of experimental psychology, pp. 1 -49. New York: Wiley, 1951

**Wendt, H.W.:** Über Ermüdungs-Indikatoren. Psychologie und Praxis, 1, 129 - 146, 1956

#### ANSCHRIFT DES VERFASSERS

**Prof. Dr. Friedhelm Nachreiner**  
AE Arbeits- und Organisationspsychologie  
Carl von Ossietzky Universität Oldenburg  
26111 Oldenburg  
E-Mail: [friedhelm.nachreiner@uni-oldenburg.de](mailto:friedhelm.nachreiner@uni-oldenburg.de)