

Er det måling?

Datakonstruktion og resultatformidling i psykisk arbejdsmiljøforskning

Jacob Hilden Winsløw

Resultaterne fra spørgeskemaundersøgelser betegnes sædvanligvis som 'målinger'. Det gøres sjældent eksplicit hvilken teori om måling, der ligger til grund herfor. I denne artikel viser jeg, at mens data fra undersøgelser af det psykiske arbejdsmiljø som regel produceres ud fra én teori om hvad måling er, så kommunikerer resultaterne i den internationale litteratur, som om data var fremkommet på baggrund af en helt anden teori om måling. I slutningen af artiklen skitserer jeg nogle af de faktorer, som kan tænkes at have bidraget til udviklingen af denne modsætning mellem empirisk og kommunikativ forskningspraksis.

Undersøgelsen af sammenhænge mellem psykisk arbejdsmiljø på den ene side og medarbejdernes helbred, trivsel og fortsatte arbejdsmarkedstilknytning på den anden – i det følgende betegnet 'Psykisk arbejdsmiljø og helbred' – er et forskningsfelt i vækst. Som det vil fremgå, beskrives arbejdslivet i litteraturen fra dette forskningsfelt fortrinsvis ved hjælp af medarbejdernes besvarelser af selvrapporterende strukturerede spørgeskemaer med faste svarkategorier. Disse besvarelser omsættes derefter til tal i såkaldte Likert-skalaer.¹ I litteraturen fra forskningsfeltet betegnes resultatet af denne datakonstruktion altid som 'måling'. I denne artikel undersøges denne sprogbrug ved at sammenholde den anvendte fremgangsmåde til datakonstruktion med fire etablerede teorier om måling. I det følgende gør jeg først rede for, hvorledes likertskalerede spørgeskemadata konstrueres. Dernæst dokumenterer jeg påstanden om, at likertskalerede spørgeskemadata spiller en domi-

nerende rolle på forskningsfeltet 'Psykisk arbejdsmiljø og helbred'. I artiklens hovedafsnit sammenholder jeg konstruktionen af likertskalerede spørgeskemadata med de fire normative teorier om måling. Til sidst rejser jeg det videnskabshistoriske spørgsmål, hvorledes den nuværende datakonstruktionspraksis inden for forskningsfeltet 'Psykisk arbejdsmiljø og helbred' udviklede sig.

Hvordan konstrueres likertskalerede spørgeskemadata?

En likertskaleret variabel, dannet på baggrund af spørgeskemadata, forudsætter et spørgeskema med to² eller flere spørgsmål med faste svarkategorier. Som eksempel på en likertskaleret variabel kan vi tage variabelen 'Indflydelse i arbejdet' i den korte version af NFA's Tredækkerspørgeskema (se Faktaboks 1).

Som det ses i faktaboksen er der fire spørgsmål, som hver har fem, indbyrdes

32. Disse spørgsmål handler om indflydelse og forskellige krav i arbejdet.
 (Besvar venligst alle spørgsmål, og sæt kun ét kryds ud for hvert spørgsmål)

	Altid	Ofte	Somme- tider	Sjældent	Aldrig/ næsten aldrig
4. Har du stor indflydelse på beslutninger om dit arbejde?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Har du indflydelse på, hvem du arbejder sammen med?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. Har du indflydelse på, hvad du laver på dit arbejde?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24. Har du indflydelse på mængden af dit arbejde?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Kilde: Spørgeskema om psykisk arbejdsmiljø tilgængeligt på www.arbejdsmiljøforskning.dk

rangordnede svarmuligheder. Målingen af den enkelte medarbejders (oplevelse af) 'Indflydelse i arbejdet' kan beskrives som bestående af tre trin:

1. Medarbejderen (respondenten) placerer sig selv i én af de $5^4 = 675$ forskellige svarkombinationer, som de fire spørgsmål og fem svarkategorier gør mulige. Konkret kunne medarbejderens svar på de fire spørgsmål være (Altid, Aldrig, Sjældent, Sommetider).
2. Forskeren oversætter herefter medarbejderens svarsæt til et ordnet sæt af tal. Oversættelsen foretages som regel således, at talværdierne er naturlige tal med lige stor indbyrdes afstand. I tilfældet 'Indflydelse i arbejdet' i NFA's Tredækker er valgt tallene 0, 25, 50, 75, 100, hvor '0' svarer til 'Aldrig/næsten aldrig' og '100' til 'Altid'. Det ovenfor repræsenterede svarsæt oversættes således til talsættet (100, 100, 25, 50).
3. Forskeren foretager en afbildning af talsættet på en talakse, dvs. foretager en

reduktion af de fire tal i talsættet til et enkelt tal. I tilfældet NFA's Tredækker er afbildningsreglen summation (sammenlægning) med efterfølgende division med antallet af spørgsmål, som indgår i skalakonstruktionen. I det valgte eksempel 'Indflydelse i arbejdet' er antallet af spørgsmål fire, og medarbejderens beskrivelse af sit arbejdsmiljø i svarsættet (Altid, Aldrig, Sjældent, Sommetider) resulterer således i skalaværdien $(100 + 100 + 25 + 50)/4 = 68,75$.

En likertskaleret spørgeskemabaseret variabel som 'Indflydelse i arbejdet' er et eksempel på en større gruppe af variable, som dannes ved at svar på spørgeskemaspørgsmål med faste svarkategorier oversættes til talværdier, som derefter vægtes og summeres. I det valgte eksempel bliver alle spørgsmål tildelt samme vægt, og den talmæssige oversættelse af svarkategorierne er til et sæt af tal, som dels er fælles for alle fire spørgsmål, og dels er lige langt fra hinanden (0, 25, 50, o.s.v.). Når disse betingelser er op-

fyldt, taler man om Likertskalering. Det er denne type variable, som dominerer den beskrivelse af arbejdslivet, som finder sted på forskningsfeltet 'Psykisk arbejdsmiljø og helbred'.

Likertskalerede spørgeskemavariablers dominans

Den internationale litteratur om sammenhænge mellem det psykiske eller psykosociale arbejdsmiljø og medarbejdernes helbred, trivsel og arbejdsmarkedstilknytning er vokset i omfang i de seneste årtier og er stadig i vækst. En indikator herpå er, at antallet af publikationer registreret i ISI's databaser er steget fra en håndfuld om året før 1980 over 200-300 om året omkring årtusindeskiftet til mellem 400 og 500 om året siden 2006³. For at belyse likertskalerede spørgeskemavariablers dominans i forskningsfeltets beskrivelse af det psykiske arbejdsmiljø fandt jeg frem til de hyppigst citerede artikler fra forskningsfeltet i årene 2006-2009 gennem en kombineret maskinel og manuel søgning i ISI's *Web of Science*⁴:

Ved en søgning på søgeordene ((work* OR occupation*) AND psychosocial) fremkom 7.210 titler af hvilke 2.268 var publiceret mellem 1.1.2006 og 11.6.2009. 76 af disse publikationer var blevet citeret 15 gange eller hyppigere. Disse 76 publikationer blev gennemgået manuelt. 28 af dem var rapporter fra eller videre videnskabelige analyser af undersøgelser af sammenhænge mellem psykisk arbejdsmiljø og medarbejdernes helbred, trivsel eller arbejdsmarkedstilknytning. De resterende 48 var overvejende kliniske artikler om betydningen af 'psykosociale faktorer' (i en meget vid forstand) for helbredelsesprocesser eller for bestemte patientkategoriers livskvalitet, hvor 'occupation*' eller 'work*' optrådte i artiklens abstract uden at blive nærmere karakteriseret ved adjektivet 'psychosocial'.

19 af de 28 publikationer var primære forskningsrapporter fra empiriske undersøgelser. Disse 19 primære forskningsrapporter var blevet citeret mellem 15 og 89 gange i årene 2006-2009 og kan dermed med nogen ret forventes at være artikler, som metodemæssigt er karakteristiske for forskningsfeltet 'Psykisk arbejdsmiljø og helbred'. Ingen af de 19 rapporter betjente sig af kvalitative data til at beskrive arbejdsmiljøet. 1 af de 19 betjente sig af et naturligt eksperiment, som fandt sted inden for en eksisterende kohorteundersøgelse af amerikanske soldater. I de øvrige 18 blev arbejdsmiljøet beskrevet ved hjælp af spørgeskemaer med faste svarkategorier, hvor hver enkelt medarbejder beskrev sit eget arbejdsmiljø, og i 17 af disse anvendtes likertskalering til at konstruere de variable, ved hjælp af hvilke de psykosociale arbejdsmiljøforhold blev beskrevet. På denne baggrund forekommer det rimeligt at konkludere, at likertkonstruerede spørgeskemadata udgør den dominerende type data inden for forskningsfeltet 'Psykisk arbejdsmiljø og helbred'.

Er likertskalerede spørgeskemavariablers resultatet af måling?

I litteraturen om likertskalering handler uenigheden sædvanligvis om, hvad man kan tillade sig at gøre med resultatet af skalkonstruktionen, dvs. hvilke modeller og hypotesetests det er betimeligt at underkaste variable med likertskalaegenskaber (se f.eks. Göb et al. 2007; Stine 1989; Michell 1986; Townsend & Ashby 1984; Gaito 1980; Lord 1953; Stevens 1946). Jeg vil undersøge et spørgsmål, som er forbundet hermed, men som er mere grundlæggende, nemlig i hvilken forstand likertskalerede svar på selvrapporterende spørgeskemaer kan siges at være 'målinger' af (her) psykosociale arbejdsmiljøforhold. Til det formål har jeg udvalgt fire normative teorier om

måling. I hvor vel der findes andre teorier om måling, såsom Adams' (Adams 1966) og Marauns (Maraun 1998; Chow 1998), er de fire valgte de normative måleteorier, som andre forfattere hyppigst påkalder sig i disse årtier. De fire teorier⁵ er:

1. Michells teori om egenskabsmåling
2. Campbells teori om genstandsmåling
3. Stevens' teori om måling som en beskrivelig fremgangsmåde
4. Luce og Tukeys teori om flerdimensionel måling.

Jeg vil i min fremstilling af teorierne ikke forholde mig til dem eller til den indbyrdes kritik, som deres fortalere har publiceret, men alene sammenholde den ovenfor beskrevne fremgangsmåde til konstruktion af likertskalerede spørgeskemadata med hver enkelt teori, for at konstatere om teorien kan danne et normativt fundament for påstanden om, at fremgangsmåden resulterer i en 'måling'.

Michells teori om egenskabsmåling

Michell hævder at kunne spore sin teori om egenskabsmåling tilbage til Euklid og Aristoteles, og at den svarer til det syn på, hvad måling er, som dominerede fra det klassiske Grækenland op til det 20. århundredes begyndelse (Michell 1990, 1999). Denne påstand har jeg ikke set anfægtet i den del af litteraturen om måling, som jeg er bekendt med, og jeg vil derfor behandle Michells teori først.

Ifølge Michell er påstanden om en variabels målbarhed en hypotese, som skal efterprøves på samme måde som alle andre hypoteser. At hævde, at 'Indflydelse i arbejdet' er målbart, er at hævde en hypotese om verden: At der findes en egenskab (indflydelse) ved en klasse af objekter (arbejde, jobs), og at denne egenskab tilfredsstillers (mindst) to kriterier: Stærk orden og additivitet.

At 'Indflydelse i arbejdet' som en egenskab tilfredsstillers kravet om stærk orden betyder, at den indflydelse, der knytter sig til et bestemt arbejde, kan sammenlignes med den indflydelse, der knytter sig til ethvert andet arbejde, og at resultatet af sammenligningen i alle tilfælde er, at der er mere indflydelse i det ene arbejde end i det andet. At 'Indflydelse i arbejdet' opfylder kravet om additivitet betyder, at hvis der er to job, af hvilke der er mere indflydelse i det ene end i det andet, så er det muligt at finde et tredje job, hvor indflydelsen i det er præcis så stor, at summen af indflydelsen i det første og det tredje job er lig indflydelsen i det andet job. Hvis vi betegner indflydelsen i et tilfældigt job x med I_x så kan vi med fordel sammenfatte disse noget knudrede sætninger i symbolsprog:

1. Kravet om stærk orden: At alle I_x kan stilles op i en rækkefølge således at $I_1 > I_2 > I_3 > \dots > I_n$ (kravet om stærk orden⁶) og
2. At hvis $I_z > I_x$ så findes der et I_y således at $I_x + I_y = I_z$ (kravet om additivitet).

Hvorvidt vores eksempel fra NFA's Tredækker på en likertkonstrueret spørgeskemavariabel, 'Indflydelse i arbejdet', tilfredsstillers Michells krav til måling er ikke godt at vide. For der er endnu ingen, som eksplicit har fremsat hypotesen, at 'Indflydelse i arbejdet' er noget, som findes ude i verden på en sådan måde, at det kan rangordnes og adderes, endsige har efterprøvet hypotesen. Forskellige grader af indflydelse er noget, som *antages* at karakterisere forskellige former for arbejde. Så Michells teori kan ikke danne normativt grundlag for at betragte den ovenfor beskrevne fremgangsmåde til datakonstruktion som en 'måling'.

Campbells teori om genstandsmåling

Ifølge Michell blev det ovenfor skitserede 'klassiske' syn på måling forladt omkring

år 1900 på grund af en fejlagtig opfattelse af (blandt andre) Bertrand Russells kritik af det realistiske syn på det naturlige talsystem, altså det system, som udgøres af tallene 1, 2, 3 ... osv. (Michell 1999, kapitel 3). Russells kritik førte angiveligt til, at fysikerne opgav opfattelsen af, at der fandtes et naturligt forhold mellem mængde og tal, som blev udtrykt i måleenheden. Dermed blev det nødvendigt at knytte nye forbindelser mellem det empirisk observerbare på den ene side og talsystemerne på den anden side. Dette var, hvad fysikeren Norman Campbell gjorde i bogen *Physics. The Elements* (Campbell 1920). Mens jeg har kaldt Michells teori for en teori om egenskabsmåling, kunne man kalde Campbells teori for en teori om måling af genstande.

I Campbells optik bliver måling mulig ved, at man definerer to systemer uafhængigt af hinanden: Et system af empiriske relationer mellem genstande og et talsystem, hvor relationerne mellem tallene spejler relationerne mellem de empiriske objekter i det empiriske relationssystem. Den konkrete målingsteori specificerer så de krav, som en afbildning fra det empiriske system til talsystemet skal leve op til for at være en måling. Disse krav kan f.eks. være, at afbildningen i talsystemet bevarer den stærke eller svage orden og additiviteten mellem elementerne i det empiriske system.

En Campbell'sk måleteori for 'Indflydelse i arbejdet' kunne tage udgangspunkt i en antagelse af, at alle jobs med den samme stillingsbeskrivelse havde om ikke identisk indflydelse så dog tilstrækkelig ens indflydelse til, at man kunne lægge dem i samme klasse i et empirisk relationssystem. Bestemmelsen af, i hvilken klasse hvert enkelt job skulle placeres, kunne man foretage på baggrund af selvrapporterende spørgeskemaer, ekspertvurderinger og/eller andre informationer om hvert enkelt job. Lad os betegne klasserne i dette empiriske system som x, y, z, \dots

og vælge et talsystem, i hvilket vi vil afbilde x, y, z, \dots . Hvis vi så for alle x definerer afbildningen $I(x)$ som det element i talsystemet (det tal), som opfylder betingelsen, at $I(x)$ kun er mindre end/lig med/større end $I(z)$, såfremt x er hhv. mindre end/lig med/større end z , så har vi defineret en afbildning fra det empiriske system til det numeriske, som bevarer rangordenen i det empiriske system. Hvis vi ydermere antager, at relationerne mellem klasserne x, y, z m.fl. i det empiriske system – dvs. relationerne mellem forskellige jobs 'indflydelsesfylde' – forholder sig til hinanden således, at indflydelsen i job A i klasse x og job B i klasse y kombineret svarer til indflydelsen i job C i klasse z for alle jobs og dermed for alle vore klasser i det empiriske relationssystem, så kan vi endvidere definere vores afbildning således, at $I(x) + I(y) = I(z)$. Hermed ville vi have skabt en skala til måling af 'Indflydelse i arbejdet', som levede op til Campbells krav til (genstands)måling. Forskellen fra Michells teori til Campbells er, at mens Michells handler om relationerne mellem *egenskaber ved genstandene*, så handler Campbells om relationerne mellem *genstandene selv*.

Læser man tidlige studier af sammenhængen mellem psykisk arbejdsmiljø og medarbejdernes helbred, er det åbenbart, at ambitionen i 1960'erne og 1970'erne var, at nå frem til en måling i en Campbell'sk forstand af det, vi i dag i Danmark kalder psykisk arbejdsmiljø. Det gælder for Karaseks første præsentation af 'Job Strain' modellen (Karasek 1979) og forgængerne fra Herzbergs Job Enrichment-skole, såsom artiklerne om udviklingen af det såkaldte 'Job Diagnostic Survey' (Hackman & Lawler 1971, Hackman & Oldham 1975). Denne ambition blev siden opgivet (jf. f.eks. Frese & Zapf 1988 for nogle af de samtidige begrundelser herfor og Kasl 1998 for den senere diskussion). Men Likertskaleringen af den enkelte medarbejders besvarelse af spørgsmålene om

'Indflydelse i arbejdet' i faktaboksen ovenfor kan ikke passes ind i den Campbell'ske målemodels form. Hver af de 675 distinkte svarkombinationer, som spørgsmålene gør mulige, kan ganske vist svare til en klasse i det empiriske relationssystem. Lad os antage at der mellem disse 675 klasser kan etableres en rangorden og additivitet. I givet fald ville vi have et ERS i fire dimensioner (ét svarende til hvert spørgsmål), hvis struktur ville blive bevaret i oversættelsen af svarsættet til talsættet i trin 2 i likertkonstruktionsprocessen. Men den efterfølgende reduktion til én dimension (summationen i processens trin 3) ville ikke bevare disse relationer. Så selv hvis vi vidste, at de mulige svarkombinationer udgjorde et ERS, så ville datakonstruktionsprocessen ikke være en måling i Campbell'sk forstand.

Stevens' teori om måling som en beskrivelig fremgangsmåde

Syv år efter Norman Campbells *Physics. The Elements* udkom Percy W. Bridgmans *The Logic of Modern Physics* (Bridgman 1927). Bridgman havde allerede i 1914 udtrykt sympati for Ernest Machs instrumentalistiske syn på fysiske teorier (Boyd 2002), men hævdede i 1927, at det først og fremmest var udviklingerne i årene 1925-1926 i den specielle relativitetsteori og kvantemekanikken, der stillede fysikken overfor et krav om at give en ny fortolkning af de fysiske begrebers ontologiske status. Bridgmans konstruktivistiske synspunkt er måske mest prægnant udtrykt i Herbert Dingles formandsberetning til the British Society for the Philosophy of Science i 1949:

"It is generally supposed that a measurement is a determination of the magnitude of some inherent property of a body... 'Bodies' are assumed, having 'properties' which have 'magnitudes'. All that 'exists', so to speak, before we begin to measure. ... I want to sug-

gest that all this is incompatible with the view of physical science which the theory of relativity forces us to take. Instead of supposing a pre-existing 'property' which our operation measures, we should begin with the operation and its result and then, if we wish to speak of a property (which I do not think we shall do) define it in terms of that. ... Formerly science was regarded as the study of the external world, independent of the observer ... [but now] any world that we may contemplate is no longer an independent existence whose nature demands or determines [our data] but rather a logical construct ..." (Dingle 1950, 7).

Hovedrepræsentanten for dette konstruktivistiske syn på måling inden for samfunds- og menneskevidenskaberne var den amerikanske perceptionspsykolog Stanley Smith Stevens (Stevens 1935a, 1935b, 1946). I sin ofte citerede 1946-artikel definerede Stevens måling som *"the assignment of numerals to objects or events according to rules"* (Stevens 1946, 677). I denne definition af måling er der ingen henvisning til nødvendige relationer mellem de egenskaber eller genstande, som er objekt for målingen. En hvilken som helst fremgangsmåde, som kan beskrives nøjagtigt, og som resulterer i et tal, er en måling. Likertskalering af spørgeskemadata, som beskrevet i trin 1-3 ovenfor, er helt klart et eksempel på måling i denne forstand. Ud fra Stevens' teori fremkommer værdierne af Tredækkervariablen 'Indflydelse i arbejdet' som et resultat af måling.

Luce og Tukeys teori om flerdimensionel måling

Michells og Campbells teorier kan begge betegnes som 'repræsentationalistiske' måleteorier, fordi deres udgangspunkt er 'noget' – egenskaber eller genstande – som eksisterer uden for måleprocessen, mens Stevens' teori som nævnt er 'konstruktivistisk' eller

'ikke-repræsentationalistisk'. Den nyeste af de fire måleteorier, som vi ser på i denne artikel, Luce og Tukeys teori fra 1964, kan ses som en videreudvikling af den repræsentationalistiske teori. Mens Hand (1996) og Kyngdom (2008) betragter Luce og Tukeys teori som en forlængelse af Campbells, ser Michell (1990) Luce og Tukeys teori som en variant af den klassiske (egenskabsmålings) model. Rigtigt er det i hvert fald, at teorien om flerdimensionel måling betragter kravet om additivitet som en afgørende forudsætning for måling og kan siges at generalisere dette fra ét aspekt eller én type genstand i verden til flere. Michells (1999) beskrivelse af flerdimensionel måling som *"a way of identifying the additive structure of attributes indirectly, via trade-offs in the way that two attributes relate to a third"* giver idéen i denne måleteori i en nøddeskal. En nærmere redegørelse for teorien sker bedst gennem et konkret eksempel:

Teorien om flerdimensionel måling har ikke været anvendt inden for forskningsfeltet 'Psykosocialt arbejdsmiljø og helbred'. Det er imidlertid muligt at forlænge argumentationen i Robert Karaseks artikel fra 1979, som kan betragtes som både teoretisk og metodisk konstituerende for feltet, på en sådan måde, at den kan illustrere principperne bag flerdimensionelle målinger i to dimensioner. Ifølge Karaseks 'Job Strain' model hersker der følgende relation mellem psykisk arbejdsbelastning ('Job Strain') og dets to underkomponenter Jobkrav ('demands of a work situation' i citatet nedenfor) og 'Indflydelse i arbejdet' ('the range of decision making freedom'):

"psychological strain results not from a single aspect of the work environment, but from the joint effects of the demands of a work situation and the range of decision-making freedom (discretion) available to the worker facing those demands. These two aspects of

the job situation represent, respectively, the instigators of action (work load demands, conflicts or other stressors which place the individual in a motivated or energized state of "stress") and the constraints on the alternative resulting actions. The individual's job decision latitude is the constraint which modulates the release or transformation of "stress" (potential energy) into the energy of action. ... If no action can be taken ... the unreleased energy may manifest itself internally as mental strain." (Karasek 1979, 287)

Karasek undersøgte denne model på en række forskellige helbredsindikatorer såsom hyppigheden af udmattelse, depression, sygedage og medicinforbrug blandt medarbejdere i forskellige jobs. Det var ikke muligt for ham at måle 'Job Strain' eller dens komponenter Jobkrav og 'Indflydelse i arbejdet' direkte. Men han mente at kunne rangordne de undersøgte jobs med hensyn til Jobkrav og med hensyn til 'Indflydelse i arbejdet'. Dermed definerede han et to-dimensionalt rum for 'Job Strain', hvor de undersøgte helbredsudfald kunne ses som en tredje dimension.

Ulempen ved kun at kunne rangordne jobbene med hensyn til de to komponenter var selvfølgelig, at det i det rum, som blev dannet af de to rangordninger, ikke var muligt at definere afstanden mellem to jobs, og dermed addition. Altså var det heller ikke muligt at sige, at der forelå en egentlig måling af 'Job Strain'. Her kommer Luce og Tukeys teori ind.

En teori for måling af 'Job Strain' i de to dimensioner Jobkrav (J) og 'Indflydelse i arbejdet' (I) kunne opstilles på følgende måde: Antag at ikke kun de af Karasek undersøgte, men alle tænkelige jobs kan rangordnes med hensyn til Jobkrav og med hensyn til 'Indflydelse i arbejdet'. Antag endvidere at vi kan bestemme en funktion h , således

1. At vi for en række helbredsudfald (udmattelse, depression, medicinforbrug m.v.) kan konstatere, at de varierer på en 'meningsfuld' måde med h, og
2. At det for relationerne mellem J, I og h gælder,
 - a. at den svage ordensrelation \leq for ordnede par af J og I tilfredsstillende en række betingelser,
 - b. at der er indbyrdes substitutionalitet mellem J og I i h. Det vil sige, at der er tilstrækkeligt mange tilstrækkeligt forskellige jobs til, at det for alle kombinationer af J og I og den tilhørende værdi af h er muligt, hvis man ændrer på værdien af J, at finde en ny værdi af I, således at værdien af h er den samme som i den oprindelige kombination af J og I (og vice versa).
 - c. Lad der være givet to efter hinanden givne tilvækster langs J-aksen og to efter hinanden givne tilvækster langs I-aksen. Hvis det for begge tilvækster langs J gælder, at den tilhørende tilvækst i h er større end tilvæksten i h givet tilvæksterne langs I, så skal effekten på h af den kombinerede tilvækst langs J være større end effekten på h af den kombinerede tilvækst langs I (og vice versa).
 - d. At alle tilvækster i h som følge af tilvækster i J og I er sammenlignelige, dvs. ingen af dem er uendeligt store eller uendeligt små.

Hvis en sådan funktion h kunne bestemmes, ville det være muligt ikke blot at måle 'Job Strain', men også hver af komponenterne Jobkrav og 'Indflydelse i arbejdet'. Det ville blive muligt at sige ikke kun, at "Postbude har jobs som er dobbelt så Psykisk belastende (straining) som hundeluftere", men også, at "Direktører har dobbelt så meget 'Indflydelse i arbejdet' som hundeluftere", og "Postbudes Jobkrav er 30% højere

end direktørers" – i hvert fald inden for den ramme, som vedrører forholdet mellem psykisk arbejdsbelastning og helbredsforhold.

Konstruktionen af likertskalerede variable på baggrund af spørgeskemadata lever ikke op til de krav, som teorien om flerdimensionel måling opstiller til måling i to dimensioner. Dette kan umiddelbart ses af den tilfældighed, som er bestemmende for oversættelsen fra svarsæt til talsæt i konstruktionsprocessens trin 2: Adspurgt, hvorfor han vælger tallene 0, 25, 50, 75 og 100 som mulige værdier for (Aldrig, ..., Altid) i stedet for 1, 2, 7, 52 og 521 er forskerens svarmuligheder begrænset til overvejelser om talværdiernes indbyrdes relationer. Om relationen mellem talværdierne på den ene side og verden udenfor på den anden side må han være tavs. Så sammenholdt med modellen for flerdimensionel måling er likertskalering af spørgeskemadata ikke 'måling'.

Hvordan har praksis udviklet sig?

Ovenfor har jeg sammenholdt den fremgangsmåde til datakonstruktion, som dominerer feltet 'Psykisk arbejdsmiljø og helbred' med fire normative teorier om måling. Fælles for tre af disse teorier (Michells, Campbells og Luce & Tukeys) er et krav om, at noget ude i verden – det som måles – kan opstilles i en ordnet rækkefølge (stærk eller svag orden), og at de elementer af dette noget, som kan opstilles i ordnet række, ydermere kan sættes sammen på en sådan måde, at resultatet er 'mere af det samme' (en større pære) og ikke en kombination af to forskellige ting (et æble og en pære, f.eks.) – kravet om additivitet. Fordi de tre teorier tager udgangspunkt i nogle egenskaber ved verden, som kræves afbildet i måleresultatet for, at fremgangsmåden kan kaldes måling, kan Michells, Campbells og Luce & Tukeys teorier kaldes 'repræsentationalistiske'. Den måde, hvorpå de tre teo-

rier udmønter kravene om orden og additivitet, er forskellig. Men i intet tilfælde lever likertskalerede spørgeskemavariabler op til dem. Derfor er likertskalerede spørgeskemadata ikke målinger i disse teoris forstand.

Den fjerde af de gennemgåede teorier – Stevens' – er radikalt forskellig fra de tre førstnævnte. I den stilles ingen krav til det, som ønskes målt, eller til relationen mellem dette og måleresultaterne, men alene til fremgangsmåden. Er denne regelbunden og resulterer den i et tal, så har vi med måling at gøre. Følgelig er likertskalering af spørgeskemadata 'måling' i Stevens' forstand.

Valget af Stevens' måleteori er imidlertid ikke gratis, som det også fremgår af citatet fra Herbert Dingle ovenfor. Den pris, man må betale for at begrunde sine målemetoder med den, er, at man ophører med at tale om f.eks. 'Indflydelse i arbejdet' som en egenskab ved verden eller noget i verden, såsom jobs. 'Indflydelse i arbejdet' bliver til navnet på en variabel, ligesom 'Jacob' er et navn på (blandt andre) mig. Og ligesom der ikke er megen jacobhed i mig, bare fordi jeg lyder navnet 'Jacob', behøver der ikke være megen indflydelse i 'Indflydelse i arbejdet'. Man kunne også sige: Hvad enten der findes jacobhed i verden eller ej, så er det at kalde en dataproducerende fremgangsmåde Jacobhed ikke tilstrækkeligt til at godtgøre, at det er jacobhed, den opfanger. Det ses måske lettest ved at sammenholde indholdet af Faktaboks 1 med indholdet i Faktaboks 2. I Faktaboks 1 er indholdet, som allerede nævnt, spørgsmål til skalaen 'Indflydelse i arbejdet'. Men hvad skal indholdet i Faktaboks 2, som også er hentet i Tredækkeren, bruges til? Er spørgsmålene til 1, 2 eller flere skalaer? Hvilke spørgsmål hører sammen? Og hvad er navnet på skalaen – hvad 'måler' den? Svarene på disse spørgsmål er ikke indlysende for den, som ikke har et indgående kendskab til Tredækterspørgeskemaet⁷.

Prisen for at definere sine datakonstruktionsmetoder som måling i den konstruktivistiske forstand er, at de relationer, vi specificerer mellem variable, som vi navngiver på denne måde, ikke længere kan hævdes at være afspejlinger af forhold ude i verden, som eksisterer uafhængigt af vore måleoperationer. Tværtimod: De er *frembragt* af vore måleoperationer. Ud fra den konstruktivistiske synsvinkel kan vi ikke bruge Tredækker- eller andre likertkonstruerede spørgeskemaresultater til at begrunde påstande som, at "Indflydelse i arbejdet" resulterer i X", eller "Manglende Indflydelse i arbejdet" forårsager Y", eller blot "Z samvarierer med graden af 'Indflydelse i arbejdet'", hvor X, Y og Z er forhold ude i verden som f.eks. tilfredse medarbejdere, fraværdsdage pga. sygdom eller kortisolindholdet i spyt. Vi må nøjes med at tale om, at X, Y og Z "samvarierer med den variabel, vi betegner 'Indflydelse i arbejdet'".

De 17 hyppigst citerede forskningsrapporter fra forskningsfeltet 'Psykisk arbejdsmiljø og helbred', som anvender likertskalerede spørgeskemadata, betaler ikke den pris, som skal betales, hvis man ønsker at definere sine dataproduktionsmetoder som målinger i Stevens' forstand. De psykosociale arbejdsmiljøvariable omtales konsekvent som noget ude i verden uafhængigt af målingen, f.eks. som 'risikofaktorer' (f.eks. Chandola et al. 2006) eller 'arbejdsstressorer' (f.eks. Kivimäki et al. 2006 c), som har et 'dosis-response'-forhold til overvægt (Brunner et al. 2007), og hævdes at 'påvirke' eller ligefrem 'øge' risikoen for langvarig sygdom (Head et al. 2006; Lund et al. 2006) og 'forårsage' dødsfald fra kardiovaskulære sygdomme (Elovaino et al. 2006). Så forskerne inden for forskningsfeltet 'Psykisk arbejdsmiljø og helbred' skriver ud fra en forudsætning om, at likertkonstruerede spørgeskemadata er målinger af 'noget' uden for de konstruerede data.

35. Disse spørgsmål handler om det psykiske arbejdsmiljø og samarbejdet på din arbejdsplads.
 (Besvar venligst alle spørgsmål, og sæt kun ét kryds ud for hvert spørgsmål)

	I meget høj grad	I høj grad	Delvist	I ringe grad	I meget ringe grad
4. Kræver dit arbejde, at du er initiativrig?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Er dine arbejdsopgaver meningsfulde?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. Føler du, at du yder en vigtig arbejdsindsats?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. Kan du bruge din kunnen eller dine færdigheder i dit arbejde?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
31. Har du mulighed for at lære noget nyt gennem dit arbejde?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
32. Føler du dig motiveret og engageret i dit arbejde?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
33. Må du sommetider foretage dig ting i dit arbejde, som forekommer dig unødvendige?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Kilde: Spørgeskema om psykisk arbejdsmiljø tilgængeligt på www.arbejdsmiljøforskning.dk

Det er i denne sammenhæng interessant, at der på forskningsfeltet 'Psykisk arbejdsmiljø og helbred' er en betydelig interesse for det, som kaldes 'validering' af skalaer. De fleste af de 17 hyppigst citerede artikler rapporterede resultater fra valideringen af deres skalaer på deres egne data, og/eller henviste til andre artikler, hvori disse skalaer var blevet valideret på de samme eller andre rådata. 10 af de 17 artikler gengav den såkaldte Cronbachs α for deres skalaer, mens en enkelt (Jian Li 2006) henviste til separate artikler, hvor resultaterne af psykometriske valideringer af de anvendte skalaer blev gengivet. Cronbach's α (Cronbach 1951) er, let forenklet, en beregning af, hvor stærkt svarene på de enkelte items (spørgeskema-spørgsmål), som anvendes i konstruktionen

af en skala, er korrelerede med hinanden. Tankegangen bag beregningen af α er, at hvis α er meget tæt på 1 så er der redundans eller overflødige spørgsmål, mens hvis α er 'lav' er det tegn på, at spørgsmålene ikke 'måler samme ting'. Dvs. at målingens gyldighed, ud fra en Cronbach α -synsvinkel er et spørgsmål om relationen mellem svarene på forskellige spørgsmål i sættet af spørgsmål i (f.eks.) 'Indflydelse i arbejdet', altså om indbyrdes relationer mellem dele af måleinstrumentet, ikke om relationer mellem måleinstrumentet og noget uden for dette eksisterende. Polemisk formuleret spørger Cronbachs α , om længde kan måles på en gyldig måde, hvis ét stykke måles med Bentes centimetermål og et andet stykke med Bents lineal, men ikke om det, hvis

længde skal måles, har udstrækning eller ej. Den samme indvending kan rettes mod statistiske modeller som Konfirmatorisk faktoranalyse (f.eks. Siegrist et al 2004, Sale & Kerr 2002) og Item response-teori (f.eks. Christensen et al 2002), når de bruges som eneste redskab til valideringsformål.

Et resultat af min undersøgelse er, at forskningsfeltet 'Psykisk arbejdsmiljø og helbred' placerer sig mellem to stole med hensyn til sin målepraksis. På den ene side konstrueres data, som om man på feltet støttede sig på en konstruktivistisk måleteori. Og på den anden side afrapporteres de samme data, som om de var frembragt i overensstemmelse med en repræsentationalistisk måleteori, altså som om målingerne var afbildninger af noget ude i verden, givet før målingerne og eksisterende uafhængigt af måleproceduren. Men disse to teorier om måling er ikke forenelige, og det bliver derfor et videnskabshistorisk interessant spørgsmål, hvorledes de kunne blive bragt sammen inden for én og samme forskningspraksis. Et bare tilnærmelsesvist tilstrækkeligt svar på dette spørgsmål går ud over denne artikels rammer. Jeg skal derfor indskrænke mig til fem mere eller mindre spekulative betragtninger:

1. Likertskalering har i de seneste 3 årtier spredt sig inden for en lang række samfunds- og folkesundhedsvidenskabelige forskningsområder, men med mærkbart forskellig gennemslagskraft. Der er således god grund til at undersøge såvel generelle som specifikke betingelser for denne spredning.
2. Som ovenfor nævnt synes det åbenbart, at der i 1960'erne og 1970'erne var en ambition på forskningsfeltet om at nå frem til såkaldt 'objektive' kvantitative beskrivelser af de psykosociale belastninger og ressourcer i enkelte jobs, og at denne ambition blev opgivet i løbet af

1980'erne, jf. det tidligere citerede arbejde af Frese og Zapf (Frese & Zapf 1988). Det vil sige, at der var en periode, hvor i hvert fald nogle forskere på feltet stræbte efter at opnå repræsentationalistisk måling af disse variable.

3. Hvis man betragter Karaseks 1979-artikel (Karasek 1979) som den konstituerende artikel på feltet, så er det en kendsgerning, at Karaseks oprindelige model af psykisk arbejdsbelastning har to typer udkomme: Helbredseffekter og læring. Dette styrker en fortolkning af den tidlige diskurs om 'Job Strain', som bl.a. understøttes af Wilkinson (Wilkinson 2000, kapitel 5), at diskursen opstår i et møde mellem en stærkt socialdemokratisk præget svensk arbejdsfysiologisk teoretiseringen om stress, repræsenteret ved Bertil Gardell m.fl., og en amerikansk Human Ressource-inspireret arbejdsmotivationspsykologi. Hvis denne formodning er rigtig, så bragte den arbejdsfysiologiske stressforskning en stærk analytisk metafor i form af den toksikologiske dosis-respons-model til det nye forskningsfelt, mens den amerikanske motivationspsykologi medbragte en stærk tradition for at spørge medarbejdere om holdninger (attitudes) og bearbejde svarene til likertskalaer til feltet.
4. Mødet mellem svensk arbejdsfysiologi og amerikansk motivationspsykologi fandt sted på et tidspunkt, hvor en eksplosion i computerteknologien muliggjorde behandlingen af store datamaterialer (som f.eks. store spørgeskemaundersøgelser) med sofistikerede statistiske modeller, herunder specielt den multiple lineære regressionsmodel, som er den model, der udtrykker dosis-response-relationen i den enkleste form. Sammenlignet med 10 år tidligere har der i begyndelsen af 1980'erne været rigelige muligheder for at producere det materielle input (sta-

tistiske analyser af likertskalerede spørgeskemadata) til en diskurs om selvrapporterende spørgeskemadata som 'risikofaktorer for' eller ligefrem 'årsager til' iskæmisk hjertesygdom og andre helbredsforhold. Disse muligheder blev fra 1980'ernes slutning udvidet betragteligt i og med, at patient- og andre registre verden over blev EDBficerede og Dunns idé om kobling af personregistre fra 1946 fik praktisk anvendelighed (se f. eks. Belloc & Arellano, 1973 for et tidligt eksempel).

5. Som en sidste betragtning om betingelserne for spredningen og accepten af likertskalerede selvrapporterende spørgeskemadata på feltet 'Psykisk arbejdsmiljø og helbred' vil jeg nævne feltets annektering af miljømetaforen på et tidspunkt, hvor det fysiske miljøes politiske betydning var veletableret⁸. Karasek anvendte 'work environment' eller 'environmental' 18 gange i sin 1979-artikel (Karasek 1979). Og den toksikologiske model var allerede i 1930'erne blevet koblet med begrebet arbejdsmiljø i betydningen det fysiske rum, hvori arbejdet udføres,⁹ ligesom begrebet genfindes i toksikologiske artikler om grænseværdier for luftbårne forureninger i 1950'erne og 1960'erne (f.eks. Roach 1954, Valic & Jacobs 1965, jf. også May R. Mayers lærebog for arbejdsmiljø-

professionelle fra 1969 (Mayers 1969)). I 1960'erne spredtes begrebet arbejdsmiljø til den systemteoretiske ergonomiske produktivetsforskning, jf. f.eks. Murrell (1965), ligesom der forekom mindre serier af artikler om f.eks. betydningen af 'office automation' for 'work environment and employee satisfaction' (Hardin 1960, Mann & Williams 1962) og om betydningen af 'the industrial work environment' for mental sundhed (sammenfattet i French & Kahn 1962)¹⁰.

Mens disse fem betragtninger synes at pege på faktorer, som kan bidrage til forståelsen af, at der på forskningsfeltet 'Psykisk arbejdsmiljø og helbred' i dag findes en konstruktivistisk målepraksis, hvis resultater kommunikerer, som om de var fremkommet på baggrund af en repræsentationalistisk målepraksis, er de langt fra at kunne levere en sammenhængende historisk beretning om udviklingen. Så her er, om ikke andet, en interessant videnskabshistorisk forskningsopgave.

Tak til Jørgen Burchardt, Adda Hilden, Jørgen Hilden, Roger Persson, Ole Henning Sørensen og de anonyme bedømmere for gode kommentarer og anden hjælp i forbindelse med denne artikel.

NOTER

- 1 Efter den amerikanske organisationspsykolog Rensis Likert, som opfandt skaleringsmetoden (Likert 1932).
- 2 Der er forfattere, som betegner selve det at oversætte standardiserede svarværdier til talværdier som likertskalering – et eksempel herpå findes således også blandt de 19 hyppigst citerede artikler fra forskningsfeltet Psykisk arbejdsmiljø og helbred, nemlig Li, Gignac & Anis 2006.
- 3 Søgning i ISI Web of Science per 3.6.09 med søgeordene psychosocial AND (health OR disease OR injury) AND (work* OR occupation*).
- 4 De detaljerede resultater fra denne del af undersøgelsen findes i et bilag til artiklen på nettet: www.nyt-om-arbejdsliv.dk/tfa09-4.asp
- 5 Teorierne har andre navne på engelsk: Mitchell kalder sin teori for 'den klassiske', Campbell sin teori om 'fundamental mea-

- surement', Stevens' omtales ofte(st) som 'operational measurement' og Luce & Tukey betegner selv deres model som 'conjoint measurement'. Jeg har valgt karakteriserende betegnelser til teorierne, som efter min mening bedre belyser forskellene mellem dem.
- 6 Man kalder ordningsrelationen $>$ ("større end") for stærk, og ordningsrelationen \geq ("større end eller lig med") for svag.
 - 7 Spørgsmålene 4, 20 og 31 bruges til at konstruere den korte version af 'Udviklingsmuligheder' (hvorfor ikke 'udviklingsbelastninger?') mens 5, 13 og 32 bruges til den korte version af 'Mening i arbejdet'. Spørgsmål 33 hører *ikke* til 'Mening i arbejdet' (hvorfor ikke?), men til skalaen 'Rollekonflikt'.
 - 8 I Danmark oprettes Ministeriet for Forureningsbekæmpelse i 1971 og omdøbes til Miljøministeriet i 1973.
 - 9 Sammenlign således Grays 1937-artikel om 'working environment' med Newmans fra 1931, hvor 'work environment' er de boligområder, som fabriksarbejdere er henvist til at bo i (Gray 1937, Newman 1931).
 - 10 Mens udtrykket 'psykisk arbejdsmiljø' har bidt sig fast i Norden forekommer det ikke som en fast forbindelse i de af de hyppigst citerede forskningsrapporter fra feltet i årene 2006-2009, som stammer fra angelsaksiske lande, ligesom 'work environment' synes at være blevet erstattet med 'job characteristics'.

REFERENCER

- Adams, Ernest W. (1966): On the nature and purpose of measurement, i *Synthese*, 16, 125-169.
- Belloc, Nedra B. & Max G. Arellano (1973): Computer record linkage on a survey population, i *Health Services Reports*, 88 (4), 344-350.
- Boyd, Richard (2002): Scientific Realism, *Stanford Encyclopedia of Philosophy* (available at <http://plato.stanford.edu/entries/scientific-realism>).
- Bridgman, Percy W. (1960/1927): *The Logic of Modern Physics*, New York, MacMillan.
- Campbell, Norman R. (1920): *Physics. The Elements*, Cambridge, Cambridge University Press.
- Chow, Siu L. (1998): In defence of the construct validity approach, i *Theory and Psychology*, 8(4), 481-487.
- Christensen, Karl Bang (2002): Testing unidimensionality in polytomous Rasch models, i *Psychometrika*, 67 (4), 563-574.
- Cronbach, Lee J. (1951): Coefficient Alpha and the Internal Structure of Tests, i *Psychometrika*, 16 (3), 297-334.
- Dingle, Herbert (1950): A Theory of Measurement, i *The British Journal for the Philosophy of Science*, 1(1), 5-26.
- Frese, Michael & Dieter Zapf (1988): Methodological issues in the study of work stress: Objective vs. subjective measurement of work stress and the question of longitudinal studies, i C.L.Cooper & R. Payne: *Causes, Coping and Consequences of Stress at Work*, Chichester, Wiley, 375-411.
- Gaito, John (1980): Measurement scales and statistics: Resurgence of an old misconception, i *Psychological Bulletin*, 87(3), 564-567.
- Göb, Rainer, Christopher McCollin & Maria Fernanda Ramalhoto (2007): Ordinal methodology in the analysis of Likert scales, i *Quality & Quantity*, 41, 601-626.
- Gray, Albert S. (1937): Working Environment, i *American Journal of Public Health*, 28 (4), 378-380.
- Hackman, J. Richard & Edward E. Lawler III (1971): Employee reactions to job characteristics, i *Journal of Applied Psychology Monograph*, 55 (3), 259-286.
- Hackman, J. Richard & Greg R. Oldham (1975): Development of the Job Diagnostic Survey, i *Journal of Applied Psychology*, 60 (2), 159-170.
- Hand, D.J. (1996): Statistics and the Theory of Measurement, i *Journal of the Royal Statistical Society, A*, 159, Part 3, 445-492.
- Hardin, Einar (1960): Computer automation, work environment, and employee satisfac-

- tion: A Case Study, i *Industrial and Labor Relations Review*, 13 (4), 559-567.
- Hurrell, Joseph J., Debra L. Nelson & Bret L. Simmons (1998): Measuring job stressors and strains: Where we have been, where we are, and where we need to go, i *Journal of Occupational Health Psychology*, 3 (4), 368-389.
- Johnson, Jeffrey V. & Ellen M. Hall (1988): Job Strain, Work Place Social Support, and Cardiovascular Disease, i *American Journal of Public Health*, 78 (19), 1336-1342.
- Karasek, Robert A. (1979): Job demands, job decision latitude, and mental strain: Implications for job redesign, i *Administrative Science Quarterly*, 24, 285-308.
- Kasl, Stanislav V. (1998): Measuring Job Stressors and Studying the Health Impact of the Work Environment, i *Journal of Occupational Health Psychology*, 3 (4), 390-401.
- Kyngdom, Andrew (2008): The Rasch Model from the perspective of Representational Theory of Measurement, i *Theory and Psychology*, 18(1), 89-109.
- Li, Xin, Monique A.M. Gignac & Aslam H. Anis (2006): The indirect costs of arthritis resulting from unemployment, reduced performance, and occupational changes while at work, i *Medical Care*, 44(4), 304-310.
- Likert, Rensis (1932): *A Technique for the Measurement of Attitudes*, New York, Archives of Psychology (R.S.Woodworth, editor), no. 140.
- Lord, Frederic M. (1953): On the statistical treatment of football numbers, i *The American Psychologist*, 8, 750-751.
- Luce, R. Duncan & John W. Tukey (1964): Simultaneous conjoint measurement: A new type of fundamental measurement, i *Journal of Mathematical Psychology*, 1(1), 1-27.
- Maraun, Michael D. (1998): Measurement as a normative practice, i *Theory and Psychology*, 8(4), 435-461.
- Mayers, May R. (1969): *Occupational Health. Hazards of the Work Environment*, Baltimore, The Williams and Wilkins Company.
- Michell, Joel (1986): Measurement scales and statistics: A clash of paradigms, i *Psychological Bulletin*, 100(3), 398-407.
- Michell, Joel (1990): *An introduction to the logic of psychological measurement*, Hillsdale, N.J., Lawrence Erlbaum.
- Michell, Joel (1999): *Measurement in Psychology. A Critical History of a Methodological Concept*, Cambridge, Cambridge University Press.
- Murrell, K.F.H. (1965): *Ergonomics. Man in His Working Environment*, London, Chapman & Hall.
- Newman, Bernard J. (1931): Work Environment as a Factor in the General Health of Workers. i *American Journal of Public Health*, 21 (12), 1354-1359.
- Roach, S.A. (1954): Measuring the Workers' Environment, i *British Journal of Industrial Medicine*, 11 (1), 54-57.
- Sale, Joanna E.M. & Michael S. Kerr (2002): The psychometric properties of Karasek's demand and control scales within a single sector: data from a large teaching hospital, i *International Archives of Occupational and Environmental Health*, 75, 145-152.
- Sanderson, Kristy & Gavin Andrews (2006): Common mental disorders in the workforce: Recent findings from descriptive and social epidemiology, i *Canadian Journal of Psychiatry*, 51(2), 63-75.
- Siegrist, Johannes m.fl. (2004): The measurement of effort-reward imbalance at work: European comparisons, i *Social Science & Medicine*, 58, 1483-1499.
- Stevens, S.S. (1935a): The operational basis of psychology, i *American Journal of Psychology*, 47(2), 323-330.
- Stevens, S.S. (1935b): The operational definition of psychological concepts, i *Psychological Review*, 42(6), 517-527.
- Stevens, S.S. (1946): On the theory of scales of measurement, i *Science*, 103(2684), 677-680.
- Stine, Wm Wren (1989): Meaningful Inference: The Role of Measurement in Statistics, i *Psychological Bulletin* 105(1), 147-155.
- Townsend, James T. & F. Gregory Ashby (1984): Measurement scales and statistics: The misconception misconceived, i *Psychological Bulletin*, 96(2), 294-401.
- Valic, Fedor & Morris B. Jacobs (1965): Assessment of mercury air concentrations in a work environment, i *American Industrial*

Hygiene Association Journal, 26 (3), 266-269,
1965.

Wilkinson, Carol (2000): *Fundamentals of Health
at Work*, London, Taylor & Francis.

Jacob Hilden Winsløw, dr.scient.soc., er seniorforsker på Det nationale forskningscenter for arbejdsmiljø.
e-mail: jhw@arbejdsmiljoforskning.dk