

Om strukturerade frågor och tillförlitliga svar i missbruks- och beroendevård

– en reliabilitetsstudie av intervjumetoden DOK

Mikael Dahlberg och Mats Anderberg

Strukturerade intervjuer har fått en ökad användning inom missbruks- och beroendevård. Den här artikeln beskriver en studie av intervjumetoden DOK, där tillförlitligheten avseende självrapporterade uppgifter analyseras. Några metodologiska frågor lyfts fram och diskuteras.

Mikael Dahlberg & Mats Anderberg är doktorander i socialt arbete på Växjö universitet vid Institutionen för vårdvetenskap och socialt arbete (IVOSA).

Kontakt: Korrespondens till Mats Anderberg, IVOSA, Växjö universitet, 351 95 Växjö. Tel. 0470-708000, Mobil: 0730-687046, Fax. 0470-36310
E-post: mats.anderberg@vxu.se

Introduktion

Ett ökande intresse för standardiserade frågeformulär inom olika vårdområden har kunnat ses det senaste decenniet men förhållandevis liten uppmärksamhet har ägnats i Sverige åt studier kring den metodologiska kvaliteten. Statliga myndigheter har pläderat för ett ökat användande av systematiska intervjumetoder för förbättrad dokumentation och har producerat flera skrifter på temat. Inte minst missbrukarvården har varit föremål för denna utveckling, men även barn- och ungdomsvården samt äldre/handikappomsorgen har omfattats.

Huvudsyftet har varit att strukturerade intervjuer med klienter ska utgöra en bättre grund för bedömningar och beslut om olika interventioner. Ett annat syfte är att den samlade dokumentationen också kan ligga till grund för utvärderings- och forskningsaktiviteter samt för nationella och internationella jämförelser. I de nyligen publicerade Nationella riktlinjerna för missbruks- och beroendevård påtalas bl.a. att det är av stor vikt att de arbetsätt och metoder som används av praktiker och forskare inom området är av god kvalitet och uppfyller veten-

skapliga krav. Ett kapitel i riktlinjerna avhandlar särskilt kvalitetsaspekter strukturerade intervjumetoder, som t.ex. reliabilitet och validitet (Alling m.fl. 2006; Socialstyrelsen 2007).

Under år 1994 inleddes ett projekt¹ inom missbruks- och beroendevården i Sverige – DOK. Utvecklingsarbetet, som påbörjades 1991, var inspirerat av andra dokumentationssystem i Sverige, övriga Europa och USA. Syftet var att initiera kvalitets- och metodutveckling på en vetenskaplig grund. Under de senaste åren har också DOK-systemet bidragit till sammanställning av nationell statistik om klienter i missbrukarvård (KIM)² som i sin tur vidarebefordras till EU (Jenner & Segraeus 2005).

Tillförlitligheten i självrapporterad droganvändning, en aspekt som är föremål för diskussion, kan variera betydligt mellan olika populationer och situationer. I en amerikansk forskningsöversikt konstateras t.ex. att yngre klienter med missbruksproblem ofta lämnar osäkra uppgifter kring sitt droganvändande, särskilt kring droger som anses tyngre och särskilt stigmatiserande, t.ex. crack och opiater. Tilltron i svar kring droganvändande har även konstaterats vara högre i inledning av vården och försämras därefter i samband med uppföljning, troligen för att man vill förmedla en mer positiv bild av sig själv och sin utveckling. Tillförlitligheten styrs således också av i vilket sammanhang undersökningspersonerna befinner sig, t.ex. om de väntar på rättegång eller när det kan finnas oro för att lämnade uppgifter kan få negativa konsekvenser för dem, men även om klienten är abstinent el-

ler drogpåverkad vid intervjutillfället (Harrison 1997).

Många av intervjumetoderna som används inom missbruksområdet har importerats från andra länder (främst USA) men det finns också exempel på intervjumetoder som skapats i Sverige, t.ex. DOK och DUDIT (Drug Use Disorders Identification Test). Flera av de värden som angivits som mått på metodologisk kvalitet har visat på god uppfyllelse, när det gäller såväl reliabilitet som validitet för olika typer av intervjumetoder. Många gånger anges dock enstaka värden som indikatorer på kvaliteten för ett helt instrument, i andra fall saknas uppgifter om instrumentens validitet och reliabilitet (Alling m.fl. 2006). Användare ifrågasätter sällan intervjumetodernas kvalitet och utgår ofta från att de värden som presenteras när det gäller tillförlitligheten är tillräckliga och korrekta.

Men det finns också studier som visar på att intervjumetoders tillförlitlighet inte alltid är tillfredsställande. I en forskningsöversikt, som omfattar totalt 37 olika studier av validitet och reliabilitet för det vanligast förekommande instrumentet ASI (Addiction Severity Index), konstateras stor variation avseende korrelation inom vissa frågeområden t.ex. kring arbete, droger, familj, social situation och psykisk hälsa. Dessa resultat har påvisats när instrumentet använts för specifika missbrukspopulationer, t.ex. klienter som har psykiska sjukdomar eller som är hemlösa (Mäkelä 2004). Andra aspekter som lyfts fram i det här sammanhanget är att de statistiska metoderna för beräkning av tillförlitlighet, när det gäller skalor och bedömnings-

instrument, kan vara problematiska (Altman 1991; Svensson 2001).

Inom ramen för ett pågående forskningsprojekt, med DOK som exempel, studeras på vilket sätt dokumentationssystem och strukturerade intervjuer kan bidra till kunskapsutveckling inom missbruks- och beroendevård. Här granskas bl.a. kvaliteten i DOK beträffande reliabilitet och validitet. I två tidigare studier har interbedömarreliabilitet och validitet undersökts (Anderberg & Dahlberg 2007; Dahlberg & Anderberg submitted).

Syftet med denna artikel är att beskriva och värdera reliabiliteten, avseende stabiliteten i uppgifter inhämtade med DOK:s inskrivningsintervju, vilket har skett genom analys beträffande graden av överensstämmelse mellan två olika intervjutillfällen. Stabilitet är en aspekt av reliabilitet och avser reproducerbarheten av det fenomen som mäts, d.v.s. i vilken grad de självrapporterade uppgifterna är stabila över tid och huruvida DOK kan mäta dessa på ett tillförlitligt sätt. Därutöver presenteras och diskuteras olika statistiska metoder för prövning av reliabilitet.

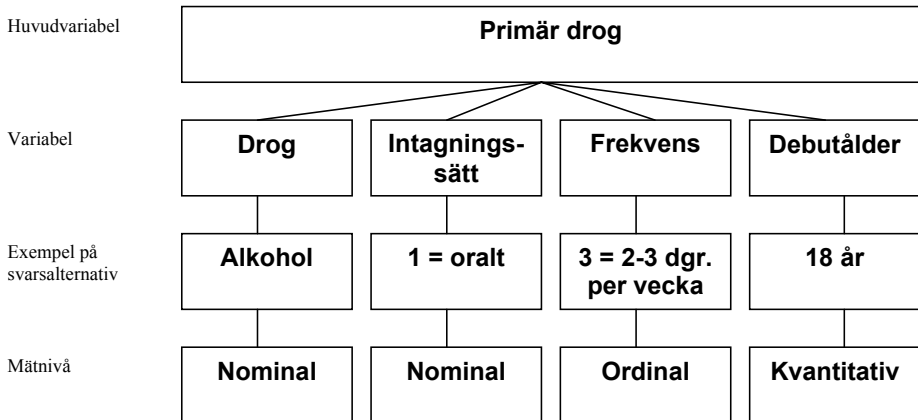
Metod

I DOK intervjuas klienten vid *inskrivning*, *avstämning* (under behandling), *utskrivning* och vid *uppföljning* efter avslutad behandling. Det finns även en kortversion av intervjun som är avsedd för en slags basdokumentation (BasDOK) av vårdtillfället och som även kan utföras utan intervju genom att redan befintliga uppgifter dokumenteras, t.ex. via aktmaterial. En lång

DOK-intervju tar ca en timme i anspråk och frågeformulärens innehåll bygger på en struktur med kartläggning av olika livsområden: uppväxt, social situation, familj och umgänge, psykisk och fysisk hälsa, kriminalitet samt användning av alkohol och droger (Jenner & Segraeus 2005). En huvudsaklig variabel för att dokumentera alkohol- och droganvändning är primär drog, dvs. det preparat som orsakar individen mest problem eller utgör anledning till behandling. Variabeln är i sin tur, uppdelad på fyra andra variabler med olika former av svarsalternativ med varierande mätintervaller (se figur 1).

Design och urval

Studiens syfte var att pröva reliabiliteten avseende stabilitet. Stabilitet är en aspekt av reliabilitet och avser reproducerbarheten av det fenomen som mäts, d.v.s. i vilken grad de självrapporterade uppgifterna är stabila över tid och huruvida DOK kan mäta dessa på ett tillförlitligt sätt (AERA, APA & NCME 1999). När det gäller val av variabler har endast de frågor som härrör från BasDOK:s inskrivningsformulär³ inkluderats. BasDOK är den grundläggande nivån på dokumentation som krävs för deltagande i DOK och innehåller flera centrala frågor. I studien analyseras 28 av totalt 82 ingående variabler i BasDOK. Variabler som innehåller aktuella uppgifter vilka sannolikt förändras över tid, t.ex. boendeform de senaste 30 dagarna och administrativa uppgifter har tagits bort. De operationellt definierade variablerna har delats in i två kategorier – *kvantitativa* och *kva-*



Figur 1. Exempel på statistisk/matematisk operationalisering, utvecklad av Svensson (2005).

*litativa*⁴. Uppdelningen i nämnda två kategorier ger vägledning om val av statistiska metoder för beräkning och analys (se även figur 1).

Genomförda DOK-intervjuer överförs till en gemensam databas via de behandlingsenheter som deltar i systemet. Undersökningens urval baseras på samtliga 2.271 påbörjade vårdtillfällen (2.112 personer) under år 2003 där inskrivningsintervju utfördes med hjälp av DOK. En undersökningsgrupp har valts ut med kriteriet att de ska ha vårdats för missbruksproblem vid minst två tillfällen under en ettårsperiod. Det fanns 136 personer som uppfyllde urvalskriteriet. Av det ursprungliga antalet föll 53 individer bort p.g.a. att inskrivningsuppgifterna vid något tillfälle baserats på aktmaterial, dvs. att dokumentationen i sig enbart baseras på redan befintliga uppgifter och inte på en intervju. Ytterligare 5 personer exkluderades då uppgift om insamlingsmetod saknades respektive otillräckligt ifyllda formulär. Undersökningsgruppen omfattar således 78 personer. I de fall

där fler än två intervjuer gjorts valdes de intervjutillfällen som låg närmast varandra i tid. Den genomsnittliga tiden mellan intervjuerna är 112 dagar (2-322 dagar). Totalt har 25 av de 78 individerna intervjuats av samma intervjuare vid båda tillfällena, resterande 53 av olika.

Statistisk analys

I ett andra steg utfördes statistisk bearbetning och analys via SPSS, differenser och överensstämmelser av varje variabel beräknades för de båda mättillfällena. Som framgår av figur 1 består datamaterialet från intervjuformulärets svarsalternativ av olika mättnivåer, vilket leder till val av relevant statistisk metod. För de kvantitativa variablerna har vi beräknat procentuell överensstämmelse (Percentage Agreement, PA), medelvärden för de båda intervjutillfällena och differensmedelvärden. De kvalitativa variablerna består av nominal eller dikotom data vilket medger beräkningar av procentuell överensstämmelse och i vissa fall Cohen's kappa (K). En pro-

centuell överensstämmelse på ca 70 % betraktas vanligtvis som acceptabel (Bordens & Abbott, 2005). Metoden kappa anses justera för slumpmässig enighet mellan mättillfällena, men är emellertid inte lämplig för dikotoma data eftersom kappavärdet ökar då antalet alternativ minskar, vilket gör att värdet med automatik blir högt. Vid bedömning av kappavärden används ofta nedanstående tabell (Altman 1991).

Tabell 1. Kappa-värden, enligt Landis & Koch, 1977.

Value of K	Strength of agreement
< 0,20	Poor
0,21 - 0,40	Fair
0,41 - 0,60	Moderate
0,61 - 0,80	Good
0,81 - 1,00	Very good

För den enda variabel som innehåller ordinaldata (utbildningsnivå) har vi förutom PA och kappa använt en statistisk metod framtagen av Svensson (2005), som även identifierar eventuell systematisk eller tillfällig oenighet mellan de båda intervjutillfällena.

Resultat

Här följer en redovisning av data om undersökningsgruppen och resultat när det gäller överensstämmelser och skillnader från de två mättillfällena. Nedanstående tabell (tabell 2) beskriver könsfördelning, medelålder, vårdform och primär drog vid första behandlingstillfället för undersökningsgruppen i relation till samtliga nyinskrivna klienter år 2003 hämtade ur DOK:s databas.

Tabell 2. Undersökningsgruppen och vårdtillfällena 2003 (nyinskrivna) med avseende på antal, kön, ålder, vårdform och primär drog.

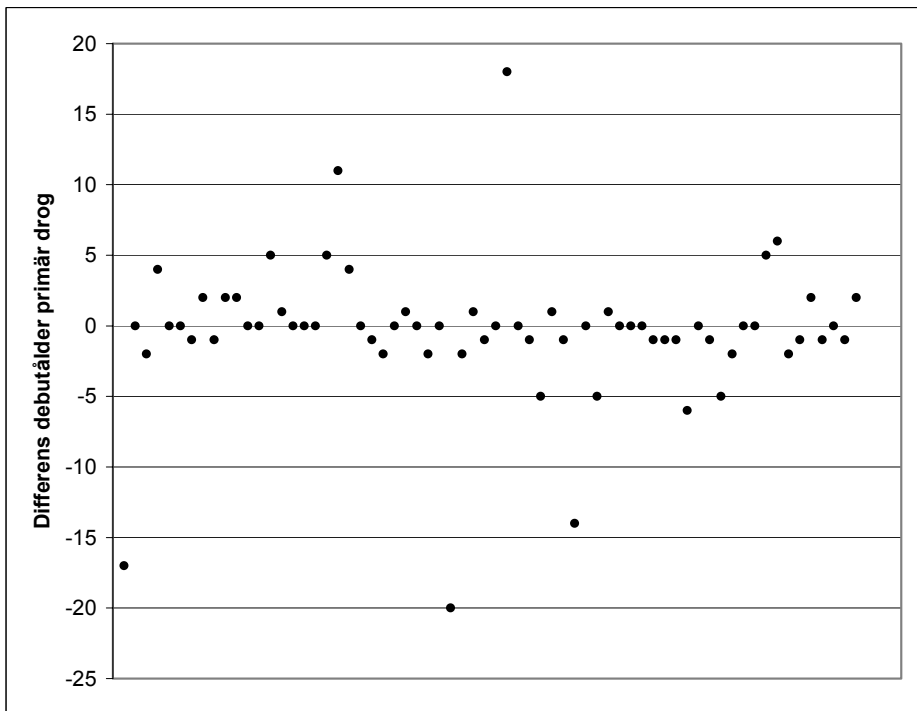
	Undersökningsgruppen	Vårdtillfällena 2003
Antal:	78	2271
Könsfördelning (%):		
Kvinnor	24	29
Män	76	71
Medelålder (år och SD):	38 (13)	37 (11)
Vårdform (%):		
Öppenvård	32	41
Slutenvård	68	59
Primär drog (%):		
Alkohol	46	49
Narkotika	53	48
Läkemedel	1	3

När det gäller de kvantitativa variablerna åskådliggör tabell 3 överensstämmelsen mellan de båda intervjuerna. Variablerna har en procentuell överensstämmelse (PA) mellan 0,95 – 1,00 med undantag av debutålder för primär drog som har PA 0,33.

Differensmedelvärde är ett genomsnittsmått för skillnader mellan behandlingstillfällena för varje par, dvs. differensmedelvärde och standardavvikelse (SD) beskriver variationen för samtliga värden.

Tabell 3. Kvantitativa variabler från intervju 1 och 2 där graden av överensstämmelse mellan två bedömningar beskrivs med procentuell överensstämmelse, medelvärde och differensmedelvärde. KIM-frågor är markerade med *.

	n	PA	M 1 (SD)	Differensmedelvärde (SD)
Antal egna barn under 18 år	77	1,00	0,57 (0,91)	0
Antal egna barn	77	0,99	1,06 (1,18)	0,03 (0,23)
Antal barn med egen/gemensam vårdnad	76	0,95	0,54 (0,97)	-0,09 (0,44)
Debutålder för primär drog*	66	0,33	18,50 (6,75)	-0,36 (5,02)



Figur 2. Differensen mellan intervju 1 och 2 för varje person med avseende på debutålder för primär drog (n=66).

För att närmare åskådliggöra hur skillnaderna fördelar sig när det gäller rapporterad debutålder för primär drog redovisar vi även variabeln med ett punktdiagram (se figur 2). Varje punkt representeras av differensen mellan de båda intervjuerna för 66 individer. Merparten av värdena fördelar sig på eller nära 0-linjen, d.v.s. ingen eller liten avvikelse. För uppgiften kan noteras ett internt bortfall om 12 personer där frågan vid något av intervjutillfällena inte besvarats.

Tabell 4 beskriver överensstämmelsen för de kvalitativa variablerna med de statistiska måtten Percentage agreement (PA) och kappa (K) och redovisas, precis som för de kvantitativa variablerna, i en fallande skala. Överensstämmelsen enligt PA varierar mellan 0,66 – 0,99. När det gäller överensstämmelse enligt K uppnår variabeln medborgarskap det högsta värdet (0,93) och variabeln initiativtagare till behandling det lägsta värdet (0,33).

Tabell 4. Kvalitativa variabler från intervju 1 och 2 där graden av överensstämmelse beskrivs med måtten percentage agreement och kappa. KIM-frågor är markerade med *.

	n	PA	K
Medborgarskap*	78	0,99	0,93
Dömd till rättspsykiatrisk vård	77	0,97	
HIV-testad	70	0,96	
Dömd till fängelse	77	0,95	
Behandling enligt lagrum	78	0,92	0,88
Tidigare vård för narkotika-, alkohol- och eller läkemedelsmissbruk*	74	0,91	
Intagningsätt för primär drog*	71	0,88	0,78
Gravid (om kvinna)	15	0,87	
Kontakt med kriminalvård senaste 6 månaderna	77	0,87	
Kontakt med socialtjänst senaste 6 månaderna	77	0,86	
Primär drog*	74	0,85	0,77
Dömd för brott	76	0,84	
Utbildningsnivå*	76	0,83	0,73
Injektionsmissbruk*	74	0,81	0,68
Kontakt med arbetsförmedling senaste 6 månaderna	77	0,79	
Problem med fysisk hälsa	77	0,77	
Försörjning senaste 6 månaderna*	77	0,73	0,66
Boendeform senaste 6 månaderna*	78	0,73	0,60
Initiativtagare till behandling*	78	0,71	0,33
Kontakt med försäkringskassa senaste 6 månaderna	77	0,70	
Kontakt med kronofogde senaste 6 månaderna	77	0,69	
Levnadsform senaste 6 månaderna*	77	0,66	0,50
Kontakt med sjukvård senaste 6 månaderna	77	0,66	

Den enda variabeln med ordinaldata – utbildningsnivå – återges också med en kontingenstabell för att beskriva hur stora skillnaderna mellan mättillfällena är och var i skalan de återfinns. Nedanstående kontingenstabell beskriver marginalfördelningen mellan mättillfällena för variabeln utbildningsnivå. I diagonalen (se tabell 7) återfinns de värden som uppvisar total enighet, i detta fall är PA 0,83. Marginalfördelningarna skiljer sig mycket lite, vilket betyder att det är en nästan försumbar systematisk instabilitet mellan de båda bedömningarna.

Tabell 5. Variabeln utbildningsnivå och överensstämmelse mellan intervju 1 och 2. Svarsalternativen är 1=Ej fullföljd folkskola, grundskola eller motsvarande, 2=Fullföljd folkskola, grundskola eller motsvarande, 3=Fullföljd gymnasium, fackskola eller motsvarande samt 4=Fullföljd eftergymnasial utbildning, högskola.

		Intervju 1				
		1	2	3	4	
Intervju 2	4		1	2	6	9
	3			27	4	31
	2	1	29	1		31
	1	1	3	1		5
		2	33	31	10	76

Diskussion

Inledningsvis diskuteras studiens resultat och därefter följer ett avsnitt om implikationer vad gäller val av statistisk metod vid studier av strukturerade intervjuers tillförlitlighet.

Som framgår av tabell 3 påvisas överlag en hög grad av stabilitet beträffande de *kvantitativa* variablerna. Överensstämmelsen är mycket hög när det gäller de tre variablerna om

barn. För uppgiften om debutålder för primär drog är graden av procentuell överensstämmelse (PA) svagast, men trots det ligger ålderuppgifterna för de båda intervjuerna nära varandra, skillnaden beskrivs tydligast i figur 2. Resultatet belyser därigenom svårigheten att efterfråga en uppgift som debutålder, där svaret ges i exakt ålder t.ex. 17 år jämfört med ordnade kategorier där åldern istället anges i t.ex. 10-15 år, 15-20 år etc. Detta faktum illustrerar betydelsen av vilken mätnivå som används när uppgifter samlas in, respondenterna svarar utifrån bästa förmåga men kommer kanske inte alltid ihåg om debutåldern är 15 eller 16 år.

För 20 av 24 *kvantitativa* variabler (tabell 3) är den procentuella överensstämmelsen 0,70 eller högre, dvs. minst 70 % av svaren är identiska vid intervju 1 och 2, vilket är en gräns för acceptabel reliabilitet som använts vid tidigare studier (Bordens & Abbott 2005). För de resterande 4 variablerna varierar PA mellan 0,64 och 0,69. När det gäller överensstämmelse enligt kappa är bedömningen följande: *very good* (se tabell 1) för variablerna medborgarskap och behandling enligt lagrum, *good* för intagningsätt för primär drog, primär drog, utbildningsnivå, injektionsmissbruk och försörjning senaste 6 månaderna, *moderate* för boendeform senaste 6 månaderna, levnadsform senaste 6 månaderna och sysselsättning senaste 6 månaderna samt *fair* för initiativtagare till behandling.

De frågor som är relativt stabila (t.ex. antal egna barn, dömd till fängelse, primär drog) uppvisar en hög

grad av överensstämmelse mellan intervjutillfällena. När det gäller variabler som är mer känsliga för förändringar kan större differenser överlag skönjas (t.ex. kontakt med sjukvård senaste 6 månaderna, sysselsättning 6 månader). Av det följer att resultaten därmed indikerar en hög stabilitet avseende de självrapporterade uppgifterna. I annat fall borde skillnader i överensstämmelse ha varit mer jämnt fördelade mellan stabila och mer instabila variabler. Eftersom 11 av de variabler som vi undersökt även finns representerade i KIM, innebär det även en viss prövning av de centrala uppgifter som lämnas för nationell statistik kring klienter i missbrukarvård.

Det är svårt att värdera resultaten i relation till andra liknande undersökningar dels då reliabilitetsstudier av den här typen är studiespecifika, dels då de statistiska metoder som används uppvisar en stor variation. Det finns några problem som är förenade med det traditionella sättet att studera tillförlitlighet hos standardiserade intervjuformulär.

Ett problem är att poängsummeringar från flera olika variabler, s.k. *sumscores*, ofta används vid beräkning av reliabilitet. Det diskutabla med sammanfattande mått är att vissa av de variabler som ingår i sådana sammanvägningar endast är relevanta för vissa typer av missbruksproblem. Följden blir att irrelevanta frågor kan få oproportionerligt stor vikt vid poängsummering – eller omvänt – relevanta frågor får för låg vikt (Altman 1991; Jansson 2001; Svensson 2005).

Ett annat problem är att vissa av de

ingående värdena i poängsummeringarna är av kvantitativ art medan andra utgörs av skattningar och personliga bedömningar, alltså ordinaldata. Mäkelä (2004) poängterar också att det saknas relevanta studier om strukturerade intervjuer där metodens tillförlitlighet granskas på variabelnivå.

Ytterligare ett problem är att forskare i reliabilitetsstudier av den här typen ofta använder sig av korrelationskoefficient vid jämförelse mellan mätresultat, vilket är diskutabelt då det mäter graden av samvariation och inte – som är mer relevant för reliabiliteten – graden av överensstämmelse. En korrelationskoefficient kan vara hög trots att överensstämmelsen är låg (Altman 1991).

Sammantaget uppvisar de undersökta variablerna i DOK:s inskrivningsformulär en övervägande god stabilitet, den procentuella överensstämmelsen (PA) för 23 av 28 studerade variabler varierar mellan 0,70 och 1,00. Överensstämmelsen kan tillskrivas de självrapporterade uppgifterna och intervju metodens tillförlitlighet, men även till viss del vara ett uttryck för undersökningsgruppens relativt oförändrade livssituation.

Referenser

- AERA (American Education Research Association), APA (American Psychological Association) & NCME (The National Council on Measurement in Education). (1999). *The Standards of Educational and Psychological Testing*.
- Alling, C., Beck, O., Bergman, H., Helander, A., Jenner, H., Nyström, S. & Wennberg, P. (2006). *Markörer och diagnostiska test. I: Faktaunderlag till Nationella riktlinjer för missbruks- och beroendevård*. Stockholm, Socialstyrelsen.
- Altman, D.G. (1991). *Practical statistics for medical research*. Florida, Chapman and Hall/CRC.
- Anderberg, M. & Dahlberg, M. (2007). Interbedömrareliabilitet – ett tillförlitligt mått på standardiserade intervjuer? En studie av DOK-intervjun. *Nordisk alkohol- och narkotikatidskrift* (1): 45-58.
- Bordens, K.S. & Abbott, B.B. (2005). *Research, design and methods: a process approach*. 6th edition. New York, McGraw Hill Higher Education.
- Dahlberg, M. & Anderberg, M. (submitted). You get what you ask for – a study of the validity of the DOK interview.
- Harrison, L. (1997). *The Validity of Self-Reported Drug Use in Survey Research: An Overview and Critique of Research Methods*. Rockville, NIDA (National Institute on Drug Abuse).
- Jansson, I. (2001). *Problem i sammanfattande mått i ASI*. Stockholm, Statens institutionsstyrelse.
- Jenner, H. & Segraeus V. (2005). The Swedish DOC system - an attempt to combine documentation and self-evaluation. *European Addiction Research* 11(4): 186-92.
- Mäkelä, K. (2004). Studies of the reliability and validity of the Addiction Severity Index. *Addiction* 99: 398-410.
- Socialstyrelsen (2007). *Nationella riktlinjer för missbruks- och beroendevård*. Stockholm, Socialstyrelsen.
- Svensson, E. (2001). Guidelines to Statistical Evaluation of Data from Rating Scales and Questionnaires. *Journal of Rehabilitation Medicine* 33: 47-48.
- Svensson, E. (2005). Val och konsekvens: mätnivån avgör den statistiska verktygslådan. *Läkartidningen* Nr 17, Vol. 102.

Noter

- 1 Projektet initierades inom den frivilliga missbruks- och beroendevården av Institutet för Kunskaps- och Metodutveckling inom ungdoms- och missbrukarvården (IKM), Växjö universitet, och för tvångsvården av Statens Institutionsstyrelse (SiS). Härav följer förkortningarna IKM-DOK och SiS-DOK, formulären är i det närmaste identiska.
- 2 De 29 KIM-frågor som ingår i DOK:s inskrivningsintervju är markerade med * . Denna symbol används även i resultatredovisningen.
- 3 Den presenterade studien avser material och formulär från IKM-DOK.
- 4 Med kvantitativa data avses metriska/numeriska data, vilket motsvaras av begreppen intervall- eller kvotskala enligt Stevens klassificering av mätnivåer, övriga mätnivåer är kvalitativa. Se Svensson (2005) för utförligare resonemang.

Summary in English

About structured questions and reliable answers in abuse and dependency care – a reliability study of the interview method DOK.

This paper presents results from a study of reliability concerning stability in self-reported information obtained from a standardized interview from a documentation system in Sweden (the DOK system). The paper also presents and discusses the choice of relevant statistical methods for testing and analyzing results from this type of reliability study. The degree of agreement between two different interviews for a total of 78 individuals from the database of DOK has been calculated and analyzed. The study shows a high level of stability for the self-reported information. Satisfactory or good reliability was achieved for 27 of the 28 studied variables. Another implication is that it is important to use relevant statistical methods when identifying strengths and weaknesses in this type of instruments and in order to make reliable conclusions from the results.

Key words: structured interview, reliability, stability, DOK System, substance abuse, documentation