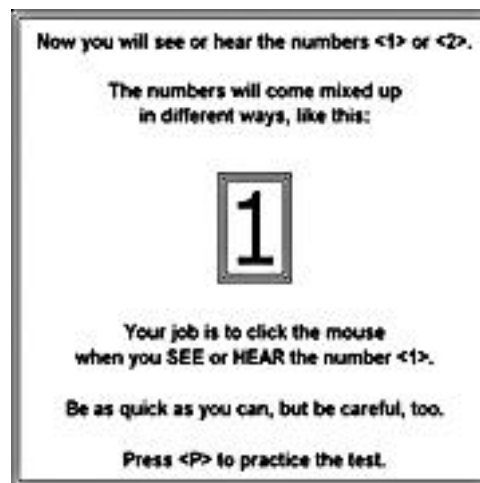


Test av ett datoriserat uppmärksamhetstest: Påverkar språket resultaten?

En jämförelse av amerikansk originalversion och svensk version av IVA+



Sara Thomée, med dr, leg psykolog
Adnan Noor, statistiker
Christina Ahlstrand, BSc, leg BMA
Linda Åhlström, MSc, leg ssk
Mats Hagberg, professor, överläkare

Arbets- och miljömedicin, Sahlgrenska Universitetssjukhuset

Göteborg, 2014-11-19

ISBN 978-91-7876-149-4

Förord

Standardiserad administration och testledarutbildning är led i att öka tillförlitligheten vid användning av psykologiska test i klinik och forskning. Datoriserade test kan ytterligare minska den variation i resultat som beror på den mänskliga faktorn, relationen mellan testperson och testledare, miljöfaktorer eller andra svårkontrollerade förhållanden. En nackdel är att man istället är beroende av de tekniska förutsättningarna och kan få mindre insyn i beräknings- och bedömningsgrunder.

Denna studie genomfördes för att jämföra en svensk anpassning av det datoriserade uppmärksamhetstestet IVA+ med den amerikanska originalversionen. Detta efter misstanke om att ett systematiskt fel smugit sig in i resultaten i samband med datainsamling för forskningsprojektet *GODArbetsförmåga* vid forskargruppen Arbetsförmåga och ergonomi, AMM Göteborg.

Studien gjordes i anslutning till den s k PiP-studien, en test-reteststudie av Purdue Pegboard, som genomfördes bland personal och anhöriga på AMM under hösten 2010 och våren 2011. Studien finansierades av FORTE.

Sara Thomée
Leg psykolog, Med dr
Arbets- och miljömedicin
Sahlgrenska universitetssjukhuset/
Göteborgs universitet

Innehåll

<i>Bakgrund och syfte</i>	1
IVA+	1
IVA+ validitet och reliabilitet	2
Användning av IVA+ i GODA-projektet.....	2
Preliminära resultat (n=38) i GODA visar tecken på snedfördelning	3
Syfte	4
<i>Material och metoder</i>	5
Studiedesign	5
Undersökningsgrupp	5
Genomförande.....	5
Undersökningsvariabler	6
Analys	6
<i>Resultat</i>	7
Undersökningsgruppen	7
Skillnader mellan originalversion och svensk version av IVA+	7
Skillnader mellan auditiva och visuella delskalor i båda språkversionerna	8
<i>Diskussion</i>	10
Konklusion	11
<i>Referenser</i>	12

Sammanfattning

IVA+Plus Integrated Visual and Auditory Continuous Performance Test (IVA+) är ett amerikanskt datoriserat test som mäter auditiv och visuell uppmärksamhet. I samband med forskningsprojektet GODA testades försökspersoner och deras kontroller med ett IVA+ som anpassats till svenska. Preliminära resultat i studien tydde på en snedfördelning mellan auditiva och visuella delskalor.

En studie genomfördes därför med syfte att jämföra den amerikanska originalversionen och den svenska anpassade versionen. Tjugosju personer i åldrarna 18-60 år testades med IVA+ två gånger med 7-20 dagars mellanrum (mv 8,08), ena gången med den amerikanska originalversionen och andra gången med den svenska versionen. Hälften började med den amerikanska och hälften med den svenska versionen.

De två språkversionerna skiljde sig åt genom att auditiv responstid var något snabbare i den svenska versionen (33 ms). Desto större skillnader framkom dock mellan auditiv och visuell responstid inom både den amerikanska och den svenska versionen av IVA+. Auditiv responstid var i båda språkversionerna mer än 1 standardavvikelse bättre än det normerade genomsnittet, medan visuell responstid låg nära förväntat värde. Skillnader framkom också i ytterligare några subskalor vilket dock svårigen kunde förklaras av en snabbare responstid. Bl a testledareffekt och selektionsbias diskuteras som tänkbara förklaringar till skillnaderna i resultat. Mest sannolikt är dock att något i tekniken påverkar resultaten, t ex genom en felklockning av stimuli och/eller respons.

Slutsatsen är att både originalversionen och den anpassade svenska versionen av IVA+ tycks ha innehållit ett systematiskt fel som ger ett bättre än förväntat resultat avseende auditiv snabbhet vid våra datainsamlingar. Detta gör att de standardiserade normerna knappast kan tillämpas på den insamlade datan. På grund av osäkerheten i vad felet består krävs stor försiktighet även vid framtida datainsamling och särskilt vid klinisk tillämpning av testresultaten. I forskningsstudier bör det vara mindre problematiskt så länge testningarna används för att jämföra grupper eller för att göra för- och eftermätningar för att t ex utvärdera en intervention, dvs där resultaten kan vara oberoende av normeringen. Det är dock troligtvis viktigt att samma hård- och mjukvara används till samtliga testningar. Eftersom resultaten avvek mer i den svenska anpassade versionen rekommenderas att den amerikanska originalversionen används om så är möjligt.

Bakgrund och syfte

IVA+

IVA+Plus Integrated Visual and Auditory Continuous Performance Test (IVA+) från amerikanska företaget Braintrain Inc, är ett datoriserat test som mäter auditiv och visuell uppmärksamhet (Sandford & Turner, 2005). IVA+ har utvecklats och validerats för att bidra till diagnosticering av ADHD samt differentiering mellan olika former av ADHD-problematik. IVA+ har också använts inom forskning av kognitiv funktion vid t ex förvärvad hjärnskada (Tinius, 2003) och utmattningssyndrom (Sandström, Rhodin, Lundberg, Olsson, & Nyberg, 2005; Österberg, Skogsliden, & Karlson, 2014). Testuppgiften är att svara eller inhibera svar på visuella och auditiva stimuli i 13 minuter, vilket kräver bibehållen uppmärksamhet respektive responskontroll. Bjudningen av IVA+ är standardiserad och utöver testledarens inledande instruktioner ges testinstruktionerna av en amerikansk digitaliserad röst. I testuppgiften finns visuella stimuli ("1" och "2") och auditiva stimuli ("one" och "two"). Som komplement till IVA+ finns också en ljudfil där det auditiva stimuli läses på en mängd olika språk, däribland på svenska ("ett" och "två").

De grundläggande mätvariablerna är responstid, korrekta svar och variabilitet i svar. Huvudskalorna Response Control och Attention är sammansatta skalor som byggs upp till lika delar av sina auditiva respektive visuella motsvarigheter (se Figur 1). Full Scale Response Control består alltså av Auditory Response Control och Visual Response Control medan Full Scale Attention består av Auditory Attention och Visual Attention. Auditory och Visual Response Control byggs sedan upp av subskalorna Prudence (ett mått på impulsivitet och förmåga till svarsinhibition), Consistency (konsistens i responser över tid), och Stamina (uthållighet). Auditory och Visual Attention byggs på samma sätt upp av subskalorna Vigilance (mäter ouppmärksamhet), Focus (variabilitet i responstider) och Speed (responstid för korrekta svar). Råpoängen som genereras i de olika skalorna omvandlas till normerade kvotpoäng; Quotients (Q), med medelvärdet 100 och standardavvikelsen 15. IVA+ består av ytterligare två grupper av skalor (Attribute och Symptomatic) som dock inte kommer att beröras här.

Figur 1. IVA+ skalor

Full scale Response Control Quotient FRCQ		Full scale Attention Quotient FAQ	
Auditory Response Control Quotient ARCQ	Visual Response Control Quotient VRCQ	Auditory Attention Quotient AAQ	Visual Attention Quotient VAQ
Prudence PRA	Prudence PRV	Vigilance VIA	Vigilance VIV
Consistency CONA	Consistency CONV	Focus FOCA	Focus FOCV
Stamina STA	Stamina STV	Speed SPDA	Speed SPDV

Testresultaten redovisas bl a i form av staplar och man kan få ut färdigskrivna utlåtanden. Testet har normerats baserat på 1700 individer utan kända neurologiska problem, medicinering, psykoterapi, kända inlärnings- eller uppmärksamhetsproblem eller hyperaktivitet. Normering är gjord för olika åldrar samt uppdelat på kön.

Vid normeringen fann man en hög korrelation mellan auditiv och visuell modalitet, dvs om man presterar bra i den ena modaliteten tenderar man att prestera bra i den andra (Sandford & Turner, 1995). Korrelationerna mellan huvudskalorna var mellan .53 och .86. Män hade en genomsnittligt snabbare responstid än kvinnorna, en skillnad om 23 millisekunder. Kvinnorna å andra sidan hade 3 % färre felaktiga impulsiva responser. Genomsnittlig responstid följde en U-formad kurva för ålder, med optimal prestation från mitten på tonåren till ung vuxen ålder, för att sedan vara ganska stabil fram till 45 års ålder då responstiden åter ökade.

Enligt manualen så resulterar IVA+ stimuli nästan alltid i långsammare rå auditiv responstid mätt i millisekunder jämfört med rå visuell responstid. Det visuella stimuli varar i 0,167 sekunder och det auditiva stimuli i 0,5 sekunder. De flesta människor tenderar således att analysera det visuella stimuli snabbare, trots att man antar att de flesta kan urskilja om det rör sig om en "ett" eller "två" innan hela ordet är uttalat. Skillnaden i responstid mellan modaliteterna är korrigerat för genom den beräknade kvotpoängen.

IVA+ validitet och reliabilitet

IVA+ har validerats för klinisk användning och diagnosticering av ADHD hos barn med 92 % sensitivitet och 90 % specificitet samt med 89 % positiv prediktionskraft och 93 % negativ prediktionskraft (Sandford & Turner, 1995).

Reliabilitetstestning har skett genom en test-reteststudie (Sandford & Turner, 1995) där 70 friska (neurologiskt, psykiatriskt mm) utförde IVA+ vid två tillfällen med 1-4 veckors mellanrum. Testpersonerna var 5-70 år, mv 21,8 år, 60 % kvinnor och 40 % män. Syftet var att bedöma test-retest reliabiliteten för 22 råskalor och 6 sammansatta skalor med hjälp av korrelationsanalyser. Alla sammansatta skalor visade statistiskt signifikanta korrelationer mellan test och retest. Responskontrollskalorna visade en måttlig korrelation (.37 till .41) medan uppmärksamhetsskalorna hade en stark korrelation (.66 till .75). Även råpoängsskalorna hade signifikant positiva korrelationer där 18 av 20 hade en måttlig till stark relation (.46 - .88). Vid re-testning förekom svagt positiva förändringar i skalorna i storleksordningen 1-3%. Rå visuell reaktionstid (MNV) hade viss inlärningseffekt (-3,25 %) och resulterade i en något förbättrad Visual Attention Quotient (+3.03 %). En förbättrad rå auditiv reaktionstid (MNA -3,88 %) märktes också. IVA+ bedömdes dock vara stabilt över tid, både globalt och för specifika skalor, med mycket liten övnings- eller inlärningseffekt, och ansågs därför användbart för att mäta interventionseffekter (Sandford and Turner 1995).

Användning av IVA+ i GODA-projektet

IVA+ användes som ett mått på mental funktion ("uppmärksamhet") i forskningsprojektet *GODArbetsförmåga - Utveckling och validering av nya metoder för att mäta arbetsförmåga och arbetsförmågeutrymme hos arbetstagare med nack-armsmäta*, 2008-2010 vid AMM Göteborg (FAS Dnr 2008-0755). Projektet syftade till att skapa ny kunskap för att behålla människor med nack- och armbesvär i arbete och förbättra deras livskvalitet genom utveckling och validering av nya metoder att bedöma arbetsförmåga. I studien testades forskningspersoner med olika metoder för att mäta olika funktioner relaterade till arbetsförmåga. Valet av IVA+ för mental funktion, baserades på att test av uppmärksamhet/visuomotorisk snabbhet har visat högst korrelation med rapporterad arbetsförmåga jämfört med andra neuropsykologiska test i en studie av Nygård et al (Nygård, Eskelinen, Suvanto, Tuomi, & Ilmarinen, 1991) samt att det är känt att

uppmärksamhet, processhastighet och visuo-motorisk snabbhet kan påverkas av smärta, (Hart, Martelli, & Zasler, 2000). Vidare korrelerar reaktionstid med generell kognitiv förmåga i form av IQ (Sandford & Turner, 1995). IVA+ hade också den fördelen att vara datoriserat och att testet använts tidigare i svensk forskning kring kognitiva funktioner vid bl a vid Arbets- och miljömedicin i Lund (Österberg et al., 2014).

Vid bjudning av IVA+ i GODA projektet inaktiverades den amerikanska digitaliserade testinstruktionsrösten och den svenska ljudfilen med auditiva stimuli ”ett” och ”två” användes. En svensk version av testinstruktionerna lästes upp av testledaren. Denna procedur har också använts av psykolog/docent Kai Österberg i forskningsstudien vid AMM Lund (personlig kommunikation K. Österberg).

Preliminära resultat (n=38) i GODA visar tecken på snedfördelning

Vid en preliminär sammanställning av testresultat tidigt i GODA-projektet (drygt en tredjedel hade undersökts) syntes med blotta ögat skillnader mellan Auditory Attention Quotient och Visual Attention Quotient, där den auditiva uppmärksamheten översteg den visuella med en differens om 15 poäng (Tabell 1). Båda skalorna förväntades genom de normerande värdena ha medelvärdet 100 och standardavvikelsen 15. Motsvarande snedfördelning syntes inte mellan Auditory och Visual Response Control (Tabell 2). Vid närmare genomgång av Attentions subskalor syntes inga större skillnader i Focus, däremot en skillnad på 11 poäng mellan auditiv och visuell Vigilance (som beräknas på antal missade responser), och en stor skillnad (25 poäng!) mellan auditiv och visuell Speed (Tabell 3). Speedskalan bygger på det normerade värdet av genomsnittlig responstid för korrekta responser. Samtliga testpersoner hade ett bättre resultat på auditiv Speed jämfört med visuell Speed (Tabell 4).

Tabell 1. Preliminära data ur GODA-studien (n=38): Attention.

Full scale Attention Quotient (FAQ)	Auditory Attention Quotient (AAQ)	Visual Attention Quotient (VAQ)	Differens AAQ-VAQ
M=101 Sd=17	M=108 Sd=13	M=93 Sd=22	M=15 Sd=14

Medelvärden (M) och standardavvikelser (Sd).

Tabell 2. Preliminära data ur GODA-studien (n=38): Response Control.

Full scale Response Control Quotient (FRCQ)	Auditory Response Control Quotient (ARCQ)	Visual Response Control Quotient (VRCQ)	Differens ARCQ-VRCQ
M=96 Sd=20	M=96 Sd=19	M=97 Sd=18	M= -1,1 Sd=16

Medelvärden (M) och standardavvikelser (Sd).

Tabell 3. Preliminära data ur GODA-studien (n=38): Attention subskalor.

Auditory Vigilance (VIA)	Visual Vigilance (VIV)	Diff VIA-VIV	Auditory Focus (FOCA)	Visual Focus (FOCV)	Diff FOCA-FOCV	Auditory Speed (SPDA)	Visual Speed (SPDV)	Diff SPDA-SPDV
M=104 Sd=6	M=94 Sd=21	M=11 Sd=21	M=91 Sd=19	M=95 Sd=21	M= -3 Sd=18	M=123 Sd=18	M=98 Sd=18	M=25 Sd=11

Medelvärden (M), standardavvikelser (Sd) och differens (Diff).

Tabell 4. Preliminära data ur GODA-studien (n=38): Antal med bättre, lika, eller sämre resultat avseende auditiv (A) jämfört med visuell (V) modalitet.

	Attention	Vigilance	Focus	Speed
A > V	33	28	14	38
A = V	4	5	3	0
A < V	1	5	21	0

Eftersom vi inte hade anledning att anta att just vår studiepopulation skulle prestera bättre när det gäller auditiv uppmärksamhet jämfört med visuell fanns det alltså anledning att misstänka att ett systematiskt fel smugit sig in i resultaten. Frågan som väcktes var betydelsen av den icke-standardiserade bjudningen, där ju förutsättningarna var annorlunda än vid normeringen av IVA+; 1) Instruktionerna hade översatts till svenska och lästes för försökspersonen eftersom det hade bedömts att de datoriserade instruktionerna på engelska inte var lämpliga för aktuell studiepopulation. Detta kan ha inneburit en ökad variabilitet i bjudning samt testledarpåverkan. 2) En särskild ljudfil med svenskt tal för de auditiva stimuli "ett" och "två" användes vid testningen. Det är möjligt att man reagerar annorlunda på de svenska ljuden jämfört med de amerikanska. Det är också möjligt att något i klockningen i den svenska ljudfilen påverkade resultatet.

En annan möjlighet var naturligtvis att de tekniska förutsättningarna kan ha påverkat resultaten. Vi använde en laptop med en tillkopplad LCD-skärm, istället för en stationär PC. Musen var av samma typ som den rekommenderade, men inte av samma märke. Även om mjukvara och inställningar användes enligt IVA+ instruktioner, kan något, t ex ljudkort, processor osv, ha påverkat mätresultaten. Den initiala klockningskontrollen (*timing check*) i samband med installationen av programmet hade visserligen visat responshastigheter inom normalområdet, men detta kan ha varit ett grovt mått. Eftersom studien redan pågick gjordes inga förändringar i administrationen av IVA+. Istället planerades en studie för att jämföra den svenska anpassade versionen med den amerikanska originalversionen. Detta kunde genomföras i samband med en test-retest-studie av Purdue Pegboard vid Arbets- och miljömedicin, Göteborg (Ahlstrand, Ekman, Noor, Åhlström, & Hagberg, 2012).

Syfte

Det huvudsakliga syftet var att jämföra den amerikanska originalversionen av IVA+ med den anpassade versionen med instruktioner och stimuli på svenska. Frågeställningen var huruvida resultaten skiljde sig åt beroende på vilken språkversion som används. Ytterligare frågeställning var om det förekom obalans mellan visuella och auditiva delskalor oavsett språkversion?

Material och metoder

Studiedesign

Testning med IVA+ vid två tillfällen med 1-2 veckors mellanrum, ena tillfället med svensk version och andra tillfället med amerikansk originalversion, i latinsk kvadrat så att hälften började med svensk version och hälften med amerikansk.

Undersökningsgrupp

Undersökningsgruppen bestod initialt av 28 personer, i åldrarna 18-60 år, som rekryterades bland personal på AMM samt några av dessas anhöriga eller vänner, för att ingå i en test-retest-studie av Purdue Pegboard samt uppmärksamhetstestning. Målet hade varit att rekrytera 36 personer i åldrarna 18-75, med en jämn fördelning avseende kön och de tre åldersgrupperna <40, 40-55, >55, vilket dock inte uppnåddes helt. En förutsättning för att delta i studien var att förstå engelska språket i tal och skrift. Ytterligare villkor för deltagande var att man var normalfrisk och i fullt arbete, utan hämmande besvär från nacke, skuldror eller händer, högerhänt samt hade normal syn (glasögon inget hinder). Försökspersonerna ersattes med 2 biobiljetter.

Genomförande

Försökspersonerna informerades om studien och lämnade skriftligt godkännande för deltagande. De undersöktes först med Purdue Pegboard (sammanlagt 5 omgångar med respektive hand) under ca 30 min (Ahlstrand, Ekman et al. 2012). Därefter hade försökspersonerna en paus om ca 5 minuter där de uppmuntrades att ta en promenad.

Uppmärksamhetstestning genomfördes i ett avskilt och tyst rum. Först testades försökspersonerna med Digit Symbol, ett papper-och-penna-test ur testbatteriet WAIS-R, med en testtid om 90 sek samt några minuters instruktioner (resultat redovisas ej här). Därefter vidtog testning med IVA+. IVA+ (version 2008.1) var installerad på en bärbar dator (HP Intelcore 2 duo, 2.1 GHz med Vista SP1, 2 GB ramminne) till vilken en 15-tums LCD-skärm, en USB-mus (HP) och en enklare hörlursmodell (Sony) kopplades, enligt inställningar och systemkrav för användning av IVA+ (Sandford and Turner 2005). Försökspersonerna placerades ca 40-60 cm från skärmen, på en höj- och sänkbar stol med blickpunkten i centrum av skärmen ca 5 cm under tps ögonhöjd. Baserat på ID-nr (udda eller jämnt) testades personen med antingen svensk eller amerikansk version av IVA+. Testningen inleddes med några verbala instruktioner från testledaren där försökspersonen placerades vid datorn, prövade ut lämplig sittställning och satte på sig hörlurar. När svensk version genomfördes, läste testledaren upp svenska instruktioner enligt en internt utvecklad manual (baserad på den som tagits fram vid AMM Lund) och som var en direkt översättning av den amerikanska. När amerikansk originalversion genomfördes gavs instruktionerna på engelska av en digitaliserad kvinnlig röst. Sammanlagd testtid inklusive instruktioner var ca 20 minuter.

Efter ca 1 vecka gjordes testning nummer två, denna gång med den alternativa versionen av IVA+ jämfört med den som gjordes vid testning nummer ett. Proceduren var i övrigt densamma som vid första testillfället och med samma testledare. Testledare vid samtliga testillfällen var

Christina Ahlstrand, leg biomedicinsk analytiker (19 testningar) eller Linda Åhlström, leg sjuksköterska (9 testningar). Dessa hade först instruerats och tränats i metoden.

Undersökningsvariabler

Skalorna som undersöktes (18 st) var Full scale Response Control Quotient (FRCQ), Auditory Response Control Quotient (ARCQ), Visual Response Control Quotient (VRCQ), och deras auditiva (A) och visuella (V) subskalor Prudence (PRA och PRV), Consistency (CONA och CONV) och Stamina (STMA och STMV), samt Full Scale Attention Quotient (FAQ), Auditory Attention Quotient (AAQ) och Visual Attention Quotient (VAQ), och deras subskalor Vigilance (VIA och VIV), Focus (FOCA och FOCV) och Speed (SPDA och SPDV). (se Figur 1).

För alla ovan nämnda variabler användes skalornas normerade kvotpoäng. Även rå reaktionstid i millisekunder (ms) för auditiva respektive visuella stimuli (MNA och MNV) undersöktes avseende skillnader mellan språkversionerna.

Analys

Deskriptiva data för undersökningsgruppen: ålderskategori, kön och utbildningsnivå redovisas. För att undersöka om skalorna i de två språkversionerna skiljde sig åt gjordes t-test som tog hänsyn till upprepad mätning med PROC MIXED proceduren i SAS och med justering för testordning och dess interaktion med språkversion. Detta gjordes för skalorna FRCQ, ARCQ, VRCQ respektive FAQ, AAQ och VAQ samt subskalorna PRA, PRV, CONA, CONV, STA, STV respektive VIA, VIV, FOCA, FOCV, SPDA och SPDV inklusive rå reaktionstid MNA och MNV.

Parvisa t-test gjordes för att undersöka om resultat i visuell och auditiv modalitet skiljde sig åt inom respektive skala. Dessa analyser gjordes uppdelat för språkversion med sammanlagt resultat oberoende av testtillfälle för samtliga ovan nämnda skalor bortsett från MNA och MNV (eftersom det redan är känt att rå auditiv och visuell responshastighet skiljer sig åt).

Statistisk signifikans baseras på $\alpha=0.05$. Alla analyser gjordes med SAS 9.3 (SAS Institute Inc., Cary, NC, USA).

Resultat

Undersökningsgruppen

Totalt genomförde 28 personer första testestomgången och 27 båda testomgångarna av IVA+. Samtliga testningar var valida enligt testets interna validitetskriterium (*Validity check*), men i en persons testprofil fanns tecken på bristande förståelse för testets genomförande enligt kontrollskalan *Comprehension* vid båda testtillfällena, varför hans testresultat uteslöts ur analysen. Detta innebär att 27 personer, varav 15 kvinnor och 12 män, bidrog med resultat från första testtillfället och 26 personer från andra testtillfället. Könsfördelningen var relativt jämn i åldersgrupperna <40 år och >55 år. I åldersgruppen 40-55 år var det övervikt på kvinnor (Tabell 5). Utbildningsnivå varierade mellan 11 och 20 år, median 17 år (M=16,30, SD= 2,67).

Fjorton personer genomförde amerikansk originalversion vid första testtillfället och 13 den svenska versionen (Tabell 6). Andra testtillfället genomfördes efter 7-20 dagar (mv 8,08).

Tabell 5. Undersökningsgrupp vid första testtillfället (n=27) och vid andra testtillfället (n=26) kursivt inom parantes.

	<40 år	40-55 år	>55	Totalt
Kvinnor	6 (5)	6 (6)	3 (3)	15 (14)
Män	7 (7)	1 (1)	4 (4)	12 (12)
Totalt	13 (12)	7 (7)	7 (7)	27 (26)

Tabell 6: Resultat av latinsk kvadrering

IVA+	Testtillfälle 1 n (män/kvinnor)	Testtillfälle 2 n (män/kvinnor)	Totalt
Original IVA+	14 (6/8)	12 (6/6)	26
Svensk IVA+	13 (6/7)	14 (6/8)	27
Totalt	27 (12/15)	26 (12/14)	53

Skillnader mellan originalversion och svensk version av IVA+

När skalorna jämfördes mellan de två språkversionerna framkom statistiskt signifikanta skillnader avseende auditiv reaktionstid. Genomsnittlig rå reaktionstid (MNA) var i den amerikanska versionen ca 502 ms och i den svenska versionen ca 469 ms, dvs en skillnad på ca 33 ms (Tabell 7). Vid t-test för upprepade mätningar var den estimerade medelskillnaden (differences of least squares means) 31.89, $p < 0.05$. Medelvärde i den kvotpoängberäknade responstiden SPDA var ca 121 i amerikansk version och ca 127 i den svenska, dvs en differens på ca 6 kvotpoäng (estimerad skillnad -6.12, $p < 0.05$). I övrigt fanns inga statistiskt signifikanta skillnader mellan språkversionerna.

Tabell 7. Jämförelse av skalorna mellan original IVA+ och svensk IVA+

Medelvärden av skalpoäng i originalversion (n=26) och svensk version (n=27) av IVA+, estimerad medeldifferens, 95 % konfidensintervall (KI) samt p-värden.

IVA+ SKALA	Original IVA+ mv	Svensk IVA+ mv	Estimerad differens ¹	95 % KI	p-värde
FRCQ	100.65	99.67	1.03	-8.02;10.08	0.82
ARCQ	101.62	99.96	1.63	-7.87;11.12	0.73
VRCQ	99.50	99.78	-0.15	-8.35;8.05	0.97
PRA	103.12	106.56	-3.43	-11.60;4.74	0.40
PRV	98.08	96.41	1.65	-8.50;11.80	0.75
CONA	94.69	90.11	4.72	-6.11;15.55	0.39
CONV	100.46	103.30	-2.56	-9.91;4.79	0.49
STMA	105.62	103.78	1.64	-7.40;10.67	0.72
STMV	100.38	99.48	0.85	-6.45;8.15	0.82
FAQ	105.81	107.85	-1.83	-9.83;5.72	0.63
AAQ	107.92	109.52	-1.42	-9.47;6.63	0.72
VAQ	102.69	104.63	-1.70	-8.76;5.35	0.63
VIA	100.65	103.81	-3.09	-11.64;5.46	0.47
VIV	103.73	102.74	1.20	-5.54;7.94	0.72
FOCA	95.88	89.89	6.06	-5.90;19.02	0.31
FOCV	99.69	103.78	-3.90	-12.56;4.76	0.37
SPDA	120.77	127.07	-6.12	-12.18;-0.06	0.05
SPDV	103.31	104.30	-0.88	-6.21;4.44	0.74
MNA ms ²	501.58	469.04	31.89	1.80;61.98	0.04
MNV ms ²	417.08	412.11	4.48	-18.69;27.65	0.70

¹Differences of least squares means

²ms=millisekunder, råpoäng responstid

Skillnader mellan auditiva och visuella delskalor i båda språkversionerna

I samtliga testningar (n=53), oavsett språkversion, var kvotpoängen högre avseende auditiv reaktionstid (SPDA) jämfört med visuell reaktionstid (SPDV). Medeldifferensen mellan SPDA och SPDV var över 17 kvotpoäng i amerikansk version (Tabell 8) och närmare 23 kvotpoäng i svensk version (Tabell 8). Medelvärdena för SPDA (120,8 respektive 127,1) var mer än en standardavvikelse över förväntat värde (dvs mv 100, SD 15), medan SPDVs medelvärden (ca 103 respektive ca 104) låg nära normvärdena.

Vidare var skillnaden statistiskt signifikant mellan de sammansatta skalorna Auditory och Visual Attention (AAQ och VAQ), vilket kan förklaras av att Attention till en tredjedel består av Speed. Skillnader förekom också mellan auditiv och visuell Prudence (PR) i båda språkversionerna, med bättre resultat i auditiv modalitet (Tabell 8). PR beräknas på antalet felaktiga responser i form av bristande responsinhibition.

Vidare, i svensk version framkom dessutom statistiskt signifikanta skillnader mellan auditiv och visuell Consistency (CON) och Focus (FOC), denna gång med bättre resultat i visuell modalitet. Auditivt resultat i CON och FOC låg ca 10 kvotpoäng under det förväntade genomsnittet (Tabell 7). I amerikansk version var skillnaderna mellan auditiv och visuell modalitet i dessa två delskalor inte statistisk signifikant, men auditiv modalitet var även här under det normerade genomsnittet. Både CON och FOC handlar om variabilitet i responstider, där CON beräknas på kvoten mellan de snabbaste och långsammaste reaktionstiderna medan FOC baseras på kvoten mellan standardavvikelsen och medelvärdet.

Tabell 8. Skillnader mellan auditiva och visuella skalor i original IVA+ (n=26) respektive svensk IVA+ (n=27)

Genomsnittlig differens (diff) mellan auditiv (A) och visuell (V) modalitet i kvotpoängsskalorna, 95 % konfidensintervall (KI) samt p-värden.

Skala	Original IVA+			Svensk IVA+		
	Diff A-V	95 % KI	p-värde	Diff A-V	95 % KI	p-värde
Response Control Quotient RCQ	2.12	-4.21;8.44	0.50	0.19	-6.02;6.39	0.95
Prudence PR	5.04	0.37;9.71	0.04	10.15	1.81;18.48	0.02
Consistency CON	-5.77	-14.79;3.25	0.20	-13.19	-22.28;-4.09	<.01
Stamina STM	5.23	-0.13;10.59	0.06	4.30	-1.54;10.13	0.14
Attention Quotient AQ	5.23	0.90;9.56	0.02	4.89	0.81;8.97	0.02
Vigilance VI	-3.08	-8.00;1.84	0.21	1.07	-4.79;6.93	0.71
Focus FOC	-3.81	-12.27;4.65	0.36	-13.89	-20.64;-7.14	<.001
Speed SPD	17.46	14.69;20.23	<.0001	22.78	19.82;25.73	<.0001

Diskussion

Denna studie gjordes mot bakgrund av att vi använde det amerikanska datoriserade uppmärksamhetstestet IVA+ med svensk anpassning i ett forskningsprojekt (GODA) och tidigt såg att resultaten tycktes snedfördelade mellan auditiv och visuell modalitet i vissa delskalor. Misstanken föll på att något i den svenska anpassningen påverkade testresultaten. Ett av huvudsyftena var därför att göra en systematisk jämförelse mellan den amerikanska originalversionen och den svenska anpassade versionen av IVA+. De två språkversionerna skiljde sig åt i endast ett avseende; auditiv responstid. Genomsnittlig auditiv responstid var 33 ms snabbare i den svenska versionen jämfört med i den amerikanska, vilket motsvarade en skillnad på drygt 6 kvotpoäng i Speed-skalan, dvs mindre än en halv standardavvikelse. Desto större skillnader framkom mellan auditiv och visuell responstid (Speed) inom både den amerikanska och den svenska versionen; ca 17 respektive 23 kvotpoängs skillnad. Auditiv responstid var i båda språkversionerna mer än 1 standardavvikelse bättre än det normerade genomsnittet, medan visuell responstid låg nära det normerade genomsnittet. Också Prudence-skalan skiljde sig avseende auditiv och visuell modalitet i båda språkversionerna, med bättre resultat auditivt. Prudence beräknas på antalet inkorrekta responser, där svar borde ha inhiberats, och har alltså inte med responstid att göra. Vidare framkom skillnader mellan auditiv och visuell modalitet i Consistency (CON) och Focus (FOC) i den svenska versionen, denna gång till den visuella modalitetens fördel. Både CON och FOC baseras på variabiliteten i responstider, där CON beräknas på kvoten mellan de snabbaste och de långsammaste responstiderna och FOC baseras på kvoten mellan standardavvikelse och medelvärde av responstiderna. Detta tyder alltså på en större variation i responstider i den svenska versionen. Skillnaden i skalan Vigilance som framkom i GODA-projektets preliminära resultat (med en skillnad om ca 10 kvotpoäng) replikerades inte i denna studie.

Att den svenska versionen gav något snabbare auditiva responstider än den amerikanska originalversionen kan kanske förklaras av att man reagerar snabbare när stimuli uttalas på modersmålet. Bjudningen var dock också icke-standardiserad i det att testledaren gav muntliga instruktioner istället för att låta IVA+ ge de datoriserade instruktionerna, samt bruket av en tillagd ljudfil för de svenska ljuden ”ett” och ”två”. Testledareffekt kan således ha bidragit till skillnader i resultaten, t ex att den mer personliga kontakten vid den svenska versionen kan ha ökat motivationen att prestera bättre. Testledarna var dessutom kända för de flesta försökspersonerna vilket också kan ha betydelse (Fuchs & Fuchs, 1986). Någon testledareffekt bör dock inte ha förekommit i själva testuppgiften eller i poängberäkningen eftersom den sköts helt datoriserat. En annan förklaring skulle kunna vara att de fonetiska ljuden ”ett” och ”två” uppfattas snabbare eller på ett annat sätt än ljuden för ”one” och ”two”. Ytterligare en förklaring kan vara att den svenska ljudfilen innehåller ett klockningsfel.

Det får nog vara ställt utom tvivel att ett systematiskt fel förekommit som genererade snabbare responstid än förväntat för auditiv modalitet oavsett språkversion. Man kan misstänka att något i de tekniska förutsättningarna - t ex modernare dator, display, operativsystem, ljudkort osv än vid normeringen av IVA+ - gör att det auditiva stimuli och/eller dess respons klockas fel, trots korrekta inställningar och uppfyllda systemkrav. Samtidigt borde detta inte förklara skillnaden i

Prudence, eftersom responstid inte ingår i beräkningen av den skalan (som handlar om svarsinhibition). En enbart snabbare responstid kan inte heller förklara den förhöjda variabiliteten i responshastigheter som förekom i den svenska versionen, om inte den möjliga felklockningen dessutom varierade under testets gång?

En faktor som skulle kunna bidra till skillnaderna är om det fanns en selektionsbias som gjorde att studiegruppen presterade snabbare och mer adekvat auditivt (men med större variabilitet), jämfört med normeringsgruppen. Studiegruppen var selekterad så till vida att de flesta var välutbildade och därmed kan förväntas prestera något över genomsnittet. Det är dock okänt att detta skulle gälla selektivt just auditiv förmåga. Ytterligare en förklaring till de funna skillnaderna, vilket naturligtvis också är en av studiens begränsningar, är att vi har gjort multipla analyser utan s.k. bonferroni-korrektion. Det innebär att en del av skillnaderna som framkom kan vara genererade av slumpen. Vi har inte heller kontrollerat för de två testledarna.

Ett nästa steg skulle kunna vara en studie som jämför resultat av IVA+ som utförts på flera olika hårdvara, t ex PC, laptop med och utan kopplad skärm osv, och kanske också med olika operativsystem. Idealt skulle en korrigerande faktor kunna beräknas för att möjliggöra tillämpning av standardnormerna. Ett alternativ är att samla in lokala norm- eller referensdata, kopplat till just den språkversion och teknik som används.

Konklusion

Den svenska språkversionen av IVA+ genererade ett något bättre resultat avseende auditiv snabbhet än den amerikanska originalversionen. I vår datainsamling synes dock både amerikansk originalversion och svensk version innehålla ett systematiskt fel som ger bättre än förväntat resultat avseende auditiv snabbhet. Detta gör att testets standardnormer knappast kan tillämpas på insamlade data. På grund av osäkerheten i vad felet består krävs stor försiktighet även vid framtida datainsamling och särskilt vid klinisk tillämpning. I forskningsstudier bör det vara mindre problematiskt så länge testningarna används för att jämföra grupper eller för att göra för- och eftermätningar för att utvärdera t ex en intervention, dvs oberoende av hänvisning till normer. Det är dock viktigt att samma hård- och mjukvara används till samtliga testningar. Eftersom resultaten avvek mer i den svenska anpassade versionen rekommenderas att originalversionen används om så är möjligt. Även den som inte talar god engelska kan antagligen förstå och skilja på ”one” och ”two”. Förtydligande instruktioner bör kunna ges på svenska vid behov.

Referenser

- Ahlstrand, C., Ekman, A., Noor, A., Åhlström, L., & Hagberg, M. (2012). Reliabilitetstestning av Purdue Pegboard (finmotoriktest) *Rapport från Arbets- och miljömedicin nr 140*. Göteborg: Sahlgrenska universitetssjukhuset.
- Fuchs, D., & Fuchs, L. (1986). Test procedure bias: A meta-analysis of examiner familiarity effects. *Review of Educational Research*, 56 (2), 243-262
- Hart, R. P., Martelli, M. F., & Zasler, N. D. (2000). Chronic Pain and Neuropsychological Functioning. *Neuropsychology Review*, 10, 131-149.
- Nygaard, C. H., Eskelinen, L., Suvanto, S., Tuomi, K., & Ilmarinen, J. (1991). Associations between functional capacity and work ability among elderly municipal employees. *Scand J Work Environ Health*, 17 Suppl 1, 122-127.
- Sandford, J., & Turner, A. (1995). IVA+Plus™. Integrated Visual and Auditory Continuous Performance Test. Interpretation Manual. : Braintrain Inc. .
- Sandford, J., & Turner, A. (2005). IVA+Plus™. Integrated Visual and Auditory Continuous Performance Test. Administration manual: BrainTrain, Inc.
- Sandström, A., Rhodin, I. N., Lundberg, M., Olsson, T., & Nyberg, L. (2005). Impaired cognitive performance in patients with chronic burnout syndrome. *Biological Psychology*, 69(3), 271-279.
- Tinius, T. P. (2003). The Integrated Visual and Auditory Continuous Performance Test as a neuropsychological measure. *Archives of Clinical Neuropsychology*, 18(5), 439-454.
- Österberg, K., Skogsliden, S., & Karlson, B. (2014). Neuropsychological sequelae of work-stress-related exhaustion. *Stress*, 17(1), 59–69.