



MEDICINSKA FAKULTETEN
Lunds universitet

**Avdelningen för Logopedi, Foniatri och Audiologi,
Institutionen för Kliniska Vetenskaper Lund**

Utprovning av tre svenska versioner av Token Test på en vuxen normalgrupp

Pia Apt

**Vetenskapligt arbete
i logopedi, 20 poäng, 2008**

Handledare: Ulrika Nettelbladt och Karl Radeborg

Sammanfattning

Token test, som konstruerades för 45 år sedan av E. de Renzi och A. Vignolo, är ett afasitest avsett att diagnostisera lätta, auditivt impressiva störningar vid afasi. Det väl validerade testet har i ett flertal studier visat sig mer sensitivt än traditionella afasitest för denna uppgift. Idag används Token test i många länder i flera omarbetade och även förkortade versioner.

Trots att tre versioner av Token test sedan många år tillbaka finns översatta till svenska och används av logopedier i Sverige, existerar inga svenska normalvärden. Syftet med denna studie har varit att fastställa normalvärden för de svenska versionerna av Token test samt att jämföra testversionerna avseende dessa värden.

De tre testversionerna prövades ut på en grupp om 96 slumpvis utvalda, friska personer mellan 21 och 80 år med jämn köns- och åldersfördelning.

Gränsvärdena för normalprestation för varje testversion fastställdes. Dessa värden var i samtliga testversioner högre än i motsvarande eller jämförbara testversioner på andra språk. Resultaten understryker vikten av att pröva ut översatta, utländska hjärnskadetest på svenska normalgrupper. Det fanns inga könsskillnader i testresultat i någon version. I de två längsta versionerna medförde lägre ålder respektive högre utbildningsnivå signifikant bättre testresultat än högre ålder respektive lägre utbildningsnivå. De två längsta testversionerna hade en jämförbar och högre reliabilitet än den kortaste testversionen. Det fanns ett högre samband mellan testresultaten i dessa två versioner än mellan resultatet av den kortaste versionen och de två övriga.

Slutsatserna av resultaten är att de erhållna normalvärdena leder till en säkrare diagnostik av lätta störningar i talförståelse vid afasi när Token test används. Den kortaste testversionen rekommenderas endast för screeningändamål medan de längre testversionerna utan större problem kan ersätta varandra vid noggrannare diagnostik, om det mer detaljerade poängsättningsalternativet används.

Nyckelord: Token test, afasi, impressiva störningar, normalvärden

Innehållsförteckning

Inledning	1
Afasi	1
Afasitest	1
Token test	1
<i>Beskrivning</i>	1
<i>Syfte och bakomliggande principer</i>	2
<i>Olika testversioner</i>	3
<i>Svenska testversioner</i>	4
Utprovning av Token test	5
<i>Validitet</i>	5
<i>Reliabilitet</i>	6
<i>Icke-språkliga faktorerers påverkan på testresultatet</i>	6
<i>Utprovning på normala personer</i>	7
<i>Förhållandet mellan testresultat och kön, ålder respektive utbildning</i>	8
Syfte	8
Metod	9
Försökspersoner	9
Material	10
Tillvägagångssätt	11
Poängsättning	11
Statistiska beräkningar	12
Etikprovning	12
Resultat	12
Version A: Token test (de Renzi & Vignolo, 1962)	12
<i>Kvantitativ analys av testresultatet</i>	12
<i>Kvalitativ analys av testresultatet</i>	13
<i>Förhållandet mellan testresultat och kön, ålder respektive utbildning</i>	15
<i>Reliabilitet</i>	16
<i>Tidsåtgång och förhållandet mellan tidsåtgång och kön, ålder respektive utbildning</i>	16
Version B: Token test – kortversion ur testet MAE (Benton, Hamsher & Sivan, 1978)	16
<i>Kvantitativ analys av testresultatet</i>	16
<i>Kvalitativ analys av testresultatet</i>	17
<i>Förhållandet mellan testresultat och kön, ålder respektive utbildning</i>	18
<i>Reliabilitet</i>	18
<i>Tidsåtgång och förhållandet mellan tidsåtgång och kön, ålder respektive utbildning</i>	18
Version C: Token test – förkortad version (de Renzi & Faglioni, 1978)	19
<i>Kvantitativ analys av testresultatet</i>	19
<i>Kvalitativ analys av testresultatet</i>	21
<i>Förhållandet mellan testresultat och kön, ålder respektive utbildning</i>	22
<i>Reliabilitet</i>	23
<i>Tidsåtgång och förhållandet mellan tidsåtgång och kön, ålder respektive utbildning</i>	23

Förhållandet mellan testversionernas totalpoäng.....	23
Diskussion.....	24
Metod.....	24
<i>Försökspersoner.....</i>	24
<i>Poängsättning, material och tillvägagångssätt.....</i>	24
Resultat.....	25
<i>Kvantitativ analys av testresultatet.....</i>	25
<i>Kvalitativ analys av testresultatet.....</i>	26
<i>Förhållandet mellan testresultat och kön, ålder respektive utbildning.....</i>	27
<i>Tidsåtgång och förhållandet mellan tidsåtgång och kön, ålder respektive utbildning.....</i>	28
<i>Jämförelse mellan testversionernas testpoäng och tillförlitlighet.....</i>	28
Slutsatser.....	29
 Referenser.....	 31

Bilaga 1: Inbjudningsbrev

Bilaga 2: Blankett till testversion A

Bilaga 3: Blankett till testversion B

Bilaga 4: Blankett till testversion C

Bilaga 5: Testledarens instruktioner till försökspersonerna

Inledning

Afasi

Afasi brukar definieras som en språklig störning, som är orsakad av en förvärvad hjärnskada (Ahlsén, 2006). Störningen kan yttra sig både i svårigheter att uttrycka sig språkligt i t.ex. tal och skrift (expressiva störningar) och i svårigheter att förstå språk, t.ex. i form av tal och skrift (impressiva störningar). Beroende på skadans lokalisering uppträder olika afasisymptom, t.ex. problem med att komma på eller att förstå betydelsebärande ord, att formulera eller att förstå satser eller att producera eller uppfatta olika språkljud. Beroende på skadans storlek kan den afasidrabbades svårigheter variera från en nästan total oförmåga att tala, förstå tal, läsa och skriva till endast diskreta problem att komma på ovanliga ord eller att förstå långa och grammatiskt komplexa utsagor (Apt, 2001). Olika afasiologiska skolor klassificerar afasisymptomen på något olika sätt i s.k. afasisyndrom. De internationellt mest kända skolorna är den av den amerikanske neurologen Norman Geschwind starkt influerade Bostonskolan och den riktning som bygger på den ryske neuropsykologen Alexander Lurias s.k. dynamiska lokaliseringsteori. Bostonskolan framhåller betydelsen av avbrott i förbindelser mellan olika områden i hjärnan som huvudorsaken till afasi (Albert, Goodglass, Helm, Rubens & Alexander, 1981). Enligt Lurias teori är språkförmågan, liksom alla andra mentala funktioner, uppbyggd av specifika förmågor, var och en lokaliserad till en viss del av hjärnan, vilka på ett flexibelt sätt samverkar med varandra i väldiga neuronala nätverk vid olika språkliga aktiviteter (Luria, 1973).

Afasitest

Vid diagnostisering av afasi används olika typer av test. I ett tidigt skede efter insjuknandet används ofta screeningartade test, t.ex. Grundläggande Neuropsykologisk Undersökning (GNU; Apt, 2001). Inför ställningstagande till och utvärdering av behandling i ett senare skede efter insjuknandet används ofta mer omfattande test eller testbatterier. Internationellt mycket använda test inom denna kategori är t.ex. Boston Diagnostic Aphasia Examination (BDAA; Goodglass & Kaplan, 1983), Western Aphasia Battery (WAB; Kertesz, 1982), Multilingual Aphasia Examination (MAE; Benton, Hamsher & Sivan, 1978) och Aachener Aphasia Test (AAT; Huber, Poeck, Weniger & Willmes, 1983). I Sverige är de mest använda testen inom denna kategori A-ning (Lindström & Werner, 1995) och Pia Apts AfasiPrövning (PAPAP; Apt, 1997). Vid behov av en noggrannare undersökning av en viss aspekt av språket används test, som endast prövar *en* specifik språklig funktion. Exempel på sådana test är Boston Naming Test (BNT; Kaplan, Goodglass & Weintraub, 1983) och Svensk BenämningsPrövning (SBP; Apt, 1999), vilka båda prövar substantivmobilisering. Till denna kategori hör också Token Test.

Token test

Beskrivning

Token test konstruerades år 1962 av de italienska forskarna Ennio de Renzi och Luigi Amadeo Vignolo (1962). Testmaterialet består av tjugo figurer (kallade tokens), som är tillverkade av plast, kartong eller trä (de Renzi & Faglioni, 1978). Dessa skiljer sig åt avseende de tre dimensionerna färg (vit, gul, grön, röd samt blå eller svart), form (cirkel, fyrkant) och storlek (stor, liten). De arrangeras på bordet framför testpersonen horisontellt i fyra parallella rader med tokens av samma storlek och form i varje rad.

Testet består av en serie muntliga uppmaningar till testpersonen, som går ut på han/hon ska vidröra, ta upp, flytta eller peka på olika tokens. Uppmaningarna är grupperade i avsnitt. I de första avsnitten ökar längden på uppmaningarna för varje avsnitt från uppmaningar av typen "Peka på den röda cirkeln!" till uppmaningar av typen "Peka på den lilla, gula cirkeln och den stora, gröna fyrkanten!". Det sista avsnittet består av grammatiskt komplexa uppmaningar, t.ex. "Innan Du rör vid den gula cirkeln, tag upp den röda fyrkanten!".

Syfte och bakomliggande principer

Syftet med Token test är att identifiera subtila auditiva, impressiva störningar vid afasi, vilka inte går att upptäcka med traditionella afasitest (de Renzi & Vignolo, 1962; Wertz, 1979). Testet kan användas för att planera träning av impressiv förmåga vid afasi och för att utvärdera resultatet av träningen. Det lämpar sig dock inte för patienter med stora impressiva störningar, eftersom svårighetsgraden på testet är hög (Wertz, 1979).

Token test utvecklades ursprungligen som en reaktion mot den typ av test, som vid den aktuella tidpunkten användes för att pröva lättare impressiva störningar. Ett exempel på ett sådant test är Pierre Maries Tre Papper Test: "Här är tre papper: ett stort, ett medelstort och ett litet. Tag det största, skrynkla ihop det och kasta det på golvet! Ge mig det medelstora! Stoppa det minsta i Er ficka!" (Marie, 1902). De Renzi och Vignolo anser att denna typ av test snarare prövar minnesförmågan än talförståelseförmågan.

Konstruktörerna till Token test ställde upp ett antal kriterier för ett bra test med ovan beskrivna syfte och försökte uppfylla dem med sitt nya test:

- Testet ska ta rimlig tid att genomföra och inte kräva avancerad apparatur eller annat svårtillgängligt material. När det gäller material kräver Token test endast en uppsättning tokens. Det anses också gå snabbt att genomföra och vara lätt att administrera, utvärdera och poängsätta (Lezak, 1995; Orgass & Poeck, 1966).
- Testet ska innehålla så korta uppmaningar att normala, vuxna personer i alla åldrar ska minnas dem. Den längsta uppmaningen i Token test innehåller på originalspråket italienska endast nitton stavelser mot fyrtioåtta i Maries test.
- Den intellektuella svårighetsgraden på testet ska vara så låg som möjligt. Vuxna personer ska - inom rimliga gränser - kunna utföra uppgifterna oberoende av IQ (de Renzi & Vignolo, 1962). Token test innehåller inga ovanliga ord eller ovanliga syntaktiska former, eftersom sådan kunskap påverkas av personens allmänna kognitiva nivå. Det kräver endast att man förstår de ord, som betecknar färg, storlek och form på figurerna i testet.
- Testet ska istället besitta en avsevärd svårighetsgrad på språklig nivå. Att en utsaga har en hög språklig svårighetsgrad innebär enligt testförfattarna att den är redundansfri eller åtminstone att den innehåller minimal redundans. Utsagan ska inte kunna förstås genom person- eller situationskännedom, genom ledtrådar i den språkliga kontexten eller genom kännedom om de i utsagan ingående objektens natur. I t.ex. utsagan "Titta ut genom..." är det utelämnade ordet sannolikt "fönstret". Ber man t.ex. en person att utföra en handling med en penna, finns det endast ett fåtal sannolika alternativ (t.ex. "skriva", "rita", "vassa"). Av det sistnämnda skälet används i Token test de mer abstrakta orden "fyrkant" och "cirkel" istället för konkreta ord som t.ex. "hus" och "blomma". Man måste förstå varje ord som bedöms i Token test för att kunna identifiera en token eller för att kunna utföra en handling med den. I det sista avsnittet har

särskild tonvikt lagts på prepositioner, då dessa enligt testkonstruktörerna utgör en speciell svårighet för de flesta afasidrabbade. Detta avsnitt innehåller även andra s.k. logiko-grammatiska strukturer, dvs. grammatiska konstruktioner som kräver rumsligt tänkande, t.ex. inverterade satser. Det är särskilt vanligt med svårigheter att förstå logiko-grammatiska strukturer hos individer med skador i skärnings-området mellan temporal-, parietal- och occipitalloben i vänster hjärnhemisfär (Luria, 1980a).

Olika testversioner

Token test har under årens lopp vidareutvecklats och modifierats av de ursprungliga konstruktörerna och även av andra forskare. Dessutom har flera förkortade versioner av testet utarbetats.

Antalet uppgifter varierar mellan de olika testversionerna. Den ursprungliga versionen, liksom en version av Orgass och Poeck (1966), innehåller 61 uppgifter, fördelade på fyra avsnitt om tio uppgifter vardera och ett sista avsnitt med 21 grammatiskt komplexa uppgifter. En version av Boller och Vignolo (1966) innehåller 62 uppgifter, då uppmaningen ”Rör vid den vita cirkeln utan att använda höger hand” har lagts till i det sista avsnittet. En förkortad version, som ingår i testet Neurosensory Center Comprehensive Examination for Aphasia (Spreen & Benton, 1969), består av 39 uppgifter. Med denna version som utgångspunkt konstruerade Spellacy och Spreen (1969) en kortversion med 16 uppgifter. En ännu kortare version av Token test har utarbetats av van Harskamp och van Dongen (1977). Detta test är en holländsk, förkortad version av Boller och Vignolos version och det innehåller bara 10 uppgifter. År 1976 reviderade Orgass Token test och skapade en version med 50 uppgifter uppdelade på fem avsnitt med samma antal uppgifter i varje. De Renzi tog år 1978 tillsammans med Pietro Faglioni fram en förkortad version med 36 uppgifter uppdelade på sex avsnitt. I testet MAE finns en kortform av Token test (engelsk version: Benton, Hamsher & Sivan, 1978; spansk version: Rey & Benton, 1991) med 22 uppgifter uppdelade på två avsnitt.

De olika versionerna skiljer sig något åt när det gäller testmaterialets utformning och placeringen av tokens vid testningen. I versionerna av de Renzi och Faglioni och av Benton m.fl. har den blå tokenbit, som finns i övriga versioner, ersatts av en svart. De Renzi och Faglioni anger som skäl till detta att det i en artikel av Scotti och Spinnler (1970) framgår att en del såväl normala som hjärnskadade personer har svårt att skilja mellan blå och grön färg. Eftersom en annan tokenbit i testet är grön ansåg de att den blå tokenbiten skulle bytas ut.¹ I versionerna av de Renzi och Faglioni, av Spreen och Benton, av Spellacy och Spreen och av Benton m.fl. har de ursprungligen rektangulärt formade tokenbitarna ersatts av kvadrater och formen benämns ”fyrkant”. Orsaken är att detta ord i många språk är kortare och vanligare än ordet ”rektangel” (Lezak, 1995). I den första versionen av Token test ska tokens med olika färg läggas ut slumpvis med undantag av att gul och grön fyrkant måste ligga intill varandra i det sista avsnittet. I alla övriga versioner läggs alla tokens ut i en fastställd ordning, som dock varierar något mellan versionerna.

Även uppmaningarna i testet skiljer sig något åt mellan de olika versionerna. I originalversionen ges uppgifterna i de fyra första avsnitten slumpvis efter en given struktur och ordningsföljden mellan ord som betecknar storlek och färg i avsnitt IV varierar slumpmässigt.

¹ Anmärkas kan att den vanligaste formen av färgblindhet är röd-grönfärgblindhet medan gul-blåfärgblindhet är betydligt ovanligare (Nationalencyklopedin, 1992). En rimligare förändring borde därför ha varit att byta ut den gröna tokenbiten mot en svart.

I de två första avsnitten används uppmaningen ”pick up” (plocka upp) medan ordet ”touch” (rör vid) används i de två följande. I alla andra versioner av testet är samtliga uppgifter fastställda i nedskrivna form. I alla andra versioner - utom i den av Orgass och Poeck - är också ordet som betecknar storlek placerat före ordet som betecknar färg i uppgifter av den typ som finns i avsnitt IV i originalversionen. Det verb, som används i den uppgiftstyp som finns i avsnitten I-IV i det ursprungliga testet, varierar mellan olika versioner. I versionerna av Boller och Vignolo och av de Renzi och Faglioni används genomgående ordet ”touch” (rör vid). I andra versioner förekommer även orden ”take” (tag), ”point to” (peka på) och ”pick up” (plocka upp). I tre versioner av testet (Spreeen och Benton, de Renzi och Faglioni, Benton m.fl.) är de inledande uppgifterna ännu enklare än de i den ursprungliga testversionen. Syftet med att inkludera uppgifter av typen ”Rör vid en cirkel” är att göra testet genomförbart även vid svårare afasi (de Renzi & Faglioni, 1978).

Poängsättningen skiljer sig också åt mellan de olika testversionerna. I originalversionen specificeras inte poängsättningen. Av poängangivelser i den artikel, där de Renzi och Vignolo (1962) beskrev testet, kan man dock dra slutsatsen att varje relevant del av uppmaningen (i fortsättningen kallad item) – och inte uppmaningen som helhet – poängsätts. Denna metod används också i versionerna av Orgass och Poeck, av Spreeen och Benton och av Spellacy och Spreeen. I versionerna av Boller och Vignolo, av de Renzi och Faglioni, av Benton m.fl. och av Orgass poängsätts varje uppgift som en helhet. I några av versionerna ska en uppmaning, som besvaras felaktigt eller inte alls, upprepas. Korrekt respons vid upprepad stimulus leder antingen till full poäng (versionen av Boller & Vignolo), 0,5 poäng (versionen av de Renzi & Faglioni) eller 1 poäng av 2 möjliga (versionerna av Benton m.fl. och Rey & Benton).

Förutom ovan beskrivna testversioner, vilka i stort sett följer testets ursprungliga uppbyggnad, syfte och målgrupp, har ett antal versioner utvecklats, som skiljer sig ganska mycket från Token test i dess originalutförande. Exempel på sådana test är den psykometriskt bearbetade versionen av Token test, Revised Token Test (McNeil & Prescott, 1978), en testversion avsedd för barn, The Token Test for Children (DiSimoni, 1978), en version med konkreta föremål istället för tokens (Martino, Pizzamiglio & Razzano, 1976) samt en version där uppgifterna ges skriftligt istället för muntligt (Kiernan, 1986).

Olika versioner av Token test finns idag översatta till många språk. Förutom på originalspråket italienska och på engelska finns testet t.ex. på så skilda språk som tyska (Orgass & Poeck, 1966; Orgass, 1976), holländska (van Harskamp & van Dongen, 1977), spanska (Rey & Benton, 1991), polska (Podemski, Jarzebska & Bakierowska, 1984), danska (Spreeen & Benton, 1977), norska (Wold & Reinvang, 1990) thai (Pracharitpukdee, Phantumchinda, Huber & Willmes, 1998), kinesiska (Wang & Niu, 2000) och hebreiska (Cohen-Mimran & Sapir, 2007).

Svenska testversioner

Tre versioner av Token test är översatta till svenska: originalversionen av de Renzi och Vignolo, den förkortade versionen av de Renzi och Faglioni och kortversionen, som ingår i testet MAE av Benton m.fl. Enligt muntligt besked från översättarna har testkonstruktörerna till de två förstnämnda testversionerna givit sitt tillstånd till att testen översätts till svenska och sprids i Sverige. Omständigheterna kring översättningen av den sistnämnda testversionen och vem som har översatt den är okänt. Alla tre versionerna består i svensk översättning av nedskrivna testuppgifter samt mycket knapphändiga instruktioner om testprocedur och poängsättning. Tokenbitar bifogas ej utan användaren måste själv (låta) tillverka dem. Till de

Renzi och Faglioni's testversion bifogas en färgbild, som illustrerar placeringen av tokens. Ingen av testversionerna går att köpa i Sverige utan blanketterna kopieras fritt.

Översättningen av originalversionen gjordes i början av 1970-talet av docent Sören Fex, Lunds universitet, med testkonstruktörernas engelska översättning som utgångspunkt (Lindström & Revstedt, 1975). Denna översättning modifierades något av Lindström och Revstedt år 1975 i samband med författandet av ett examensarbete i logopedi, där Token test användes. Den svenska versionen av testet skiljer sig på några punkter från originalversionen. Blå tokenbitar ersatta med svarta. Tokens arrangeras med rektanglarna överst. Uppgifterna är fastställda i nedskrivna form. Ordet "rectangle" har översatts med det svenska ordet "fyrkant". I avsnitten I-IV används genomgående ordet "tag". I uppgift V6 har "pick up" översatts med ordet "tag". I uppgift V15 ges leden i uppmaningen i omvänd ordning jämfört med den engelska versionen.

Översättningen av den förkortade testversionen av de Renzi och Faglioni är gjord av logoped Eva Holmberg. Den enda förändring som har gjorts av utgångsmaterialet, som utgjordes av testkonstruktörernas engelska version av testet, är att leden i uppgift VI35 har kastats om. Den enda ändringen som gjorts i översättningen av kortversionen av Benton m.fl. är att ordet "take" genomgående har översatts med orden "ta upp".

Utprovning av Token test

Validitet

Ett antal studier har utförts för att undersöka Token tests validitet. Van Dongen och van Harskamp (1972) jämförde testresultaten från 46 normala personer, 27 hjärnskadade personer med afasi och 59 hjärnskadade personer utan afasi, som utförde originalversionen av Token test. Av personerna med afasi klassificerades 88 % korrekt som afasidrabbade. Samma version av testet användes i en studie av Hartje, Kerschensteiner, Poeck och Orgass (1973). I denna studie klassificerades 91 % av personerna med afasi korrekt när testresultaten från 57 hjärnskadade personer med afasi och 63 hjärnskadade personer utan afasi jämfördes. Liknande resultat framkom i en studie av Poeck, Hartje, Kerschensteiner och Orgass (1973), där 92 % av personerna med afasi klassificerades korrekt när resultaten av testning med originalversionen av testet från 41 hjärnskadade personer med afasi jämfördes med resultaten från 56 hjärnskadade personer utan afasi. I en studie av Orgass och Poeck (1966), där resultaten av testning med deras testversion från 51 normala, vuxna personer, 26 hjärnskadade personer med afasi och 49 hjärnskadade personer utan afasi jämfördes, klassificerades 84 % av personerna med afasi korrekt. Boller och Vignolo (1966) jämförde resultaten av testning med sin testversion från 30 normala personer, 30 högersidigt hjärnskadade personer utan afasi, 26 vänstersidigt hjärnskadade personer utan afasi och 34 vänstersidigt hjärnskadade personer med afasi. Totalt klassificerades 88 % av personerna korrekt. Hjärnskadade personer med afasi presterade signifikant sämre än hjärnskadade personer utan afasi. Vänstersidigt hjärnskadade personer utan afasi presterade signifikant sämre än högersidigt hjärnskadade utan afasi, vilket kan tolkas som att Token test förmådde identifiera språkliga svårigheter, som inte gick att upptäcka med andra afasitest. Swisher och Sarno (1969) jämförde resultaten av testning med samma version av Token test från 22 normala personer, 22 vänstersidigt hjärnskadade personer med afasi och 22 högersidigt hjärnskadade personer utan afasi. De fann att testet väl förmådde skilja ut afasidrabbade personer från övriga grupper. I en studie av de Renzi och Faglioni (1978) jämfördes testresultaten av deras förkortade version av testet från 215 normala personer och 200 vänstersidigt hjärnskadade personer med afasi. Det visade sig att 93 % av individerna med afasi klassificerades rätt. Vid valideringen av den version av Token test, som ingår i MAE

(Benton m.fl., 1978) fick 360 normala personer i åldern 16-69 år och 50 hjärnskadade personer med afasi utföra testet. Av personerna med afasi klassificerades 84 % korrekt. Spellacy och Spreen (1969) jämförde resultaten av testning med Spreen och Bentons testversion från 67 hjärnskadade med afasi med resultaten från 37 hjärnskadade utan afasi och fann att 89 % av personerna med afasi klassificerades korrekt. Då deras egen förkortade version av testet utfördes på samma försökspersoner klassificerades 82 % av de afasidrabbade personerna korrekt. Coupar (1976) använde också denna testversion när han jämförde testresultaten från 15 vänstersidigt hjärnskadade personer med afasi, 9 högersidigt hjärnskadade personer utan afasi, 9 lobotomerade personer och 15 normala personer. Han fann att Token test signifikant skilde de olika grupperna åt. De studier som har gjorts visar således samstämmigt att Token test i dess olika versioner väl förmår skilja ut personer med afasi från personer utan afasi och att Token test är ett känsligt test för att upptäcka lätta auditiva, impressiva, språkliga störningar.

Reliabilitet

Testets reliabilitet har prövats i några studier. Gallaher (1970) undersökte test-retestreliabiliteten genom att 30 afasidrabbade individer testades tre gånger med originalversionen av Token test under ett åttadagarsintervall. Korrelationskoefficienten varierade mellan .91 och .98, vilket visar att test-retestreliabiliteten är hög. Orgass (1976) rapporterade en test-retestreliabilitet på .96 för sin reviderade testversion. I en studie av Wertz (1979), där interbedömarreliabiliteten i Boller och Vignolos testversion undersöktes, var sambandet mellan bedömare .99, vilket innebär att interbedömarreliabiliteten är mycket hög.

Icke-språkliga faktorerers påverkan på testresultatet

Ett av de kriterier de Renzi och Vignolo satte upp för Token test var att testresultatet inte ska påverkas av en individs intelligens och övriga icke-språkliga kognitiva förmågor. Huruvida så verkligen är fallet har undersökts i ett antal studier. Van Dongen och van Harskamp (1972) fann ett signifikant samband mellan resultatet på Token test och intelligens mätt med det icke-språkliga intelligenstestet Ravens progressiva matriser (Raven, 1938) hos 46 normala personer. Steinberg, Bieliauskas, Smith, Langellotti och Ivnik (2005) fann ett starkt samband mellan prestationen på Benton, Hamsher och Sivans version av Token test och intelligens mätt med intelligenstestet WAIS-R (Wechsler, 1981) hos 437 normala personer i åldern 56-99 år. Coupar (1976) fann att korrelationen mellan den sista delen i Token test och Ravens progressiva matriser var betydligt högre än för testet som helhet. Detta tolkade han som att intelligensen har en viss betydelse för resultatet på Token test. Däremot fann Boller och Vignolo (1966) inget signifikant samband mellan resultatet på Token test och resultatet på Ravens progressiva matriser. Flera studier (Boller & Vignolo, 1966; Spellacy & Spreen, 1969; Orgass m.fl., 1973) har visat att den sista delen av Token test är lika sensitivt som testet som helhet för att upptäcka lätta impressiva störningar, vilket tyder på att det är den språkliga förmågan och inte den allmänna intelligensen, som mäts med avsnittet ifråga. Således föreligger något motsägelsefulla resultat mellan olika studier när det gäller intelligensens påverkan på testresultatet på Token test.

Andra icke-språkliga kognitiva störningar, som har föreslagits kunna påverka resultatet på Token test är agnosi för färg och form (de Renzi & Vignolo, 1962), perseverationer och mental inflexibilitet (Swihart, Panisset, Becker, Beyer & Boller, 1989), störningar i koncentration eller uppmärksamhet (Leischner, 1974), störningar i korttidsminne (Lesser, 1976; Mack & Boller, 1979), visuospatiala störningar (Agrell, Dehlin och Nilsson, 1995; Swisher & Sarno, 1969) och neglekt (Swisher & Sarno, 1969). Forskningsresultaten kring icke-språkliga kognitiva

störningars påverkan på testresultatet på Token test är dock inte helt entydiga. Agrell m.fl. (1995) fann t.ex. inga signifikanta skillnader i resultat på Token test mellan hjärnskadade personer med och utan neglekt.

Utprovning på normala personer

För att kunna avgöra om en testprestation faller inom eller utom normala gränser, är det viktigt att det finns fastställda normalvärden för test, som är avsedda för hjärnskadade. För Token test är detta särskilt viktigt, eftersom testet är avsett att upptäcka mycket lätta störningar. Olika versioner av Token test har prövats ut på normalgrupper i ett antal studier men antalet försökspersoner, urvalskriterier för försökspersonerna och poängsättningen av testet skiljer sig åt mellan studierna.

Den holländska versionen av originaltestet prövades ut på en normalgrupp av van Dongen och van Harskamp (1972). Man rekryterade 46 vuxna personer bland patienter på sjukhuset, vilka inte hade skador eller sjukdomar i centrala nervsystemet. Gruppen bestod av 29 män och 17 kvinnor. Medelåldern var 42 år och gruppens medel-IQ var 102. Ingen försöksperson hade mer än 10 minuspoäng när uppgiftspoängen (maxpoäng 61) beräknades. Den svenska versionen av denna testversion har prövats ut på en liten normalgrupp om 15 personer (Lindström & Revstedt, 1975). Åldern på försökspersonerna varierade mellan 25 och 76 år. Åtta personer var under 50 år och fem personer hade studentexamen. Ingen försöksperson hade mer än 7 minuspoäng när itempoängen beräknades på testet.

Orgass och Poeck (1966) prövade ut sin tyska testversion på 51 friska individer äldre än 15 år. Knappt hälften av gruppen rekryterades bland patienter utan skador eller sjukdomar i centrala nervsystemet. Hur de resterande försökspersonerna rekryterades framgår inte. Könsfördelningen i gruppen var lika och medelåldern var 44 år. Sju personer hade 12 års utbildning eller mer, 13 personer hade ungefär 10 års utbildning och 31 personer hade ungefär åtta års utbildning. Ingen försöksperson hade mer än 11 minuspoäng då itempoängen beräknades.

Testversionen av Boller och Vignolo har prövats ut på normalgrupper i två studier. Boller och Vignolo (1966) testade personer, vilka rekryterades bland patienter utan skador eller sjukdomar i centrala nervsystemet, med den italienska testversionen. Medelåldern i gruppen var 41 år och medelvärdet av antalet utbildningsår var 8.35 år. Ingen försöksperson hade mer än 4 minuspoäng när uppgiftspoängen på testet beräknades (maxpoäng 62). Swisher och Sarno (1969) testade 22 normala personer med den engelska versionen av testet. Könsfördelningen i gruppen var lika och medelåldern var 56 år. Fyra personer hade mer än 12 års utbildning, 15 personer hade 9-12 års utbildning och 3 personer hade 6-8 års utbildning. Ingen försöksperson hade mer än 14 minuspoäng när uppgiftspoängen på testet beräknades.

De Renzi och Faglioni (1978) rekryterade en normalgrupp om 215 personer från patienter utan skador eller sjukdomar i centrala nervsystemet, anhöriga till patienter, sjuksköterskor och bekanta till forskarna. Försökspersonernas kön, ålder och utbildningsnivå anges inte i studien. Försökspersonerna fick utföra den italienska versionen av forskarnas förkortade testversion. Ingen försöksperson hade mer än 15 minuspoäng när uppgiftspoängen på testet beräknades (maxpoäng 36).

Den testversion, som ingår i testet MAE, prövades ut på en normalgrupp i en studie av Benton m.fl. (1978). Gruppen bestod av 360 individer mellan 16 och 69 år, vilka hade engelska som

modersmål, inte uppvisade tecken på hjärnskada eller hade någon psykiatrisk sjukdomshistoria. Maxpoängen på testet, som består av 22 uppgifter, var 44 eftersom 2 poäng gavs om en uppgift utfördes korrekt vid första försöket och 1 poäng om den utfördes korrekt vid andra försöket. Ingen försöksperson hade mer än 8 minuspoäng på testet.

Förhållandet mellan testresultat och kön, ålder respektive utbildning

I ovan beskrivna och även i ytterligare några studier har också undersökts hur faktorer som kön, ålder och utbildning påverkar testresultatet hos normala personer.

I studierna av Orgass och Poeck (1966), van Dongen och van Harskamp (1972), Benton m.fl. (1978) samt i en studie av Rey och Benton (1991) fann man inga könsskillnader i testresultat. I den sistnämnda studien prövades den spanska testversionen i MAE ut i USA på 234 personer, vilka hade spanska som huvudspråk och som var mellan 18 och 70 år. Det finns inga studier som visar på könsskillnader i testresultat på Token test.

När det gäller ålderns påverkan på testresultatet är forskningsresultaten inte entydiga. I ett flertal studier (Orgass & Poeck, 1966; Swisher & Sarno, 1969; van Dongen & van Harskamp, 1972; de Renzi & Faglioni, 1978; Benton m.fl., 1978) påverkade åldern inte resultatet. I andra studier har motstridiga resultat rapporterats. Wertz fann i en studie (1971) refererad av Brookshire och Manthie (1980), en gradvis ökning av antalet fel med ökad ålder. Emery (1986) fann att normala personer i åldersgruppen 75-93 år presterade signifikant sämre än personer i åldersgruppen 30-42 år vid utförandet av testversionen av Spreen och Benton. I studien av Rey och Benton (1991) påverkade åldern också resultatet signifikant på så sätt att ökad ålder medförde ett sämre testresultat.

Samtliga studier (förutom Benton m.fl., 1978) där utbildningsnivåns betydelse för testresultatet har undersökts (Orgass & Poeck, 1966; Boller & Vignolo, 1966; Swisher & Sarno, 1969; de Renzi & Faglioni, 1978; Lindström & Revstedt, 1975; Rey & Benton, 1991) har visat på ett signifikant samband mellan utbildningsnivå och testresultat på så sätt att högre utbildningsnivå innebär ett bättre testresultat.

Syfte

Förutom i den mycket begränsade studie, som utfördes av Lindström och Revstedt (1975), har de svenska versionerna av Token test, trots att de har använts i flera decennier i Sverige, inte prövats ut på normala, vuxna personer. Även om Token test har prövats ut på normalgrupper i flera andra länder och på flera andra språk, finns det flera skäl att anta att resultaten från dessa studier inte problemfritt kan tillämpas på svenska förhållanden. Som påpekas både av testkonstruktörerna själva (de Renzi & Vignolo, 1962; de Renzi & Faglioni, 1978) och av Swisher och Sarno (1969) skiljer sig olika språk åt i sin uppbyggnad. Token test på olika språk är därför inte helt jämförbara. Utbildningsnivåns betydelse för testresultatet har fastställts i ett antal studier. I Sverige är utbildningsnivån högre idag än den var vid den tidpunkt då Token test prövades ut i andra länder på andra språk. Hypotesen är därför att gränsvärdena för de svenska versionerna är högre än i de motsvarande utländska. De gränsvärden för normalprestation som anges i utländska studier är inte heller helt tillämpbara på de svenska testversionerna eftersom poängsättningen i flera fall skiljer sig åt. Dessutom är tillförlitligheten i flera av de refererade studierna inte helt tillfredsställande eftersom antalet försökspersoner är få. I flera studier understiger antalet försökspersoner 50 stycken (Orgass & Poeck, 1966; Boller & Vignolo, 1966; Swisher & Sarno, 1969; van Dongen & van Harskamp, 1972; Lindström & Revstedt,

1975). Inte i någon av de tidigare studierna har heller rekryteringen till normalgrupperna skett genom slumpvis urval.

Syftena med föreliggande studie är:

- att fastställa normalvärden för de versioner av Token test, som är översatta till svenska och som används i Sverige
- att jämföra de undersökta testversionerna avseende dessa värden.

En förhoppning är att studien ska bidra till en mer tillförlitlig diagnostik av auditiva, impressiva, språkliga störningar vid skador och sjukdomar i centrala nervsystemet, vilket är av stor betydelse i såväl klinisk verksamhet som forskning.

Metod

Försökspersoner

En grupp om 96 försökspersoner i åldern 21-80 år valdes slumpmässigt ut från Malmö kommuns befolkningsregister. Försökspersonerna delades in i sex åldersgrupper (21-30 år, 31-40 år osv.) och varje grupp om 16 personer hade lika könsfördelning. Rekryteringen till varje grupp skedde konsekutivt tills den var fylld. Inklusionskriterierna var svenska som modersmål, normal syn (vid behov med glasögon), normalt färgseende, normal hörsel, avsaknad av sjukdomar och skador i hjärnan inklusive psykiatriska sjukdomar, avsaknad av medfött intellektuellt funktionshinder och avsaknad av missbruk av alkohol och/eller narkotika. Att försökspersonerna uppfyllde inklusionskriterierna prövades på följande sätt. I svarsformuläret till det inbjudningsbrev som skickades ut, fanns fem frågor avseende inklusionskriterierna. Utifrån svaren på dessa frågor kallades försökspersonerna per telefon till en läkarundersökning. Vid denna utfördes en formaliserad intervju avseende relevanta medicinska, psykologiska och sociala frågor, en allmänmedicinsk och neurologisk läkarundersökning, en hörselprövning och en synundersökning. Undersökningen utfördes av en legitimerad läkare på Neurologiska kliniken, Universitetssjukhuset MAS i Malmö. Försökspersonerna fick också genomgå den neuropsykologiska screeningundersökningen Mini Mental Status Examination (Folstein, Folstein & McHugh, 1975). Lägre än 24/30 poäng ansågs som patologiskt, vilket medförde exklusion ur studien. Dessutom fick försökspersonerna utföra en depressionsbedömning med självskattningsinstrumentet Beck Depression Inventory (Beck & Steer, 1996). ≥ 15 poäng ansågs som tecken på depression, vilket medförde exklusion ur studien.

Fyra personer exkluderades vid läkarundersökningen (två p.g.a. allvarligt missbruk, en p.g.a. vid läkarundersökningen upptäckt MS och en p.g.a. svår depression). Medelåldern för hela gruppen om 96 personer var 50 år och 110 dagar. Det fanns ingen signifikant skillnad i ålder mellan kvinnor och män ($t(94) = 0.02$, $p = .958$). Fördelningen av deltagarnas högsta utbildning framgår av tabell 1.

Tabell 1 Försökspersonernas högsta utbildning

Utbildning	Antal fpp
Oavslutad folk- eller grundskola	0
Folkskola (6-8 år)	14
Grundskola, flickskola, realexamen	17
Studentexamen eller 2-4 årigt gymnasium	31
Utbildning minst 1 år utöver gymnasium	11
Examen från universitet eller högskola	23

Det fanns ingen signifikant skillnad mellan kvinnor och män när det gäller utbildningsnivå ($t(94) = -0.30, p = .765$). Däremot fanns ett signifikant negativt samband mellan ålder och utbildningsnivå på så sätt att en högre ålder medförde en lägre utbildningsnivå ($r = -.55, p < .0001$).

Material

De tre versioner av Token Test, som är översatta till svenska, ingick i studien. De tre versionerna var:

A: Token Test av de Renzi och Vignolo (1962). Svensk översättning och bearbetning av Fex (årtal okänt), Lindström och Revstedt (1975).

B: Token Test – kortversion ur testet Multilingual Aphasia Examination (MAE) av Benton, Hamsher och Sivan (1978). Svensk översättning och bearbetning av icke känd person.

C: Token Test – förkortad version av de Renzi och Faglioni (1978). Svensk översättning och bearbetning av Holmberg (årtal okänt).

Token Test av de Renzi och Vignolo består av fem avsnitt. De fyra första avsnitten, vilka består av 10 uppgifter vardera, har en likartad och enkel grammatisk struktur. För varje avsnitt ökar uppgifterna i längd från uppgifter av typen ”Peka på den vita cirkeln” till uppgifter av typen ”Peka på den lilla gröna cirkeln och den stora röda fyrkanten”. De 21 uppgifterna i det femte avsnittet prövar förståelse av grammatiskt komplexa utsagor. Testet består av totalt 61 uppgifter, som i sin tur består av 261 items.

Inför studien gjordes följande förändringar av uppgifterna: I avsnitt IV placerades ordet, som betecknar storlek alltid före ordet, som betecknar färg. I alla uppgifter i avsnitten I-IV och i uppgifterna V11, V17 och V18 ersattes ordet ”tag” med orden ”peka på”. I uppgift V6 ersattes ordet ”tag” med orden ”rör vid”. I uppgift V19 ändrades ”plockat upp” till ”tagit upp”. Genomgående ändrades tilltalsordet ”ni” till det numera oftare använda ordet ”du”.

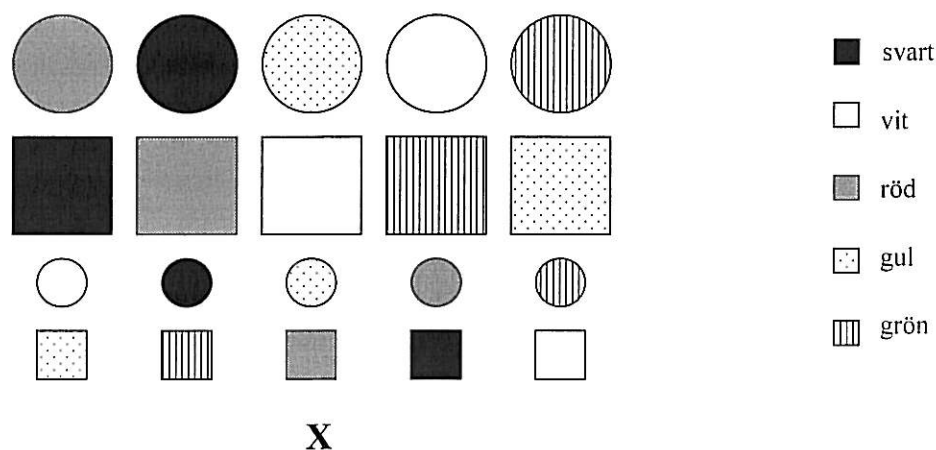
Token Test – kortversion ur testet MAE av Benton, Hamsher och Sivan består av två avsnitt. Det första avsnittet består av 10 uppgifter med en likartad och enkel grammatisk struktur. Uppgifterna ökar successivt i längd från uppgifter av typen ”Peka på en cirkel” till uppgifter av typen ”Tag upp den lilla gula cirkeln och den stora svarta fyrkanten”. De 12 uppgifterna i det andra avsnittet prövar förståelse av grammatiskt komplexa utsagor. Testet består av totalt 22 uppgifter, som i sin tur består av 108 items.

Token Test – förkortad version av de Renzi och Faglioni består av sex avsnitt. Det första avsnittet består av sju uppgifter och de fyra följande avsnitten av fyra uppgifter vardera, vilka alla har en likartad och enkel grammatisk struktur. För varje avsnitt ökar uppgifterna i längd från uppgifter av typen ”Peka på en cirkel” till uppgifter av typen ”Rör vid den lilla gröna cirkeln och den stora röda fyrkanten”. De 13 uppgifterna i det sjätte avsnittet prövar förståelse av grammatiskt komplexa utsagor. Testet består av totalt 36 uppgifter, som i sin tur består av 134 items.

De tokenbitar som användes vid testningen var tillverkade av 3 mm tjock plast. Sidan på de stora fyrkanterna och diametern på de stora cirkelarna var 25 mm; sidan på de små fyrkanterna och diametern på de små cirkelarna var 19 mm.

Tillvägagångssätt

Testningen med de tre versionerna av Token test utfördes av två erfarna logopeders i ett tyst rum utan störande inslag i samband med läkarundersökningen. Försökspersonen satt mitt emot testledaren vid ett bord, där tokenbitarna var arrangerade framför försökspersonen på samma sätt i alla testversionerna för alla försökspersoner (fig. 1). Placeringen av tokens överensstämde med den i den förkortade versionen av Token test av de Renzi och Faglioni. De tre olika testversionerna gavs i en motbalanserad ordning mellan de olika försökspersonerna. Före testningen fick alla försökspersoner samma muntliga information efter en skriven instruktion.



Figur 1 Tokenbitarnas placering framför försökspersonen, som är markerad med ett X.

Testningen utfördes enligt ett fastställt tillvägagångssätt: Uppgifterna gavs muntligt av testledaren, som talade med en normal, konstant talhastighet och utan att använda emfatisk betoning. Testledarna tränade detta moment tillsammans före testningarna. Varje uppmaning gavs bara en gång men om försökspersonen inte uppfattade en uppmaning p.g.a. ett störande moment i omgivningen (t.ex. en nysning) gick testledaren vidare till nästa uppgift och sedan tillbaka till den förra. Efter den långa versionen av testet gjordes 10 minuters paus och efter de korta versionerna gjordes 5 minuters paus. Testledaren uppmanade försökspersonerna att vara tysta och att gärna röra sig i rummet under pausen.

Poängsättning

Vad som poängsattes skilde sig något åt mellan de olika testversionerna. I kortversionen av Benton m.fl. poängsattes även de inledande uppmaningarna "peka på" och "tag upp" i det första avsnittet, vilket inte var fallet i motsvarande avsnitt i de två andra versionerna. I de två förkortade versionerna poängsattes både varje item och varje uppgift; i den långa versionen poängsattes endast varje item.

På testprotokollet till varje version av testet noterades en korrekt respons med "+" och en felaktig respons med "-". I de uppgifter som innehåller prepositioner markerades försökspersonens respons i för ändamålet avsedda rutor i testprotokollet. För varje test mättes tidsåtgången med ett tidtagarur från det att den första uppmaningen började ges till dess att försökspersonen hade utfört den sista uppgiften i testet. Den erhållna tiden avrundades till närmaste hela eller halva minut.

Poängsättningen av testet skedde på följande sätt. Korrekt respons (+) gav 1 poäng; felaktig eller utebliven respons (-) gav 0 poäng. I de två förkortade versionerna av testet, där varje uppgift också poängsattes som en helhet, gavs 1 poäng om hela uppgiften utfördes korrekt. Utfördes ett eller flera moment i uppgiften felaktigt gavs 0 poäng. I de uppgifter som går ut på att röra token X med token Y gavs -1 poäng för omkastning. Både förfarandet att lyfta upp en token och röra med den vid den andra och att skjuta den ena tokenbiten mot den andra gav pluspoäng. I de uppgifter där man ska röra vid tokens "långsamt" respektive "fort" gav både tillvägagångssättet att röra vid en token i taget på angivet sätt och att vidröra alla tokens i en viss kategori samtidigt på angivet sätt full poäng. I de uppgifter där uppmaningen är att manipulera en token "samtidigt med" en annan eller att "förutom" att manipulera en token även en annan, gav både tillvägagångssättet att manipulera bitarna på en gång och att manipulera dem omedelbart efter varandra pluspoäng. I del V i den långa versionen av testet bedömdes inte uppgifterna 1 och 2 förrän uppgift 8 hade utförts. För att få pluspoäng krävdes ett konsekvent utförande, dvs. utförandet i uppgifterna 1 och 2 måste vara lika och motsatt det i uppgift 8.

För alla versionerna av testet beräknades en "itemtotalsumma", dvs. den totalpoäng på testet som erhålls när poängen på alla items i testet summeras och en "itemavsnittsumma", dvs. poängen på varje avsnitt i testet när poängen på alla items i ett visst avsnitt summeras. I de två förkortade versionerna beräknades dessutom en "uppgifttotalsumma", dvs. den totalpoäng på testet som erhålls när poängen på alla uppgifter i testet summeras och en "uppgiftavsnittsumma", dvs. den poäng på varje avsnitt i testet som erhålls när poängen på alla uppgifter i ett visst avsnitt summeras.

Statistiska beräkningar

Statistiska beräkningar gjordes med StatView. Skillnader beräknades med tvåsidiga t-test. Skillnader mellan flera kategorier utvärderades med variansanalys, följda av posthoc-test (Fishers PLSD). Samband mellan variabler beräknades med Pearsonkorrelation. Reliabilitet beräknades med Cronbachs alpha. Signifikansgräns sattes genomgående till $p < .05$.

Etikprövning

Projektet är godkänt av Forskningsetikkommittén, Medicinska Fakulteten, Lunds Universitet (LU 319-02). En begäran att utföra datortomografi av hjärnan på försökspersoner äldre än 50 år avslogs.

Resultat

Version A: Token Test (de Renzi & Vignolo, 1962)

Kvantitativ analys av testresultatet

Den högsta möjliga itempängsumman (maxpoäng), försökspersonernas medelpoäng, lägsta och högsta poäng samt standardavvikelsen på testet som helhet och på varje avsnitt i testet framgår av tabell 2.

Tabell 2 Maxpoäng (itempoängssumma), medelpoäng, lägsta poäng, högsta poäng, standardavvikelse på testet som helhet och på varje avsnitt

	maxpoäng	M	min	max	S
hela testet	261	259.5	253	261	1.9
avsnitt I	20	20.0	20	20	0.0
avsnitt II	30	30.0	29	30	0.1
avsnitt III	40	39.9	37	40	0.4
avsnitt IV	60	60.0	59	60	0.2
avsnitt V	111	109.7	103	111	1.5

Den procentuella fördelningen av försökspersonernas itempoäng på testet som helhet och på varje avsnitt framgår av tabell 3.

Tabell 3 Itempoängfördelning på testet som helhet och på varje avsnitt

hela testet		avsnitt I		avsnitt II		avsnitt III		avsnitt IV		avsnitt V	
poäng	% fpp	poäng	% fpp	poäng	% fpp	poäng	% fpp	poäng	% fpp	poäng	% fpp
261	34	20	100	30	98	40	94	60	96	111	38
260	32			29	2	39	3	59	4	110	33
259	14					38	2			109	10
258	8					37	1			108	8
257	2									107	5
256	3									106	1
255	3									105	2
254	1									104	0
253	2									103	2
Σ	99*		100		100		100		100		99*

*Pga avrundning blir värdet ej 100

Som framgår av tabellerna hade 34 % av försökspersonerna full poäng på testet. Sextiosex procent av försökspersonerna hade högst en minuspoäng på hela testet. Samtliga försökspersoner utförde avsnitt I felfritt och ingen försöksperson hade mer än åtta minuspoäng på hela testet.

Försökspersonerna gjorde tillsammans fel på 0.6 % av testets samtliga items.

Försökspersonerna gjorde tillsammans fel på 0 % av samtliga items i avsnitt I, på 0.1 % av samtliga items i avsnitt II, på 0.3 % av samtliga items i avsnitt III, på 0.1 % av samtliga items i avsnitt IV och på 1.2 % av samtliga items i avsnitt V. På avsnitten I-IV tillsammans gjorde försökspersonerna fel på 0,1 % av samtliga items. Således begicks en betydligt större andel fel i avsnitt V än i de övriga avsnitten.

Kvalitativ analys av testresultatet

I avsnitten I-IV fanns ingen skillnad mellan andelen fel i förhållande till antalet möjliga, som gjordes på de items som betecknar färg (0.12 %) och andelen fel i förhållande till antalet möjliga, som gjordes på de items som betecknar form (0.16 %). Inga fel begicks på items som betecknar storlek.

Det fanns ingen tydlig tendens att en uppgifts placering i avsnittet eller ett items placering i en uppgift hade någon betydelse för resultatet.

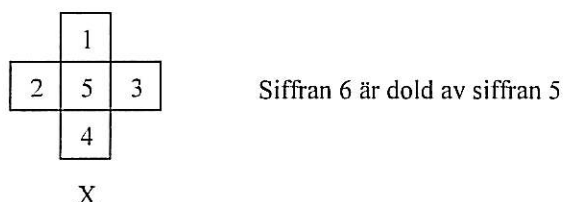
I avsnitt V gjordes sammanlagt fel på 4.7 % av de items som betecknar logiko-grammatiska strukturer (inkl. orden ”och” och ”eller”), på 0.7 % av de items som betecknar färg, på 0.5 % av de items som betecknar form, på 0.4 % av de items som är verb och på 0.7 % av övriga items (t.ex. ”långsamt”, ”fort”, ”nej”). Den största andelen fel gjordes således på logiko-grammatiska strukturer samt på orden ”och” och ”eller” när de förekommer i uppgifter i omedelbar anslutning till uppgifter, som har en likartad struktur förutom att de innehåller en logiko-grammatisk struktur. Övriga typer av fel fördelar sig ungefär lika.

De items (understrukna) som flest försökspersoner misslyckades med var:

- I uppgift V5: ”Rör vid den svarta cirkeln och den röda fyrkanten.” (22 % av fpp)
- I uppgift V2: ”Lägg den vita fyrkanten bakom den gula cirkeln.” (12 % av fpp)
- I uppgift V1: ”Lägg den röda cirkeln bakom den gröna fyrkanten.” (10 % av fpp)
- I uppgift V8: ”Lägg den vita cirkeln framför den svarta fyrkanten.” (9 % av fpp)
- I uppgift V6: ”Rör vid den svarta cirkeln eller den röda fyrkanten.” (7 % av fpp)
- I uppgift V17: ”Peka på den gula cirkeln istället för den vita fyrkanten.” (7 % av fpp)

Den enda uppgift som samtliga försökspersoner klarade felfritt i avsnitt V var V20: ”Lägg den svarta cirkeln under den vita fyrkanten.”.

I de uppgifter i avsnitt V som innehåller uppmaningar att placera en tokenbit ”bakom”, ”framför”, ”bredvid” och ”under” en annan tokenbit finns alternativa lösningar. Av tabell 4 framgår hur de försökspersoner som fått ”+” på aktuellt item placerade biten i fråga. Av figur 2 framgår vad numreringen av de olika placeringarna innebär.



Figur 2 X:et visar försökspersonens placering i förhållande till tokenbitarna. Siffran 5 visar placeringen av det token försökspersonen agerade mot och att försökspersonen lade aktuellt token ovanpå det token han agerade mot. Siffran 6 innebär motsatsen, dvs. att försökspersonen lade aktuellt token så att det doldes av det token han agerade mot.

Tabell 4 Procentuell fördelning av försökspersonernas placering av tokenbitar i de uppgifter i avsnitt V som innehåller prepositionerna ”bakom”, ”framför”, ”bredvid” och ”under”

uppgift/item	V1: bakom	V2: bakom	V8: framför	V12: bredvid	V20: under
placering	% försökspersoner				
1	64	60	31	0	0
2	0	1	6	68	0
3	2	1	1	32	0
4	27	31	56	0	9
5	0	0	6	0	0
6	7	7	0	0	91

Förhållandet mellan testresultat och kön, ålder respektive utbildning

Det fanns ingen signifikant skillnad mellan kvinnor och män när det gäller itemtotalpoängen på testet ($t(94) = -0.27, p = .785$). Det fanns inte heller någon signifikant skillnad mellan kvinnor och män när det gäller den sammanlagda poängen på avsnitten I – IV ($t(94) = 0.79, p = .432$) eller på avsnitt V ($t(94) = -0.54, p = .592$).

Det fanns ett signifikant negativt samband mellan itemtotalpoäng på testet och ålder på så sätt att ju högre ålder försökspersonen hade desto lägre var totalpoängen ($r = -.38, p < .0001$).

En ensidig variansanalys visade att det fanns ett generellt inflytande av ålder på så sätt att högre ålder medförde såväl lägre sammanlagd poäng på avsnitten I-IV ($F(5,90) = 3.09, p = .013$) som lägre poäng på avsnitt V ($F(5,90) = 4.40, p < .01$). I avsnitten I-IV visade parvisa jämförelser med Fishers PLSD att det fanns signifikanta skillnader i testresultat mellan den äldsta åldersgruppen (71-80 år) och de åldersgrupper som redovisas i tabell 5. Övriga parvisa jämförelser uppvisade inga signifikanta skillnader.

Tabell 5 Parvisa jämförelser med Fishers PLSD mellan de åldersgrupper som visade signifikanta skillnader i testresultat i avsnitten I-IV

åldersgrupp	åldersgrupp	p-värden
21 – 30	71 – 80	.005
31 – 40	71 – 80	.002
51 – 60	71 – 80	.013
61 – 70	71 – 80	.016

I avsnitt V visade parvisa jämförelser med Fishers PLSD att det fanns signifikanta skillnader i testresultat mellan de åldersgrupper som redovisas i tabell 6. Övriga parvisa jämförelser uppvisade inga signifikanta skillnader.

Tabell 6 Parvisa jämförelser med Fishers PLSD mellan de åldersgrupper som visade signifikanta skillnader i testresultat i avsnitt V

åldersgrupp	åldersgrupp	p-värden
21 – 30	61 – 70	.011
21 – 30	71 – 80	.046
31 – 40	61 – 70	.001
31 – 40	71 – 80	.0003
41 – 50	61 – 70	.001
41 – 50	71 – 80	.006
51 – 60	61 – 70	.020

Det fanns ett signifikant samband mellan itemtotalpoängen på testet och utbildningsnivå på så sätt att högre utbildningsnivå medförde högre totalpoäng ($r = .49, p < .0001$).

En ensidig variansanalys visade att utbildningsnivån inte hade något inflytande på den sammanlagda poängen på avsnitten I-IV ($F(4,91) = 1.60, p = .181$). Däremot fanns det ett signifikant inflytande av utbildningsnivå på poängen i avsnitt V på så sätt att högre utbildning medförde en högre poäng på avsnittet ($F(4,91) = 10.16, p < .0001$).

I avsnitt V visade parvisa jämförelser med Fishers PLSD att det fanns signifikanta skillnader i testresultat mellan de utbildningsnivåer som redovisas i tabell 7. Övriga parvisa jämförelser uppvisade inga signifikanta skillnader.

Tabell 7 Parvisa jämförelser med Fishers PLSD mellan de utbildningsnivåer som visade signifikanta skillnader i testresultat i avsnitt V

utbildning	utbildning	p-värden
Folkskola (6-8 år)	Grundskola, flickskola, realexamen	.001
Folkskola (6-8 år)	Studentexamen eller 2-4 årigt gymnasium	< .0001
Folkskola (6-8 år)	Utbildning minst 1 år utöver gymnasium	< .0001
Folkskola (6-8 år)	Examen från universitet eller högskola	< .0001
Grundskola, flickskola, realexamen	Examen från universitet eller högskola	.013

Reliabilitet

Testets reliabilitet var .60.

Tidsåtgång och förhållandet mellan tidsåtgång och kön, ålder respektive utbildning

Medeltidsåtgången för hela testet var 9 min 6 sek (kortaste tiden var 7 min; längsta tiden var 12 min 30 sek), $S = 1.3$.

Det fanns ingen signifikant skillnad i tidsåtgång mellan könen ($t(94) = 0.38$, $p = .702$).

Det fanns ett signifikant samband mellan den tid det tog att utföra testet och ålder på så sätt att ju högre ålder försökspersonen hade desto längre tid tog det att utföra testet ($r = .62$, $p < .0001$).

Det fanns ett signifikant negativt samband mellan tidsåtgång och utbildningsnivå på så sätt att ju högre utbildningsnivå försökspersonen hade desto kortare tid behövdes för att utföra testet ($r = -.46$, $p < .0001$).

Version B: Token Test – kortversion ur testet MAE (Benton, Hamsher & Sivan, 1978)

Kvantitativ analys av testresultatet

Den högsta möjliga itempoängssumman (maxpoäng), försökspersonernas medelpoäng, lägsta och högsta poäng samt standardavvikelsen på testet som helhet och på varje avsnitt i testet framgår av tabell 8.

Tabell 8 Maxpoäng (itempoängssumma), medelpoäng, lägsta poäng, högsta poäng, standardavvikelse på testet som helhet och på varje avsnitt

	maxpoäng	M	min	max	S
hela testet	108	107.3	105	108	0.6
avsnitt I	45	44.9	43	45	0.2
avsnitt II	63	62.8	60	63	0.5

Den högsta möjliga uppgiftpoängssumman (maxpoäng), försökspersonernas medelpoäng, lägsta och högsta poäng samt standardavvikelsen på testet som helhet och på varje avsnitt i testet framgår av tabell 9.

Tabell 9 Maxpoäng (uppgiftpoängssumma), medelpoäng, lägsta poäng, högsta poäng, standardavvikelse på testet som helhet och på varje avsnitt

	maxpoäng	M	min	max	S
hela testet	22	21.8	20	22	0.5
avsnitt I	10	10.0	9	10	0.3
avsnitt II	12	11.8	10	12	0.6

Den procentuella fördelningen av försökspersonernas itempoäng på testet som helhet och på varje avsnitt framgår av tabell 10.

Tabell 10 Itempoängfördelning på testet som helhet och på varje avsnitt

hela testet		avsnitt I		avsnitt II	
poäng	% fpp	poäng	% fpp	poäng	% fpp
108	85	45	96	63	90
107	5	44	3	62	2
106	8	43	1	61	7
105	1			60	1
Σ	99*	100		100	

* Pga avrundning blir värdet ej 100

Den procentuella fördelningen av försökspersonernas uppgiftpoäng på testet som helhet och på varje avsnitt framgår av tabell 11.

Tabell 11 Uppgiftpoängfördelning på testet som helhet och på varje avsnitt

hela testet		avsnitt I		avsnitt II	
poäng	% fpp	poäng	% fpp	poäng	% fpp
22	85	10	96	12	90
21	10	9	4	11	6
20	4			10	4
Σ	99*	100		100	

* Pga avrundning blir värdet ej 100

Som framgår av tabellerna hade 85 % av försökspersonerna full poäng på testet. Ingen försöksperson hade mer än tre minuspoäng på hela testet då summan av itempoängen beräknades och ingen försöksperson hade mer än två minuspoäng då summan av uppgiftpoängen beräknades.

Försökspersonerna gjorde tillsammans fel på 0.2 % av samtliga items i testet, på 0.1 % av samtliga items i avsnitt I och på 0.3 % av samtliga items i avsnitt II. Försökspersonerna gjorde sammanlagt fel på 0.9 % av samtliga uppgifter i testet, på 0.4 % av samtliga uppgifter i avsnitt I och på 1.3 % av samtliga uppgifter i avsnitt II. Således begicks en större andel fel i avsnitt II än i avsnitt I.

Kvalitativ analys av testresultatet

I avsnitt I gjordes sammanlagt endast fel på fem items och andelen fel som gjordes på items som betecknar färg, form och storlek skilde sig inte åt.

Inga fel gjordes i de sex första uppgifterna i avsnitt I, vilka endast innehåller uppmaningar att agera mot en tokenbit. I övrigt fanns ingen tydlig tendens att en uppgifts placering i avsnittet eller att ett items placering i en uppgift hade någon betydelse för resultatet.

I avsnitt II gjordes fler fel (0.8 %) i förhållande till antalet möjliga fel på items som betecknar logiko-grammatiska strukturer (inkl. ordet "eller") än på övriga items (0.3 %).

Det items (understruket) som flest försökspersoner misslyckades med var:

I uppgift II13: "Rör vid den vita cirkeln med den gröna fyrkanten." (4 % av fpp)

De uppgifter som flest försökspersoner misslyckades med var:

Uppgift II13: "Rör vid den vita cirkeln med den gröna fyrkanten." (4 % av fpp)

Uppgift II19: "Lägg den gula fyrkanten på den vita cirkeln." (4 % av fpp)

Uppgift II20: "Rör vid den svarta cirkeln med den röda fyrkanten." (4 % av fpp)

Samtliga försökspersoner klarade uppgifterna II11, II12, II14, II16, II17 och II18 felfritt.

Förhållandet mellan testresultat och kön, ålder respektive utbildning

Det fanns ingen signifikant skillnad mellan kvinnor och män varken när det gällde itemtotalpoängen på testet ($t(94) = -1.59, p = .116$) eller uppgifttotalpoängen ($t(94) = -1.50, p = .138$).

Det fanns inget signifikant samband mellan itemtotalpoäng på testet och ålder ($r = -.08, p = .426$) eller mellan uppgifttotalpoäng och ålder ($r = -.16, p = .123$).

Det fanns inget signifikant samband mellan itemtotalpoängen på testet och utbildningsnivå ($r = .11, p = .297$). Det fanns inte heller något signifikant samband mellan uppgifttotalpoängen och utbildningsnivå men här föreligger ett gränsvärde ($r = .20, p = .059$).

Reliabilitet

Testets reliabilitet var .45 när itempoängssumman beräknades och .22 när uppgiftpoängssumman beräknades.

Tidsåtgång och förhållandet mellan tidsåtgång och kön, ålder respektive utbildning.

Medeltidsåtgången för hela testet var 3 min 24 sek (kortaste tiden var 2 min 30 sek; längsta tiden var 18 min), $S = 1.6$. Anmärkas kan att den längsta tiden härrör från en försöksperson, som i alla uppgifter i denna testversion tog mycket god tid på sig att tänka igenom sin respons innan han gav den. Det näst högsta värdet var 7 min.

Det fanns ingen signifikant skillnad i tidsåtgång mellan könen ($t(94) = -1.01, p = .317$).

Det fanns inget signifikant samband mellan den tid det tog att utföra testet och ålder ($r = .02, p = .845$).

Det fanns inget signifikant samband mellan tidsåtgång och utbildningsnivå ($r = -.13, p = .198$).

Version C: Token Test – förkortad version (de Renzi & Faglioni, 1978)

Kvantitativ analys av testresultatet

Den högsta möjliga itempoängssumman (maxpoäng), försökspersonernas medelpoäng, lägsta och högsta poäng samt standardavvikelsen på testet som helhet och på varje avsnitt i testet framgår av tabell 12.

Tabell 12 Maxpoäng (itempoängssumma), medelpoäng, lägsta poäng, högsta poäng, standardavvikelse på testet som helhet och på varje avsnitt

	maxpoäng	M	min	max	S
hela testet	134	132.9	128	134	1.7
avsnitt I	7	7.0	7	7	0.0
avsnitt II	8	8.0	7	8	0.2
avsnitt III	12	12.0	12	12	0.0
avsnitt IV	16	16.0	15	16	0.1
avsnitt V	24	24.0	23	24	0.2
avsnitt VI	67	66.0	61	67	0.8

Den högsta möjliga uppgiftspoängssumman (maxpoäng), försökspersonernas medelpoäng, lägsta och högsta poäng samt standardavvikelsen på testet som helhet och på varje avsnitt i testet framgår av tabell 13.

Tabell 13 Maxpoäng (uppgiftspoängssumma), medelpoäng, lägsta poäng, högsta poäng, standardavvikelse på testet som helhet och på varje avsnitt

	maxpoäng	M	min	max	S
hela testet	36	35.4	33	36	0.9
avsnitt I	7	7.0	7	7	0.0
avsnitt II	4	4.0	3	4	0.1
avsnitt III	4	4.0	4	4	0.0
avsnitt IV	4	4.0	3	4	0.1
avsnitt V	4	4.0	3	4	0.2
avsnitt VI	13	12.5	10	13	1.6

Den procentuella fördelningen av försökspersonernas itempoäng på testet som helhet och på varje avsnitt framgår av tabell 14.

Tabell 14 Itempoängfördelning på testet som helhet och på varje avsnitt

hela testet		avsnitt I		avsnitt II		avsnitt III		avsnitt IV		avsnitt V		avsnitt VI	
poäng	% fpp	poäng	% fpp	poäng	% fpp	poäng	% fpp	poäng	% fpp	poäng	% fpp	poäng	% fpp
134	60	7	100	8	99	12	100	16	99	24	97	67	65
133	14			7	1			15	1	23	3	66	10
132	7											65	8
131	5											64	3
130	8											63	8
129	2											62	3
128	3											61	2
Σ	99*		100		100		100		100		100		99*

*Pga avrundning blir värdet ej 100

Den procentuella fördelningen av försökspersonernas uppgiftspoäng på testet som helhet och på varje avsnitt framgår av tabell 15.

Tabell 15 Uppgiftspoängfördelning på testet som helhet och på varje avsnitt

hela testet		avsnitt I		avsnitt II		avsnitt III		avsnitt IV		avsnitt V		avsnitt VI	
poäng	% fpp	poäng	% fpp	poäng	% fpp	poäng	% fpp	poäng	% fpp	poäng	% fpp	poäng	% fpp
36	60	7	100	4	99	4	100	4	99	4	97	13	65
35	23			3	1			3	1	3	3	12	20
34	11											11	13
33	5											10	3
Σ	99*		100		100		100		100		100		101*

* Pga avrundning blir värdet ej 100

Som framgår av tabellerna hade 60 % av försökspersonerna full poäng på testet, 74 % av försökspersonerna hade högst en minuspoäng när itemtotalsumman beräknades och 83 % av försökspersonerna hade högst en minuspoäng när uppgifttotalsumman beräknades. Ingen försöksperson hade mer än sex minuspoäng på hela testet då summan av itempoängen beräknades och ingen försöksperson hade mer än tre minuspoäng då summan av uppgiftspoängen beräknades. Samtliga försökspersoner utförde avsnitten I och III felfritt.

Försökspersonerna gjorde tillsammans fel på 0.8 % av testets samtliga items.

Försökspersonerna gjorde tillsammans fel på 0 % av samtliga items i avsnitt I, på 0.1 % av samtliga items i avsnitt II, på 0 % av samtliga items i avsnitt III, på 0.1 % av samtliga items i avsnitt IV, på 0.1 % av samtliga items i avsnitt V och på 1.5 % av samtliga items i avsnitt VI. På avsnitten I-V tillsammans gjorde försökspersonerna fel på 0,1 % av samtliga items.

Försökspersonerna gjorde tillsammans fel på 1.7 % av testets samtliga uppgifter.

Försökspersonerna gjorde tillsammans fel på 0 % av samtliga uppgifter i avsnitt I, på 0.3 % av samtliga uppgifter i avsnitt II, på 0 % av samtliga uppgifter i avsnitt III, på 0.3 % av samtliga uppgifter i avsnitt IV, på 0.8 % av samtliga uppgifter i avsnitt V och på 4.2 % av samtliga uppgifter i avsnitt VI. På avsnitten I-V tillsammans gjorde försökspersonerna fel på 0,1 % av samtliga uppgifter. Således begicks en avsevärt större andel fel i avsnitt VI än i de övriga avsnitten.

Kvalitativ analys av testresultatet

I avsnitten I–V gjorde försökspersonerna sammanlagt endast fem fel. Inga fel gjordes avseende färg.

Det fanns ingen tydlig tendens att en uppgifts placering i avsnittet eller ett items placering i en uppgift hade någon betydelse för resultatet.

I avsnitt VI gjordes sammanlagt fel på 3.8 % av de items som betecknar logiko-grammatiska strukturer (inkl orden ”och” och ”eller”), på 1.2 % av de items som är verb, på 1.0 % av de items som betecknar färg, på 1.0 % av de items som betecknar form och på 0.7 % av övriga items (t.ex. ”långsamt”, ”fort”, ”nej”). Den största andelen fel gjordes således på logiko-grammatiska strukturer samt på orden ”och” och ”eller” när de förekommer i uppgifter i omedelbar anslutning till uppgifter, som har en likartad struktur förutom att de innehåller en logiko-grammatisk struktur. Övriga typer av fel fördelar sig ungefär lika.

De items (understrukna) som flest försökspersoner misslyckades med var:

I uppgift VI26: ”Rör vid den svarta cirkeln och den röda fyrkanten.” (16 % av fpp)

I uppgift VI36: ”Förutom att röra vid den gula cirkeln ska Du röra vid den svarta cirkeln.” (13 % av fpp)

I uppgift VI36: ”Förutom att röra vid den gula cirkeln ska Du röra vid den svarta cirkeln.” (12 % av fpp)

I uppgift VI36: ”Förutom att röra vid den gula cirkeln ska Du röra vid den svarta cirkeln.” (12 % av fpp)

I uppgift VI36: ”Förutom att röra vid den gula cirkeln ska Du röra vid den svarta cirkeln.” (12 % av fpp)

Orsaken till de många felen på olika items i uppgift VI36 var att flera försökspersoner inte utförde det första ledet i uppmaningen.

De uppgifter som flest försökspersoner misslyckades med var:

VI26: ”Rör vid den svarta cirkeln och den röda fyrkanten.” (19 % av fpp)

VI36: ”Förutom att röra vid den gula cirkeln ska Du röra vid den svarta cirkeln.” (13 % av fpp)

VI27: ”Rör vid den svarta cirkeln eller den röda fyrkanten.” (7 % av fpp)

De uppgifter som samtliga försökspersoner klarade felfritt i avsnitt VI var:

VI29: ”Om det finns en blå cirkel, rör vid den röda fyrkanten.”

VI35: ”Rör vid den gula cirkeln istället för den vita fyrkanten.”

I den uppgift i avsnitt VI som innehåller uppmaningen att placera en tokenbit ”bredvid” en annan tokenbit finns alternativa lösningar. Av tabell 16 framgår hur de försökspersoner som fått ”+” på aktuellt item placerade biten i fråga. Vad numreringen av de olika placeringarna innebär framgår av figur 2.

Tabell 16 Procentuell fördelning av försökspersonernas placering av tokens i uppgiften i avsnitt VI innehållande prepositionen "bredvid" (uppgift VI30)

placering	% fpp
1	5
2	69
3	20
4	6
5	0
6	0

Förhållandet mellan testresultat och kön, ålder respektive utbildning

Det fanns ingen signifikant skillnad mellan kvinnor och män vare sig när det gäller itemtotalpoäng på testet ($t(94) = -0.49, p = .627$) eller uppgifttotalpoäng på testet ($t(94) = 0.34, p = .732$).

Det fanns inget signifikant samband mellan itemtotalpoängen på testet och ålder ($r = -.15, p = .138$). Däremot fanns ett signifikant negativt samband mellan den sammanlagda uppgifttotalsumman på testet och ålder ($r = -.31, p = .002$) på så sätt att ju högre ålder försökspersonen hade desto lägre var totalpoängen.

En ensidig variansanalys visade att åldern inte hade något inflytande på den sammanlagda uppgiftpoängen på avsnitten I-V ($F(5,90) = 1.78, p = .124$) eller på uppgiftpoängen i avsnitt VI ($F(5,90) = 1.67, p = .150$).

Det fanns ett signifikant samband mellan såväl itemtotalpoängen på testet och utbildningsnivå ($r = .34, p = .001$) som uppgifttotalpoängen på testet och utbildningsnivå ($r = .38, p = .0001$) på så sätt att högre utbildningsnivå medförde högre totalpoäng.

En ensidig variansanalys visade att utbildningsnivån inte hade något inflytande på den sammanlagda itempoängen på avsnitten I-V ($F(4,91) = 0.50, p = .738$) eller på itempoängen i avsnitt VI ($F(4,91) = 1.05, p = .386$).

Utbildningsnivån hade inte heller något inflytande på den sammanlagda uppgiftpoängen i avsnitten I-V ($F(4,91) = 0.50, p = .738$). Däremot uppvisade den ensidiga variansanalysen att utbildningsnivån hade ett signifikant inflytande på uppgiftpoängen i avsnitt VI ($F(4,91) = 3.09, p = .02$).

Parvisa jämförelser med Fishers PLSD visade att det fanns signifikanta skillnader i testresultat mellan de utbildningsnivåer som redovisas i tabell 17. Övriga parvisa jämförelser uppvisade inga signifikanta skillnader.

Tabell 17 Parvisa jämförelser med Fishers PLSD mellan de utbildningsnivåer som visade signifikanta skillnader i testresultat i avsnitt VI

utbildning	utbildning	p-värden
Folkskola (6-8 år)	Studentexamen eller 2-4 årigt gymnasium	.009
Folkskola (6-8 år)	Utbildning minst 1 år utöver gymnasium	.008
Folkskola (6-8 år)	Examen från universitet eller högskola	.003

Reliabilitet

Testets reliabilitet var .65 när itempoängsumman beräknades och .31 när uppgiftpoängsumman beräknades.

Tidsåtgång och förhållandet mellan tidsåtgång och kön, ålder respektive utbildning

Medeltidsåtgången för hela testet var 5 min 6 sek (kortaste tiden var 4 min; längsta tiden var 8 min 30 sek), $S = 0.7$.

Det fanns ingen signifikant skillnad i tidsåtgång mellan könen ($t(94) = 0.00$, $p \approx 1$).

Det fanns ett signifikant samband mellan den tid det tog att utföra testet och ålder ($r = .56$, $p < .0001$) på så sätt att ju högre ålder försökspersonen hade desto längre tid tog det att utföra testet.

Det fanns ett signifikant negativt samband mellan tidsåtgång och utbildningsnivå ($r = -.40$, $p < .0001$) på så sätt att ju högre utbildningsnivå försökspersonen hade desto kortare tid behövdes för att utföra testet.

Förhållandet mellan testversionernas totalpoäng

Sambandet mellan de tre testversionernas itemtotalpoäng framgår av tabell 18.

Tabell 18 Sambandet mellan testversionernas itemtotalpoäng

	A	B	C
A		.23*	.43***
B			.06
C			

* $p < .05$

** $p < .01$

*** $p < .0001$

Det fanns inget signifikant samband mellan uppgifttotalpoängen i version B och den i version C ($r = .14$, $p = .184$).

Diskussion

De tre versioner av Token test, som är översatta till svenska och som används av logopedier i Sverige, prövades ut på 96 normala individer i åldern 21-80 år med syftena att erhålla normalvärden för de olika testversionerna och att jämföra de tre versionerna avseende dessa värden.

Metod

Försökspersoner

Stor vikt lades vid att erhålla så tillförlitliga normalvärden som möjligt. Försökspersonerna valdes ut slumpvis ur befolkningen, vilket inte har varit fallet i någon tidigare motsvarande studie. Tillvägagångssättet gör resultaten mer generaliserbara än om t.ex. patienter, anhöriga till patienter, bekanta eller arbetskamrater till den aktuella forskaren rekryteras som försökspersoner. Det senare har varit fallet i tidigare publicerade studier av normalgruppers resultat på Token test. Stor vikt lades också vid att genom läkarundersökning, anamnes och testning försäkra sig om att försökspersonerna verkligen var friska ur relevant synvinkel. Bergman, Blomberg & Almkvist (2007) undersökte 118 försökspersoner mellan 26 och 91 år, vilka hade rekryterats och godkänts som friska kontroller i olika studier. Försökspersonerna fick genomgå en noggrann medicinsk undersökning med särskild inriktning på sjukdomar och skador i centrala nervsystemet samt systemsjukdomar. De fick också genomföra en serie kognitiva inklusive språkliga test. Det visade sig att fysiska hälsoproblem, vilka hade upptäckts vid den medicinska undersökningen, påverkade resultaten på de kognitiva testen mer än både utbildningsnivå och ålder. Vikten av att undersöka försökspersonernas hälsotillstånd framgick också av att fyra försökspersoner i föreliggande studie fick exkluderas p.g.a. fynd vid den medicinska undersökningen. Antalet försökspersoner var också förhållandevis stort. Endast vid utprövningen av de Renzis och Faglioni's förkortade, italienska version av testet (1978) och vid utprövningarna av den engelska och den spanska kortversionen av Token test i MAE (Benton m.fl., 1978; Rey & Benton, 1991) användes fler försökspersoner. För att göra resultaten ännu säkrare hade det dock varit önskvärt med ett ännu större antal försökspersoner, vilket dock omöjliggjordes av resursskäl.

Personer, som var äldre än 80 år, ingick inte i studien. Orsaken är att de lätta störningar i talförståelse, som Token test är avsett att diagnostisera, sällan har relevans för personer i denna åldersgrupp. Uppsatsförfattarens erfarenhet är att Token test framför allt är användbart för att testa yngre personer, där det finns krav på en hög förmåga att förstå det talade språket i arbetslivet eller vid fritidsaktiviteter.

Poängsättning, material och tillvägagångssätt

Vid utprövningen av de olika testversionerna användes samma tillvägagångssätt och samma poängsättningsmetod för samtliga testversioner. Syftet var att underlätta jämförelser mellan testversionerna, att öka möjligheten att låta testversionerna ersätta varandra vid t.ex. upprepad testning och att underlätta proceduren för den som utför testningen.

Något som dock skilde sig åt mellan testversionerna var att uppgiftspoängen inte beräknades för version A. Orsaken är att det i ett så omfattande test inte är meningsfullt att använda denna mindre noggranna poängsättningsmetod. Önskar man göra en mer screeningartad bedömning, bör man istället välja en kortare version av testet. I den ursprungliga testversionen av de Renzi's

och Faglioni och i den ursprungliga versionen i MAE används poängsättningsprincipen att en korrekt respons vid upprepad stimulus ger halva poängen jämfört med en korrekt respons vid första försöket. I föreliggande studie poängsattes endast ett försök. (Undantag var om ett störande inslag i omgivningen förhindrade försökspersonen att uppfatta den första uppmaningen. Detta inträffade dock inte vid något tillfälle.) Orsaken är att Token test inte är avsett att undersöka talförståelse av enskilda ord, vilket lätt blir fallet om testpersonen vid andra försöket endast koncentrerar sig på det ord han/hon var osäker på vid det första försöket. Trots att studier har visat att afatikens prestationer inte påverkas av hur tokens är placerade (Cohen, Gutbrod, Meier & Römer, 1987; Wold & Reinvang, 1990), arrangerades tokens på samma sätt vid alla testningar för att uppnå ett så standardiserat testningsförfarande som möjligt. Inför studien reviderades några uppmaningar i testversion A. I avsnitt IV placerades ordet, som betecknar storlek alltid före ordet, som betecknar färg, dels för att detta är en naturligare ordföljd i svenska, dels för att motsvarande uppgifter i de olika testversionerna skulle få en likartad uppbyggnad. I uppgift V19 ändrades "plockat upp" till "tagit upp" för att uppgifterna i alla testversionerna skulle få ett likartat ordval. I avsnitten I-IV och i några uppgifter i avsnitt V ersattes ordet "tag" med orden "peka på" eftersom den uppmaningen är lättare att utföra om man har nedsatt handmotorik. I uppgift V6 ersattes ordet "tag" med orden "rör vid" för att göra formuleringen av denna uppgift parallell med den i uppgifterna V3-5, vilket är fallet i motsvarande uppgifter i de andra testversionerna.

Även om fastställda uppmaningar ges och tokens är arrangerade på ett fastställt sätt kan testpersonens prestation påverkas av testledarens talhastighet. Testledaren måste därför vara uppmärksam på att han/hon inte anpassar sin talhastighet till testpersonens förmåga (Lezak, 1995). I en studie av Poeck och Pietron (1981) fick afasidrabbade personer signifikant bättre resultat på originalversionen av Token test, när uppmaningarna gavs med 1.25 gånger lägre hastighet. Liles och Brookshire (1975) rapporterade att afasidrabbade personers resultat på icke-grammatiskt komplexa uppgifter i en modifierad version av Token test förbättrades när korta pauser lades in i uppmaningarna. För att uppmaningarna i föreliggande studie skulle ges på ett så standardiserat sätt som möjligt, tränade testledarna tillsammans flera gånger före testningarna på att hålla en konstant, normal talhastighet och på att inte använda enfattig betoning i uppmaningarna. Ett alternativ hade varit att använda inspelade uppmaningar. Det finns dock flera nackdelar med detta förfaringsätt. Tal blir svårare att uppfatta om man inte samtidigt ser den person som talar. Detta gäller särskilt om man har nedsatt hörsel, vilket inte är ovanligt hos äldre afasidrabbade. Även om detta problem kan undvikas genom videoinspelade uppmaningar innebär att starta och stoppa uppspelningen ett störande och kommunikativt onaturligt moment. Inspelade uppmaningar har heller inte använts i någon tidigare studie med Token test.

Resultat

Kvantitativ analys av testresultatet

I version A var -8 poäng den högsta minuspoäng någon testperson hade när itemtotalpoäng beräknades och 34 % av försökspersonerna hade full poäng på testet. I version B var -3 poäng den högsta minuspoängen när itemtotalpoäng beräknades, -2 poäng när uppgifttotalpoäng beräknades och 85 % av försökspersonerna hade full poäng på testet. I version C var -6 poäng den högsta minuspoängen när itemtotalpoäng beräknades, -3 poäng när uppgifttotalpoäng beräknades och 60 % av försökspersonerna hade full poäng på testet. I den enda motsvarande svenska studie, som har utförts tidigare (Lindström & Revstedt, 1975), var resultaten ganska likartade. Den högsta minuspoängen på version A var -7 poäng och 40 % av försökspersonerna hade full poäng. I tidigare utländska studier av motsvarande eller jämförbara versioner av

Token test är den högsta minuspoängen högre och andelen försökspersoner med full poäng lägre än i föreliggande undersökning. Studier med testversioner som motsvarar version A är: Orgass och Poeck (1966) där -11 itempoäng var högsta minuspoäng; Boller och Vignolo (1966) där -4 uppgiftpoäng var högsta minuspoäng och 26 % av försökspersonerna hade full poäng när inga avdrag gjordes för upprepad stimulus; Swisher och Sarno (1969) där -14 uppgiftpoäng var högsta minuspoäng; van Dongen och van Harskamp (1972) där -10 uppgiftpoäng var högsta minuspoäng. Anmärkas kan, att om försökspersoner i de tre sistnämnda studierna hade fler fel än ett/uppgift, vilket förekom i föreliggande studie, skulle minuspoängen omräknad till itempoäng bli ännu högre. Studier med testversioner som motsvarar version B är: Benton m.fl. (1978) där -8 uppgiftpoäng var högsta minuspoäng och ca 30 % av försökspersonerna hade full poäng. Studier som motsvarar version C är: de Renzi och Faglioni (1978) där -15 uppgiftpoäng var högsta minuspoäng och 10 % av försökspersonerna hade full poäng.

Hypotesen att svenska informanter presterar bättre på Token test än informanter i andra länder, som utförde testet på andra språk, bekräftades således. En orsak kan vara att försökspersonerna i föreliggande studie kontrollerades noggrannare avseende sitt hälsotillstånd än i tidigare studier. En viktig orsak är sannolikt att utbildningsnivån är högre i Sverige idag än den var i många länder vid den tidpunkt Token test prövades ut där. Utbildningsnivåns påverkan på resultatet av Token test har fastställts i ett antal studier. I föreliggande studie hade 68 % av försökspersonerna 12 års utbildning eller mer, vilket är en högre andel än i någon tidigare, motsvarande studie. I flera utländska studier (Boller & Vignolo, 1966; de Renzi & Faglioni, 1978; Rey & Benton, 1991) förekom försökspersoner, som hade < 6 års utbildning. Ingen försöksperson i föreliggande studie hade så kort utbildning. En synpunkt, som har framförts av bl.a. Swisher och Sarno (1969), är att grammatiska och lexikala skillnader mellan olika språk också kan vara en bidragande orsak. För att kunna dra säkrare slutsatser om detta behövs jämförande studier mellan Token test på olika språk, där försöksgruppernas utbildningsnivå och ålder inte skiljer sig åt. Några sådana studier finns inte i dagsläget. Resultatet i föreliggande studie belyser dock vikten av att fastställa normalvärden genom att pröva ut test på försökspersoner från den aktuella populationen. Om utländska normer på Token test används i Sverige riskerar ett antal personer med afasi att felaktigt klassificeras som icke-afatiska.

Uppsatsförfattaren anser att den lägsta poäng, som någon försöksperson hade på en testversion i föreliggande studie, bör sättas som gränsvärde för normalprestation i den testversionen. Endast om speciella omständigheter föreligger när det gäller enstaka testpersoners låga testpoäng bör deras resultat exkluderas vid fastställande av gränsvärde. Så är dock inte fallet i flera studier (van Dongen & van Harskamp, 1972; Hartje m.fl., 1973; de Renzi & Faglioni, 1978), där gränsvärdet för normalprestation har satts högre än lägsta poäng utan att speciella omständigheter redovisas.

Kvalitativ analys av testresultatet

I de avsnitt, som inte innehåller några grammatiskt komplexa uppmaningar, fanns sammantaget ingen tydlig skillnad i andelen fel på items, vilka betecknar färg, form och storlek. Ett items placering i en uppgift eller en uppgifts placering i ett avsnitt hade heller ingen betydelse för resultatet. Det fanns en tendens att andelen fel i de längre uppmaningarna var större än i de kortare. Dessa resultat överensstämmer med de förväntade.

I alla versionerna var andelen fel i det sista avsnittet, som innehåller grammatiskt komplexa uppmaningar, större än i de tidigare avsnitten/det tidigare avsnittet. I alla versioner var också andelen fel på items, vilka betecknar logiko-grammatiska strukturer (samt orden "och" och

”eller” när de förekommer i uppgifter i omedelbar anslutning till uppgifter, som har en likartad struktur förutom att de innehåller en logiko-grammatisk struktur) avsevärt mycket större än andelen fel på övriga items i avsnittet. Av detta kan slutsatsen dras att den större svårighetsgraden i det sista avsnittet huvudsakligen beror på förekomsten av logiko-grammatiska strukturer. Fynden överensstämmer med dem i andra studier (Boller & Vignolo, 1966; Swisher & Sarno, 1969; Lindström & Revstedt, 1975).

Andelen fel på uppgifter av typen ”Rör vid den röda cirkeln och/eller den svarta fyrkanten” ökade avsevärt när uppgiften var placerad i omedelbar anslutning till uppgifter av typen ”Rör vid den röda cirkeln med den svarta fyrkanten”. Orsaken är sannolikt interferens från den föregående uppgiften/de föregående uppgifterna, dvs. tankemässiga perseverationer av den typ som ofta förekommer vid frontallobskador (Luria, 1980b) förekommer även hos normala personer om än i mindre omfattning.

Förhållandet mellan testresultat och kön, ålder respektive utbildning

Genomgående gäller att värdena för version B är osäkra p.g.a. att spridningen i testresultatet är så liten i denna testversion.

Det fanns inga signifikanta könsskillnader i testresultat i någon av testversionerna, vilket överensstämmer med fynden i tidigare studier (Orgass & Poeck, 1966; van Dongen & van Harskamp, 1972; Benton m.fl., 1978; Rey & Benton, 1991).

I version A presterade äldre personer signifikant sämre än yngre. Den allra äldsta åldersgruppen (71-80 år) presterade signifikant sämre på hela testet medan personer över 61 år bara presterade signifikant sämre på det sista testavsnittet. Mellan de övriga åldersgrupperna fanns inga signifikanta skillnader i resultat. Resultaten överstämmer inte med tidigare studier (Orgass & Poeck, 1966; Swisher & Sarno, 1969; van Dongen & van Harskamp, 1972), där inga signifikanta åldersskillnader i testresultat rapporterades. Åldersfördelningen bland försökspersonerna i dessa studier anges dock inte, vilket är anmärkningsvärt. En möjlig förklaring till skillnaden i resultat mellan dessa studier och föreliggande studie är att det inte ingick försökspersoner med lika hög ålder i dem. Detta kan också vara en förklaring till att resultatet i version C inte överensstämmer med resultaten i de Renzi och Faglioni studie (1978), där det inte förelåg några signifikanta ålderskillnader i testresultat när uppgiftpoängen beräknades. I version B fanns inga signifikanta åldersskillnader i testresultat, vilket överensstämmer med resultatet i den engelska testversionen. I den spanska testversionen presterade dock äldre försökspersoner signifikant sämre än yngre. Den äldsta åldersgruppen var här 61-70 år och resultatet kan inte förklaras av skillnader i utbildningsnivå mellan de olika åldersgrupperna.

I testversionerna A och C presterade försökspersoner med högre utbildning signifikant bättre än de med lägre utbildning. Detta beror framför allt på att personer med mycket låg utbildningsnivå (6-8 års utbildning) presterade signifikant sämre än övriga. Skillnaderna var också störst i det sista avsnittet i testen. Utbildningsnivån påverkade inte testresultatet i version B. Resultaten är i överensstämmelse med dem i flertalet tidigare studier (Orgass & Poeck, 1966; Boller & Vignolo, 1966; Swisher & Sarno, 1969; Lindström & Revstedt, 1975; de Renzi & Faglioni, 1978; Benton m.fl., 1978).

Anmärkas bör att det i föreliggande studie fanns ett signifikant negativt samband mellan ålder och utbildningsnivå på så sätt att en högre ålder medförde en lägre utbildningsnivå. Av detta

skäl går det inte att dra någon säker slutsats om det är åldern, utbildningsnivån eller bådadera som påverkar testresultatet i versionerna A och C. I några studier av testet (de Renzi & Faglioni, 1978; Rey & Benton, 1991) rekommenderas att testpoängen justeras efter ålder och/eller utbildning utifrån fynden om sambandet mellan testresultat och dessa faktorer. I dessa studier är dock antalet försökspersoner mer än dubbelt så stort som i föreliggande studie. Åldersgrupperna och grupperna med olika utbildningsnivå är i föreliggande undersökning för små för att säkra värden för en sådan justering ska kunna anges. Vid utvärdering av testresultatet i versionerna A och C bör man dock beakta att en frisk, lågutbildad, äldre individ sannolikt har en testpoäng i normalzonens nedre del medan en välutbildad, ung individ bör ha full eller nära full poäng på testet.

Tidsåtgång och förhållandet mellan tidsåtgång och kön, ålder respektive utbildning

Medeltidsåtgången i version A (9 min 6 s) var som förväntat ungefär dubbelt så stor som den i version C (5 min 6 s) och ungefär tre gånger så stor som den i version B (3 min 24 s). Tidsskillnaderna räknat i minuter är dock så små att det ur resurssynpunkt inte finns något skäl att välja en kortare version än en längre. Mellan det snabbaste och långsammaste utförandet av samtliga versioner skilde det bara drygt 15 minuter. Det existerar inga publicerade studier kring tidsåtgången för Token test vare sig när det gäller normala eller hjärnskadade individer. I sin artikel från 1978 anger de Renzi och Faglioni i allmänna ordalag tidsåtgången för deras förkortade version till ca 10-15 min när personer med afasi utför testet mot ca 20-30 min för den ursprungliga versionen. Wertz (1979) anger att tidsåtgången för personer med afasi på en testversion med 61 uppgifter i en ej publicerad studie varierade mellan 9 och 34 min med en medeltidsåtgång på 15 min. De angivna tiderna det tar för en individ med afasi att genomföra testet är således längre än de i föreliggande studie på normala personer. Även uppsatsförfattarens erfarenhet är att personer med afasi ofta behöver längre tid – ibland avsevärt längre tid – än friska personer för att utföra Token test. Det är möjligt att den ökade tidsåtgången i vissa fall är den enda faktor som skiljer en afasidrabbad eller annan hjärnskadad persons prestation från en normal prestation. Detta bör dock studeras närmare innan några säkra slutsatser kan dras.

I versionerna A och C behövde äldre personer signifikant längre tid än yngre och lågutbildade signifikant längre tid än högutbildade. Detta bör man ta hänsyn till vid utvärderingen av testresultatet.

Jämförelse mellan testversionernas testpoäng och tillförlitlighet

Sambandet mellan de olika testversionernas itemtotalpoäng undersöktes. Korrelationen mellan itemtotalpoängen i versionerna A och C var högre (.43) än mellan versionerna A och B (.23). Korrelationen mellan itemtotalpoängen resp. uppgifttotalpoängen i versionerna B och C var inte ens signifikant.

Att reliabiliteten i studien var tämligen låg beror på att spridningen i testpoäng blir liten när Token test utförs på en normalgrupp, dvs. en målgrupp testet egentligen inte är avsedd för. Särskilt i version B är spridningen i testpoäng mycket liten. Reliabiliteten för versionerna A och C, när itempoäng beräknades, var ungefär jämförbar (.60 respektive .65). Reliabiliteten i version B, när itempoäng beräknades, var lägre (.45). När uppgiftpoängen för versionerna B och C beräknades, var reliabiliteten ännu lägre (.22 respektive .31).

Även om korrelationerna blir förhållandevis låga beroende på den begränsade spridningen i en normalgrupp uppvisar versionerna A och C vid itempoängsättning högst och mest jämförbar reliabilitet och högst samband när det gäller testresultat. Dessa versioner är om itempoängsättning används därför mer utbytbara mot varandra än mot version B vid t.ex. ett test-retestförfarande.

Slutsatser

De viktigaste resultaten som framkom i studien var:

- Gränsvärden för normalprestation för de tre versioner av Token test, som är översatta till svenska och används i Sverige, fastställdes. Dessa var i samtliga testversioner högre än i motsvarande eller jämförbara testversioner på andra språk.
- I alla testversioner var andelen fel i det sista avsnittet, som innehåller grammatiskt komplexa utsagor, större än i övriga avsnitt. I alla versioner var andelen fel på items, som betecknar logiko-grammatiska strukturer större än på övriga items i detta avsnitt. Fynden överensstämmer med tidigare studier.
- Som i tidigare studier av motsvarande eller jämförbara testversioner på andra språk fanns inga könsskillnader i testresultat i någon av versionerna. I versionerna A och C medförde lägre ålder signifikant bättre testresultat än högre (när uppgifttotalpoängen beräknades i version C) och en högre utbildningsnivå signifikant bättre testresultat än en lägre. I version B fanns inga skillnader i testresultat beroende på ålder eller utbildningsnivå. Resultaten avseende ålderns påverkan på testresultatet skiljer sig från dem i flertalet tidigare studier medan resultaten av utbildningens påverkan överensstämmer med dem i flertalet tidigare studier.
- Medeltidsåtgången för version A var ungefär dubbelt så stor som för version C och tre gånger så stor som för version B. Skillnaden i minuter räknat mellan medeltidsåtgången för den kortaste och den längsta testversionen var dock mindre än 6 minuter, dvs. en försumbar skillnad ur resursåtgångssynpunkt. I versionerna A och C behövde äldre personer signifikant längre tid än yngre och lågutbildade signifikant längre tid än högutbildade.
- Testversionerna A och C hade en jämförbar och högre reliabilitet än version B, när itemtotalpoängen beräknades. Det fanns ett högre samband mellan itemtotalpoängen i version A och i version C än mellan denna poäng i version B och i de övriga versionerna. Versionerna A och C är således jämförbara medan version B skiljer sig från de övriga.

Genom att använda resultaten i denna studie vid utvärdering av testning med de olika svenska versionerna av Token test blir diagnostisering av lätta, auditivt impressiva störningar vid afasi säkrare än tidigare. För att göra en så tillförlitlig bedömning som möjligt bör version A eller C av testet användas. Version B lämpar sig mer för screeningändamål. Om man förutom att göra en kvantitativ bedömning även önskar göra en noggrann kvalitativ bedömning av afasi bör itempoängsättning och inte uppgiftpoängsättning användas. Denna metod rekommenderas också vid bedömning av personer med afasi, om man önskar göra en noggrann kvantitativ bedömning. I denna studie utfördes Token test på normala personer. Normala personer gör få fel på testet och då blir skillnaden i totalpoäng mellan bedömning med itempoäng och uppgiftpoäng liten. För personer med afasi, vilka ofta gör mer än ett fel per uppgift blir skillnaden naturligtvis större. Uppgiftpoängsättning bör därför endast användas vid screeningbedömningar.

I framtida studier vore det önskvärt att jämföra tidsåtgången vid utförandet av de tre versionerna av Token test mellan hjärnskadade personer utan afasi, hjärnskadade personer med afasi och normala personer. Grammatiska och lexikala skillnader mellan olika språk som en orsak till att normalgruppers lägsta poäng på Token test skiljer sig åt mellan studier gjorda på olika språk skulle också kunna undersökas i jämförande studier, där faktorerna utbildningsnivå och ålder hålls lika i normalgrupper, som talar olika språk.

Referenser

- Agrell B., Dehlin O., Nilsson I.L. (1995). Receptive aphasia in elderly stroke patients as assessed by the Token Test. *Scandinavian Journal of Logopedics and Phoniatrics*, 20, 27-34.
- Ahlsén E. (2006). *Introduction to neurolinguistics*. Amsterdam/Philadelphia: John Benjamins Publishing Company (s.101).
- Albert M. L., Goodglass H., Helm N. A., Rubens A.B., Alexander M. P. (1981). *Clinical aspects of dysphasia*. Wien: Springer Verlag.
- Apt P. (1997). *Manual till Pia APTs AfasiPrövning (PAPAP)*. Stockholm: Psykologiförlaget.
- Apt P. (1999). *Svensk BenämningsPrövning (SBP)*. Malmö: Neurologiska kliniken, Universitetssjukhuset MAS.
- Apt P. (2001). *Manual till Grundläggande Neuropsykologisk Undersökning (GNU)*. Malmö: Neurologiska kliniken, Universitetssjukhuset MAS.
- Beck A.T., Steer R.A. (1996). *Manual till Beck Depression Inventory*. (Svensk översättning av Åkerlund M.) Stockholm: Psykologiförlaget.
- Benton A.L., Hamsher K.deS., Sivan A.B. (1978). *Multilingual Aphasia Examination (MAE). Manual of Instructions*, 3rd ed. Lutz, Florida: Psychological Assessment Resources, Inc.
- Bergman I., Blomberg M., Almkvist O. (2007). The importance of impaired physical health and age in normal cognitive aging. *Scandinavian Journal of Psychology*, 48, 115-125.
- Boller F., Vignolo L. (1966). Latent sensory aphasia in hemisphere-damaged patients: An experimental study with the Token Test. *Brain*, 89, 815-831.
- Brookshire R.H., Manthie M.A. (1980). Speech and language disturbances in the elderly. I Maletta G.J., Prozzolo F.J. (Eds.). *The aging nervous system*. New York: Praeger Publishers.
- Cohen R., Gutbrod K., Meier E., Römer P. (1987). Visual search processes in the Token Test performance of aphasics. *Neuropsychologia*, 25, 983-987.
- Cohen-Mimran R., Sapir S. (2007). Deficits in working memory in young adults with reading disabilities. *Journal of Communication Disorders*, 40, 168-183.
- Coupar A. M. (1976). Detection of mild aphasia: a study using the Token Test. *British Journal of Medical Psychology*, 49, 141-144.
- De Renzi E., Faglioni P. (1978). Normative data and screening power of a shortened version of the Token Test. *Cortex*, 14, 41-49.
- De Renzi E., Vignolo L. A. (1962). The Token Test: A sensitive test to detect receptive disturbances in aphasics. *Brain*, 85, 665-678.

- DiSimoni F. (1978). *The Token Test for children*. Hingham, MA: Teaching Resources Corporation.
- Emery O. B. (1986). Linguistic decrement in normal aging. *Language and Communication*, 6, 47-64.
- Folstein M.F., Folstein S.E., McHugh P.R. (1975). Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-198.
- Gallagher A.J. (1979). Temporal reliability of aphasic performance on the Token Test. *Brain and Language*, 7, 34-41.
- Goodglass H., Kaplan E. (1983). *Manual for The Boston Diagnostic Aphasia Examination (BDAE)*, Philadelphia: Lea & Febiger.
- Hartje W., Kerschensteiner M., Poeck K., Orgass B. (1973). A cross-validation study on the Token Test. *Neuropsychologia*, 11, 119-121.
- Huber W., Poeck D., Weniger D., Willmes K. (1983). *Aachener Aphasie Test (AAT)*. Göttingen: Verlag für Psychologie.
- Kaplan E.F., Goodglass H., Weintraub S. (1983). *Boston Naming Test (BNT)*, 2nd ed. Philadelphia: Lea & Febiger.
- Kertesz A. (1982). *Western Aphasia Battery (WAB)*. New York: Grune & Stratton.
- Kiernan J. (1986). Visual presentation of the Revised Token Test: Some normative data and use in modality independence testing. *Folia Phoniatrica*, 38, 25-30.
- Leischner A. (1974). Die neuropsychologisch-hirnpatologische Untersuchung. *Archiv für Psychiatrie und Nervenkrankheiten*, 219, 53-77.
- Lesser R. (1976). Verbal and non-verbal memory components in the Token Test. *Neuropsychologia*, 14, 79-85.
- Lezak M. D. (1995). Token Test. I *Neuropsychological Assessment* 3rd ed. New York: University Press, 532-536.
- Liles B. Z., Brookshire R. H. (1975). The effects of pause time on auditory comprehension of aphasic subjects. *Journal of Communication Disorders*, 8, 221-235.
- Lindström E., Revstedt E. K. (1975). En undersökning av förståelsen av de språkliga strukturerna i Token testet hos afatiker och talförsenade barn. *Examensarbete i logopedi*, 10 p. Lund: Institutionen för logopedi, Lunds Universitet.
- Lindström E., Werner C. (1995). *A-ning: Neurolingvistisk afasiundersökning*. Stockholm: Ersta Utbildningsinstitut.

- Luria A. R. (1973). *The working brain. An Introduction to neuropsychology*. New York: Basic Books.
- Luria A.R. (1980a). *Higher cortical functions in man*, 2nd ed. New York: Basic Books (s.177-182).
- Luria A.R. (1980b). *Higher cortical functions in man*, 2nd ed. New York: Basic Books (s. 246-365).
- Mack J. L., Boller F. (1979). Components of auditory comprehension: Analysis of errors in a revised Token Test, 45-69. I Boller F., Dennis M. (Eds.) *Auditory comprehension: Clinical and experimental studies with the Token Test*. New York: Academic.
- Marie P. (1902). Marie's paper test. *Revue Neurologique*, 16, 611-974.
- Martino A.A., Pizzamiglio L., Razzano C. (1976). A new version of the "Token Test" for aphasics: A concrete object form. *Journal of Communication Disorders*, 9, 1-5.
- McNeil M. M., Prescott T.E. (1978) *Revised Token Test*. Austin, Texas: Pro-Ed.
- Nationalencyklopedin, 7:e bandet (1992). Höganäs: Bra Böcker AB (s.162).
- Orgass B. (1976). Eine Revision des Tokens Tests, Teil I und II. *Diagnostica*, 22, 70-87, 141-156.
- Orgass B., Poeck K. (1966). Clinical validation of a new test for aphasia: an experimental study on the Token Test. *Cortex*, 2, 222-243.
- Orgass B., Poeck K., Hartje W., Kerschensteiner M. (1973). Zum Vorschlag einer Kurzform des Token Tests Zur Auslese von Aphasikern. *Nervenarzt*, 44.
- Podemski R., Jarzebska E., Bakierowska A. (1984). Usefulness of the Token Test in the diagnosis of aphasia. *Neurologia Neurochirurgia Pol.*, 18, 547-551.
- Poeck K., Hartje W., Kerschensteiner M., Orgass B. (1973). Sprachverständnisstörungen bei aphasischen und nicht aphasischen Hirnkranken. *Deutsche Medizinische Wochenschrift*, 98, 139-147.
- Poeck K., Pietron H.P. (1981). The influence of stretched speech presentation on Token Test performance of aphasic and right brain damaged patients. *Neuropsychologia*, 19, 133-136.
- Pracharitpukdee N., Phanthumchinda K., Huber W., Willmes K. (1998). The Thai version of the German Aachen Test (AAT): description of the test and performance in normal subjects. *Journal of Medical Association Thai*, 81, 402-412.
- Raven J.C. (1938). *Progressive Matrices: A perceptual test of intelligence*. Individual Form. Oxford: Oxford Psychologists Press Ltd.
- Rey G.J., Benton A. (1991). *Exámen de Afasia Multilingüe (Multilingual Aphasia Examination-Spanish)*. *Manual de Instrucciones*. Iowa City, Iowa: AJA Associates, Inc.

Scotti G., Spinnler H. (1970). Colour imperception in unilateral hemisphere damaged patients. *Journal of Neurology Neurosurgery and Psychiatry*, 33, 22-28.

Spellacy F.J., Spreen O. (1969). A short form of the Token Test. *Cortex*, 5, 390-397.

Spreen O., Benton A.L. (1969, 1977). *Neurosensory Center Comprehensive Examination for Aphasia*. Victoria, British Columbia: University of Victoria, Neuropsychology Laboratory.

Steinberg B. A., Bieliauskas L. A., Smith G. E., Langellotti C., Ivnik R.J. (2005). Mayo's older Americans normative studies: Age- and IQ-adjusted norms for the Boston Naming Test, The MAE Token Test, and the Judgment of Line Orientation Test. *The Clinical Neuropsychologist*, 19, 280-328.

Swihart A.A., Panisset M., Becker J.T., Beyer J.R., Boller F. (1989). The Token Test: Validity and diagnostic power in Alzheimer's disease. *Developmental Neuropsychology*, 5, 69-78.

Swisher L. P., Sarno M. T. (1969). Token Test scores of three matched patient groups: left brain-damaged with aphasia; right brain-damaged without aphasia; non brain-damaged. *Cortex*, 5, 264-273.

van Dongen H. R., van Harskamp F. (1972). The Token Test. A preliminary evaluation of a method to detect aphasia. *Psychiatria, Neurologia, Neurochirurgia*, 75, 129-134.

van Harskamp F., van Dongen H. R. (1977). Construction and validation of different short forms of the Token Test. *Neuropsychologia*, 15, 467-470.

Wang Y., Niu J. (2000). Relationship between the Token Test and Chinese aphasia with different types due to left hemispheric damages. *Chinese Journal of Rehabilitation theory & Practice*, 6, 49-52, 61.

Wechsler D. (1981). *Wechsler Adult Intelligence Scale – Revised: Manual*. New York: The Psychological Corporation.

Wertz R. T. (1979). Review of the Token Test (TT). I Darley F. L (Ed.). *Evaluation of appraisal techniques in speech and language pathology*. Reading, Maine: Addison – Wesley.

Wold A. H., Reinvang I. (1990). The relation between integration, sequence of information, short-term memory, and Token Test performance of aphasic subjects. *Journal of Communication Disorders*, 23, 31-59.

Internetkälla

<http://www.dpf.dk> 07 05 17. *Dansk Psykologisk Forlag*. Spreen O., Benton A.L. (1977). Token Test - dansk översättning.