



LUNDS
UNIVERSITET

Institutionen för psykologi
Psykologprogrammet

En utprovning av Hayling Test i svensk version

Psykometriska data och jämförelse med patienter med
frontotemporal demens

Erik Blenow Nordström

Psykologexamensuppsats. 2013

Handledare: Susanna Vestberg
Examinator: Sven Ingmar Andersson

Clinicians should be not only knowledgeable about traditional content areas in psychology and the various contexts of assessment, but also able to integrate the test data into a relevant description of the person. This description, although focusing on the individual, should take into account the complexity of his or her social environment, personal history, and behavioural observations. Yet, the goal is not merely to describe the person, but rather to develop relevant answers to specific questions, aid in problem solving, and facilitate decision making.

(Groth-Marnat, 2003, s. 5)

Tack!

Författaren vill tacka handledaren Susanna Vestberg som bidragit med en svårslagen kombination av kunskap, entusiasm och lugn. Resten av LUPROFS-gruppen och övrig personal på Minneskliniken tackas för gedigna fackkunskaper, inte minst tack till den fenomenala Anna-Karin Thulin för noggrann interbedömerskattning och värdefulla synpunkter. Författaren vill tacka de personer som medverkat i studien och på så vis bidragit till ökad kunskap inom fältet. Tack till Marie Liljekvist på Päronskogens vårdboende i Malmö stad för storsint lån av lokal. Tack även till Karl Nordström för synpunkter på manus. Till Ylva Blennow, Björn Nordström, vänner, kollegor och personal på träffpunkter samt mötesplatser riktas ett varmt tack för hjälp med kontaktvägar till målgruppen för studien. Slutligen tackas Stiftelsen Carl Axel Bergstrand och Folke Ljungdahls stiftelse för generösa bidrag.

Sammanfattning

Studiens syfte var att psykometriskt utprova en svenskspråkig version av Hayling Test, ett neuropsykologiskt test som framför allt avser att mäta verbal inhibitionsförmåga. Testet utprovades på en frisk kontrollgrupp ($N = 76$) i åldern 40-95 år. Studien fann ett signifikant negativt samband mellan ålder och prestationer på testet. Män presterade signifikant bättre än kvinnor på ett av delmått. Det förelåg en signifikant korrelation mellan hög utbildningsnivå och bättre testprestation på ett av delmått. Homogenitetsreliabiliteten var otillräcklig till bra, beroende på deltest. Interbedömarreliabiliteten var god till utmärkt. Begreppsvaliditeten var god då testet korrelerade signifikant positivt med ett annat test som avser att mäta inhibition, Color Word Interference Test (CWIT). Den diagnostiska validiteten undersöktes genom jämförelse mellan kontrollgruppen och en patientgrupp med frontotemporal demens (FTD; $N = 30$). Patientgruppen presterade signifikant sämre på testet, även om testet inte helt kunde skilja mellan grupperna. Det rekommenderas att några item och bedömningskriterier justeras för att förbättra användbarheten. Studiens slutsats är att den svenska versionen av Hayling Test kan användas vid neuropsykologisk utredning av inhibitorisk förmåga vid misstänkt FTD.

Nyckelord: inhibition, exekutiva funktioner, Hayling Test, frontotemporal demens, neuropsykologisk bedömning, validitet, reliabilitet

Abstract

The aim of this study was to examine the psychometric properties of an adapted Swedish version of the Hayling Test, a neuropsychological test which mainly intends to measure response inhibition. The test was administered to a healthy control group ($N = 76$) ranging in age from 40 to 95 years. The study found a significant negative correlation between age and test performance. Men performed significantly better than women in one sub-scale. A high education level correlated significantly with better test performance in one sub-scale. The internal consistency reliability varied between unacceptable to good, depending on the different sub-scales. Inter-rater reliability was good to excellent. Construct validity was good, as the test showed significant positive correlations with another test designed to measure inhibition, Color Word Interference Test (CWIT). The diagnostic validity was examined by comparing the control group with a group of patients with frontotemporal dementia (FTD; $N = 30$). The patient group performed significantly worse on the Hayling Test than the control group, although a slight overlap between groups was noted. Some modifications of items and scoring criteria are advised. This study concludes that the Swedish version of the Hayling Test can be of use within neuropsychological assessment of inhibition when FTD is suspected.

Keywords: inhibition, executive function, Hayling Test, frontotemporal dementia, neuropsychological assessment, validity, reliability

Innehållsförteckning

Inledning.....	1
Teori.....	2
Exekutiva funktioner.....	2
Exekutiva funktioner och åldrande.....	2
Modeller över exekutiva funktioner.....	3
Kärnprocessmodellen	3
Modellen över det uppmärksamhetskontrollerande systemet	4
Inhibition.....	5
Psykometri	6
Att mäta verbal initiering och inhibition genom Hayling Test	8
Administrering av Hayling Test	8
Standardisering av Hayling Test	9
Reliabilitet för Hayling Test.....	11
Neurala korrelat till verbal initiering och inhibition mätt genom Hayling Test.....	11
Resultat på Hayling Test för friska deltagare	12
Kliniska användningsområden för Hayling Test.....	13
Icke-engelskspråkiga versioner av Hayling Test.....	14
FTD	14
Incidens och prevalens vid FTD.....	15
Lokalisation av hjärnpåverkan och kliniska symtom vid FTD.....	15
Resultat på Hayling Test för patienter med FTD	16
Syfte.....	17
Hypoteser.....	18
Metod	18
Deltagare.....	18
Kontrollgrupp.....	18
Patientgrupp	18
Instrument.....	19
Hayling-testet	19
CWIT	19
Administrering av CWIT betingelse 3	20
Procedur.....	21

Kontrollgrupp.....	21
Patientgrupp	21
Rättning.....	21
Etik.....	22
Dataanalys.....	22
Resultat.....	24
Deltagarnas karaktäristika	24
Kontrollgruppens resultat	25
Jämförelse mellan resultat på Hayling-testet och de demografiska variablerna.....	25
Homogenitetsreliabilitet	26
Interbedömarreliabilitet	27
Begreppsvaliditet	28
Diagnostisk validitet	28
Diskussion	29
Resultatdiskussion	29
Demografiska variabler	29
Homogenitetsreliabilitet.....	30
Interbedömarreliabilitet.....	31
Begreppsvaliditet	31
Diagnostisk validitet	32
Metoddiskussion.....	34
Deltagare	34
Instrument	36
Etik.....	36
Dataanalys	37
Avslutande diskussion	38
Utvecklande av Hayling-testet	38
Förslag på framtida forskning	39
Slutsatser	40
Referenser.....	41
Bilaga.....	i

Figur- och tabellförteckning

Figur 1: Modellen över det uppmärksamhetskontrollerande systemet.....	4
Figur 2: Hjärnan ur ett lateralt respektive medialt perspektiv med relevanta BA angivna	12
Figur 3: Låddiagram över Hayling-testet Skalpoäng totalt för kontrollgruppen och patientgruppen	29
Tabell 1: Skalpoäng totalt på Hayling Test, tillhörande percentiler och klassifikation	10
Tabell 2: Resultat på Hayling Test för patienter med bvFTD	17
Tabell 3: Kontrollgruppens och patientgruppens karaktäristika	24
Tabell 4: Kontrollgruppens resultat på Hayling-testet	25
Tabell 5: Jämförelse mellan män och kvinnors resultat på Hayling-testet.....	26
Tabell 6: Samband mellan resultat på Hayling-testet och ålder respektive utbildningsnivå....	26
Tabell 7: Intraklasskoefficienter för Hayling-testet Felpoäng.....	27
Tabell 8: Jämförelse mellan kontrollgruppens och patientgruppens resultat på Hayling-testet	28

Människan har en unik förmåga att långt i förväg planera och sedan utföra viljestyrda aktiviteter. Denna förmåga gör det möjligt för många att hejda sig och tänka efter både en och två gånger innan de berättar sitt hjärtas mening för chefen om dennes tvivelaktiga egenskaper. Det är också denna förmåga som gör att många kan lägga upp en strategi för att lösa ett klurigt problem, som ett sudoku eller varför inte att köra och hämta barnen på förskolan i rusningstrafik samtidigt som veckans middagsmat planeras. Egentligen handlar detta inte om en enda förmåga, utan om flera förmågor som samordnas. Begreppet exekutiva funktioner används för att beskriva denna samordnande funktion, oberoende vad det är för typ aktivitet som utförs. Allt fler studier ägnar sig åt att undersöka exekutiva funktioner, eftersom de är en viktig aspekt av människans kognition, påverkar beteendet samt interagerar på ett komplext sätt med andra förmågor (de Sousa Siqueira, Scherer, Reppold & Fonseca, 2010). Ett exempel på en exekutiv funktion är inhibition. Inhibition kan innebära olika saker men kan exemplifieras med förmågan att hejda sig från att utföra en nästintill automatiserad handling, som att inte gå i försvar eller skratta bort situationen när man blir kritiserad, om det nu är det man vanligen gör.

Exekutiva funktioner är svårare att mäta än andra kognitiva förmågor som logisk slutledningsförmåga. Detta beror bland annat på att den psykologiska testsituationen är styrd och inte med lätthet kan generaliseras till vardagliga situationer. Det finns således inget neuropsykologiskt test som går ut på att inte skälla ut chefen. Liksom andra högre kognitiva förmågor är exekutiva funktioner dessutom svårt, för att inte säga omöjligt, att korrelera till ett enskilt område i hjärnan.

Icke desto mindre finns behov av neuropsykologiska test som mäter olika aspekter av exekutiva funktioner, däribland inhibitorisk förmåga. Generellt sett är nämligen flera patientgrupper, såsom personer med schizofreni och frontotemporal demens (FTD), drabbade av nedsatt funktion i hjärnområden som ofta relateras till exekutiva funktioner. Traditionella test som ämnar mäta exekutiva funktioner är tyvärr inte alltid tillräckliga, till exempel vid utredning av misstänkt FTD (Hodges & Patterson, 2007). Testen kan mäta flera exekutiva funktioner samtidigt eller därutöver oavsiktligt mäta kognitiva förmågor som inte alltid är lika försämrade. Detta försvårar tolkningen av resultaten. Test som mäter rätt och tillförlitligt är en viktig pusselbit i undersökning och bedömning samt bidrar till att underlätta utredningen. Med bättre test ges en tydligare bild av personens styrkor och svagheter. I en klinisk kontext är detta till gagn för patienten och omgivningen, som med kunskap kan anpassa miljö, förväntningar och bemötande.

I föreliggande studie kommer ett neuropsykologiskt test som inte tidigare har använts i Sverige att utvärderas. Testet avser att mäta två aspekter av exekutiva funktioner, verbal initiering och verbal inhibition. Inhibition kommer att vara fokus för studien. Testet kommer att appliceras på en frisk kontrollgrupp och på en patientgrupp drabbad av FTD.

Teori

Exekutiva funktioner. Ett av de senaste forskningsområdena inom neuropsykologin är exekutiva funktioner (Burgess & Alderman, 2013). I en reviewartikel konstaterar Jurado och Rosselli (2007) att en samstämmig definition av begreppet exekutiva funktioner saknas. Exekutiva funktioner räknas oftast, så även i föreliggande uppsats, som en viktig aspekt av kognitiva förmågor som samtidigt påverkar människans beteende i hög utsträckning (Steverson, 2011). Lezak, Howieson och Loring (2004) beskriver exekutiva funktioner som "the ability to engage oneself in independent, purposive and self-serving behavior" (s. 35). Lezak et al. (2004) delar också in exekutiva funktioner i fyra underkategorier: *Volition* (vilja/avsikt), *planning* (planering), *purposive function* (ändamålsenlig funktion/utförande av viljestyrda handlingar) och *effective performance* (effektivitet). Bielak, Mansueti, Strauss och Dixon (2006) uttrycker ungefär detsamma, fast med andra ord; de kallar exekutiva funktioner för en uppsättning kognitiva processer av den högre ordningen, vilka organiserar och justerar mänskligt beteende samt kognition så att det passar kontext och avsikt. Beskrivningen av exekutiva funktioner som en uppsättning relaterade men självständiga processer står inte oemotsagd. Duncan, Emslie, Williams, Johnson och Freer (1996) argumenterar för att det finns en underliggande förmåga som starkt korrelerar med exekutiva funktioner: allmänintelligens, det vill säga Spearman's *g*. Ardila, Pineda och Rosselli (2000) drar likväl slutsatsen att enbart traditionella intelligenstest inte är lämpade för att undersöka exekutiva funktioner. Rör det sig om ett tidigt skede av neurodegenerativ sjukdom kan patienter prestera adekvat på traditionella intelligenstest, trots nedsättning av exekutiva funktioner. Friedman et al. (2008) menar att exekutiva funktioner är något annat än generell intelligens och att graden av exekutiva funktioner till stor del är genetiskt ärftliga.

Exekutiva funktioner och åldrande. Människans exekutiva funktioner är sent färdigutvecklade, men kan också riskera att försämrats tidigt jämfört med andra förmågor. Exekutiva funktioner kopplas oftast till hjärnans frontallober. Utvecklingsmässigt utvecklas exekutiva funktioner ungefär samtidigt som frontalloberna; mellan födseln och tvåårsåldern,

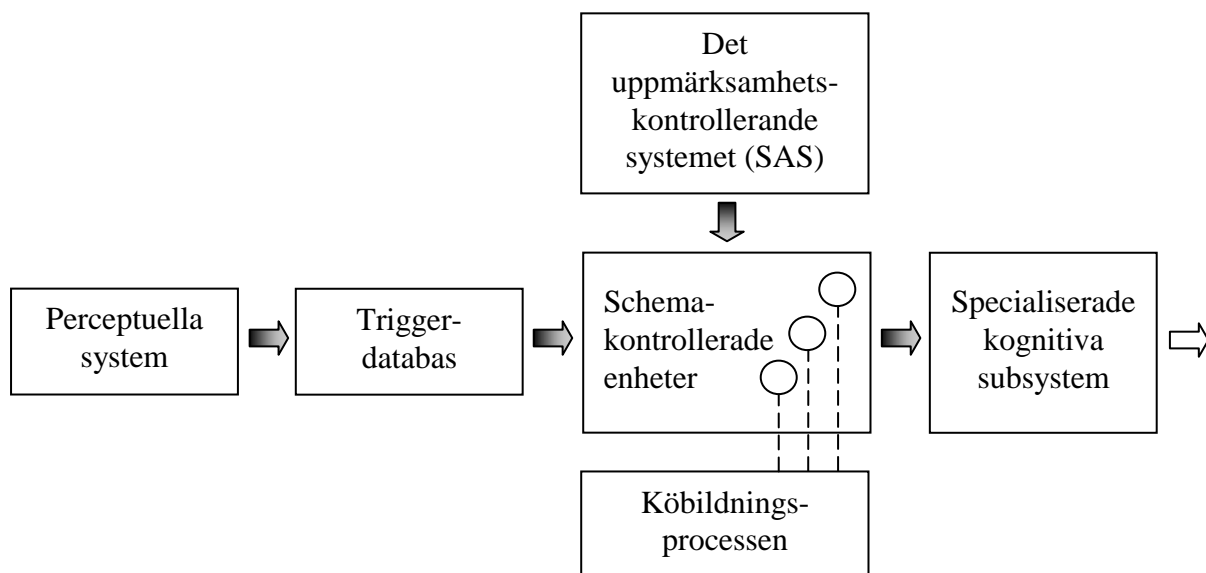
mellan sju- och nioårsåldern samt mellan 16-årsåldern och början av vuxenlivet (Anderson, Northam, Hendy & Wrennall, 2001). Därmed skulle människans exekutiva funktioner vara färdigutvecklade senare än andra kognitiva förmågor. Vid ökad ålder drabbas flera exekutiva funktioner relativt tidigt av försämrad funktion. Park och Reuter-Lorenz (2009) finner visserligen att frontallobsaktiviteten ökar vid högre ålder, men de sätter detta i förbindelse med att prefrontala cortex kompenserar för funktionsbortfall inom andra hjärnregioner. Exekutiva funktioner såsom processhastighet, arbetsminne och inhibition försämras med åldern (Park & Reuter-Lorenz, 2009). Det är vanligen så att flera exekutiva funktioner försämras samtidigt (Lezak et al., 2004). Bielak et al. (2006) menar att det är noterbart hur flera exekutiva funktioner försämras med åldern och att de eventuellt spelar en central roll för att förstå ett flertal åldersbetingade kognitiva försämringar. Utöver normalt åldrande kan exekutiva funktioner också försämras vid psykisk sjukdom eller sjukdom i centrala nervsystemet, till exempel demenssjukdom.

Modeller över exekutiva funktioner. Det finns flera teoretiska modeller som behandlar hur människans exekutiva funktioner fungerar. Två av dessa är av större relevans för föreliggande studie. Miyake, Friedman, Emerson, Witzki och Howerter (2000) modell delar in exekutiva funktioner i tre kärnprocesser: växling mellan olika arbetsuppgifter, uppdatering av arbetsminnet samt inhibition. Norman och Shallice (1986) modell över det uppmärksamhetskontrollerande systemet, som har många likheter med exekutiva funktioner, är också inflytelserik.

Kärnprocessmodellen. Enligt Miyake et al. (2000) består exekutiva funktioner av tre kärnprocesser som är måttligt korrelerade med varandra, men ändå distinkta. Den första processen är *shifting*, förmågan att flexibelt växla mellan flera arbetsuppgifter. Den andra processen är *updating*, förmågan att uppdatera och kontrollera representationer i arbetsminnet. Här måste ny information utvärderas för att sedan kunna ersätta gammal information i arbetsminnet. Den tredje processen är *inhibition*, förmågan att inhibera dominanta automatiska eller överinlärda responser. Här måste arbetsminnet skyddas från distraherande stimuli. Aktuell forskning om exekutiva funktioner är mest inriktad på ovanstående tre kärnprocesser (Friedman et al., 2008). Collette et al. (2005) undersökte neurala korrelat till kärnprocesserna och fick resultatet att de alla aktiverade hjärnområden i vänstra mediala och inferiora frontala vindlingen, men att varje kärnprocess också aktiverade separata

hjärnområden. Intressant nog aktiverades inte bara frontala områden, utan även områden i parietalloben och cerebellum (Collette et al., 2005).

Modellen över det uppmärksamhetskontrollerande systemet. I Norman och Shallices (1986) modell över kognitiv kontroll och målinriktat beteende tar schemakontrollerade enheter in perceptuell information och avgör vilken handling som är rimlig att utföra utifrån invanda scheman, se figur 1. Gazzaniga, Ivry och Mangun (2009) ger som exempel ett skrivet ord; när man ser ordet så aktiverar ett schema för handling den artikulation som krävs för att kunna uttala det, och när man läser ordet så aktiveras dess semantiska mening och närliggande representationer. Modellen gör en viktig åtskillnad mellan automatiserade rutinhandlingar och handlingar som kräver mer kognitiv kontroll. Enligt Norman och Shallice (1986) får de schemakontrollerade enheterna hjälp av köbildningsprocessen och av det uppmärksamhetskontrollerande systemet (*supervisory attentional system, SAS*). Köbildningsprocessen modererar ordningen som rutinbeteenden utförs i. Om mer komplext beslutsfattande krävs är det istället det uppmärksamhetskontrollerande systemet som tar över från de schemakontrollerade enheterna.



Figur 1. Modellen över det uppmärksamhetskontrollerande systemet. Figuren är anpassad från Norman och Shallice (1986) med svenska begrepp från Groome et al. (2006/2010).

Modellen över det uppmärksamhetskontrollerande systemet erbjuder kanske "the most detailed specification of the multiple processes considered under the title of executive

functions” (Steverson, 2011, s. 41). Enligt teorin behövs det uppmärksamhetskontrollerande systemet för att reglera målinriktade handlingar vid nya situationer, eller då invanda scheman inte räcker till. Systemet aktiveras också vid planering, beslutsfattande, problemlösning och inhibition av starka invanda responser (Norman & Shallice, 1986). För att kunna göra detta samordnar det uppmärksamhetskontrollerande systemet processer som utförs av olika subsystem, till exempel arbetsminne och framplockning från det episodiska minnet (Shallice & Burgess, 1998). När behov av flexibelt beteende uppstår, som när starka invanda responser måste inhiberas, är det uppmärksamhetskontrollerande systemet till stor hjälp. I den mån det uppmärksamhetskontrollerande systemet inte fungerar som det ska kan detta leda till svårigheter med planering, målformulering och sekvensering som krävs för att uppnå ett specifikt mål (Norman & Shallice, 1986), det vill säga bristande exekutiva funktioner.

Det uppmärksamhetskontrollerande systemet är en psykologisk modell för att illustrera kognitiv kontroll och motsvaras inte av aktivitet i en enstaka hjärnstruktur. Det är snarare så att flera regioner i prefrontala cortex samspelar för att förverkliga processerna som beskrivs av detta teoretiska ramverk.

Inhibition. En förmåga som enligt Miyake et al. (2000) är en av kärnprocesserna i exekutiva funktioner är inhibition. Friedman och Miyake (2004) menar att begreppet inhibition har funnits i över hundra år; de nämner Freuds teorier om bortträngning vid psykopatologi som ett tidigt exempel. På senare tid har Clark (1996) utifrån en neurovetenskaplig synvinkel definierat inhibition som en mekanism som reducerar eller dämpar neuronal, mental eller beteendemässig aktivitet. Från ett mer psykologiskt perspektiv delar Nigg (2000) in exekutiva inhibitoriska processer i fyra klasser: interferenskontroll där interferens undertrycks, kognitiv inhibition där irrelevant arbetsminnedata undertrycks, beteendehinhibition där automatiska, förberedda eller överinlärda responser undertrycks och okulomotorisk inhibition där sackadiska rörelser i ögat undertrycks. Som man kan förstå av detta är begreppet inhibition brett och definieras olika beroende på vem som står bakom studien. Beroende på studie används olika begrepp såsom inhibition, disinhibition, inhibitorisk kontroll eller *response suppression* – disinhibition kan dessutom användas som beteckning på ohämmat beteende. I föreliggande studie används begreppet inhibition. Enligt Niggs (2000) indelning kommer studien att fokusera på beteendehinhibition, eftersom den kräver att överinlärda responser måste inhiberas. Detta ligger också i linje med Norman och Shallices (1986) modell över det uppmärksamhetskontrollerande systemet.

Precis som för andra exekutiva funktioner är det svårt att hitta tester och experimentella betingelser som går att isolera till funktionen inhibition. Detta kallas för *task impurity problem* (Friedman et al., 2008) och går ut på att ett instrument som ska mäta en exekutiv funktion dessutom mäter närliggande funktioner: Beroende på vilket test som används kan perception, motorik eller andra kognitiva förmågor aktiveras.

Det har genomförts några studier om inhibitionsförmåga och åldrande. Schnitzspahn, Stahl, Zeintl, Kaller och Kliegel (2013) lät yngre vuxna mellan 18 och 39 år samt äldre vuxna mellan 57 och 77 år genomföra tester för att mäta exekutiva funktioner. De fann att gruppen med äldre vuxna hade signifikant sämre inhibitionsförmåga än yngre vuxna. De fann också att testpersonernas inhibitionsförmåga starkt predicerade hur väl de skulle prestera vid test av prospektivt minne, det vill säga förmågan att komma ihåg att göra saker i framtiden, där god inhibitionsförmåga predicerade bra prospektivt minne.

Många studier pekar på att inhibitionsförmågan, däribland den verbala, styrs av frontalloberna men exakt vilken eller vilka subregioner som är nödvändiga diskuteras fortfarande. Detta kan bero på att termen inhibition refererar till flera skilda processer, att flera olika hjärnregioner aktiveras vid inhibition eller att vissa uppgifter som avser att mäta inhibition faktiskt inte gör det (Collette et al., 2005).

Psykometri. Den del av psykologin som undersöker funktionen hos ett instrument eller test som avser att mäta psykologiska variabler kallas för psykometri. Ett psykologiskt test måste besitta goda psykometriska egenskaper för att resultatet ska kunna tolkas och därmed vara till nytta. Samtidigt måste den som administrerar ett test för att genomföra en klinisk bedömning vara medveten om testets begränsningar och vid behov komplettera psykometriska resultat med beteendeobservationer, anamnesupptagning eller andra test (Groth-Marnat, 2003). Ett psykologiskt test bör vara standardiserat med normering och tydliga anvisningar för hur testet ska administreras (Cicchetti, 1994). Råpoängen ska kunna omvandlas till standardiserade poäng och percentiler, så att jämförelser mellan olika individers testprestation kan ske och avvikande värden lättare kan upptäckas. (Crawford, 2013). Wechsler-skalpoäng ($M = 10$, $SD = 3$) är vanliga som standardiserade poäng, liksom Wechsler-indexpoäng ($M = 100$, $SD = 15$) och T-poäng ($M = 50$, $SD = 10$). Ett exempel på standardiserade poäng som sällan används är sten-poäng ($M = 5,5$, $SD = 2$), en förkortning för *standard ten*. Enligt Crawford (2013) är neuropsykologiska test som baseras på sten-skalan alldeles för oprecisa; ett för stort omfång råpoäng ryms i ett enda skalpoäng.

Inom neuropsykologin sker teoretiska framsteg i snabb takt medan det är dyrt och tidskrävande att ta fram normer, varför många test bygger på blygsamma underlag för normativa data (Crawford, 2013). Ett bra test har särskilda normer för subgrupper om de skiljer sig från den huvudsakliga normeringen (Groth-Marnat, 2003), där normer för olika åldergrupper är vanligast. I synnerhet när det krävs subjektiva bedömningar för att mäta prestation på testet måste bedömningskriterierna vara exakta, det vill säga både valida och reliabla (Cooper, 1999). Reliabilitet och validitet är viktiga begrepp för att undersöka psykometriska egenskaper. Med reliabilitet menas att testets resultat är konsekvent eller tillförlitligt. Med validitet menas att testet mäter den variabel som det avser att mäta. Det finns flera sätt att konceptualisera reliabilitet och validitet. Härnäst kommer olika former av reliabilitet och validitet som är relevanta för föreliggande studie att definieras.

Homogenitetsreliabilitet (*internal consistency reliability*) betyder hur väl testets olika item hänger ihop, förutsatt att alla item antas mäta samma begrepp (Colton & Covert, 2007). Denna sorts reliabilitet kallas också intern konsistens, men då detta är en felaktig översättning används i föreliggande studie begreppet homogenitetsreliabilitet. Den undersöks genom Cronbachs alfa, som ger ett exaktare mått än Spearman's split-half-reliabilitet; svagheter i enskilda item upptäcks inte med den senare metoden (Egidius, 2008).

Interbedömarreliabilitet (*inter-rater reliability*) betyder hur stabilt testet är om mer än en bedömare rättar testet. Interbedömarreliabilitet undersöks genom att låta två eller fler bedömare rätta testet och sedan beräkna korrelation, eller ännu hellre intraklasskorrelation. När många tänkbara sorters svar ska poängsättas ökar risken för låg samstämmighet mellan bedömarna, men genom en förstudie kan man åtgärda detta (Colton & Covert, 2007), till exempel genom utvecklade bedömningskriterier.

Begreppsvaliditet (*construct validity*) betyder hur väl ett test mäter begreppet det utger sig för att mäta. Begreppsvaliditet kan undersökas på många olika sätt, till exempel genom att mäta andra närliggande begrepp, konvergent validitet – testet *ska* korrelera med det andra testet, samt icke-närliggande begrepp, diskriminativ validitet – testet *ska inte* korrelera med det andra testet (Groth-Marnat, 2003). Både konvergent och diskriminativ validitet bör finnas för god begreppsvaliditet (Colton & Covert, 2007). Korrelationen ska vara måttlig till hög men inte för hög. Om korrelationen är för hög utan att det nya testet förenklar administreringen, behövs inte det nya testet (Groth-Marnat, 2003).

Diagnostisk validitet (*diagnostic validity*) betyder hur väl ett test mäter psykiska störningar eller egenskaper som intelligens (Egidius, 2008). Till exempel hur säkert ett test

som avser att mäta verbal inhibitionsförmåga identifierar personer med just bristande verbal inhibitionsförmåga. Diagnostisk validitet undersöks genom att låta olika patientgrupper alternativt patientgrupp och kontrollgrupp utföra testet, för att mäta testets förmåga att skilja mellan grupperna. I föreliggande studie används testet inte för att ställa diagnos utan för att mäta en egenskap.

Att mäta verbal initiering och inhibition genom Hayling Test. Hayling Test eller The Hayling Sentence Completion Test är ett relativt nytt engelskspråkigt neuropsykologiskt test som är utvecklat av Burgess och Shallice (1997). Hayling Test syftar till att mäta verbal initiations- och inhibitions hastighet samt verbal inhibitionsförmåga. Testet förefaller vara mest använt för att undersöka verbal inhibitionsförmåga. I testet ska testpersonerna komplettera meningar där det sista ordet saknas. Hayling Test sektion 1 ska mäta initiationshastighet. Här ska testpersonen avsluta meningar på ett adekvat sätt. Därpå följer sektion 2 som ska mäta inhibition, men även strategi och processhastighet. Här ska nya meningar avslutas på ett inadekvat sätt. Shallice och Burgess (1998) knyter testet till modellen om det uppmärksamhetskontrollerande systemet beskrivet ovan. De menar att testpersonerna i sektion 1 använder sig av scheman via köbildningsprocessen för att initiera korrekt svar, vilket är en rutinhandling. De tänker sig också att testpersonerna i sektion 2 använder sig av det uppmärksamhetskontrollerande systemet för att inhibera den invanda responsen och för att utveckla en ny strategi. Meningarna kommer från Bloom och Fischlers (1980) normeringsförteckning; de meningar där det var troligast att man skulle få ett specifikt ord som svar valdes ut till Hayling Test (Burgess & Shallice, 1996).

Administrering av Hayling Test. Hayling Test utgörs av två sektioner som administreras direkt efter varandra (Burgess & Shallice, 1997). I sektion 1 som består av 15 item ska testpersonen komplettera en mening i taget med ett ord som avslutar meningen på ett passande sätt. Exempel på detta är ”The rich child attended a private ...” – exempel på svar: ”school”. I sektion 2 som också består av 15 item ska testpersonen komplettera en mening i taget, denna gång med ett ord som inte alls hör ihop med meningen. Exempel på detta är ”London is a very lively...” – exempel på fel svar: ”city”, exempel på korrekt svar: ”banana”. Testledaren läser varje mening högt för testpersonen som ombeds att muntligen avge svar så snabbt som möjligt. Varje sektion inleds med två övningsmeningar som inte poängsätts. Efter att övningsmeningarna genomförts upprepas uppgiftens betingelser. Tidtagning påbörjas

direkt efter att meningen har lästs upp och avslutas när testpersonen har avgett svar.

Testledaren skriver ner svaret och använd tid i hela sekunder, avrundade nedåt, för varje item.

Hayling Test består av tre delmått som avser att mäta olika exekutiva funktioner.

Delmåttet Tid sektion 1 utgörs av summan av svarslatenserna i sektion 1 och ska ge ett mått på testpersonens initiationshastighet. Ju mindre svarslatens, desto högre initiationshastighet.

Delmåttet Tid sektion 2 utgörs av summan av svarslatenserna i sektion 2. Detta och följande delmått ska ge ett mått på testpersonens inhibitionsförmåga. Varje avgivet svar i sektion 2 klassificeras i en av tre möjliga kategorier. När svaret inte alls hör ihop med meningen, vilket är enligt uppgiftens betingelser, ger det noll felpoäng. När svaret hör lite ihop med meningen men inte helt och hållet är en meningskomplettering bedöms det som ett så kallat kategori B-fel. Till kategori B-fel räknas också motsatsord till det förväntade svaret och avsett komiska eller bisarra svar. När svaret hör ihop med meningen och testpersonen kompletterar meningen på ett adekvat sätt bedöms det som ett så kallat kategori A-fel. Till kategori A-fel räknas också uppenbart vulgära svar som svärord. Kategori A-fel medför fler felpoäng än kategori B-fel. Antalet felpoäng stiger i takt med antalet kategori A- och B-fel, exempelvis ger ett kategori A-fel tre felpoäng medan tre kategori A-fel ger tio felpoäng. Summan av felpoäng utgör råpoängen i delmåttet Felpoäng.

Om testpersonen avger flera svar ska det första svaret räknas. Avges inte svar inom 60 sekunder avslutas försöket och man fortsätter till nästa item. Svarar testpersonen med tre eller fler ord så påminner testledaren om att endast ett ord önskas. Testpersonen påminns om detta endast första gången som regeln överträds. Svar med flera ord bedöms i sektion 2 inte som fel, utan utifrån de vanliga bedömningskriterierna. Väljer testpersonen istället strategin att i sektion 2 säga samma ord i slutet av flera meningar så påpekas att detta inte är tillåtet. Fortsatta upprepningar bedöms som fel. Alla andra strategier för att utföra sektion 2 är tillåtna, men testledaren ger aldrig exempel på strategier. Om testpersonen i sektion 2:s första övningsmening avslutar meningen på ett adekvat sätt, vilket inte är korrekt, föreslår testledaren emellertid för första och enda gången ett ord som avslutar meningen på ett inadekvat sätt. Testpersonen ombeds sedan att komma på ett nytt ord som avslutar meningen på ett inadekvat sätt. Varje gång testpersonen under sektion 2 avslutar meningen på ett helt adekvat sätt påminner testledaren om att ett inadekvat svar önskas.

Standardisering av Hayling Test. På varje delmått får man råpoäng som konverteras till skalpoäng. Ju högre råpoäng på delmåtten, desto sämre resultat, men ju högre skalpoäng,

desto bättre resultat. På Tid sektion 1 räknas råpoängen om till skalpoäng på en skala mellan ett och sju. På Tid sektion 2 och Felpoäng räknas råpoängen om till skalpoäng på en skala mellan ett och åtta. Skalpoängen från delmåtten kan läggas samman till Sammanlagda skalpoäng och ge ett värde från 3 till 23, vilka i sin tur konverteras till Skalpoäng totalt på en skala mellan ett och tio. I tabell 1 åskådliggörs percentilnivå och klassifikation för respektive skalpoäng på Skalpoäng totalt.

Tabell 1

Skalpoäng totalt på Hayling Test, tillhörande percentiler och klassifikation. Sammanställt utifrån Burgess och Shallice (1997).

Skalpoäng totalt	Percentil	Klassifikation
10	99	Very superior
9	95	Superior
8	90	Good
7	75	High average
6	50	Average
5	25	Moderate average
4	10	Low average
3	5	Poor
2	1	Abnormal
1	< 1	Impaired

Den använda skalan är en lätt modifikation av sten-skalan ($M = 5,5$, $SD = 2$). Modifikationen består i att Hayling Tests skalpoäng inte tillåter decimaler varför medelvärdet avrundats uppåt till sex. Burgess och Shallice (1997) motiverar valet av skala med att den är intuitivt lättare att förstå och att den är tydligt kopplad till percentiler.

Skalpoängen i Hayling Test är densamma oberoende av testpersonens ålder. Resultat på den femte percentilen redovisas dock för åldrarna 18-40, 41-65 och 66-80, baserat på 118 friska personers resultat. Den femte percentilen för olika åldersgrupper redovisas dessvärre inte på Skalpoäng totalt, men väl på måttet Sammanlagda skalpoäng som kan ha ett värde mellan 3 och 23. För Sammanlagda skalpoäng var den femte percentilen 12 skalpoäng för 41-65 år och 11 skalpoäng för 66-80 år. Även på delmåttsnivå rapporteras den femte percentilen för åldersgrupperna och högre ålder förknippas med sämre prestation på alla delmått. Hayling Test utprovades på såväl en frisk kontrollgrupp som patienter med frontala lesioner, men normeringen baserades på kontrollgruppens resultat. Kontrollgruppen bestod av 118 personer,

varav 61 kvinnor, mellan 18 och 80 år i Storbritannien. Medelåldern var 45,3 år med en standardavvikelse på 18,1.

Reliabilitet för Hayling Test. Split-half-reliabiliteten är enligt Hayling Test-manualen 0,35 ($p < 0,001$) för Tid sektion 1; 0,83 ($p < 0,001$) för Tid sektion 2 och 0,41 ($p < 0,001$) för Felpoäng, undersökt på 118 friska kontrollpersoner (Burgess & Shallice, 1997). Värdena är generellt låga, med undantag för Tid sektion 2. I en patientgrupp bestående av 47 patienter med frontala lesioner är split-half-reliabiliteten högre: 0,93 ($p < 0,0001$) för Tid sektion 1; 0,80 för Tid sektion 2 och 0,72 för Felpoäng.

Test-retestreliabiliteten, det vill säga hur stabilt testet är över tid, är enligt manualen 0,62 ($p < 0,001$) för Tid sektion 1; 0,78 ($p < 0,001$) för Tid sektion 2 och 0,52 ($p < 0,01$) för Felpoäng. För Skalpoäng totalt är test-retestreliabiliteten 0,76 ($p < 0,001$). Test-retestreliabiliteten undersöktes på 31 friska kontrollpersoner som testades på nytt mellan två dagar och fyra veckor efter första undersökningstillfället.

Manualen redovisar ingen homogenitets- eller interbedömarreliabilitet.

Neurala korrelat till verbal initiering och inhibition mätt genom Hayling Test.

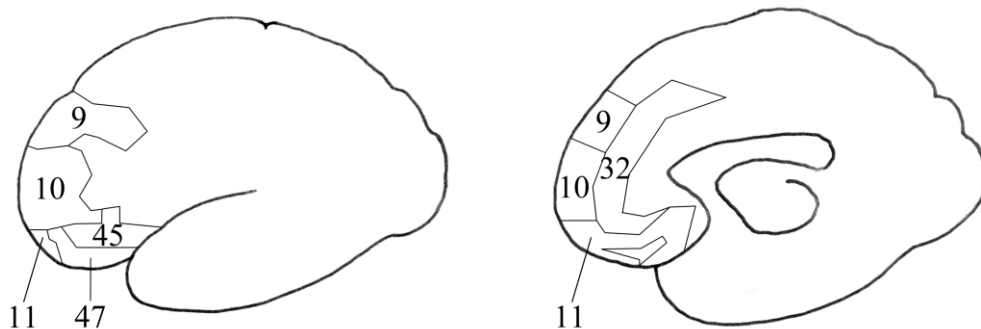
Burgess och Alderman (2013) menar att låga poäng på Hayling Test eller andra test som avser att mäta exekutiva funktioner inte omedelbart indikerar frontal dysfunktion, utan en funktionell försämring som ofta associeras med frontal dysfunktion. Neuropsykologisk undersökning ska mäta funktion – hur nedsatt en hjärnregion är mäts bäst genom hjärnabbildningstekniker. Likväl har flera studier använt sig av Hayling Test för att undersöka neurala korrelat till verbal initiering och inhibition. Initiering mättes genom sektion 1 och inhibition mättes genom sektion 2. I mån av tillgängliga uppgifter redovisas Brodmann-områden (*Brodman areas*, BA) som är ett sätt att dela in hjärnans områden (för översikt, se till exempel Gazzaniga et al., 2009). BA med relevans för föreliggande studie visas i figur 2.

Nathaniel-James, Fletcher och Frith (1997) fann att BA 45, vänstra inferiora frontala vindlingen, samt BA 32, högra anteriora gyrus cingulum, aktiverades i Hayling Test. Initiering och inhibition aktiverade här samma hjärnområden och ingen extra aktivitet förelåg vid inhibition. Detta kan enligt Collette et al. (2001) bero på att studien var okänslig för inhibition eftersom meningarna presenterades i sexsekundersintervall. Testpersonerna hade möjlighet att tänka efter om de svarade snabbare än sex sekunder per mening, utan att hjärnområdesaktiviteten undersöktes under detta intervall. Collette et al. (2001) fann vid en

modifiering av experimentet att initiering aktiverade vänstra BA 45/47 och att inhibition aktiverade ett nätverk av vänstersidiga frontala områden, däribland mediala (BA 9 och 10) och inferiora (BA 45) frontala vindlingarna. Hornberger, Geng & Hodges (2011) visade att atrofi i orbitofrontala cortex, mediala prefrontala cortex och anteriora temporalloben hade samband med inhibition genom delmättet Felpoäng, antalet felpoäng i sektion 2. Volle et al. (2012) fann att initiering och inhibition aktiverade dels samma, dels distinkta hjärnområden. BA 45/47 aktiverades i båda betingelser, men mediala rostrala frontala vindlingen (BA 10) aktiverades mest vid initiering. Nedsatt funktion i högra orbitofrontala vindlingen (BA 11) korrelerade med delmättet Felpoäng.

Följaktligen visar forskningen att ett nätverk bestående av anteriora och mediala frontala regioner är engagerade under utförandet av Hayling Test, men om dessa är vänster- eller högersidiga eller även samverkar med andra hjärnområden varierar mellan studier.

Enligt Shallice och Burgess (1998) kan de neurala korrelaten se olika ut beroende på om friska personer eller patienter testas med Hayling Test. Friska testpersoner utvecklar ofta en strategi; de förbereder ett svar innan meningen läses upp, exempelvis genom att titta i rummet efter objekt. Patienter med allvarlig frontal nedsättning använder ingen strategi, troligen snarare för att de inte kan *utveckla* den än att de inte kan *genomföra* den (Shallice & Burgess, 1998).



Figur 2. Hjärnan ur ett lateralt respektive mediant perspektiv med relevanta BA angivna.

Resultat på Hayling Test för friska deltagare. Bielak et al. (2006) fann efter att ha testat 457 personer mellan 53 och 90 år med Hayling Test att testpersonerna var snabbare på Tid sektion 1 än på Tid sektion 2 och att de gjorde fler kategori B-fel än kategori A-fel. Ökad ålder respektive färre antal utbildningsår ledde till sämre totalresultat, men testpersonernas kön påverkade inte resultatet. De redovisade normdata baserat på testpersonernas ålder. Antal utbildningsår påverkade resultatet på Tid sektion 1, där högre utbildningsnivå korrelerade med

snabbare svarstid, men inte på Tid sektion 2. Andrés och Van der Linden (2000) samt Belleville, Rouleau och Van der Linden (2006) kom fram till att äldre personer presterade signifikant sämre än yngre personer på Tid sektion 2, men inte på Tid sektion 1, efter att ha använt sig av två olika franskspråkiga versioner av Hayling Test. Däremot fann Borella, Ludwig, Fagot och De Ribaupierre (2011) ingen signifikant skillnad mellan yngre vuxnas och äldre vuxnas resultat på något av delmåttan, efter att ha testat 204 personer med ännu en franskspråkig version av Hayling Test.

Kliniska användningsområden för Hayling Test. Hayling Test är utvecklat för kliniskt bruk, det vill säga för att kunna vara till hjälp vid neuropsykologisk och psykopatologisk utredning och bedömning. Det tycks finnas mest vetenskapligt stöd för att Hayling Test sektion 2 som avser att mäta inhibitionsförmåga är till hjälp vid utredning. Hayling Test har använts vid studier av patientgrupper med möjliga bristande exekutiva funktioner, däribland patienter med FTD vilket belyses i ett eget avsnitt nedan. Därutöver har testet använts för att utforska andra patientgruppers initierande och inhibitoriska förmågor, såsom unga med ADHD-diagnos (Clark, Prior & Kinsella, 2000; Bayliss & Roodenrys, 2000), patienter med schizofrenidiagnos (Groom et al., 2008; Chan et al., 2012), tvångssyndrom (Van der Linden, Ceschi, Zermatten, Dunker & Perroud, 2005), Parkinsons sjukdom (Bouquet, Bonnaud & Gil, 2003), egentlig depression (Gohier et al., 2009) och Tourettes syndrom (Channon, Sinclair, Waller, Healy & Robertson, 2004). Wang et al. (2013) visade i en metaanalys att patienter med bipolär sjukdom samt patienter med schizofreni presterade ungefär likvärdigt och signifikant sämre än psykiskt friska personer på Hayling Tests båda sektioner. En modifierad version av Hayling Test har också utvecklats för att mäta förvrängd uppmärksamhet och inhibitionsförmåga hos personer med kronisk alkoholism (Rose, Mason-Li, Nicholas & Hobbs, 2010).

Belleville et al. (2006) visade att en franskspråkig version av Hayling Test kan vara till hjälp vid utredning av Alzheimers sjukdom (*Alzheimer's disease, AD*). AD-patienter var signifikant långsammare på Tid sektion 2 och hade fler fel än friska yngre och äldre personer på Felpoäng, men ingen signifikant skillnad mellan grupperna uppmättes på Tid sektion 1.

Burgess och Shallice (1996) visade att patienter med frontala lesioner som föranletts av till exempel tumörer eller vaskulära orsaker presterade långsammare på både Tid sektion 1 och Tid sektion 2, samt gjorde fler fel på Felpoäng, jämfört med patienter med lesioner i andra delar av hjärnan. Enligt dessa fynd presterar patienter med frontala lesioner sämre än friska

personer på samtliga delmått i Hayling Test. Detta skiljer sig från Andrés och Van der Lindens (2001) resultat. De jämförde friska personers resultat på en franskspråkig version av Hayling Test med en patientgrupp med frontala lesioner. Patientgruppen presterade inte signifikant sämre på Hayling Test i sin helhet, utan presterade endast signifikant långsammare på Tid sektion 2. Författarna tror trots litet antal deltagare ($N = 13$ i respektive grupp) att studiens statistiska power är tillräcklig, eftersom signifikanta skillnader ändå kunnat hittas.

Icke-engelskspråkiga versioner av Hayling Test. Flera franskspråkiga versioner av Hayling Test har tagits fram. I tre av dem valdes meningarna ut på liknande sätt som i den engelska originalversionen, det vill säga efter genomförda förstudier som pekade på vilka meningar som sannolikt skulle avslutas med ett specifikt ord (Collette et al., 2001; Belleville et al., 2006; Borella et al., 2011). En av versionerna (Borella et al., 2011) hade fem övningsmeningar per sektion istället för originalets två per sektion. Man behöll också meningarna från sektion 1 i sektion 2, men bad testpersonerna att avsluta dem inadekvat denna gång, eftersom förstudien inte visat skillnad i testprestation mellan nya och gamla meningar i sektion 2. Författarna menade att samma meningar i båda sektionerna skulle öka sannolikheten att inhibition av tidigare aktiverade ord krävdes. Dessutom har en brasiliansk version av Hayling Test tagits fram (Fonseca, Oliveira, Gindri, Zimmerman & Reppold, 2010, refererat i de Sousa Siqueira et al., 2010), liksom en kinesisk (Chan, Chen, Cheung & Cheung, 2004) och en tyskspråkig (Jantscher, Willinger, Schmöger, Müller & Auff, 2011). Den tyskspråkiga versionen testades på 74 friska vuxna mellan 19 och 66 år. Författarna jämförde testpersonernas resultat i Hayling Test sektion 2 med deras resultat på ett test som också avser att mäta verbal inhibitionsförmåga, Stroop deltest 3, men resultatet var icke signifikant. Det går inte att utläsa exakt vilka mått som användes i korrelationsanalysen.

FTD. FTD är en grupp demenssjukdomar som alla är progressivt neurodegenerativa och drabbar frontala samt temporala hjärnområden (Elfgren, Englund, Nilsson & Passant, 2011). Sjukdomsgruppen är heterogen och behandling saknas i dagsläget. Sjukdomar inom FTD-komplexet som baseras på försämrade högre kortikal funktion kan delas in i följande tre sjukdomar: FTD beteendevariant (bvFTD) och språkvarianterna progressiv icke-flytande afasi (PNFA) samt semantisk demens (SD). Andra sjukdomar inom FTD-komplexet är FTD med motorneuronsjukdom (FTD-MND), progressiv supranukleär pares (PSP) och kortikobasal degeneration (CBD). De två sistnämnda sjukdomarna överlappar ofta varandra avseende den

kliniska bilden och drabbar i högre utsträckning basala ganglier (PSP och CBD), parietalloberna (CBD) och hjärnstammen (PSP) utöver frontotemporala områden (Elfgren et al., 2011). En definitiv FTD-diagnos kan ställas först efter döden med en neuropatologisk undersökning, precis som för flertalet neurodegenerativa sjukdomar. Dessförinnan ställs klinisk diagnos som baseras på extensiv utredning. FTD utreds genom anamnesupptagning, statusundersökning, hjärnabbildande undersökning, laboratorieprov och cerebrospinalvätskeprov. Neuropsykologisk undersökning bör genomföras tidigt i sjukdomsförloppet, innan kognitiva och exekutiva brister gör att patienten inte längre kan medverka.

Incidens och prevalens vid FTD. FTD är mindre vanlig än AD och vaskulär demens, men underdiagnosticering är tänkbar. Symtomen kan misstas för psykos eller depression och bland äldre misstolkas som utslag av andra demenssjukdomar (Passant, Elfgren, Englund & Gustafson, 2005). Beteendevarianten står för 60-75 procent av alla FTD-diagnoser (Elfgren et al., 2011). Majoriteten forskare menar att FTD oftast debuterar mellan 45 och 65 års ålder, men insjuknande före och efter kan också förekomma. Sendebuterande FTD tycks vara vanligare än vad man tidigare trott (Graham, 2007). Durationen varierar mellan ett par år upp till 25 år (Passant et al., 2005). De flesta studier visar att män och kvinnor drabbas lika (Elfgren et al., 2011). Etiologi och riskfaktorer är oklara, men drygt 40 procent av FTD-patienter har en nära släkting som drabbats (Graham, 2007).

Lokalisation av hjärnpåverkan och kliniska symtom vid FTD. I detta avsnitt beskrivs sjukdomstillstånd inom FTD-komplexet som är relevanta för föreliggande studie.

I bvFTD sker bilateral atrofi framför allt i frontalloberna och anteriora temporalloberna, men även hippocampus i inferiora mediala temporalloberna samt de basala ganglierna kan drabbas av atrofi (Elfgren et al., 2011). I ett sent skede av sjukdomen kan nervcellsdegenerationen sprida sig till delar av parietalloberna. Kliniska kännetecken för bvFTD är smygande debut och långsam progress, tidig försämring av sociala relationer, tidig försämring av personligt uppträdande, tidig emotionell avtrubbning och tidig insiktsförlust (Neary et al., 1998). Språkförmågan är vanligen intakt även om patienterna kan komma att konversera mindre på grund av socialt tillbakadragande eller apati (Hodges & Patterson, 2007). Patienterna har relativt långt in i sjukdomsförloppet välbevarad minnesförmåga, orientering och perception, men exekutiva funktioner försämras tidigt (Elfgren et al., 2011)

och försämringen är utmärkande för sjukdomen. Till skillnad från patienter som drabbats av språkvarianterna av FTD har bvFTD-patienterna ofta bristande sjukdomsinsikt (Kipps, Knibb & Hodges, 2007).

I SD sker assymmetrisk atrofi i temporalloberna, framför allt de inferiora och mediala vindlingarna (Snowden, 2013). Kliniska kännetecken för SD är smygande debut och långsam progress liksom språkstörning som karaktäriseras av flytande men innehållslöst tal, förlust av ords betydelse – försämrade benämningsförmåga och förståelse – samt semantisk parafasi (Neary et al., 1998). Därtill kan komma en perceptuell störning som karaktäriseras av prosopagnosi, oförmåga att känna igen ansikten, och/eller associativ agnosi, oförmåga att känna igen objekt. Kopieringsförmågan ska dock vara bevarad, enstaka ord ska kunna repeteras och vanliga ord ska kunna läsas högt och nedtecknas efter diktamen. Exekutiva funktioner är längre bevarade hos patienter med språkvarianterna av FTD jämfört med bvFTD (Hodges & Patterson, 2007).

I PSP sker nervcellsdegeneration i de basala ganglierna, övre delen av hjärnstammen och cerebellum samt motorcortex och posteriora frontalloberna (Elfgren et al., 2011). Kliniska kännetecken för PSP är balansstörning där fallolyckor ofta förekommer tidigt, tilltagande okulomotoriska störningar som bland annat försvårar frivillig vertikal ögonrörelse nedåt, motoriska störningar som rigiditet och nedsatt ansiktsmimik, kognitiv förlångsamning och försämring av exekutiva funktioner (Lovestone, 2009). En utmärkande egenskap för sjukdomen är bristande exekutiva funktioner som visas genom apati, nedsatt omdöme, nedsatt sekvenseringsförmåga, nedsatt mental flexibilitet och växlingssvårigheter.

Resultat på Hayling Test för patienter med FTD. Flera studier har jämfört friska personers resultat på Hayling Test med patientgrupper som har diagnosticerats med FTD. Hayling Test förefaller inte ha utprovats på andra sjukdomar inom FTD-komplexet än bvFTD.

Nästan alla studier (Lough et al., 2006; Hornberger et al., 2010; Hornberger et al., 2011; O’Callaghan, Naismith, Hodges, Lewis & Hornberger, 2013) fann signifikanta skillnader i testprestation på måttet Skalpoäng totalt, där patientgrupperna presterade sämre än de friska kontrollgrupperna. Detta är undantaget Hornberger, Piguet, Kipps och Hodges (2008) som undersökte bvFTD-patienter med snabb progress samt bvFTD-patienter med långsam eller ingen progress. De fann signifikant sämre testprestation endast vid jämförelse mellan patientgruppen med snabb progress och den friska kontrollgruppen. I tabell 2 redovisas en sammanställning över patientgruppernas resultat.

Tabell 2

Resultat på Hayling Test för patienter med bvFTD

Studie	N	Skalpoäng totalt M (SD) ^a
Lough et al. (2006)	18	2,7 (2,0)
Hornberger et al. (2008)	27 ^b	3,1 (1,7)
Hornberger et al. (2008)	23 ^c	4,4 (1,7)
Hornberger et al. (2010)	11	2,3 (1,8)
Hornberger et al. (2011)	14	2,6 (1,9)
O'Callaghan et al. (2013)	11	2,0 (1,8)

Not. ^a På skalan 1-10; ^b bvFTD-patienter med snabb progress; ^c bvFTD-patienter med långsam eller ingen progress.

Hornberger et al. (2011) jämförde friska personers resultat på Hayling Test med dels en patientgrupp som var diagnosticerad med AD, dels en patientgrupp som var diagnosticerad med bvFTD. De fann att framför allt råpoängen på delmättet Felpoäng var mycket användbar för att skilja mellan grupperna (bvFTD > AD > friska). Hornberger et al. (2010) noterade emellertid att en liten grupp (8,3 %) av patienterna i studien med bvFTD-diagnos uppvisade normal testprestation på Hayling Test måttet Skalpoäng totalt och normal hjärnabbildning av orbitofrontala cortex. Flera andra studier (Hornberger et al., 2008; Hornberger et al., 2011; O'Callaghan et al., 2013) som bara presenterade resultat gällande detta på delmättet Felpoäng fann ett liknande mönster; en liten grupp patienter presterade på ungefär samma nivå som genomsnittsvärdet i kontrollgruppen.

Syfte

Studiens övergripande syfte är att psykometriskt utvärdera en svenskspråkig översättning av Hayling Test, för att kunna anpassa testet till svenska förhållanden. Den svenskspråkiga översättningen kommer fortsättningsvis att benämnas som Hayling-testet för att lättare kunna differentiera mellan svenska och internationella versioner.

I studien undersöks huruvida resultatet på Hayling-testet skiljer sig mellan män och kvinnor och om det finns samband med ålder respektive utbildningsnivå. Den psykometriska utvärderingen sker genom att undersöka testets homogenitetsreliabilitet, interbedömarreliabilitet, begreppsvaliditet och diagnostiska validitet. Syftet är också att undersöka nyttan av Hayling-testet vid utredning av patienter med misstänkt FTD. Detta är ett sätt att undersöka Hayling-testets diagnostiska validitet.

Hypoteser

Studiens riktade hypoteser är:

- Det föreligger ett signifikant negativt samband mellan ålder och prestation på Hayling-testet.
- En patientgrupp med klinisk FTD-diagnos presterar signifikant sämre än en frisk kontrollgrupp på Hayling-testet.

Metod

Deltagare

Kontrollgrupp. Rekrytering och testning av kontrollgruppen pågick under september och oktober 2013 samt skedde i Malmö och Lund. Skriftlig information om studien förmedlades till potentiella deltagare, se bilaga. Om möjligt kompletterades detta med muntlig information. Rekryteringen skedde genom träffpunkter och mötesplatser för pensionärer samt genom privat kontaktnät. Träffpunkter och mötesplatser kontaktades per telefon och fick sedan skriftlig information, se bilaga, för att undersöka intresset bland besökarna. Om personalen önskade besökte författaren träffpunkter och mötesplatser för att själv informera potentiella deltagare. Privat kontaktnät fick information om studien genom bekanta eller författaren och kunde sedan höra av sig för att delta. Ambitionen var att uppnå spridning vad gäller variablerna ålder, kön och utbildning. Inklusionskriterierna var att testpersonerna för närvarande skulle vara psykiskt och neurologiskt friska, tala och förstå flytande svenska samt vara över 40 år. Psykiskt och neurologiskt frisk definierades som avsaknad av diagnos. Två deltagare plockades bort efter genomförd testning då dessa inte uppfyllde inklusionskriterierna. En deltagare avbröt testningen då denna upplevde att ett deltest var för svårt (Hayling-testet sektion 2). Kontrollgruppen innefattade efter detta 76 personer.

Patientgrupp. Patientgruppen bestod av 30 patienter. Patienterna hade erhållit en klinisk FTD-diagnos på Minneskliniken vid Skånes universitetssjukhus i Lund enligt Internationell statistisk klassifikation av sjukdomar och relaterade hälsoproblem, ICD-10-SE (Socialstyrelsen, 2010). Patienterna deltog i Lund Prospective Frontotemporal Dementia Study, LUPROFS. LUPROFS är en pågående longitudinell studie på Enheten för geriatrisk psykiatri, Institutionen för kliniska vetenskaper vid Lunds universitet. För att uppfylla inklusionskriterierna i föreliggande studie skulle patienten vara inkluderad i LUPROFS, ha diagnosticerats med en sjukdom inom FTD-komplexet och ha testats med och fullföljt

Hayling-testet till och med oktober 2013. Ett exklusionskriterium var att man inte skulle ha diagnosticerats med en klinisk PNFA-diagnos. PNFA kännetecknas bland annat av språkstörning som karaktäriseras av icke-flytande tal med minst en av följande: agrammatism, fonemisk parafasi eller anomi (Neary et al., 1998). Eftersom Hayling-testet är ett verbalt test och PNFA-patienter ofta har svårt med talproduktionen är det osäkert om man genom Hayling-testet får ett tillförlitligt mått på denna patientgrupps inhibitoriska förmåga.

Instrument

Två neuropsykologiska test användes i studien. Hayling-testet var det ena testet, vilket skulle utvärderas i svensk kontext utifrån syfte och hypoteser ovan. Color Word Interference Test (CWIT; Delis, Kaplan & Kramer, 2001) var det andra testet, vilket användes som jämförelsematerial för att kunna utvärdera Hayling-testets begreppsvaliditet. Snowden (2013) och Belleville et al. (2006) nämner Hayling Test och Stroop, vilket CWIT bygger på, som exempel på test som mäter ungefär samma funktion – verbal inhibitionsförmåga.

Hayling-testet. Det engelskspråkiga Hayling Test samt bakomliggande teori och undersökningsresultat har redan beskrivits i avsnittet Att mäta verbal initiering och inhibition genom Hayling Test. I detta avsnitt redogörs för det svenskspråkiga Hayling-testet.

Hayling Test översattes till svenska av docent och psykolog Christina Elfren samt forskningsadministratör och logoped Helena Andersson för användning i LUPROFS (Personlig kommunikation, H. Andersson, 23 maj, 2013). Den svenskspråkiga översättningen utgick från de engelskspråkiga meningarna. Svenskspråkiga bedömningskriterier togs inte fram, utan bedömare använde sig av den engelskspråkiga manualen. Testets premisser, administration och bedömning är alltså desamma i Hayling Test och Hayling-testet. I de fall de svenska och engelska meningarna skiljde sig åt använde bedömare sig av de generella bedömningskriterierna i den engelskspråkiga manualen. En mindre skillnad mellan versionerna var försvenskningar av namn i den svenska versionen. Större skillnad utgjordes av byte av preposition i mening där ”up the” blev ”in i”. En annan större skillnad var att vissa meningar i den svenska versionen inte hade ett lika självklart avslutningsord som de engelska, till exempel översattes ”At last the time for action had...” till ”Det var äntligen dags att...”.

CWIT. Detta test ingår i Delis Kaplan Executive Function System, D-KEFS (Delis et al., 2001), ett av de mest välanvända testbatterierna avsedda att mäta exekutiva funktioner.

CWIT bygger på Stroop-testet från 1935. CWIT består av fyra betingelser eller deltest. I betingelse 1 och betingelse 2 ska testpersonerna benämna färger respektive läsa färgord som röd, grön och blå. I betingelse 3 ska testpersonerna benämna färgen som färgorden är tryckta med, en annan färg än själva färgordet. Detta test avser att mäta inhibitionsförmågan av en överinlärd verbal respons. I betingelse 4 ska testpersonerna växla mellan att benämna färgen som färgorden är tryckta med och att läsa själva färgordet. Detta test avser att mäta inhibitionsförmåga och växling. Betingelse 3 som avser att mäta en form av verbal inhibitionsförmåga valdes ut som jämförelsematerial till Hayling-testet. Kontrollpersonerna genomförde emellertid alla fyra betingelser för att testprestationen i betingelse 3 inte skulle riskera att vara missvisande.

Administrering av CWIT betingelse 3. Enligt instruktionerna i det svenska manualsupplementet (Delis, Kaplan & Kramer, 2001/2005) som använts i studien lägger testledaren fram ett stimuluskort med flera rader färgord som är tryckta i andra färger än själva ordet, och säger att betingelsen kommer att vara lite svårare jämfört med tidigare betingelser. Testledaren berättar att betingelsen går ut på att säga den tryckta färgen, inte att läsa färgordet, och ber testpersonen att läsa de tre första itemen. Här kan testledaren korrigera eventuella fel. Sedan ombeds testpersonen att öva på resten av övningsraderna. Om testpersonen inte har förstått utför testledaren den första övningsraden som hjälp och ber testpersonen att utföra den andra övningsraden. Läser testpersonen fel på fyra eller fler item under övningen så administreras inte testet. Om övningen klaras av ombeds testpersonen att fortsätta på resten av pappret, från vänster till höger ända fram till slutet. Uppgiften ska utföras så snabbt men så korrekt som möjligt. Tidtagning påbörjas när testpersonen är beredd. Testpersonen får följa raderna på stimuluskortet med ett finger. Testledaren antecknar eventuella felsägningar och om de korrigeras av testpersonen eller inte.

Om testpersonen råkar hoppa över en rad påpekar testledaren detta genast utan att stanna tidtagningen. Om testpersonen inte är färdig med betingelsen efter 180 sekunder avbryts uppgiften och 180 antecknas som råpoäng för Tidsåtgång. Item som inte hinns med bedöms inte som fel.

CWIT betingelse 3 består av det primära måttet Tidsåtgång som kan konverteras till Wechsler-skalpoäng från ett till nitton ($M = 10$, $SD = 3$). Som tilläggsmått finns möjlighet att göra felanalyser, där antal korrigerade respektive okorrigerade fel i råpoäng kan omvandlas till percentiler. Korrigerade samt okorrigerade fel utgör måttet Fel totalt och kan summeras i

råpoäng samt omvandlas till skalpoäng. Den svenska versionen av CWIT har använt den amerikanska normeringen. Normer för olika åldersgrupper finns och skalpoängen baseras på ålderskategori. Hela D-KEFS-batteriet, således även CWIT, utprovades på ett stratifierat stickprov bestående av 1 750 personer mellan 8 och 89 år. I CWIT betingelse 3 sker enligt normeringen en markant försämring i testresultat bland äldre personer avseende både Tidsåtgång och Fel totalt (Delis et al., 2001).

Procedur

Kontrollgrupp. I testsituationen kontrollerades muntligt att inklusionskriterierna för att medverka i studien var uppfyllda. Samtliga testpersoner fick möjlighet att genomföra Hayling-testet och CWIT. Testningen av kontrollgruppen tog 20-30 minuter per person. För att inte ordningen som testen administrerades i skulle påverka resultatet testades ena hälften av kontrollgruppen först med Hayling-testet och sedan med CWIT, andra hälften först med CWIT och sedan med Hayling-testet. Testningen utfördes av författaren. I kontrollgruppen testades ett flertal seniorer på träffpunkter och mötesplatser i Malmö och Lund, då det bedömdes som svårt att motivera dem till att testas på annan plats. Vissa testpersoner i kontrollgruppen deltog också i testning i samtalsrum på deras arbetsplats. Resterande personer i kontrollgruppen testades på Minneskliniken i Lund och Päronskogens vårdboende i Malmö. Samtliga testningar skedde i avskild miljö för att optimera prestationen och i görligaste mån standardisera förfarandet.

Patientgrupp. Patientgruppen genomgick en första neuropsykologisk undersökning och bedömning på cirka tre timmar som utfördes av psykologer vid Minneskliniken. Därefter deltog patienterna i kompletterande neuropsykologisk undersökning på cirka 1,5 timme inom ramen för LUPROFS som genomfördes av psykologer vid Minneskliniken eller författaren. Patientgruppen testades mellan juni 2009 och oktober 2013. Därifrån inhämtades rådata från Hayling-testet. Samtliga patienter testades på Minneskliniken i Lund i för ändamålet avsedda testrum.

Rättning. CWIT rättades enligt manual (Delis et al., 2001). Hayling-testet i både kontroll- och patientgrupp rättades av författaren enligt bedömningskriterier på itemnivå i den engelska manualen (Burgess & Shallice, 1997). I de fall översättningen till svenska skilde sig

från den engelska betydelsen så att specifika bedömningskriterier inte var tillämpliga användes istället de generella kriterierna i den engelska manualen.

För att utvärdera interbedömarreliabiliteten rättade en mångårigt erfaren psykolog vid Minneskliniken de 45 första Hayling-protokollen i kontrollgruppen. Psykologen rättade delmåttet Felpoäng, då endast detta delmått kräver subjektiv bedömning av många tänkbara sorters svar. Rättningen skedde enligt samma princip som ovan, men för tre item utarbetades i förväg mindre tillägg till manualen för poängsättning av deltagarnas svar.

Två deltagare ur kontrollgruppen kunde på minst en av meningarna i Hayling-testet inte svara inom angiven tid, 60 sekunder. Enligt manualen (Burgess & Shallice, 1997) ska då 60 sekunder noteras som svarslatens för aktuellt item. Det står emellertid inte om överskriden tidsgräns ska bedömas även som ett kategori A- eller B-fel i Felpoäng. I föreliggande studie påverkade överskriden tidsgräns bara tidsmåten (Tid sektion 1 och Tid sektion 2).

Etik

Föreliggande studie gick under etikansökan för LUPROFS, som är godkänd av Regionala etikprövningsnämnden i Lund, registreringsnummer 617/2008. Alla patienter skrev under informerat samtycke i samband med inklusion i LUPROFS. Deltagarna i kontrollgruppen gav muntligt samtycke i samband med testningen och informerades om att de när som helst kunde avbryta sitt deltagande.

I de fall rekryteringen av kontrollgruppen skedde genom privat kontaktnät testades inte personer som författaren hade en närmare personlig relation till. Till denna grupp räknades släktingar, vänner, kollegor och före detta kollegor till författaren.

Analys av materialet skedde på gruppnivå och deltagarna garanterades anonymitet. För att kunna göra statistiska beräkningar av de demografiska variablerna samlades dock data om deltagarnas ålder, kön och utbildningsnivå in, vilket godtogs av deltagarna.

Dataanalys

Statistiska analyser utfördes i SPSS 21.0.

Råpoängen på delmåtten samt antalet kategori A- respektive B-fel användes i analyserna. Dessutom konverterades råpoängen till skalpoäng utifrån den engelska normeringen och lades samman till Sammanlagda skalpoäng och Skalpoäng totalt. Måttet Sammanlagda skalpoäng användes enbart för att kunna jämföra kontrollgruppens resultat med den engelska normgruppens åldersbundna resultat på den femte percentilen. Måttet Skalpoäng

totalt användes enbart för att jämföra kontrollgruppens resultat med den engelska normeringen och för att undersöka skillnader i testresultat mellan kontrollgrupp och patientgrupp.

Eventuella skillnader i testprestation mellan könen undersöktes genom oberoende t -test för olika varianser. Effektstorleken undersöktes genom Cohens d . Sambandet mellan ålder och testprestation undersöktes genom Pearsons r . Sambandet mellan utbildningsnivå och testprestation undersöktes genom det icke-parametriska korrelationstestet Spearmans rho, r_s , eftersom data var på ordinalskalenivå.

Homogenitetsreliabiliteten undersöktes på kontrollgruppens resultat på delmåtten genom Cronbachs alfa.

Interbedömarreliabiliteten undersöktes på kontrollgruppens resultat på Felpoäng genom intraklasskorrelation, ρ_r , för genomsnittliga mätvärden. Intraklasskorrelationen var av tvåvägs blandad typ med krav på absolut överensstämmelse.

Begreppsvaliditeten undersöktes i termer av konvergent validitet genom att granska sambandet mellan kontrollgruppens resultat på Hayling-testet Tid sektion 2 och CWIT betingelse 3 Tidsåtgång. Här jämfördes korrelationen mellan råpoängen på de båda måtten, som båda mäter använd tid i sekunder. Sambandet mellan kontrollgruppens resultat på Hayling-testet Felpoäng och CWIT betingelse 3 Fel totalt granskades också. Här jämfördes korrelationen mellan råpoängen på de båda måtten, som båda mäter antal fel (i Hayling-testets fall antalet felpoäng, som stiger i takt med antalet kategori A- och B-fel). Dessa test avser båda att mäta inhibition. Sambandet undersöktes genom Pearsons r .

Den diagnostiska validiteten undersöktes genom att jämföra kontrollgruppens och patientgruppens resultat på Hayling-testet. Eventuella skillnader mellan gruppernas testprestation undersöktes genom oberoende t -test för olika varianser.

Valet av oberoende t -test som analysmetod skedde på basis av att stickproven ansågs vara tillräckligt stora för att vara normalfördelade enligt centrala gränsvärdessatsen, där cirka trettio personer per grupp är en tumregel för tillräckligt stort stickprov (Field, 2013). Grupperna som skulle jämföras var ojämnt fördelade med fler kvinnor än män i kontrollgruppen och fler personer i kontrollgruppen än i patientgruppen. Därför användes konsekvent t -test för olika varianser (*equal variances not assumed*), även kallat Welch-Satterthwaite t -test, enligt Fields (2013) rekommendationer.

Vid tolkning av Cronbachs alfa användes Cicchetti och Sparrows (1990, refererat i Cicchetti, 1994) riktlinjer, där Cronbachs $\alpha < 0,7$ anses vara dåligt eller otillräckligt; $\geq 0,7$ acceptabelt; $\geq 0,8$ bra och $\geq 0,9$ utmärkt. Vid tolkning av intraklasskorrelation användes

Cicchettis (1994) riktlinjer, där $\rho_I < 0,40$ anses vara dåligt; $\geq 0,40$ acceptabelt; $\geq 0,60$ bra och $\geq 0,75$ utmärkt. Vid tolkning av effektstorlek användes Cohens (1988) riktlinjer, där $d \geq 0,20$ anses utgöra en liten effekt; $\geq 0,50$ en måttlig effekt och $\geq 0,80$ en stor effekt. Vid tolkning av korrelationer användes Cohens (1988) riktlinjer, där r respektive $r_s \geq 0,10$ anses vara en svag korrelation; $\geq 0,30$ anses vara en måttlig korrelation och $\geq 0,50$ anses vara en stark korrelation.

Resultat

Inledningsvis redovisas kontrollgruppens och patientgruppens karaktäristika. Därefter följer kontrollgruppens resultat i sin helhet samt de demografiska variabelernas betydelse. Sedan redogörs för mått på homogenitetsreliabilitet och interbedömarreliabilitet. Slutligen redovisas resultat för begreppsvaliditet och diagnostisk validitet.

Deltagarnas karaktäristika

I tabell 3 redovisas deskriptiv data över kontrollgruppen och patientgruppen, där variabelerna kön, ålder och utbildningsnivå finns med.

Tabell 3

Kontrollgruppens och patientgruppens karaktäristika

Demografisk variabel	Kontrollgrupp, $N = 76$	Patientgrupp, $N = 30$
Kön		
Män (%)	26 (34,2)	10 (33,3)
Kvinnor (%)	50 (65,8)	20 (66,7)
Ålder		
M (SD)	66,97 (11,56)	69,30 (6,19)
Min-max	40-95	57-83
Utbildningsnivå		
Grundskola eller motsvarande ^a (%)	25 (32,9)	17 (56,7)
Gymnasium (%)	17 (22,4)	5 (16,7)
Högskola/universitet (%)	34 (44,7)	8 (26,7)

Not. ^a Sju- till tioårig utbildning motsvarande grundskola.

I patientgruppen var tolv personer diagnosticerade med en bvFTD-diagnos, elva personer var diagnosticerade med en PSP-diagnos, fem personer var diagnosticerade med en SD-diagnos

och en person var diagnosticerad med FTD utan närmare specifikation. En person i patientgruppen var diagnosticerad med misstänkt bvFTD, ännu ej demensdignitet.

Kontrollgruppens resultat

Kontrollgruppen uppvisade i sin helhet längre svarslatenser på Tid sektion 2 än på Tid sektion 1 och gjorde fler kategori B-fel än kategori A-fel. Kontrollgruppen presterade som mest sju skalpoäng av tio på Skalpoäng totalt. Kontrollgruppens resultat redovisas i sin helhet i tabell 4.

Tabell 4

Kontrollgruppens resultat på Hayling-testet, $N = 76$

Mått	M (SD)	Min-max	Percentil				
			5	25	50	75	95
Tid sektion 1 råpoäng	14,41 (14,59)	0-74	45,00	18,75	9,50	4,25	1,00
Tid sektion 2 råpoäng	52,17 (42,64)	2-192	152,50	67,75	38,50	22,00	5,00
Kategori A-fel	1,62 (1,67)	0-6	5,00	3,00	1,00	0,00	0,00
Kategori B-fel	4,62 (2,64)	0-9	9,00	7,00	5,00	2,00	0,00
Felpoäng råpoäng	15,82 (13,50)	0-43	42,00	27,00	12,50	3,25	1,00
Skalpoäng totalt ^a	4,20 (1,79)	1-7	1,00	3,00	5,00	6,00	6,00

Not. ^a Baserat på engelska normer för Hayling Test.

På måttet Sammanlagda skalpoäng (min: 3; max: 23) motsvarar den sammanlagda skalpoängen 7 den femte percentilen. Det fanns åtta outliers på Tid sektion 1, varav två extremvärden. Det fanns fyra outliers på Tid sektion 2, varav inga extremvärden. Det fanns inga outliers på Felpoäng.

Jämförelse mellan resultat på Hayling-testet och de demografiska variablerna

Männen tenderade att prestera bättre än kvinnorna, men inga signifikanta skillnader fanns förutom på Tid sektion 2. Se tabell 5. På Tid sektion 2 förelåg också en måttlig effektstorlek i jämförelse mellan mäns och kvinnors resultat. I övrigt var effektstorlekarna små eller obefintliga.

Tabell 5

Jämförelse mellan män och kvinnors resultat på Hayling-testet

Mått	Män, $n = 26$ $M (SD)$	Kvinnor, $n = 50$ $M (SD)$	$t(df)$	p	d
Tid sektion 1 råpoäng	10,88 (7,99)	16,24 (16,83)	-1,88(73,588)	0,06	-0,41
Tid sektion 2 råpoäng	39,35 (32,46)	58,84 (45,96)	-2,14(67,093)	0,04	-0,50
Kategori A-fel	1,42 (1,70)	1,68 (1,68)	-0,63(50,294)	0,53	-0,15
Kategori B-fel	4,19 (2,59)	4,86 (2,63)	-1,06(51,478)	0,29	-0,26
Felpoäng råpoäng	13,96 (13,29)	16,78 (13,64)	-0,87(51,931)	0,39	-0,22

Not. Ett negativt t - och d -värde innebär att män hade lägre värden än kvinnor. Ju lägre medelvärden på delmåten, desto bättre resultat.

Signifikanta positiva samband mellan ålder och råpoäng förelåg på Hayling-testets samtliga delmått, där högre råpoäng innebär ett sämre resultat. Se tabell 6. Korrelationerna var svaga till måttliga på delmåttsnivå, men en stark positiv korrelation fanns mellan ålder och antalet kategori A-fel.

Tabell 6

Samband mellan resultat på Hayling-testet och ålder respektive utbildningsnivå, $N = 76$

Mått	Ålder	Utbildningsnivå
Tid sektion 1 råpoäng	$r = 0,23^{**}$	$r_s = -0,37^{**}$
Tid sektion 2 råpoäng	$r = 0,45^{**}$	$r_s = -0,22$, ns
Kategori A-fel	$r = 0,56^{**}$	$r_s = -0,16$, ns
Kategori B-fel	$r = 0,30^{**}$	$r_s = -0,03$, ns
Felpoäng råpoäng	$r = 0,46^{**}$	$r_s = -0,09$, ns

Not. $** p < 0,01$

Det fanns ett signifikant negativt samband mellan utbildningsnivå och resultat på Tid sektion 1, det vill säga personer med högre utbildningsnivå hade kortare svarslatenser i sektion 1. Se tabell 6. Korrelationen var måttlig. I övrigt erhöles inga signifikanta korrelationer mellan utbildningsnivå och testprestation.

Homogenitetsreliabilitet

Alfakoefficienten för Tid sektion 1 var 0,37. Detta tyder på otillräcklig intern överensstämmelse. Analys vid borttagande av enskilt item visade att reliabiliteten skulle öka

till Cronbachs $\alpha = 0,51$ om man tog bort item 5 i sektion 1, men även detta värde är otillräckligt.

Alfakoefficienten för Tid sektion 2 var 0,81. Detta tyder på god intern överensstämmelse. Analys vid borttagande av enskilt item visade att reliabiliteten inte skulle öka om man tog bort något item från sektion 2.

De två deltagarna som inte kunde svara inom 60 sekunder exkluderades från beräkningen av homogenitetsreliabiliteten för Felpoäng. Detta eftersom deras svar inte kunde bedömas som vare sig rätt (utan samband), kategori A-fel eller kategori B-fel.

Alfakoefficienten för Felpoäng bestod följaktligen av 74 deltagares svar.

Alfakoefficienten för Felpoäng var 0,81. Detta tyder på god intern överensstämmelse. Analys vid borttagande av enskilt item visade att reliabiliteten skulle öka marginellt, till Cronbachs $\alpha = 0,82$, om man tog bort item 1 i sektion 2.

Interbedömarreliabilitet

Interbedömarreliabiliteten var utmärkt, $\rho_I = 0,97$, för Felpoäng råpoäng. På itemnivå varierade intraklasskorrelationen mellan 0,67 och 0,98. Se tabell 7. Alla 15 item i Felpoäng hade bra interbedömarreliabilitet eftersom $\rho_I \geq 0,60$. Tretton av 15 item i Felpoäng, item 6 och 11 undantagna, uppnådde dessutom utmärkt interbedömarreliabilitet enligt riktlinjen $\rho_I \geq 0,75$.

Tabell 7

Intraklasskoefficienter för Hayling-testet Felpoäng

Mått	ρ_I	Mått	ρ_I
Felpoäng item 1	0,89	Felpoäng item 9	0,93
Felpoäng item 2	0,94	Felpoäng item 10	0,95
Felpoäng item 3	0,91	Felpoäng item 11	0,67
Felpoäng item 4	0,95	Felpoäng item 12	0,91
Felpoäng item 5	0,88	Felpoäng item 13	0,98
Felpoäng item 6	0,70	Felpoäng item 14	0,81
Felpoäng item 7	0,92	Felpoäng item 15	0,86
Felpoäng item 8	0,78	Felpoäng råpoäng	0,97

Not. Analys genomförd på två personers bedömning av de 45 första Hayling-testen i kontrollgruppen.

Begreppsvaliditet

Det fanns en svag men signifikant positiv korrelation ($r = 0,24$, $p = 0,035$) mellan Hayling-testet Tid sektion 2 råpoäng och CWIT betingelse 3 Tidsåtgång råpoäng.

Det fanns en måttligt stark och höggradigt signifikant positiv korrelation ($r = 0,43$, $p = 0,0001$) mellan Hayling-testet Fel-poäng råpoäng och CWIT betingelse 3 Fel totalt råpoäng.

Diagnostisk validitet

Patientgruppen presterade signifikant sämre än kontrollgruppen på Hayling-testets samtliga delmätt. Se tabell 8. Patientgruppen hade även signifikant fler kategori A-fel än kontrollgruppen, men skillnaden i antalet kategori B-fel var inte signifikant. Effektstorlekarna var stora på Tid sektion 1, Tid sektion 2 och på Skalsoäng totalt. I övrigt var effektstorlekarna måttliga eller obefintliga.

Det fanns en outlier på Tid sektion 1 i patientgruppen, men inga extremvärden. På Tid sektion 2 och Fel-poäng fanns inga outliers i patientgruppen.

Tabell 8

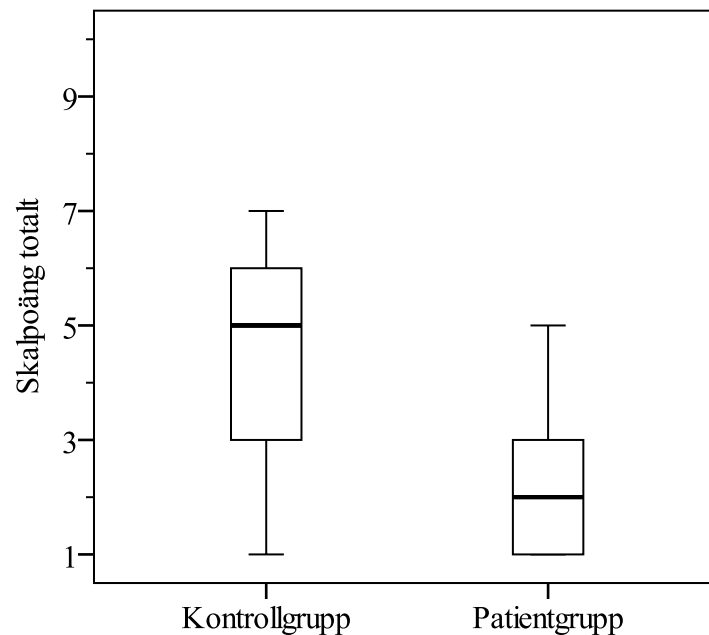
Jämförelse mellan kontrollgruppens och patientgruppens resultat på Hayling-testet

Mått	Kontrollgrupp, $N = 76$ $M (SD)$	Patientgrupp, $N = 30$ $M (SD)$	$t(df)$	p	d
Tid sektion 1 råpoäng	14,41 (14,59)	51,90 (38,18)	-5,23(32,399)	0,0001	-1,30
Tid sektion 2 råpoäng	52,17 (42,64)	96,63 (40,77)	-4,99(55,454)	0,0001	-1,07
Kategori A-fel	1,62 (1,67)	4,13 (4,43)	-3,03(32,316)	0,0001	-0,75
Kategori B-fel	4,62 (2,64)	4,13 (2,69)	0,84(52,416)	0,40	0,18
Fel-poäng råpoäng	15,82 (13,50)	27,17 (22,36)	-2,60(37,647)	0,013	-0,62
Skalsoäng totalt ^a	4,20 (1,79)	2,10 (1,21)	6,95(77,931)	0,0001	1,38

Not. Ett negativt t - och d -värde innebär att kontrollgruppen hade lägre värden än patientgruppen. Ju lägre medelvärden på delmåtten, desto bättre resultat, men ju lägre medelvärden på skalsoängen, desto sämre resultat;

^a Baserat på engelska normer för Hayling Test.

Störst skillnad mellan grupperna uppmättes på Skalsoäng totalt. Skillnaden mellan grupperna på detta mått åskådliggörs i figur 3.



Figur 3. Låddiagram över Hayling-testet Skalpoäng totalt för kontrollgruppen ($N = 76$) och patientgruppen ($N = 30$). Lådorna rymmer 25:e-75:e percentilen, där strecken genom lådorna är medianvärden. De horisontella strecken utanför lådorna indikerar minimum- och maximumvärden för respektive grupp.

Diskussion

Resultatdiskussion

Demografiska variabler. I enlighet med studiens syfte att utvärdera Hayling-testets psykometriska egenskaper utprovades testet på en frisk kontrollgrupp. Inga signifikanta skillnader mellan män och kvinnors resultat på Hayling-testet uppmättes, förutom på Tid sektion 2. Där var männen signifikant snabbare än kvinnorna med en måttlig effektstorlek. Resultatet är bara delvis förenligt Bielak et al. (2006), som inte fann några signifikanta skillnader mellan män och kvinnors resultat på det engelskspråkiga Hayling Test.

Hypotesen att det förelåg ett signifikant negativt samband mellan ålder och prestation på Hayling-testet fick stöd av resultatet. På samtliga delmått var korrelationen svag eller måttlig. Resultatet är i enlighet med Bielak et al. (2006) som också fann ett signifikant samband mellan hög ålder och sämre resultat på samtliga delmått. Föreliggande studies resultat stämmer bara delvis överens med Andrés och Van der Lindens (2000) samt Belleville et al.:s (2006) resultat, vilka kom fram till att äldre personer presterar signifikant sämre än yngre personer på Tid sektion 2 och Felpoäng men inte på Tid sektion 1. Resultatet skiljer sig från studien av Borella et al. (2011) som inte fann en signifikant skillnad mellan yngre vuxnas

och äldre vuxnas resultat på något av delmått. Skillnaderna mellan de olika studierna kan bero på att olika versioner av testet har använts.

Inga signifikanta samband mellan utbildningsnivå och prestation på Hayling-testet uppmättes, förutom på Tid sektion 1. Där var personer med högre utbildningsnivå signifikant snabbare än personer med lägre utbildningsnivå. Korrelationen var måttlig. Resultatet är förenligt med Bielak et al. (2006).

Riktlinjerna för att tolka de olika analysmetoderna ska betraktas som just riktlinjer. Hur stark en korrelation är avgörs av sammanhanget – forskning om psykologiska test kräver generellt sett mindre r -värde för att en korrelation ska betraktas som stark än naturvetenskaplig forskning. Valda riktlinjer för analysmetoderna baseras på Cohens (1988) och Cicchettis (1994) rekommendationer, namnkunniga forskare inom psykologi.

Homogenitetsreliabilitet. Otillräcklig intern överensstämmelse i Tid sektion 1, där inte heller god intern överensstämmelse uppnås vid borttagande av enskilt item, gör att hela delmåttets samstämmighet kan ifrågasättas. Resultatet är överraskande eftersom premissen är att Tid sektion 1 ska ge ett mått på testpersonens initiationsförmåga, en rutinhandling. I Shallice och Burgess (1998) termer rör det sig om en aktivering av köbildningsprocessen enligt modellen över det uppmärksamhetskontrollerande systemet. Låg intern samstämmighet gör det inte troligt att Tid sektion 1 mäter denna rutinhandling. En anledning kan vara att meningarna i Hayling-testet är direktöversatta från den engelska originalversionen. Hur meningar konstrueras är språk- och kulturbundet; att den engelska meningen ”The rich child attended a private”... kan sluta med ”school” är kanske mer självklart än den svenska översättningen ”Barnet deltog i en privat”... kan sluta med ”lektion”. Svartsvarianterna på denna övningsmening har varit många, såsom ”skola”, ”fest”, ”lek”, ”grupp” och ”sammankomst”. Variationen mellan item avseende svarslatensen beror sannolikt på att vissa item i Tid sektion 1 i den svenska versionen inte har något självklart svar och därmed kräver längre eftertanke än andra item. Flera andra versioner av Hayling Test, däribland den engelska originalversionen, är konstruerade genom förstudier där de meningar som försökspersonerna oftast avslutar med ett specifikt ord valts ut. Ett sådant förfarande hade varit önskvärt i den svenskspråkiga versionen. En annan anledning till den låga samstämmigheten kan vara att Hayling-testets reliabilitet bara är utprovad på en frisk kontrollgrupp. På Tid sektion 1 i den engelska versionen var split-half-reliabiliteten 0,35 för friska kontroller, vilket ligger i linje med föreliggande studies resultat för homogenitetsreliabilitet, men när en patientgrupp med

frontala lesioner testades förbättrades reliabiliteten (Burgess & Shallice, 1997). Eftersom homogenitetsreliabiliteten inte är undersökt på resultatet hos patientgruppen i denna studie är det oklart om en liknande effekt skulle uppstå i den svenska versionen.

God intern överensstämmelse i Tid sektion 2 och Felpoäng tyder på att item i sektion 2 mäter samma sak. Kanske blir konsekvensen av den snarlika översättningen från engelska till svenska inte lika stor i sektion 2: Om någon av meningarna kan avslutas på flera sätt måste dessa svar ändå inhiberas.

En styrka i föreliggande studie är att Hayling-testet i både kontroll- och patientgrupp rättades av en och samma person. Det gör att man med större säkerhet kan jämföra resultaten inom och mellan grupperna. Därmed bidras också med säkrare värden för homogenitetsreliabiliteten undersökt på kontrollgruppen, eftersom samma bedömningskriterier för Felpoäng är använda.

Interbedömarreliabilitet. Den utmärkta interbedömarreliabiliteten för Felpoäng råpoäng är tillfredsställande eftersom detta delmått ställer krav på subjektiv bedömning av många tänkbara svar. Colton och Covert (2007) varnar annars för att subjektiv bedömning av flera svarsalternativ riskerar att försämra interbedömarreliabiliteten, men resultatet i sin helhet tyder på att tillräckligt utvecklade bedömningskriterier använts vid bedömningen. Därmed borde dessa risker vara förebyggda. För de enskilda item, item 6 och 11, som uppnår bra men inte utmärkt interbedömarreliabilitet bör bedömningskriterierna ses över, så att itemen uppnår kraven för utmärkt.

Begreppsvaliditet. De signifikanta sambanden mellan prestationerna på Hayling-testet sektion 2 och CWIT betingelse 3 stöder Hayling-testets begreppsvaliditet, i enlighet med Groth-Marnat (2003). Den positiva korrelationen var svag mellan tidsmåttan på de båda testen och måttlig mellan felmåttan på de båda testen. Därmed uppfylls delvis rekommendationen om måttlig till hög korrelation mellan testen. Den svaga till måttliga korrelationen indikerar att både likheter och skillnader mellan testen finns. Både Hayling-testet och CWIT avser att mäta verbal inhibitionsförmåga, men Hayling-testet torde inkludera en semantisk komponent och förmågan att utveckla en strategi i högre utsträckning. Detta eftersom testpersonerna förväntas svara med olika ord och dessförinnan lyssna på meningarna så att inte orden hör samman med meningen, till skillnad från CWIT där stimulit är visuellt och antalet svarsalternativ är begränsade till tre färgord.

Jantscher et al. (2011) fann ingen signifikant korrelation mellan det tyskspråkiga Hayling Test sektion 2 och Stroop deltest 3. CWIT betingelse 3 bygger på Stroop deltest 3. Skillnaden mellan föreliggande studies resultat och Jantscher et al.:s (2011) resultat kan bero på att de olika versionerna av Hayling Test påverkar utfallet. Skillnaden kan också bero på olika mätmetoder, eftersom den tyskspråkiga studien inte har specificerat vilket av tidsmått eller felmått som har använts vid beräkningen, inte heller om skalpoäng eller råpoäng har använts.

Diagnostisk validitet. Hypotesen att patientgruppen presterade signifikant sämre än kontrollgruppen på Hayling-testet fick stöd av resultatet. Hayling-testet besitter således god diagnostisk validitet avseende att finna skillnader i egenskapen verbal inhibitionsförmåga mellan en frisk kontrollgrupp och en patientgrupp med FTD. Skillnaden i prestation mellan grupperna stämmer överens med merparten av tidigare forskning på området (Lough et al., 2006; Hornberger et al., 2010; Hornberger et al., 2011; O'Callaghan et al., 2013). Resultatet tyder på att Hayling-testet, det vill säga den svenska versionen av Hayling Test, är användbart i kliniska sammanhang.

På delmåttsnivå är det förvånande att störst skillnad mellan grupperna uppmättes på Tid sektion 1. Det stämmer inte riktigt överens med Hornberger et al. (2011), som uppmätte störst skillnader mellan bvFTD-patienter och friska kontroller på Felpoäng. Skillnaden mellan studiernas resultat kan bero på olika versioner av testet har använts och att föreliggande studie även inkluderade patienter med andra FTD-sjukdomar än bvFTD. Exempelvis är kognitiv förlångsamning ett kännetecken för PSP (Lovestone, 2009) och kanske har patienter med denna diagnos haft längre svarslatenser som sänker den genomsnittliga prestationen på Tid sektion 1 och Tid sektion 2 i föreliggande studie. Om så är fallet har inte undersökts vidare eftersom det ligger utanför studiens syfte.

Även på Tid sektion 2 och Felpoäng presterar patientgruppen signifikant sämre än kontrollgruppen. Eftersom testet primärt klassas som ett test som ska mäta verbal inhibitionsförmåga, och Tid sektion 2 och Felpoäng avser att mäta denna funktion, är det positivt för Hayling-testets diagnostiska validitet att skillnader i prestation mellan grupperna föreligger även på dessa mått. Antalet kategori A- och B-fel ligger till grund för delmättet Felpoäng. Patientgruppen hade emellertid bara signifikant fler kategori A-fel än kontrollgruppen, medan inga signifikanta skillnader i antalet kategori B-fel fanns. Det är svårt att dra slutsatser om detta är avvikande eller är vanligt förekommande, eftersom tidigare

studier som undersökt resultat på Hayling Test för patienter med FTD inte redovisar antalet kategori A- respektive B-fel.

Den stora skillnaden mellan grupperna på Skalpoäng totalt visar att delmåttens tillsammans mest effektivt skiljer mellan grupperna. Det är inte möjligt att uppmäta större skillnad mellan grupperna genom att använda sig av bara ett av delmåttens. Värt att notera är dock att en liten andel av kontrollgruppen, med resultat under den femte percentilen, uppnår endast skalpoängen 1 på Skalpoäng totalt som baseras på den engelska normeringen. Med tanke på att antalet deltagare i kontrollgruppen var relativt stort ($N = 76$) är det statistiskt inte konstigt att enstaka värden ligger betydligt under genomsnittet. På samma vis erhåller en liten andel av patientgruppen skalpoängen 5 på Skalpoäng totalt, vilket är i nivå med kontrollgruppens prestation. Resultatet avviker inte från tidigare forskning eftersom flera andra studier (Hornberger et al., 2008; Hornberger et al., 2011; O'Callaghan et al., 2013) antyder och Hornberger et al. (2010) i klartext beskriver att en liten grupp patienter presterar på ungefär samma nivå som genomsnittsvärdet i kontrollgruppen.

Det faktum att Hayling-testet inte helt säkert kan skilja mellan grupperna kan bero på att det finns enstaka patienter som inte har drabbats av försämrade verbal inhibitionsförmåga, antingen eftersom sjukdomen inte progredierat eller eftersom symtombilden är en annan. Det kan också vara så att testet är inte tillräckligt känsligt för just dessa patienters inhibitoriska problematik. I sådana fall kan testet kompletteras med test som avser att mäta ungefär samma sak, såsom Stroop eller CWIT, eller genom kvalitativa beteendeobservationer. Det är värt att komma ihåg att testet i likhet med andra neuropsykologiska test inte är utformat som ett komplett diagnostiskt verktyg utan som ett av flera test i en neuropsykologisk utredning och bedömning. Låga poäng på test som avser att mäta exekutiva funktioner indikerar en funktionell försämring som ofta associeras med, men inte kan likställas med, frontal dysfunktion, som Burgess och Alderman (2013) konstaterar. De neuropsykologiska undersökningsresultaten är sedan en av flera pusselbitar innan diagnos kan ställas. Föreliggande studies resultat stödjer inte att det går att använda endast Hayling-testet för att ställa en specifik diagnos. Konstaterade nedsättningar i verbal inhibitionsförmåga, eller bristande exekutiva funktioner i stort, räcker inte för att en specifik diagnos ska ställas, utan flera andra kriterier som kännetecknar sjukdomen måste uppvisas.

Metoddiskussion

Deltagare. Ambitionen att uppnå spridning vad gäller de demografiska variablerna i kontrollgruppen för ett representativt urval uppnåddes delvis. Det blev en god spridning i deltagarnas åldersfördelning. Däremot medverkade nästan dubbelt så många kvinnor som män i studien i kontrollgruppen, trots försök att rekrytera fler män. Även andelen högskole- eller universitetsutbildade är något hög i jämförelse med övriga utbildningsnivåer. Kontrollgruppens skevhet i könstillhörighet och utbildningsnivå är en brist eftersom det blir svårare att generalisera till befolkningen i stort utifrån deltagarnas resultat på Hayling-testet.

I den skriftliga informationen till potentiella deltagare, se bilaga, stod att testet skulle undersöka en förmåga som kunde vara påverkad i bland annat FTD. Detta försvårade rekryteringen av deltagare till kontrollgruppen. Några potentiella deltagare avböjde medverkan i studien då de uttryckte rädsla för att demenssjukdom skulle upptäckas genom testet. Avstående från att nämna testets möjliga användningsområden i den skriftliga informationen hade kanske inte väckt de potentiella deltagarnas rädsla, men det bedömdes som mer ärligt att informera om detta. Andra potentiella deltagare blev intresserade av att medverka när testets eventuella kliniska nytta lyftes fram.

Kontrollgruppen presterade snabbare på Tid sektion 1 än på Tid sektion 2 och gjorde fler kategori B-fel än kategori A-fel, vilket överensstämmer med Bielak et al. (2006). Dessa tendenser tyder visserligen på att Hayling-testet har liknande psykometriska egenskaper som den engelska originalversionen. Jämförelser mellan kontrollgruppens resultat (se tabell 4) och den ursprungliga engelska normeringen (se tabell 1) indikerar dock olikheter mellan grupperna, som utgör en begränsning avseende hur resultaten kan användas. På måttet Skalpoäng totalt motsvarar skalpoängen 3 den femte percentilen på det ursprungliga Hayling Test, men på Hayling-testet motsvarar skalpoängen 1 den femte percentilen, enligt kontrollgruppens resultat. På samma sätt motsvarar skalpoängen 9 den 95:e percentilen på det ursprungliga Hayling Test, men på Hayling-testet motsvarar skalpoängen 6 den 95:e percentilen. Skalpoängen 6 är en genomsnittlig prestation (50:e percentilen) i den engelska normeringen, vilket ytterligare åskådliggör skillnaderna. Det är därför tveksamt om den engelska normeringen kan användas på den nuvarande versionen av Hayling-testet. Kontrollgruppen är äldre än den ursprungliga engelska normgruppen. Eftersom inhibitionsförmåga (se bland annat Schnitzspahn et al., 2013) och exekutiva funktioner i stort (se bland annat Park & Reuter-Lorenz, 2009) försämras vid åldrande är det möjligt att kontrollgruppen presterar sämre än den ursprungliga engelska normgruppen på grund av högre

ålder. Något som talar emot detta är den åldersbundna engelska normeringen för den femte percentilen. Sammanlagda skalpoäng (min: 3; max: 23) på ursprungliga Hayling Test är mellan 11 och 12 vid femte percentilen för åldern 41-80 år (Burgess & Shallice, 1997). Sammanlagda skalpoäng på Hayling-testet är 7 vid femte percentilen, enligt kontrollgruppens resultat. Ingen i kontrollgruppen uppnådde heller resultat motsvarande de högre skalpoängen, trots hög utbildningsnivå och erforderlig åldersspridning. Det är således tvivelaktigt att kontrollgruppens lägre resultat skulle bero på åldersskillnader jämfört med gruppen som ligger till grund för den ursprungliga normeringen. Skillnaden mellan grupperna gör att man kan ifrågasätta om stickprovet, det vill säga kontrollgruppen, är representativt för populationen. Mest optimalt vore om kontrollgruppen hade rekryterats genom randomiserat urval utifrån demografiska variabler. Detta har inte varit möjligt i föreliggande studie.

Vad gäller deltagarna i patientgruppen gjordes valet att inte bara använda resultat på Hayling-testet för en enskild sjukdom inom FTD-komplexet, till exempel bvFTD, utan att låta alla sjukdomar i FTD-komplexet utom PNFA ingå i studien. Nackdelen med detta förfarande är att patientgruppen blir mer heterogen. Sjukdomstillstånd som delvis tar sig olika uttryck i såväl funktion som hjärnlokalisering sätts ihop i en enda grupp. Som tidigare konstaterats har patienter med SD-diagnos generellt mer välbevarade exekutiva funktioner än patienter med bvFTD-diagnos (Hodges & Patterson, 2007). En försiktighet vid tolkningen av patientgruppens resultat får därför tillämpas. Fördelen är ett betydligt större antal medverkande i patientgruppen, vilket gör studiens statistiska power högre. Alla i patientgruppen är trots allt diagnostiserade med en klinisk FTD-diagnos, där likheter i just skadelokalisering och kliniska symtom finns.

Patientgruppen var drabbad av en neurodegenerativ sjukdom. Därför är det troligt att deltagarna i patientgruppen befann sig i olika sjukdomsstadier när de genomgick den neuropsykologiska undersökningen. Kanske hade en mindre skillnad mellan patientgruppens och kontrollgruppens resultat uppmätts om alla i patientgruppen hade varit i ett tidigt sjukdomsskede. Det är inte okomplicerat att skatta sjukdomsskede. Om så ändå hade skett och till exempel endast patienter i ett tidigt sjukdomsskede hade inkluderats hade studiens statistiska power försvagats till följd av ett litet stickprov. Många studier som studerar kliniska grupper gör ingen indelning på basis av sjukdomsskede. I det avseendet utgör Hornberger et al. (2008) ett undantag bland de studier som har testat patienter med FTD med Hayling Test. Man kan emellertid anta att de flesta i patientgruppen i föreliggande studie

befann sig i ett tidigt till måttligt sjukdomsskede. I ett senare stadium hade de förmodligen inte kunnat medverka i den neuropsykologiska undersökningen.

Kontrollgruppen var större och hade större åldersspridning än patientgruppen. Fler i kontrollgruppen än i patientgruppen hade högskoleutbildning. Likväl hade kontrollgruppen och patientgruppen ungefär samma könsfördelning och medelålder, vilket är positivt för att kunna jämföra grupperna. Om kontrollgruppen hade rekryterats genom matchning av deltagarna i patientgruppen utifrån de demografiska variablerna hade en jämförelse mellan gruppernas resultat kunnat ske med ännu större säkerhet. Å andra sidan hade en exakt matchning av grupperna inneburit betydligt färre deltagare i kontrollgruppen. Det hade i sin tur lett till mindre säkra resultat på grund av lägre statistisk power vid undersökningarna av de validitets- och reliabilitetsmått som bara beräknas på kontrollgruppens resultat. Därför övervägde fördelarna med en större kontrollgrupp.

Det hade också varit bra om det i anslutning till testsituationen kontrollerades för att deltagarna i kontrollgruppen inte hade intagit läkemedel som påverkade deras testprestation. Risken för detta är troligen liten, men det borde ha utgjort ett exklusionskriterium.

Instrument. CWIT valdes för att undersöka Hayling-testets begreppsvaliditet. Valet föll på CWIT då det, liksom Stroop som CWIT bygger på, har nämnts samtidigt som Hayling Test som exempel på test som undersöker verbal inhibitionsförmåga (Snowden, 2013; Belleville et al., 2006). Jantscher et al. (2011) har dessutom använt sig av Stroop för att undersöka korrelation med Hayling Test, varför valet av CWIT i föreliggande studie föll sig naturligt.

I föreliggande studie har Hayling-testets begreppsvaliditet i termer av konvergent validitet undersökts, det vill säga om testet korrelerar positivt med annat test som avser att mäta ungefär samma sak. För bättre begreppsvaliditet borde även Hayling-testets diskriminativa validitet undersökas, det vill säga om testet korrelerar negativt med annat test som inte avser att mäta samma sak. Det har emellertid varit svårt att hitta ett test som skulle kunna korrelera negativt med Hayling-testet. Tidigare forskning tycks inte ha undersökt Hayling Tests diskriminativa validitet. Detta tillsammans med avgränsningsskäl gjorde att bara Hayling-testets konvergenta validitet undersöktes.

Etik. Personer som författaren hade en närmare personlig relation till testades inte för att undvika rollsammanblandning och potentiella etiska dilemman om dessa personers

testprestationer hade varit svaga. Om rekrytering inte hade skett genom privat kontaktnät hade antalet deltagare i studien sjunkit betydligt, framför allt bland personer under 65 år. Därför framstår avvägningen att rekrytera genom privat kontaktnät men att inte testa personer som författaren hade en närmare personlig relation till som rimlig.

Vid rekrytering av kontrollgruppen gavs information om att deltagarna skulle erhålla en mindre gåva, se bilaga, vilket kan ha gjort att fler personer ställde upp. Risken med att utlova en gåva skulle kunna vara att personer ställer upp enkom för belöningens skull och därmed inte är lika motiverade att prestera optimalt i själva testningen. Efter genomförd testning kan det tänkas att gåvan var en tillkommande motiverande faktor men att deltagarna var positiva till testning ändå. Flera hade glömt bort att en mindre gåva utlovats och uttryckte glädje när de fick gåvan. Möjligen hade studien varit lite svårare att genomföra utan gåvan, i synnerhet för personer som var tvungna att resa för att genomföra testningen.

I de fall deltagare i kontrollgruppen bedömdes ha presterat betydligt under genomsnittet fördes ett resonemang med dessa personer om tänkbar orsak. Om deltagarna var oroliga rekommenderade författaren dem att vända sig till vårdcentral.

Dataanalys. Tidigare forskning har använt olika mått vid redovisning av resultat på Hayling Test, varför det har varit svårt att välja vilka mått som skulle användas i föreliggande studie. Användandet av måtten Sammanlagda skalpoäng och Skalpoäng totalt är problematiskt. Skalpoängen baseras nämligen på den engelska normeringen som inte självklart kan appliceras på det svenskspråkiga Hayling-testet. Det finns inget bättre sätt att jämföra gruppernas resultat på Hayling-testet i sin helhet än att jämföra på skalpoängsnivå. Alternativet hade varit att lägga samman råpoängen på de olika deltesten men då hade tidsmåten (Tid sektion 1 och Tid sektion 2) potentiellt fått större betydelse än felmättet (Felpoäng). Ett sådant alternativ har heller inte använts i tidigare forskning.

Det förekom flera outliers i kontrollgruppens resultat, framför allt på Tid sektion 1. Outliers behövs i analyserna av flera skäl: För att inte manipulera insamlad data, för att en extrem outlier på Hayling-testet Tid sektion 1 presterade adekvat på CWIT och på grund av att färre antal outliers förekom på de andra delmåten i Hayling-testet. Det är svårt att säga om extremvärden i kontrollgruppen berodde på kognitiv svikt. Deltagarna gjordes uppmärksamma på inklusionskriterierna, men det är tänkbart att enstaka deltagare inte ville prata om eventuell kognitiv nedsättning eller inte var medveten om den. För att få färre outliers i kontrollgruppens resultat skulle man istället kunna genomföra en mindre kognitiv

screening innan testning med Hayling-testet och CWIT påbörjades. Ett screeningtest skulle kunna användas för att exkludera personer som presterar under ett visst cut-off-värde. Nackdelen med detta alternativa förfarande är att testningen hade tagit längre tid och att psykiskt och neurologiskt friska personer hade riskerat att exkluderas i onödan. Fördelen är att antalet outliers antagligen hade reducerats utan att behöva manipulera insamlade data. Därför hade en screening varit bra.

Avslutande diskussion

Utvecklande av Hayling-testet. För att kunna anpassa Hayling-testet till svenska förhållanden följer här ett generellare resonemang om ändringar i testet som skulle kunna genomföras utifrån studiens resultat.

För att förbättra homogenitetsreliabiliteten på Tid sektion 1 skulle man kunna utveckla nya item utifrån en förstudie, som man gjorde vid framtagandet av Hayling Test i original (Burgess & Shallice, 1996). Om flera svarsalternativ finns för enskilda item i Tid sektion 1 kan det vara så att delmålet inte bara mäter verbal initiationsförmåga utan även beslutsförmåga eller framplöckning ur långtidsminnet – kort sagt, råkar ut för det som Friedman et al. (2008) kallar för *task impurity problem*. Item med mindre risk för tvetydiga svar torde stärka begreppsvaliditeten, att man verkligen mäter verbal inhibitionsförmåga. Finns inte resurser till att genomföra en förstudie är ett alternativ att konsultera expertis inom neuropsykologi, lingvistik eller psykometri och be dessa att hitta tveksamma item samt komma med ersättningsförslag.

Valet av en modifikation av sten-skalan som skalpoäng är olyckligt. Trots det vällovliga syftet med percentiler tydligt kopplade till skalpoängen är skalan trubbig, eftersom testpersoner som presterar tämligen olika varandra på råpoängnivå ändå landar på samma skalpoäng, som Crawford (2013) konstaterar. Ett bättre alternativ vore exempelvis Wechsler-skalpoäng, som är vanligare och erbjuder något större precision, förutsatt att skalan är kopplad till konfidensintervall.

Man skulle också kunna överväga att införa fler övningsmeningar, i synnerhet i sektion 2, som Borella et al. (2011) gjorde. Övningsmeningar skulle kunna erbjuda testpersonerna större chans att förstå premisen i sektion 2 och undvika kategori B-fel. I föreliggande studie förekommer kategori B-fel ungefär lika mycket i kontrollgrupp och patientgrupp och är det enda måttet där signifikanta skillnader mellan grupperna inte förekommer. Risken för att fler övningsmeningar skulle förbättra prestationen även för en

kliniskt nedsatt grupp och göra det svårare att skilja mellan grupperna är inte troligt, enligt Shallice och Burgess (1998) resonemang om att patienter med allvarlig frontal nedsättning inte kan utveckla en strategi. Däremot skulle fler övningsmeningar göra att testet tog längre tid och utsätta personer med svårigheter för onödig prövning. Som ett alternativ till fler övningsmeningar skulle man därför kunna låta testledaren korrigera testpersonen även vid övningsmening 2 i sektion 2, istället för som idag bara i övningsmening 1 i sektion 2.

Föreliggande studies resultat visar att testprestationen på samtliga deltest försämras med ökad ålder och tidigare forskning visar att exekutiva funktioner och inhibitionsförmåga försämras med ökad ålder. Mot bakgrund av detta vore åldersbaserad normering av Hayling-testet bra, såsom Groth-Marnat (2003) rekommenderar när subgrupper skiljer sig från den huvudsakliga normeringen. Bielak et al. (2006) har redan tagit fram åldersbaserade normer för den brittiska versionen, men med tanke på testets språkbundenhet är det tvivelaktigt om dessa kan användas i den svenska versionen.

Å andra sidan avviker ovanstående förslag från den ursprungliga administreringen av Hayling Test och gör det svårt att jämföra testresultat mellan olika länder. Det är redan komplicerat att överföra resultat på verbala test där språken och itemen skiljer sig åt, men att ändra på antalet item, övningsmeningar eller korrigeringar vid övningsmeningar gör det ännu svårare att jämföra testresultat mellan olika språkversioner.

Slutligen är det relevant att diskutera varför den svenska kontrollgruppen uppnår sämre resultat än gruppen som ligger till grund för den engelska normeringen. Interbedömarreliabiliteten på Fel-poäng var god fastän svenska bedömningskriterier saknades på de flesta item, varför det är mindre sannolikt att författaren har rättat alltför strängt. Dessutom underpresterade kontrollgruppen i förhållande till den engelska normeringen även på Tid sektion 1 och Tid sektion 2, som inte är känsliga för subjektiv bedömning på samma sätt som Fel-poäng. Det låga resultatet i förhållande till den engelska normeringen kan bero på outliers som drar ner kontrollgruppens resultat. Vissa item i sektion 2 tycks också vara svåra att inte svara korrekt på oberoende av svenskspråkig eller engelskspråkig version: Ett svar som är fel men avges i rätt ordklass kan ibland bedömas som ett kategori B-fel eftersom det ofta blir bisarrt enligt de engelska generella bedömningskriterierna (Burgess & Shallice, 1997). Dessa item eller dess bedömningskriterier bör göras om.

Förslag på framtida forskning. Efter omarbetning av Hayling-testet bör det utprovas på nytt på ett större stickprov med jämnare fördelning avseende deltagarnas kön och

utbildningsnivå. I en sådan utprovning hade man även kunnat undersöka test-retestreliabiliteten och ta fram en åldersbaserad normering. För att kunna använda testet vid undersökning av ungdomar och unga vuxnas inhibitoriska förmåga skulle testet behöva utprovas även på dessa grupper.

Föreliggande studies resultat att män presterade bättre än kvinnor på Tid sektion 2 bör undersökas särskilt på ett större stickprov, då det inte överensstämmer med tidigare forskning.

En intressant aspekt för framtida forskning är att utforska Hayling-testets diagnostiska validitet i andra patientgrupper som tros ha nedsatt inhibitionsförmåga, vilket har skett med god evidens i andra språkversioner.

Det vore också intressant att undersöka hur Hayling-testet korrelerar med andra test som avser att mäta andra former av inhibitionsförmåga, såsom interferenskontroll, kognitiv inhibition och okulomotorisk inhibition samt se om dessa test aktiverar olika hjärnregioner.

Slutsatser. Sammanfattningsvis finns ett negativt samband mellan ålder och prestation på Hayling-testet. Resultat på Hayling-testet korrelerar på vissa delmätt med könstillhörighet respektive utbildningsnivå, men detta bör undersökas med ett större stickprov.

Hayling-testets homogenitetsreliabilitet är otillräcklig på Tid sektion 1 och bra på Tid sektion 2 och Felpoäng. Bland annat på grund av detta är rekommendationen inför framtiden att vissa item i Hayling-testet modifieras. Interbedömarreliabiliteten är god till utmärkt. Begreppsvaliditeten är god då test som också avser att mäta inhibition korrelerar svagt till måttligt med Hayling-testet.

Den viktigaste slutsatsen som dras av föreliggande studie är att Hayling-testet generellt sett har god diagnostisk validitet vad gäller förmågan att skilja mellan den verbala inhibitionsförmågan hos en frisk kontrollgrupp och en patientgrupp drabbad av FTD. Då testet är avsett för kliniskt bruk är den diagnostiska validiteten av betydelse för användning av Hayling-testet i större skala. Emellertid presterar ett fåtal i kontrollgruppen på patientgruppens nivå och ett fåtal i patientgruppen presterar på kontrollgruppens nivå. I likhet med andra neuropsykologiska test bör Hayling-testet därför användas i ett större testbatteri och kompletteras med anamnestiska data samt beteendeobservationer så att den kliniska bedömningen blir säkrare. Eftersom grupperna på det stora hela skiljer sig åt i testprestation kan Hayling-testet användas vid neuropsykologisk utredning av inhibitorisk förmåga vid misstänkt FTD.

Referenser

- Anderson, V., Northam, E., Hendy, J. & Wrennall, J. (2001). *Developmental neuropsychology: A clinical approach*. Hove: Psychology Press.
- Andrés, P. & Van der Linden, M. (2000). Age-related differences in supervisory attentional system functions. *Journal of Gerontology: Psychological Sciences*, 55B(6), 373-380. doi:10.1093/geronb/55.6.P373
- Andrés, P. & Van der Linden, M. (2001). Supervisory attentional system in patients with focal frontal lesions. *Journal of Clinical and Experimental Neuropsychology*, 23(2), 225-239. doi:10.1076/jcen.23.2.225.1212
- Ardila, A., Pineda, D. & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology*, 15(1), 31-36. doi:10.1016/S0887-6177(98)00159-0
- Bayliss, D. M. & Roodenrys, S. (2000). Executive processing and attention deficit hyperactivity disorder: An application of the supervisory attentional system. *Developmental Neuropsychology*, 17(2), 161-180. doi:10.1207/S15326942DN1702_02
- Belleville, S., Rouleau, N. & Van der Linden, M. (2006). Use of the Hayling task to measure inhibition of prepotent responses in normal aging and Alzheimer's disease. *Brain and Cognition*, 62(2), 113-119. doi:10.1016/j.bandc.2006.04.006
- Bielak, A. A. M., Mansueti, L., Strauss, E. & Dixon, R. A. (2006). Performance on the Hayling and Brixton tests in older adults: Norms and correlates. *Archives of Clinical Neuropsychology*, 21(2), 141-149. doi:10.1016/j.acn.2005.08.006
- Bloom, P. A. & Fischler, I. (1980). Completing norms for 329 sentence contexts. *Memory & Cognition*, 8(6), 531-642. doi:10.3758/BF03213783
- Borella, E., Ludwig, C., Fagot, D. & De Ribaupierre, A. (2011). The effect of age and individual differences in attentional control: A sample case using the Hayling test. *Archives of Gerontology and Geriatrics*, 53(1), 75-80. doi:10.1016/j.archger.2010.11.005
- Bouquet, C. A., Bonnaud, V. & Gil, R. (2003). Investigation of supervisory attentional system functions in patients with Parkinson's disease using the Hayling task. *Journal of Clinical and Experimental Neuropsychology*, 25(6), 751-760. doi:10.1076/jcen.25.6.751.16478

- Burgess, P. W. & Alderman, N. (2013). Executive dysfunction. I L. H. Goldstein & J. E. McNeil (Red.), *Clinical neuropsychology: A practical guide to assessment and management for clinicians* (2. uppl.) (ss. 209-237). Chichester: Wiley-Blackwell.
- Burgess, P. & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia*, *34*(4), 263-273.
doi:10.1016/0028-3932(95)00104-2
- Burgess, P. & Shallice, T. (1997). *The Hayling and Brixton tests*. Thurston Suffolk: Pearson.
- Chan, K. K. S., Xu, J. Q., Liu, K. C. M., Hui, C. L. M., Wong, G. H. Y. & Chen, E. Y. H. (2012). Executive function in first-episode schizophrenia: A three-year prospective study of the Hayling sentence completion test. *Schizophrenia Research*, *135*(1-3), 62-67. doi:10.1016/j.schres.2011.12.022
- Chan, R. C. K., Chen, E. Y. H., Cheung, E. F. C. & Cheung, H. K. (2004). Executive dysfunctions in schizophrenia: Relationships to clinical manifestation. *European Archives of Psychiatry and Clinical Neuroscience*, *254*(4), 256-262.
doi:10.1007/s00406-004-0492-3
- Channon, S., Sinclair, E., Waller, D., Healey, L. & Robertson, M. M. (2004). Social cognition in Tourette's syndrome: Intact theory of mind and impaired inhibitory functioning. *Journal of Autism and Developmental Disorders*, *34*(6), 669-677.
doi:10.1007/s10803-004-5287-x
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284-290. doi:10.1037/1040-3590.6.4.284
- Clark, C., Prior, M. & Kinsella, G. J. (2000). Do executive function deficits differentiate between adolescents with ADHD and oppositional defiant/conduct disorder? A neuropsychological study using the Six elements test and Hayling sentence completion test. *Journal of Abnormal Child Psychology*, *28*(5), 403-414.
doi:10.1023/A:1005176320912
- Clark, J. M. (1996). Contributions of inhibitory mechanisms to unified theory in neuroscience and psychology. *Brain and Cognition*, *30*(1), 127-152. doi:10.1006/brcg.1996.0008
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. uppl.). Hillsdale: L. Erlbaum Associates.

- Collette, F., Van der Linden, M., Delfiore, G., Degueldre, C., Luxen, A. & Salmon, E. (2001). The functional anatomy of inhibition processes investigated with the Hayling task. *NeuroImage*, 14(2), 258-267. doi:10.1006/nimg.2001.0846
- Collette, F., Van der Linden, M., Laureys, S., Delfiore, G., Degueldre, C., Luxen, A & Salmon, E. (2005). Exploring the unity and diversity of the neural substrates of executive functioning. *Human Brain Mapping*, 25(4), 409–423. doi:10.1002/hbm.20118
- Colton, D. & Covert, R. W. (2007). *Designing and constructing instruments for social research and evaluation*. San Francisco: Jossey-Bass.
- Cooper, C. (1999). *Intelligence and abilities*. Hove: Psychology Press.
- Crawford, J. R. (2013). Quantitative aspects of neuropsychological assessment. I L. H. Goldstein & J. E. McNeil (Red.), *Clinical neuropsychology: A practical guide to assessment and management for clinicians* (2. uppl.) (ss. 129-155). Chichester: Wiley-Blackwell.
- de Sousa Siqueira, L., Scherer, L. C., Reppold, C. T. & Fonseca, R. P. (2010). Hayling test – Adult version: Applicability in the assessment of executive functions in children. *Psychology & Neuroscience*, 3(2), 189-194. doi:10.3922/j.psns.2010.2.008
- Delis, D. C., Kaplan, E. & Kramer, J. H. (2001). *Delis Kaplan executive function system (D-KEFS): Examiner's manual*. San Antonio: The Psychological Corporation.
- Delis, D. C., Kaplan, E. & Kramer, J. H. (2005). *Delis Kaplan executive function system (D-KEFS): Svenskt manualsupplement* (H. Järvå, Övers.). Stockholm: Psykologiförlaget. (Originalarbete publicerat 2001).
- Duncan, J., Emslie, H., Williams, P., Johnson, R. & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology*, 30(3), 257-303. doi:10.1006/cogp.1996.0008
- Elfgrén, C., Englund, E., Nilsson, C. & Passant, U. (2011). Frontotemporal demenssjukdom. I L.-O. Wahlund, C. Nilsson & A. Wallin (Red.), *Kognitiv medicin* (ss. 206-222). Stockholm: Norstedts.
- Egidius, H. (2008). *Psykologilexikon* (4. uppl.). Stockholm: Natur och kultur.
- Field, A. (2013). *Discovering statistics using IBM SPSS Statistics* (4. uppl.). London: SAGE.
- Friedman, N. P. & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, 133(1), 101-135. doi:10.1037/0096-3445.133.1.101

- Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P. & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*(2), 201-225. doi:10.1037/0096-3445.137.2.201
- Gazzaniga, M. S., Ivry, R. B. & Mangun, G. R. (2009). *Cognitive neuroscience: The biology of the mind* (3. uppl.). New York: Norton.
- Gohier, B., Ferracci, L., Surguladze, S. A., Lawrence, E., Hage, W. E., Kefi, M. Z., ... Le Gall, D. (2009). Cognitive inhibition and working memory in unipolar depression. *Journal of Affective Disorders*, *116*(1-2), 100-105. doi:10.1016/j.jad.2008.10.028
- Graham, A. (2007). Epidemiology of frontotemporal dementia. I J. R. Hodges (Red.), *Frontotemporal dementia syndromes* (ss. 25-37). Cambridge: Cambridge University Press.
- Groom, M. J., Jackson, G. M., Calton, T. G., Andrews, H. K., Bates, A. T. & Liddle, P. F. (2008). Cognitive deficits in early-onset schizophrenia spectrum patients and their non-psychotic siblings: A comparison with ADHD. *Schizophrenia Research*, *99*(1-3), 85-95. doi:10.1016/j.schres.2007.11.008
- Groome, D., Brace, N., Dewart, H., Edgar, G., Edgar, H., Esgate, A., ... Stafford, T. (2010). *Kognitiv psykologi – processer och störningar* (K. Larsson Wentz, Övers.). Lund: Studentlitteratur. (Originalarbete publicerat 2006).
- Groth-Marnat, G. (2003). *Handbook of psychological assessment* (4. uppl.). New Jersey: Wiley.
- Hodges, J. R. & Patterson, K. (2007). The neuropsychology of frontotemporal dementia. I J. R. Hodges (Red.), *Frontotemporal dementia syndromes* (ss. 102-133). Cambridge: Cambridge University Press.
- Hornberger, M., Geng, J. & Hodges, J. R. (2011). Convergent grey and white matter evidence of orbitofrontal cortex changes related to disinhibition in behavioural variant frontotemporal dementia. *Brain*, *134*(9), 2502-2512. doi:10.1093/brain/awr173
- Hornberger, M., Piguet, O., Kipps, C. & Hodges, J. R. (2008). Executive function in progressive and nonprogressive behavioral variant frontotemporal dementia. *Neurology*, *71*(19), 1481-1488. doi:10.1212/01.wnl.0000334299.72023.c8
- Hornberger, M., Savage, S., Hsieh, S., Mioshi, E., Piguet, O. & Hodges, J. R. (2010). Orbitofrontal dysfunction discriminates behavioral variant of frontotemporal dementia

- from Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, 30(6), 547-552. doi:10.1159/000321670
- Jantscher, S., Willinger, U., Schmöger, M., Müller, C. & Auff, E. (2011). P01-417 - Validation of the Hayling sentence completion test - German version & Stroop-test. *Abstracts of the 19th European Congress of Psychiatry, European Psychiatry*, 26(Supplement 1:420). doi:10.1016/S0924-9338(11)72128-9
- Jurado, M. B. & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, 17(3), 213-233. doi:10.1007/s11065-007-9040-z
- Kipps, C. M., Knibb, J. A. & Hodges, J. R. (2007). Clinical presentation of frontotemporal dementia. I J. R. Hodges (Red.), *Frontotemporal dementia syndromes* (ss. 38-79). Cambridge: Cambridge University Press.
- Lezak, M. D., Howieson, D. B. & Loring, D. W. (2004). *Neuropsychological assessment* (4. uppl.). New York: Oxford University Press.
- Lough, S., Kipps, C. M., Treise, C., Watson, P., Blair, J. R. & Hodges, J. R. (2006). Social reasoning, emotion and empathy in frontotemporal dementia. *Neuropsychologia*, 44(6), 950-958. doi:10.1016/j.neuropsychologia.2005.08.009
- Lovestone, S. (2009). Alzheimer's disease and other dementias (including pseudodementias). I A. S. David, S. Fleminger, M. D. Kopelman, S. Lovestone & J. D. C. Mellers (Red.), *Lishman's organic psychiatry: A textbook of neuropsychiatry* (4. uppl.) (ss. 543-615). Chichester: Wiley-Blackwell.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H. & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49-100. doi:10.1006/cogp.1999.0734
- Nathaniel-James, D. A., Fletcher, P. & Frith, C. D. (1997). The functional anatomy of verbal initiation and suppression using the Hayling test. *Neuropsychologia*, 35(4), 559-566. doi:10.1016/S0028-3932(96)00104-2
- Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S., ... Benson, D. F. (1998). Frontotemporal lobar degeneration: A consensus on clinical diagnostic criteria. *Neurology*, 51(6), 1546-1554. doi:10.1212/WNL.51.6.1546

- Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, 126(2), 220-246. doi:10.1037/0033-2909.126.2.220
- Norman, D. A. & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. I R. J. Davidson, G. E. Schwartz & D. Shapiro (Red.), *Consciousness and self-regulation* (Vol. 4) (ss. 1-18). New York: Plenum.
- O'Callaghan, C., Naismith, S. L., Hodges, J. R., Lewis, S. J. G. & Hornberger, M. (2013). Fronto-striatal atrophy correlates of inhibitory dysfunction in Parkinson's disease versus behavioural variant frontotemporal dementia. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 49(7), 1833-1843. doi:10.1016/j.cortex.2012.12.003
- Park, D. C. & Reuter-Lorenz, P. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology*, 60, 173-196. doi:10.1146/annurev.psych.59.103006.093656
- Passant, U., Elfgrén, C., Englund, E. & Gustafson, L. (2005). Psychiatric symptoms and their psychosocial consequences in frontotemporal dementia. *Alzheimer Disease and Associated Disorders*, 19(Suppl1), S15-S18. doi:10.1097/01.wad.0000183084.22562.5a
- Rose, A. K., Mason-Li, M., Nicholas, D. & Hobbs, M. (2010). A new test to measure attentional bias and cognitive disinhibition in drinkers, based on the Hayling task. *Alcohol & Alcoholism*, 45(6), 501-506. doi:10.1093/alcalc/agq062
- Schnitzspahn, K. M., Stahl, C., Zeintl, M., Kaller, C. P. & Kliegel, M. (2013). The role of shifting, updating, and inhibition in prospective memory performance in young and older adults. *Developmental Psychology*, 49(8), 1544-1553. doi:10.1037/a0030579
- Shallice, T. & Burgess, P. (1998). The domain of supervisory processes and the temporal organization of behavior. I A. C. Roberts, T. W. Robbins & L. Weiskrantz (Red.), *The prefrontal cortex: Executive and cognitive functions* (ss. 22-35). doi:10.1093/acprof:oso/9780198524410.003.0003
- Snowden, J. (2013). Neuropsychology of frontotemporal lobar degeneration: Frontotemporal dementia, semantic dementia and progressive non-fluent aphasia. I L. H. Goldstein & J. E. McNeil (Red.), *Clinical neuropsychology: A practical guide to assessment and management for clinicians* (2. uppl.) (ss. 375-396). Chichester: Wiley-Blackwell.
- Socialstyrelsen (2010). *Internationell statistisk klassifikation av sjukdomar och relaterade hälsoproblem – Systematisk förteckning, svensk version 2011 (ICD-10-SE)*. Stockholm: Socialstyrelsen.

- Steverson, T. (2011). *Ecological assessment of the supervisory attentional system in people with intellectual disabilities* (Doktorsavhandling). University of East Anglia, School of social work and psychology. Hämtat från <https://ueaeprints.uea.ac.uk/39454/>
- Van der Linden, M., Ceschi, G., Zermatten, A., Dunker, D. & Perroud, A. (2005). Investigation of response inhibition in obsessive-compulsive disorder using the Hayling task. *Journal of the International Neuropsychological Society*, 11(6), 776-783. doi:10.1017/S1355617705050927
- Volle, E., de Lacy Costello, A., Coates, L. M., McGuire, C., Towgood, K., Gilbert, S., ... Burgess, P. W. (2012). Dissociation between verbal response initiation and suppression after prefrontal lesions. *Cerebral Cortex*, 22(10), 2428-2440. doi:10.1093/cercor/bhr322
- Wang, K., Song, L.-L., Cheung, E. F. C., Lui, S. S. Y., Shum, D. H. K. & Chan, R. C. K. (2013). Bipolar disorder and schizophrenia share a similar deficit in semantic inhibition: A meta-analysis based on Hayling sentence completion test performance. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 46, 153-160. doi:10.1016/j.pnpbp.2013.07.012

Skriftlig information om studien

Deltagare sökes till forskningsprojekt

Jag, snart färdig psykolog, söker deltagare till en studie som handlar om **metodutveckling av test**. Testet avser att mäta språklig inhibitionsförmåga, förmågan att hindra och reglera sitt tal, som kan vara påverkat i bland annat **pannlobsdemens** (frontotemporal demens).

Väljer du att delta i studien kommer du att genomföra själva testet och genomföra ett liknande test som avser att mäta ungefär samma sak. Testet går ut på att komplettera meningar och heter *Hayling-testet*. Det har tidigare inte använts i någon större utsträckning i Sverige. Studiens resultat kommer att bidra med mer information om huruvida testet är lämpligt att användas vid neuropsykologisk utredning av till exempel demenssjukdomar.

Studien genomförs som mitt psykologexamensarbete vid Lunds universitet och är en del i ett större forskningsprojekt, Luprofs – Lund Prospective Frontotemporal dementia study. Luprofs bedrivs på Minneskliniken vid Skånes universitetssjukhus i Lund och på Avdelningen för geriatrisk psykiatri vid Lunds universitet.

För att delta ska du:

- Vara över 40 år gammal.
- För närvarande vara psykiskt frisk och inte lida av någon neurologisk sjukdom eller hjärnskada.
- Förstå och tala flytande svenska.

Stämmer detta på dig så är du välkommen att delta i testning under **september och oktober 2013**. Du som deltar kan genomföra testet antingen i Lund eller i Malmö. Testningen tar endast 30 minuter. Som tack för hjälpen får du en mindre gåva, samtidigt som du bidrar till forskning om metodutveckling av test som förhoppningsvis ska kunna användas vid utredning av pannlobsdemens.

Är du intresserad av att delta, kontakta mig på:

E-post: [REDACTED]

Telefon: [REDACTED]

Ditt deltagande betyder väldigt mycket för min studie. Tack på förhand!

Med vänlig hälsning

Erik Blenow Nordström
Psykologkandidat
Institutionen för psykologi vid Lunds universitet



LUNDS
UNIVERSITET