



LUNDS
UNIVERSITET

Institutionen för psykologi

Psykologprogrammet

Klurisen: ett nytt sätt att mäta matematisk förmåga

Valideringsstudie av ett instrument för vetenskaplig testning
av akademiska färdigheter hos barn i lågstadiet.

Oskar Hallström Elthammar

Psykologexamensuppsats. 2014

Handledare: Jonas Bjärehed

Examinator: Karin Stjernqvist

“One does not simply walk into Mordor.”
Boromir, *Lord of the Rings: The Fellowship of the Ring*

— J.R.R. Tolkien

Tack!

Författaren vill tacka handledaren Jonas Bjärehed för stort tålamod, goda råd och handledning under uppsatsprocessens gång. Ett tack även till Martin Hassler Hallstedt som hjälpt mig framåt och gett mig möjlighet att delta i forskningsprojektet. Ett stort tack till elever, föräldrar, lärare och rektorer i berörd kommun, som genom sitt deltagande och medgivande möjliggjorde ett rikt underlag att undersöka.

Ett varmt tack till Sara Andersson, som tröstat och stöttat när det gått långsamt i uppsatsprocessen. Ett extra stort tack till Philip Pärnamets som trott på min akademiska förmåga, när jag själv tvekade. Med frågor, kommentarer och uppmuntran hjälpte du mig framåt. Tack Gustav Nilsson för hjälp med korrekturläsning, handfasta tips på hur en teoridel kan skrivas och alla supersmoothies.

Tack Åsa Hallström och Olle Elthammar för muntra tillrop under processens gång, när jag ringde och behövde prata av mig. En extra shout-out till min arbetsgivare Kasper Kollo som hjälpt mig finansiera en utdragen uppsatsprocess när CSN inte rimligtvis kunde finnas där längre.

Sammanfattning

Studiens syfte var att psykometriskt undersöka ett testverktyg för mätning av matematiskt problemlösningsförmåga. För detta syfte fick 273 elever i åk 2 göra det läsplattebaserade testet Klurisen vid två tillfällen och en svensk version av Iowa Test of Basic Skills, deltest Mathematical Problem Solving, benämnt IOWA, vid ett tillfälle. Studien fann att Klurisens deltest, A- och B-del hade otillräcklig reliabilitet som enskilda test, men acceptabel reliabilitet när de ansågs vara två delar av samma test. Begreppsvaliditeten för Klurisen var god, då testresultaten samvarierade signifikant med testresultatet från Iowa. Det fanns inga signifikanta skillnader i testresultat för Klurisen, grupperat på könsväribaler, socioekonomisk status eller etnicitet. Från resultaten dras slutsatsen att Klurisens B-del behöver undersökas vidare på frågenivå, för att öka deltestets tillförlitlighet. Vidare behöver Klurisen testas på en större grupper elever för att påbörja en fullständig normering av testresultat och därmed utvidga testverktygets användbarhet i kliniska sammanhang. Studiens resultat indikerar att Klurisen kan vara ett användbart screeningverktyg för att upptäcka elever i behov av stödinsatser. Ett viktigt fynd från teoriöversikten är att elever i riskzonen för skolmisslyckande troligtvis inte identifieras genom ett lågt testresultat vid ett enskilt testtillfälle, de kännetecknas istället av att de trots stödinsatser visar få eller uteblivna framsteg.

Nyckelord: psykometri, matematisk problemlösning, begreppsvaliditet, screeningverktyg, elever i riskzon för skolmisslyckande, Klurisen, ITBS.

Abstract

The purpose of this study was to examine the psychometric properties of Klurisen, a test measuring mathematical problem-solving ability. A total of 273 Second grade students participated in trials, testing Klurisen twice and an adapted Swedish version of a subtest of the Iowa Test of Basic Skills once. Results indicated that the reliability of the two subtests of Klurisen were low when assessed separately, but sufficient when assessed together as one test. Construct validity was good, as there were a significant strong correlation between results of Klurisen and the results of Iowa. There were no significant differences between test results when gender, socioeconomic status or ethnicity were used as grouping variables. A conclusion from this study was that subtest B of Klurisen needs revision on item level to increase reliability. In order to further the clinical usefulness of Klurisen norm data from a bigger sample of students need to be collected. Also, Klurisen could be of use when identifying students in need of supportive interventions. A finding from previous research suggests that students at-risk for school failure are not likely to be correctly identified by a single test result, but rather characterised as those students who still show a lack of progress even after receiving support.

Keywords: psychometrics, mathematical problem-solving, construct validity, screening tool, students at-risk for school failure, Klurisen, ITBS.

Att utveckla och förbättra skolundervisning är ett ständigt angeläget mål, särskilt mot bakgrund av de senaste årens trender med sjunkande skolresultat i Sverige. Psykologer är idag en central aktör, både i det dagliga arbetet inom skolans område (Schad, 2014), och för att utveckla, utforma och utvärdera undervisningsmetoder och tillvägagångssätt inom skolans område. Psykologisk kompetens har historiskt varit central för att både utveckla tillförlitliga metoder för att bedöma elevers förmågor och särskilda behov samt för att utforma åtgärder som möter dessa behov. Av naturliga skäl blir en särskilt angelägen uppgift att kunna bedöma och möta behoven hos de elever som är i riskzonen för att drabbas av långsiktigt skolmisslyckande.

Testning av akademiska färdigheter

Syften med testning av akademiska färdigheter kan sägas inkludera screening inför stödundervisning, kunna ge underlag vid utredningar, kunna kartlägga individuella styrkor och svagheter för att förbättra undervisningen samt uppföljning och utvärdering av insatser (Panter & Bracken, 2013).

I USA är testning av akademiska färdigheter utbrett, men det varierar stort hur användbara och tillförlitliga testerna är (Naglieri, 2013). Vidare är det en svår uppgift för testledare att testa barn, speciellt barn med kort uppmärksamhetsspann, låg tolerans för frustration och ojämn förmåga i ovana miljöer (Willis, Dumont & Kaufman, 2013). Dessa komplikationer gör att utvecklandet av tillförlitliga tester av akademiska färdigheter är en stor uppgift och det saknas idag konsensus kring vilket angreppssätt som är det bästa.

Enligt Martin Hassler-Hallstedt (personlig kommunikation, 19 september 2014) är vetenskaplig testning av akademiska färdigheter inte lika utbrett i Sverige som i USA, trots ingående efterforskningar går få valida mätinstrument på svenska att finna.

Nationella provet, som görs av samtliga elever i vissa årskurser, är ett test skapat för att få ett externt mått på elevernas kunskap i olika kärnämnen. Nationella proven kritiserar för låg tillförlitlighet som mått på kunskapsutveckling över tid (Gustafsson, Cliffordson & Erickson, 2014). Det främsta problemet med detta prov är låg interbedömarreliabilitet, som kan bero på att rättande lärare oftare ger sina egna elever generösare bedömningar än externa rättare (Gustafsson & Erickson, 2013).

I Sverige skulle det alltså finnas stor nytta med nya metoder för att bättre kunna utreda, utvärdera och ge stöd till elever i skolan.

Elever i riskzonen för skolmisslyckande

I litteraturen (Hindman, Erhart & Wasik 2012; Morgan, Farkas & Wu 2011) om

skolmisslyckanden finns det olika teoretiska perspektiv på elevers utveckling samt olika identifierade strukturella och individuella riskfaktorer som verkar ge olika grupper av elever olika förutsättningar att lyckas i skolan. Strukturella riskfaktorer för elevers skolprestationer har exempelvis konstaterats inkludera; att komma från hem med låg socioekonomisk status (SES), tillhöra en etnisk minoritet, samt könsvariabler.

Två utvecklingsmodeller för elever i riskzonen för skolmisslyckande. Vilka elever är i riskzonen för skolmisslyckande? I en longitudinell studie från USA undersöker Morgan, et al. (2011) två utvecklingsmodeller om inläring över tid. Den första berör, vad som benämns, kumulativ tillväxt, där små initiala skillnader i skolprestation mellan elever växer över tid. Med det här synsättet förväntas en långsam inläringstakt bero på olika akademiska förutsättningar och klyftor mellan elevgrupper förväntas växa över tid. Enligt den kumulativa modellen kommer de med goda kunskaper vid tidiga mätningar av skolprestationer, ex. vid skolstart, fortsätta lära sig saker i en hög takt och de elever som börjar med mindre kunskap vid tidiga mätningar kommer lära sig saker i en långsam takt. Avståndet mellan grupperna kommer på så vis växa över tid. Denna ”rika blir rikare, fattiga blir fattigare”-effekt, även kallad Matteus-effekten (Merton 1968), har dock inte kunnat visas gälla för elever med hög inläringstakt. Fynd från Morgan et al. (2011) pekar på att den kan gälla för elever med långsam inläringstakt (learning disabilities) i ämnet matematik, men inte för läs- och skrivfärdigheter. I en tidigare studie konstaterar Morgan, Farkas och Hibel (2008) resultat som antyder att en ”ensidig Matteus-effekt” påverkar elever med skolsvårigheter negativt, dvs. elever som börjar med mindre skolkunskap än jämnåriga lär sig ny kunskap i långsammare takt.

Den andra modellen om elever i riskzonen för skolmisslyckanden berör mognadsfaktorer, där vissa elever mognar tidigt medan andra mognar senare. Dessa kan behöva tid att bli redo och lärarstöd för att komma ikapp. En lägre initial kunskap kan med stödundervisning överkommas och kompenseras för, vilket gör att den kallas den kompensatoriska modellen. I en interventionsstudie från USA, slår Hindman et al., (2012) fast att utfallen av stödinsatser delvis beror på lärarnas skicklighet, vilket ger belägg för en kompensatorisk utveckling för elever, om rätt förutsättningar säkerställs.

Forskningsfynd till stöd för respektive av de två modellerna varierar (Leppeanen et al., 2004), men för läs- och skrivfärdigheter samt matematiska färdigheter finns belägg för den kompensatoriska modellen (Catts, Fey, Tomblin & Zhang, 2002; Fuchs et al, 2004; Hindman et al., 2012; Parrila, Aunola, Leskinen, Nurmi & Kirby, 2005;). Ett fynd som talar för den

kompensatoriska modellen är en rapport där samband jämfördes mellan bl.a. skolprestationer och barns födelsemånad. Elever födda på våren presterar bättre i skolan än elever födda senare på året, vilket tyder på att mognadsfaktorer påverkar skolresultat (Persson & Öhrvall, 2011).

Dock är det viktigt att notera att båda de två dominerande modellerna ger stöd åt tanken att det är viktigt att tidigt identifiera eleverna med låga initiala kunskaper och ge dem stödinsatser utöver vanlig undervisning. Det som skiljer utvecklingsmodellerna åt, är hur de förhåller sig till hur olika elever svarar på dessa stödinsatser. Den ena fokuserar på en inhämtning av tidiga små skillnader, en kompensatorisk utveckling och den andra fokuserar på förväntade stegvisa relativa försämringar, där små skillnader växer över tid mellan elevgrupper, den s.k. Matteus-effekten.

Socioekonomisk status (SES). Elever vars familjer lever i relativ fattigdom eller är en del av en etnisk minoritet har större sannolikhet att uppleva skolsvårigheter än jämnåriga som kommer från mer välbärgade hem eller som är en del av majoritetsbefolkning (Garbarino & Ganzel, 2006). I en longitudinell studie från Nya Zeeland undersöktes hur SES påverkar människor över tid. Ett fynd var att låg SES var en riskfaktor för skolmisslyckande, vilket antogs bero på att låg SES innebar mer stress över längre tid (Fergusson, Swain-Campbell & Horwood, 2004). Boxer, DeLorenzo, Savoy & Mercado, Goldstein (2010) drar liknande slutsatser kring att låg SES är en bidragande faktor till skolmisslyckanden. De undersökte bl.a. skillnader mellan akademiska aspirationer och förväntningar och fann att grupper som hade större avstånd mellan strävan och förväntan hade lägre akademiska prestationer över tid. Det är en betydande stressor för elever att ha höga krav på egen prestation men inte ha en tilltro till egen förmåga att nå målen. Elever med låg SES hade i större grad ett större avstånd mellan strävan och förväntan än andra grupper. Boxer et al. (2010) föreslår att elever i riskzonen behöver få stöd att utveckla en högre tilltro till sin egen förmåga, snarare än hjälp att ställa lägre krav på sig själva, som verkar resultera i självuppfyllande profetior.

Vid en svensk utredning undersöker Larheden (2012) samband mellan föräldrars utbildningsgrad och andel elever som når grundskolans mål, dvs. får betyg i alla ämnen som ingått i utbildningen i slutet av nionde klass. För gruppen barn med högutbildade föräldrar nådde 87% av barnen kunskapsmålen, för gruppen barn med lågutbildade föräldrar uppnådde endast 44% av barnen kunskapsmålen. Larheden (2012) fann att för åren 2000-2011 har skillnaden i skolprestation mellan barn till lågutbildade föräldrar och barn till högutbildade föräldrar ökat, främst i form av att elever med lägre SES har försämrat sina skolresultat

jämfört med tidigare år.

All evidens konvergerar dock inte kring att SES är en viktig faktor för elevers skolprestation. Li-Grining, Maldonado-Carreño, Votruba-Drzal och Haas (2010) undersöker bakomliggande mekanismer som leder till att elever är olika förberedda för de krav som ställs vid skolstart. De fann inget stöd för att elever från hem med låg SES eller från etniska minoriteter var signifikant sämre utrustade att möta skolans krav på eleverna än andra grupper. Istället föreslår de självreglering, dvs. förmågan att hantera egna beteende, känslor och uppmärksamhet på viljestyrda och adaptiva sätt, som förklaring till tidiga skillnader i skolprestation. Dock indikerar majoriteten av den granskade evidensen att SES är en viktig faktor för skolprestation, även om Li-Grining et als. (2010) resultat indikerar att vidare forskning behövs.

Etnicitet som riskfaktor. Studier som vill undersöka riskfaktorer får ibland problem att särskilja låg SES från omständigheten att tillhöra en etnisk minoritet, då dessa kan samvariera. Elever vars familjer växer upp i relativ fattigdom eller är en del av en etnisk minoritet har större sannolikhet att uppleva skolvårigheter än jämnåriga som kommer från mer välbärgade hem eller som är en del av majoritetsbefolkning (Garbarino & Ganzel, 2006).

I en regressionstudie (Battin-Pearson & Newcomb, 2000) från USA undersöktes olika riskfaktorer som kunde predicera skolmisslyckande. Det var nästan lika stor riskfaktor att vara pojke (0,18) som att ha afroamerikanskt påbrå (0,22), medan att ha låg SES (0,08) hade lägre prediktiv styrka.

Turton och Green (2013) lyfter i sin kunskapsöversikt fram etnicitet som riskfaktor för att ha en diskrepans mellan förmåga och prestation, men går inte in på möjliga orsaker till detta. I en grundbok av Passer och Smith (2007) diskuteras möjliga orsaker till varför etnicitet är en riskfaktor vid bl.a. skolprestation. Författarna beskriver att försiktiga slutsatser skulle kunna vara att SES samvarierar med etnicitet, när en etnisk grupp får bättre förutsättningar, presterar deras barn bättre på begåvningsstest. De menar alltså att det inte är en riskfaktor i sig att tillhöra en viss etnisk minoritet, men att samvariationen med låg SES leder till sämre förutsättningar för barn att utvecklas optimalt. Morgan et als. (2008) fynd motsäger de här resultaten visar i sin studie att risken för att uppleva en negativ Matteus-effekt i skolan ökar om eleven kommer från hem med låg SES och är en del av den afroamerikanska befolkningen i USA.

Kön som riskfaktor. McCoy och Reynolds (1999) undersöker vad som ökar risken för att gå om årskurser under skolgången, då detta är något som ökar risken för skolmisslyckande.

De finner bl.a. att pojkar i större utsträckning än flickor går om årskurser, vilket innebär att det är en riskfaktor att vara pojke (Mccoy & Reynolds, 1999). Li-Grinig et al. (2010) undersökte hur könsvariabler i olika grad gjorde elever redo för inläring. De fann att skillnader mellan könen gör sig signifikant gällande först i åk 5, där flickor visar på högre inläringstakt än pojkar med liknande förutsättningar.

Steele (1997) undersökte även hur flickor och pojkar presterar på ett identiskt test i matematik under olika testförhållanden. Om de gjorde ett matematiskt test där resultatet förväntades vara likvärdigt mellan könen, fanns inga signifikanta testskillnader. Fick de höra att testet skulle ge olika resultat för pojkar och flickor, presterade flickor signifikant sämre än pojkar. Steele drog slutsatsen att liknande ångest-skapande stereotypot påverkade flickors prestation negativt under vissa förhållanden.

Kontext. VanDerHeyden och Witt (2005) drar i en studie från USA slutsatsen att lärares förmåga att korrekt identifiera elever i behov av stödinsatser varierar med sammanhanget. Lågpresterande elever upptäcktes lättare av lärare i en högpresterande klass, men lärares träffsäkerhet att urskilja lågpresterande elever sjönk i en lågpresterande klass. Det kan alltså vara en riskfaktor att gå i en klass där många elever presterar under förväntan. En viktig upptäckt för studien är att med tillförlitliga screeningverktyg kan fler elever upptäckas i ett tidigare skede även om de går i en klass med lågpresterande elever. VanDerHeyden och Witt (2005) argumenterar för att testverktyg behövs som komplement till lärarförmåga, då det är en svår uppgift för professionella i skolan att upptäcka elever i behov av stödinsatser och svårt att upptäcka vilka som är motståndskraftiga mot stödinsatser. Andra studier drar samma slutsats (Gersten et al., 2005).

Förutom strukturella faktorer finns även individuella faktorer som påverkar elevens förmåga att lyckas i skolan. Det kan vara bl.a. intellektuell förmåga, inläringssvårigheter, psykisk ohälsa, eller dålig matchning mellan elevens inlärningsmönster och lärarens undervisningsstil (McLeod, Uemura & Shawna, 2012; Pool, Carter, Johnson & Carter 2013).

Låg begåvning som riskfaktor. Gottfredson (2008) hävdar att generell intelligens, *g*, är den enskilt viktigaste prediktorn för viktiga utfall i livet såsom akademisk framgång, framgång i arbetslivet, inkomstnivå och sannolikhet för att hamna i fängelse. Jensen (1999) menar att *g*-faktorn förväntas föregå och påverka akademiska prestationer, då inläring kräver *g*-faktor. Tidiga begåvningsstest användes för att uppskatta hur olika elever skulle klara skolgången.

Elever som får en standardavvikelse under genomsnitt (KI: 70-85) på intelligenstest,

kan kallas ”elever med långsam inläring” då dessa generella sett har en långsammare inläringstakt än genomsnittliga elever (Cooter & Cooter, 2004). Utan stödinsatser kommer de hamna mer och mer efter, och chanserna att klara hela skolgången blir låga. Fuchs, Fuchs & Speece (2002) är kritiska till att använda begåvningsstest för att upptäcka barn i riskzonen för skolmisslyckande, då inläringssvårigheter ofta är oväntade, elever uppvisar en diskrepans mellan begåvning och skolprestationer.

Begåvning beror både på genetiska och miljömässiga faktorer och i skolområden med låg SES har fler elever än förväntat långsam inläring (Snow, Burn, & Griffin, 1998).

Inläringssvårigheter. Fuchs et al. (2002) beskriver att inläringssvårigheter, traditionellt sett förklarats som det oväntade avståndet mellan uppmätt begåvning och uppvisad prestation. De argumenterar för en ny, fördjupad modell för stödinsatser, där åtskillnad görs mellan de som presterar under förväntan men som svarar bra på stödinsatser, jämfört med de som presterar under förväntan och gör små framsteg även med stödinsatser. Små eller uteblivna framsteg efter stödinsatser föreslår de ska bli kriteriet för att uppvisa/få diagnosen inläringssvårigheter. Fuchs och Fuchs (2006) utvecklar det här synsätten i en forskningsartikel där det som undersökts är olika elevers respons efter intervention (response to intervention) och de som svarar sämre på insatser kallas ”elever motståndskraftiga mot stödinsatser”.

I en interventionsstudie från Finland undersökte Niemi et al. (2011) olika insatser som når elever som är motståndskraftiga mot stödinsatser. Deras studie undersökte bakomliggande orsaker till långsam inläring och deras fynd tyder på att det finns en samverkan mellan individuella sårbarhetsfaktorer och bl.a. familjefaktorer men hälften av dessa elever hade inte föräldrar med en historia av liknande typ av problem. De fann att denna elevgrupp inte gick att upptäcka med begåvningsstester, utan istället utmärkte sig genom att de inte gjorde så många framsteg över tid medan deras jämnåriga klasskamrater lärde sig i vanlig takt. Niemi et al. (2011) finner att den enda stödinsats som ter sig verksamt med denna elevgrupp är ovanligt stora mängder träning/repetition.

Gersten, Jordan och Flojo (2005) undersökte i en studie bakomliggande orsaker till varför vissa elever gör goda framsteg i matematik och andra inte gör det. Syftet var att hitta resultat med klinisk signifikans på individnivå. De finner att elever med oväntad diskrepans mellan förväntad förmåga och uppvisad prestation har svårt att minnas och tillämpa bra lösningsstrategier. De använder i högre ålder långsamma metoder, ex att räkna på händer istället för att använda huvudräkning och verkade ha svårt att bemästra nya räknesätt eller nya

effektiva lösningar. Denna tendens kvarstod när hänsyn togs till strukturella riskfaktorer som låg SES, etnicitet och könsvariabler. Utmärkande för elever i riskgruppen var att de behövde längre tid än vuxna (som föräldrar eller lärare) uppskattade att de behövde för att ta till sig nya färdigheter. Kroesbergen och Van Luit (2003) replikerar resultatet, att elever med inlärningssvårigheter hade svårigheter att komma ihåg och tillämpa rätt Lösningsstrategier för de problem de stötte på. Gruppen hade svårt att generalisera fungerande lösningar från gamla uppgifter till nya, kallat *transfere* (Kroesbergen & Van Luit, 2003). Gersten et al. (2005) argumenterar för att nya sätt att diagnostisera inlärningssvårigheter behövs, som tar hänsyn till elevens motståndskraft mot stödinsatser. De menar att mindre tyngd ska ges till testresultat från ett enskilt mättillfälle och mer tyngd ska ges till tolkning av upprepa mätningar över tid, efter att eleven deltagit i stödinsatser. På detta sätt kan longitudinell testning skilja elever som mognar sent (dvs. de som kompenserar för tidiga brister) från elever med ihållande svårigheter. Till samma slutsats kommer också andra forskare (Fuchs & Fuch, 2006). ***Hur kan riskelever identifieras – på gruppnivå kontra individnivå?*** I en regressionsstudie med 6 olika longitudinella datamängder, undersöktes en mängd faktorer och huruvida de kunde predicera kommande skolförmåga (Duncan et al., 2007). Läsförmåga, matematiska färdigheter, uppmärksamhetsfaktorer, internaliserade problembeteenden, utåtagerande beteende samt sociala förmågor var möjliga prediktorer. Studien tog hänsyn för SES, kön och etnicitet för att få mer rättvisande mått på vilka faktorer som bäst predicerade kommande skolprestation. Duncan et al. (2007) fann att bästa mått på kommande prestation var tidiga läsfärdigheter, tidiga färdigheter i matematik samt uppmärksamhetsförmåga. Kraftfullast var tidiga matematiska färdigheter (standardiserad regressionskoefficient 0,34), följt av läsfärdigheter (0,17) och uppmärksamhetsförmåga (0,10). Fynden från regressionsstudien av Duncan et al. (2007) har replikerats i flera studier och bekräftar att tidig färdighet i matematik är en predicerande faktor för kommande prestationer, starkare än ex läsförmåga eller andra grupperande faktorer (Hooper, Roberts, Sideris, Burchinal & Zeisel, 2010). Hooper et al. (2010) menar att ett test som mäter matematisk problemlösning skulle kunna användas som ett screeningverktyg för att i låg ålder identifiera elever i riskzonen för skolmisslyckande.

Siegler et. al (2012) gjorde en uppföljningsstudie där de undersökte vilka typer av matematisk färdigheter som bäst predicerade kommande färdigheter, med elever från åk 4. De fann att kunskap i division var bästa prediktorn, följt av färdigheter i bråk. En möjlig orsak trodde Siegler et al. (2012) var att division och bråk är byggstenar till mer avancerad matematik såsom matematisk problemlösning.

Den äldre definitionen av skolmisslyckande, ”det oväntade avståndet mellan förmåga

och prestation”, beskriver svårigheten att upptäcka elever i tidigt skede. Riskfaktorer som är signifikanta på gruppnivå har låg klinisk relevans i det individuella fallet. Möjliga faktorer med högre signifikans på individnivå verkar vara hur eleven använder lösningsstrategier för att lösa okända problem, s.k. *transfere*. Både lärare och begåvningsstest har svårt att förutse vilka elever som kommer ha långsam inlärningsstakt samt vilka som kommer vara motståndskraftiga mot stödsatser. Tidiga färdigheter i matematik är bättre predicerande för kommande skolförmåga än andra faktorer, såsom SES, etnicitet, kön, begåvning eller psykisk ohälsa. Vidare kan tillförlitliga tester, utförda upprepade gånger över tid vara viktiga för att upptäcka elever i riskzon för skolmisslyckande. För att säkerställa att ett test är tillförlitligt behöver det uppfylla vissa kriterier och ha vissa egenskaper. Nedan följer en genomgång av sådana kriterier och egenskaper.

Testkonstruktion

Testning. I det skede testet provas ut, är oväntade förbättringar något som sänker tillförlitligheten i resultaten. Om deltagaren får högre poäng genom att göra testet flera gånger kallas mätfelet träningsseffekt, om deltagaren gör testet med större tidsavstånd kan en högre poäng beror på historiska effekter, att deltagaren lärt sig mer av ex undervisning (Shadish 2002).

Reliabilitet. Ett tests reliabilitet avser hur väl testets olika frågor hänger ihop, förutsatt att alla frågor antas mäta samma begrepp. För ett test med hög tillförlitlighet beror en stor del av resultatet på deltagarens förmåga och en liten del på systematiska och slumpvisa mätfel (Shadish 2002). Enligt Wells och Wollack (2003) beror mätfel på tre faktorer: deltagare (motivation, test-inducerad utmattning, tur), test (tvetydiga frågor, oklara direktiv) och rättnings specifika fel (slarvfel vid rättning, räknefel, felinmatad råpoäng inför dataanalyser).

Test-retest reliabilitet är ett mått på sambandet mellan testresultat från olika tillfällen med samma grupp deltagare. För ett test med hög test-retest reliabilitet kommer en stor del av resultatet från det ena testtillfället förklara testresultatet från det andra testtillfället, trots avstånd i tid mellan tillfällena (Shadish 2002).

Parallell-form reliabilitet är ett mått på sambandet mellan två likvärdiga deltest utförda vid ett tillfälle. Vid skapandet av parallell-form test hämtas frågorna från de båda testen ur en gemensam frågepool. En fördel med parallell-form test är att två olika deltest kan användas som mått vid för- och eftermätningar när behandlingseffekter ska utvärderas. Givet att båda testen mäter samma begrepp (ex matematisk problemlösning) och ger testresultat med

liknande medelvärden, varianser och alfa-koefficiens, behöver korrelationen vid parallell-form reliabilitetsmätningar överstiga 0,80. (DeVon et al., 2007).

Homogenitetsreliabilitet är ett mått på hur väl testets olika frågor hänger ihop. För att förbättra homogenitetsreliabiliteten går det antingen att öka antalet frågor för testet, eller förbättra kvalitén på frågorna (Wells & Wollack 2003). Ett vanlig mått på homogenitetsreliabilitet är att beräkna Cronbachs alfa, även kallat alfa-koefficienten eller α -värdet (Shadish 2002).

För s.k. high-stakes assessment tester, såsom Wechsler Intelligence Scale for Children (WISC) gäller homogenitetskoefficient mellan 0,90- 0,99, för screeningtester 0,80-0,89. För lärarskapade skoltester eller instrument på grundforskningsnivå (men ej för klinisk tillämpning) gäller reliabilitetskoefficienter mellan 0,70–0,79. (Nunnally, 1979, Wells, Wollack, 2003, Panter & Bracker, 2013).

Begreppsvaliditet. Om reliabiliteten är god för testen, kan också validiteten för testen vara det (om för stor del av resultat beror på slumpen, är vidare analyser överflödiga). Ett sätt att mäta begreppsvaliditet är att jämföra testresultat mellan två tester som avser sig mäta samma område (konvergent validitet), där testet ska korrelera med det andra testet. En hög korrelation tolkas som att testerna mäter samma sak.

Test floors. Eftersom testning i sammanhang med risk för skolmisslyckanden främsta uppgift är att upptäcka elever i det lägre spannet av prestationsskalan, är det viktigt att testet på ett meningsfullt sätt kan differentiera elever avsevärt under det genomsnittliga spannet (två standardavvikelser), från de eleverna i linje med eller något under det genomsnittliga spannet. Ju större skillnad i råpoäng mellan 5e och 25e percentilen, desto bättre förmåga att differentiera olika elevgrupper. (Panter & Bracker, 2013).

Sensitivitet. Anger ett tests förmåga att identifiera deltagare som möter testkriterier, ex ”risk för skolmisslyckande”. Hög sensitivitet betyder att fler deltagare möter testkriterierna (Shadish 2002). VanDerHeyden och Witt (2005) argumenterar för breda kriterier som hellre felaktigt identifierar fler än korrekt antal elever i riskzonen för skolmisslyckande, än att missa elever som hade behövt stödinsatser, då negativa effekter av skolmisslyckande är allvarigare än att upptäcka högpresterande elever i stödinsatsgrupper. Gersten et al. (2005) rekommenderar att screeningtester inkluderar elever som presterar under den 35e percentilen för att vara säkra på att inkludera de i möjliga behov av stödundervisning.

Specificitet. Anger träffsäkerheten mellan testkriterier och utfallsmått. Deltagare som identifierats att tillhöra riskgrupp men som sedan når kunskapsmål (falska positiva), samt deltagare som presterar godkänt vid screening men inte når kunskapsmål vid utfallsmått (falska negativa) sänker specificiteten (Shadish 2002).

Det pågår en debatt kring hur känsligt och specifikt screeningtester bör vara. Känsliga tester upptäcker fler i riskzonen för skolmisslyckande, men elever som identifieras som i riskzonen, speciellt ”falska positiva” elever, kan mötas av lägre krav och förväntningar av lärare, föräldrar och kan själv utveckla en bild av sig själv som någon som presterar under medel (Panter & Bracken, 2013).

Normering. För screeningverktyg är ovanstående egenskaper viktiga. För utökade syften, ex för att mäta utveckling över tid och i jämförelse med andra deltagare, är det viktigt att ett test normeras. På så vis kan enskilda resultat jämföras mot större gruppers resultat och tolkas i förhållande till ett större, mer rättvisande sammanhang.

En svårighet vid utvärdering av elevers framsteg i skolämnen, är frågan hur resultaten ska tolkas. En elev som deltar i stödundervisning kan vid jämförelser mellan för-, under- och eftermätningar visa på stora framsteg som kan tyda på att eleven inte längre är i behov av stödundervisning. Sätts resultaten i relation till kontrollgruppens framsteg övertid, kan bedömningen göras att eleven är i fortsatt behov av stöd (Naglieri, 2013). Genom att tolka testresultat mot normskalor kan utredare fatta bättre beslut om fortsatta insatser eller inte.

Sammantaget behöver ett testverktyg ha goda psykometriska egenskaper. Det första som krävs är tillförlitlighet i testresultat, där en stor del beror på deltagarens förmåga och en låg del beror på slumpvisa mätfel. Detta går att undersöka på statistisk väg. För att veta vad ett test mäter (validitet), går det att jämföra testverktyget med ett annat som avser mäta samma sak. Förutsatt att båda har en hög tillförlitlighet, kommer samvariation i testresultat vara ett mått på hur väl testen mäter samma sak. Testverktyget behöver korrekt kunna identifiera vilka deltagare som ska inkluderas i program (sensitivitet) och vilka som ska lämnas utan åtgärd, och vid utfallsmätningar är andelen korrekta beslut ett mått på testets träffsäkerhet (specificitet).

Sammanfattning

Vissa elever har en långsammare inläringstakt än andra. Det finns olika förklaringar för det här, varav belägg för två teorier utforskas ovan. Den ena teorin menar att skillnader beror på olik inlärningsförmåga och att små skillnader kommer växa över tid. Den andra

teorin menar att skillnader beror på att elever blir redo för inläring olika snabbt. Båda teorierna betonar vikten av tidig upptäckt av elever i riskzonen för skolmisslyckande. Under speciella omständigheter, ex när det finns fungerande interventionsprogram, bra utvärderingsmetoder och väldigt kompetenta lärare, gör lågpresterande elever framsteg. Det verkar som om olika länder är olika framgångsrika att nå elever i riskzonen, ex har Sverige en högre andel lågpresterande elever än Finland (Kupiainen, , Hautamäki, , & Karjalainen, (2009).

En förklaring av inläringssvårigheter är det oväntade avståndet mellan potential/förmåga och uppvisad prestation. På gruppnivå verkar flera faktorer samvariera med skolmisslyckande, där vissa faktorer bättre predicerar utfall än andra. På individnivå är det svårt att förutse vilka som kommer få skolsvårigheter, men en ny metod att upptäcka dessa elever utforskas. Genom att låta lågpresterande elever ta del av insatser och med testinstrument utvärdera elevernas framsteg över tid, går det att särskilja de som på sikt kan exkluderas från vidare stödinsatser från de som är motståndskraftiga mot stödinsatser och därmed i risk för skolmisslyckande. Ett lämpligt urvalsmått kan vara tidiga färdigheter i matematik, som har god prediktiv förmåga på gruppnivå.

Den här studien jämför två test som båda mäter matematisk problemlösning, där det ena testet, förutsatt att det har goda psykometriska egenskaper, kommer användas kliniskt som urvalsmått för vilka elever som behöver tidiga insatser.

Syfte

Studiens syfte är att utföra en utvärdering av ett nytt test, Klurisen (som finns i två versioner, en A-del och en B-del), som tagits fram för att mäta, jämföra och öva matematisk problemlösningsförmåga hos barn. För detta syfte undersöks och jämförs Klurisen psykometriskt med Iowa Test of Basic Skills (ITBS) (specifikt delskalan Mathematical Problem Solving [MPS]), ett välkänt amerikanskt test av akademiska förmågor, Klurisen skulle kunna ha en rad praktiska fördelar gentemot traditionella test, som att det enkelt administreras genom surfplatta, vilket medger viss automatisering av administrationen, och det även innehåller interaktiva inslag som kan antas bidra till en mer rättvisande prestation hos eleven.

Utvärderingen sker genom att undersöka Klurisens resultat grupperat på kön och skoltillhörighet, samt genom att undersöka testets begreppsvaliditet, test-retest reliabilitet, parallell-form reliabilitet, samt homogenitetsreliabilitet. Detta för att säkerställa testets användbarhet som mått på matematisk problemlösningsförmåga.

Frågeställningar

Studiens frågeställningar är:

- Är Klurisen ett tillförlitligt testverktyg?
- Mäter Klurisen matematisk problemlösning?
- Är korrelationen starkare mellan Klurisen och Matte-Iowa än Klurisen och Iowa?
- Kan testresultat förklaras av könsvariabler, föräldrars utbildningsgrad eller elevers etnicitet?

Metod

Deltagare

Studien rekryterade elever från åk 2 från 7 skolor i en medelstor kommun i södra Sverige. Samtliga testdeltagare fick klassvis möjlighet att göra samtliga tester.

I studien deltog 140 flickor (51%) och 134 pojkar (49%). Uppvisat bortfall var 7% vid IOWA, 11% vid Klurisen första mättillfälle, samt 10% vid Klurisen andra mättillfälle. Åldern på deltagarna sträckte sig från 8-9 år, med undantag för ett fåtal barn som bedömts behöva gå om eller som började skolan i senare ålder. Vid jämförelser mellan utvalda skolor och riksgenomsnitt grupperat på genomsnittligt meritvärde eller andel elever som får betyg i samtliga ämnen, var majoriteten av skolorna under genomsnittet, några i linje med genomsnittet och en skola motsvarade genomsnittet. I fem av sju utvalda skolor fanns en större andel utlandsfödda barn än riksgenomsnittet. Två skolor hade större andel högt utbildade föräldrar än riksgenomsnittet och övriga skolor hade lägre andel än riksgenomsnittet. Urvalet verkade därför inte likvärdigt med landet i stort.

Etnicitet. Riksgenomsnittet för 8-9 åringars härkomst är ca 5% utlandsfödda, och ca 10% av barnen är födda i Sverige med båda föräldrarna födda utomlands. Skolor 2 och 3 hade en större andel föräldrar med svensk härkomst än riksgenomsnittet, övriga skolor (1,4, 5,6,7) hade större andel utlandsfödda barn eller svenskfödda barn med utlandsfödda föräldrar än riksgenomsnittet.

Utbildningsgrad. Demografiska data för den aktuella kommunen lämnades ut grupperat på skolnivå av tjänstemän vid kommunen. En jämförelse mellan riksgenomsnitt och aktuella skolor visade att färre föräldrar än genomsnittet hade eftergymnasial utbildning. Utbildningsgrad grupperat för gruppen föräldrar saknades på statistiska centralbyrån (SCB),

varför utbildningsgrad för 25-44 åringar användes som referensgrupp för riksgenomsnittet. En skola hade en stor andel lågutbildade föräldrar, jämfört med riksgenomsnittet och övriga skolor hade samma andel eller lägre andel lågutbildade än genomsnittet. Skolor med en större andel lågutbildade föräldrar i större utsträckning antogs här utgöra en indikator på andelen elever som växte upp med låg socioekonomisk status.

Instrument

Två test för mätning av matematisk problemlösningsförmåga användes i studien. Klurisen var det ena testet. IOWA Test of Basic Skills (ITBS) – deltest Math Problem Solving (MPS) var det andra testet, vilket användes som jämförelsematerial för att kunna utvärdera Klurisens psykometriska egenskaper.

Iowa Test of Basic Skills. ITBS är ett testbatteri av akademiska färdigheter utvecklat av Iowa Universitet och bygger på 80 års pågående forskning (Hoover et al., 2003). Testdeltagarens råpoäng på deltester jämförs mot en nationell norm, vilket gör att testresultaten kan belysa skillnader mellan individer, klasser och skolor. Görstesterna under upprepade årskurser kan resultaten även visa på utveckling över tid. Testen är skapade för elever mellan förskoleålder och åk 8. Innehållet i testbatteriet avspeglar den nationella läroplanen i USA; förändringar i läroplan innebär förändringar i testen. Testskaparna framhåller att deras tester inte är tänkta att predicera kommande prestationer, utan utvärdera om pedagogiken når eleven, samt följa den enskildes styrkor och svagheter i skolämnen över tid (Hoover et al., 2003). Test-retest reliabilitet för deltagare med ett års intervall är för olika deltest mellan 0,70-0,90 och reliabilitetskoefficienter varierar mellan 0,80-0,90 (Gregory, 1996).

Deltest Math Problem Solving (MPS) av IOWA består av totalt 30 frågor varav fråga 1 – 22 mäter problemlösning medan resterande mäter informationsbearbetning, dvs. att tolka och förstå tabeller och grafer. Testet är inte normerat för svenska skolbarn eller anpassat efter den svenska läroplan men kan ses som ett psykometriskt test inom matematisk problemlösning med hög reliabilitet och känd validitet. Test-retest reliabilitet för MPS är 0,75 för mellan höst- och vårtermin åk 2 och reliabilitetskoefficienten 0,83. (Hoover et al., 2003)

ITBS – MPS är ett skriftligt test med flervalssalternativ som enligt amerikansk testmanual uppskattas ta 25 minuter att genomföra (Dunbar et al., 2008). Varje deltest av ITBS har medföljande instruktionsmanual, som behöver följas för att säkerställa interbedömarreliabiliteten vid testadministration. Testet görs individuellt men administreras

klassvis av en testledare. De första 8 frågorna ges helt muntligt, de resterande 22 frågorna kan eleverna också läsa själva. Testet görs i en gemensam takt, testledaren bedömer hur länge gruppen behöver för att svara på en enskild fråga innan det är dags att gå vidare.

Instruktioner, testfrågor och testformulär översattes från engelska till svenska för denna studies genomförande. World Health Organizations (WHO) rekommendationer (2013) för översättning och korsöversättning vägledde arbetet, för att minimera fel under översättningsarbetet.

En mindre för-/pilotstudie genomfördes på en första översättning av ITBS-MPS, där svarsformulär, delfrågor och instruktioner testades i halvklass. Frågor och formuleringar som översättare eller korsöversättare reagerat på under översättningsprocessen testades i fokusgrupp på elever som lärare valt ut som i behov av extra stöttning och hjälp. Efter feedback genom mötet med halvklass och fokusgrupp, från närvarande lärare, testledare och testhandledare sammanställdes en slutgiltig version av ITBS-MPS, som fortsättningsvis kallas IOWA.

Klurisen. Klurisen består av två parallella delskalor, A- och B-del, med 15 frågor vardera. Delfrågorna är konstruerade utifrån problemlösningshierarkier som börjar enkelt men progressivt blir svårare. Tidiga nivåer innehåller korta meningar och enkel huvudräkning. Senare frågor innehåller mer information som deltagaren behöver ta ställning till. Vidare förs oviktig information in som deltagaren behöver identifiera och bortse ifrån. De första frågorna är troligtvis välbekanta för att bli allt mer ovana, för att ge testpersonen tillfälle att applicera gammal kunskap för nya problem (för diskussion kring lämplig utformning av frågor se Fuchs et al., 2009; Jitendra, 2008).

Klurisens A- och B-del är parallell-form test, dvs. delskalor skapade enligt samma problemlösningshierarkier, för att mäta samma typ av färdigheter och för att ge liknande resultat. Tidigare studier har visat alfa-koefficienter på 0,83 för A-del och 0,81 för B-del (Sjöberg, 2012), vilket är en hög grad av tillförlitlighet och möjliggjort att deltesterna kan betraktas som fullgoda test var för sig.

Klurisen görs individuellt på surfplattor. Instruktioner ges genom hörlurar av en animerad figur och svaren trycker deltagarna in själva på surfplattan, s.k. dator-baserad testning (CBT). Den animerade figuren ger intermitterent beröm efter klarade uppgifter. Vid behov ger testledaren motiverande stöd, men förklarar inga begrepp eller hjälper till på andra sätt som underlättar problemlösningen för eleven. Instruktioner innan och under test är standardiserade. När både A och B-del är genomförda följer ett ”bonusspel” som uppehåller

deltagare som är färdiga med testet från att störa de som fortfarande gör testet. Testet är normerat för svenska elever och testat på barn runt om i landet, främst Stockholm, Uppsala och Lund. Beräknad testtid är 20 minuter. För att undvika uttröttnadseffekter av att samtliga elever börjar med ena testdelen följt av den andra, balanserades ordningen klassvis. Av de 15 klasser som gjorde Klurisen började 8 klasser, med att göra A-del följt av B-del. 7 klasser, började med B-del följt av A-del. Vid retest-tillfället vändes testordningen.

För att säkra procedurell reliabilitet pilottestades Klurisen med två helklasser i södra Sverige, främst för att avgöra rätt procedur att ge förberedande instruktioner, fördelning av surfplattor, hörlurar och individuella inloggningar till testet. Ett fynd var att vissa elever snarare hoppade över uppgifter istället för att ge ett svar. Frågorna blev progressivt svårare för varje fråga, så sannolikheten att kunna en senare fråga när en tidig fråga varit för svår, var låg, men efter 15 frågor började nästa, inledningsvis enkla test. För att hjälpa elever att anstränga sig på enkla frågor, infördes en informationsruta när deltagaren gjort det första testet och skulle börja på nästa del. Rutan förklarade att det skulle bli enkla frågor igen och att de skulle göra sitt bästa. Vid båda testtillfällena hade deltagare rätt till kladdpapper. Vid pilottestning framkom det att inte alla elever förstod vad de kunde användas till (ex. kontrollräkning av svar). Kontrollfrågor och korta instruktioner för vad kladdpapper kunde användas till, blev en del av muntliga instruktioner vid kommande testning av Klurisen. Efter feedback genom testförfarandet, feedback från testdeltagare och närvarande lärare antogs en procedurmall för testledare.

Procedur

Rektorer för skolor kontaktades under hösten 2012, därefter lärare till berörda klasser. Eleverna rekryterades klassvis till studien. Studien genomfördes inom ramen för ett större universitetsanknutet forskningsprojekt. I rekryteringsprocessen till forskningsprojektet hölls olika möten där föräldrar till deltagarna fick chans att veta mer om projektet, möjliga fördelar med deltagande i projektet, samt deltagarnas frivillighet att när som helst avsluta sitt deltagande och sekretesshantering av data om person och resultat. Informerat samtycke inhämtades från deltagares vårdnadshavare. Den här studien ansågs vara en underordnad men sammanhängande del av projektet, varför inga ytterligare samtycken inhämtades till dessa testtillfällen. Resultaten kommer publiceras på gruppnivå och enskilda skolor, klasser och elever anonymiseras.

Vid testtillfället av IOWA fick deltagarna frågorna upplästa för sig och fick sedan ge skriftliga svar. Detta ställde höga krav på deltagarnas förmåga att ta till sig muntliga

instruktioner. En vanlig konsekvens under testförfarandet var att vissa frågor lästes upp långsammare en andra gång (efter det fick inga vidare instruktioner ges). Under ett tillfälle reagerade elever med frustration och gav uttryck för verbalt missnöje under testet, men testningen kunde fortsätta efter uppmaningar från klasslärare. Enskilda elever behövde olika lång tid att svara på frågor och att göra 30 frågor i helklass ställde höga krav på testledare att göra avvägningar när det var dags att gå vidare.

Då koncentrationsspann för barn är begränsad i låg ålder och växer med stigande ålder (Luna, Garver, Urban, Lazar & Sweeney, 2004), ökar risken för systematiska mätfel relaterade till låg koncentrationsförmåga ju längre tid testet tar att utföra. I samtliga fall översteg testtiden 25 minuter, genomsnittlig testtid var 33 min. För IOWAs testformulär fanns 4-8 frågor per sida, där elever också markerade rätt svar. Detta möjliggjorde att elever kunde tjuvkika på varandra. Att sitta nära varandra förvärrade det fenomen, vilket gjorde administrationen av IOWA känsligt för den fysiska miljön i klassrummet.

Vid administreringen av Klurisen fick deltagarna låna surfplatta, hörlurar, kladdpapper, samt inloggningsuppgifter i form av användar-id och lösenord. Därefter fick de svara på frågor av typen ”Är det någon som inte fått en surfplatta?” och ”Är det någon som saknar hörlurar?”. Instruktioner och en demonstration följde om hur hörlurar skulle sättas i på rätt sätt. Elever blev visade hur de själva kunde höja och sänka volymen på testinstruktioner. Därefter visades hur läderskyddet till surfplattan kunde vändas för att fungera som ett stativ under testningen. Lärare och testledare gick runt i klassrummet och såg till att samtliga elever följt instruktioner. Nästa steg var att sätta på surfplattan, ta på hörlurar, fylla i användaruppgifter och påbörja testningen. Deltagarna fick höra att testet bestod av flera delar och att testledare förklarar när testningen är avslutad, så de arbetar på tills de blir avbrutna. Elever som utförde Klurisens A- och B-del snabbt skulle få sitta längre med bonusspelet, medan den sista eleven som blev klar med A- och B-del fick en kort stund att spela bonusspelet innan testledare avbröt testtillfället. Klurisen var mindre beroende av den fysiska test-miljön, då deltagare lyssnade genom hörlurar, tittade på egen skärm, arbetade i sin egen takt och svarsalternativ försvann efter att eleven gett sitt svar. Testprogrammet skickade elevernas svar via trådlösa nätverk till en server som rättade och poängsatte proven.

Forskargruppen. I projektgruppen ingår Ata Ghaderi (professor vid Karolinska Institutet, KI), Martin Hassler Hallstedt (doktorand vid Uppsala universitet), en projektledare, två programmerare (konsulter), en testexpert (konsult), och ca tio forskningsassistenter. Forskningsprojektet är finansierat av Socialstyrelsen och Uppsala universitet. I den utvidgade

för projektet ingår professor Torkel Klingberg och hans forskargrupp vid KI som är specialiserad på arbetsminnesträning.

Statistisk analys

Statistiska analyser utfördes i SPSS 20.0. Råpoängen på Klurisen A- och B-del från båda testtillfällena, samt råpoäng från IOWA användes i analyserna.

För att avgöra om testdistributionerna var normalfördelade användes Shapiro-Wilks test, som har goda statistiska egenskaper i detta syfte. För att jämföra olika empiriska testdistributioner med varandra användes Kolmogorov–Smirnov test.

Homogenitetsreliabiliteten för Klurisen och IOWA och undersöktes genom Cronbachs alfa.

Eventuella skillnader i testprestation beroende på vilken testordning Klurisens A- och B-del hade som gjorts, undersöktes med oberoende t-test med korrektion för olika varianser.

För att undersöka samband mellan testresultat för test-retest-mätningar, parallell-form-mätningar samt konvergent validitet användes Pearsons r . Vid tolkning av Cronbachs alfa användes Panter och Brackers (2013) riktlinjer, där värden under 0,70 är otillräckliga, 0,70-0,79 duger för grundforskning, 0,80-0,89 är godtagbart för screeningverktyg och alfakoefficient över 0,90 krävs för testverktyg som används vid ex diagnosättning. Vid tolkning av korrelationer användes Bishops (2008) riktlinjer, där $r = 0,20-0,39$ avser ett svagt samband, $r = 0,40-0,59$ avser ett medelstarkt samband, $r = 0,60-0,79$ avser ett starkt samband, $r = 0,80-1,00$ avser ett mycket starkt samband. För samband mellan parallell-form test, användes DeVons et als. (2007) riktlinjer, där samband $r < 0,80$ avser ett otillräckligt samband, $r > 0,80$ avser ett tillräckligt starkt samband.

Eventuella skillnader i testprestation mellan könen undersöktes genom oberoende t-test. För att undersöka hur Iowas 22 delfrågor om matematisk problemlösning samvarierade med Klurisen, skapades kategorin Matte-Iowa, vars reliabilitet undersöktes och testresultat jämfördes med Klurisen. För att undersöka hur Iowas 8 övriga delfrågor rörande informationsbearbetning, samvarierade med Klurisen skapades kategorin Info-Iowa, vars samvariation med Klurisen undersöktes på liknande sätt.

För att undersöka om strukturella faktorer såsom etnicitet eller låg SES var predicerande variabler för utfallsvariabeln testresultat, utfördes linjär regression där prediktorerna aggregerades per skola. Det som undersöktes var föräldrars utbildningsgrad, och andel elever med annat födelseland än Sverige.

Resultat

Deskriptiva resultat för problemlösningstester

I tabell 1 redovisas testresultat i sin helhet, för IOWA-testet, Klurisen A-del vid första testtillfället (Klurisen A-1), Klurisen B-del vid första testtillfället (Klurisen B-1), samt Klurisen A-del vid andra testtillfället (Klurisen A-2), Klurisen B-del vid andra testtillfället (Klurisen B-2), Klurisen A+B/B+A vid första testtillfället (Klurisen Hel-1) samt Klurisen A+B/B+A vid andra testtillfället (Klurisen Hel-2). Även testernas råpoäng vid olika percentiler, redovisas.

Tabell 1

Medelvärde, standardavvikelser, poängintervall och cut-off gränser för percentiler för IOWA och Klurisen A- och B-del vid första och andra testtillfället, samt Klurisen Hel vid första och andra tillfället.

Test	M(SD)	Min-max	Percentil				
			5	25	50	75	95
IOWA (n=254)	19,67 (5,3)	7-30	11,0	16,0	19,5	24,0	28,0
Klurisen A-1 (n=243)	5,70 (2,6)	0-14	2,0	4,0	6,0	7,0	10,0
Klurisen B-1 (n=247)	5,68 (2,6)	0-14	1,0	4,0	6,0	7,0	10,0
Klurisen A-2 (n=243)	5,96 (3,0)	0-13	1,0	4,0	6,0	8,0	11,6
Klurisen B-2 (n=247)	6,06 (2,8)	0-14	1,0	4,0	6,0	8,0	10,6
Klurisen Hel-1 (n=247)	11,99 (5,4)	0-26	3,0	8,0	12,0	16,0	21,6
Klurisen Hel-2 (n=243)	11,42 (4,8)	0-28	4,0	8,0	11,0	15,0	19,8

IOWA. Testresultaten för IOWA approximerar en normalfördelad distribution med tydlig takeffekt, med ett medelvärde på $M=19,7$ ($SD=5,3$). Trots en tydligare takeffekt för pojkar, indikerade ett Kolmogorov-Smirnov test (K-S-test) av fördelningarna ingen skillnad i testdistributionerna för pojkar och flickor ($D=0,87$, $p=0,43$). Crohnbachs alfa för IOWAs 30 delfrågor var 0,83 (som väl motsvarar amerikanska referensvärden [$K-R_{20}$]= 0,83). Alfa-koefficienten för Matte-IOWAs 22 delfrågor kring matematisk problemlösning var 0,80. Alfa-koefficienten för Info-IOWAs 8 delfrågor kring informationsbearbetning var 0,48.

Klurisen A+B-1, B+A-1 (Klurisen Hel-1). Medelvärdet för gruppen som vid första testtillfället genomförde A+B var $M=11,7$ ($SD=5,5$), och gruppen som började med B+A hade i genomsnitt 12,4 ($SD=5,3$) i medelvärde. Ett oberoende t-test visade att det inte fanns någon

signifikant skillnad mellan de som gjort Klurisen A- och B-del i ordningen A+B, jämfört med de som gjort test i ordningen B+A från första testtillfället, $t(245) = 1,02, p = 0,31$. Resultatet kan alltså betraktas som likvärdig oavsett ordningsföljd på deltester. Testresultaten för Hel-1 uppvisar normalfördelad distribution, baserat på Shapiro Wilks test ($p=0,11$), med medelvärde $M=12,0$ ($SD=5,4$), utan signifikant skillnad i testdistributionerna för pojkar och flickor, baserat på K-S test av fördelningarna ($D= 1,12, p=0,165$).

Klursien A+B-2+B+A-2 (Klurisen Hel-2). Medelvärdet för gruppen som vid andra testtillfället genomförde A+B var $M=11,2$ ($SD 4,5$), och gruppen som började med B+A hade i genomsnitt $11,7$ ($SD=5,1$). Ett oberoende t-test visade att det inte fanns någon signifikant skillnad mellan de som gjort Klurisen A- och B-del i ordningen A+B, jämfört med de som gjort test i ordningen B+A från andra testtillfället, $t(242) = 0,75, p=0,46$. Resultatet kan alltså betraktas som likvärdig oavsett ordningsföljd på deltester. Testresultaten för Hel-2 visar på en normalfördelad distribution med tendens till takeffekt, baserat på Shapiro Wilks test ($p=0,06$), med medelvärde $M=11,4$ ($SD=4,8$). Två outliers förekommer i överkant av skalan. Vidare förekommer en signifikant skillnad mellan pojkar och flickor i testdistributionerna, baserat på K-S test av fördelningarna ($D=1,57, p=0,015$). Pojkars resultat för Hel-2 visade på en normalfördelad distribution med tendens till takeffekt, baserat på Shapiro Wilks test ($p=0,125$), med medelvärde $M=11,0$ ($SD= 5,2$), se figur 1 i Appendix. Flickor resultat approximerar en normalfördelad distribution för Hel-2, baserat på Shapiro Wilk test ($p=0,009$) med tydlig takeffekt, med medelvärde $M=11,8$ ($SD=4,3$), se figur 2 i Appendix.

Klurisen A-1, Klurisen A-2, Klurisen B-1, Klurisen B-2. För samtliga deltesten är medelvärden i spannet mellan $5,70-6,06$, med standardavvikelser mellan $2,6-3,0$. Testresultaten för Klurisen A- och B-delar från båda testtillfällena approximerar normalfördelade distributioner, med fyra outliers i överkant. Medelvärden och standardavvikelser går att utläsa i tabell 1.

Interkorrelation mellan för Klurisens deltest vid olika testtillfällen.

Deltest	Klurisen A-1	Klurisen B-1	Klurisen A-2	Klurisen B-2
Klurisen A-1	1	0,70**	0,72**	-
Klurisen B-1		1	-	0,56**
Klurisen A-2			1	0,69**
Klurisen B-2				1

För att pröva test-retest reliabilitet jämfördes resultat för samma delskalor mellan olika testtillfällen, samt resultat från båda delskalorna tillsammans från olika testtillfällen. Enligt Bishops (2008) riktlinjer för r -värden visade test-retest mätningar på starka samband för Klurisens A-del mellan båda testtillfällena, samt mellan Klurisen Hel-1 och Hel-2 (Klurisen A+B jämfört mellan båda testtillfällena). Klurisens B-del visade vid test-retest mätningar på medelstarka samband. Klurisen A-1 och A-2 ($n=219$) korrelerade signifikant, $r = 0,72$ $p < 0,001$. Resultat för Klurisen B-1 och B-2 ($n=219$) korrelerade signifikant, $r = 0,56$ $p < 0,001$. Resultat för Klurisen Hel-1 och Hel-2 ($n=219$) korrelerade signifikant, $r = 0,74$ $p < 0,001$. Det är viktigt att notera att korrelationer som görs med B-delen från någon av testtillfällena (parallell-form reliabilitet, eller test-restest reliabilitet) visar på en lägre reliabilitet, som möjligen kan härledas till att B-delen är en mindre tillförlitlig testdel.

För att undersöka parallell-form reliabiliteten jämfördes resultat för båda delskalor från samma testtillfälle. Detta gjordes för båda testtillfällena. Enligt DeVons (2007) riktlinjer var sambandet otillräckligt starkt mellan deltest A och B vid både första tillfället ($r = 0,70$, $p < 0,001$) och vid andra andra testtillfället ($r = 0,69$, $p < 0,001$).

Enligt Panter och Brackers (2013) riktlinjer för homogenitetsreliabilitet, var α -värden för deltesten A-1, B-1, B-2 godkända för grundforskning (0,70-0,79) och deltestet A-2 var godkänt för screeningsyften. För Hel-1 och Hel-2 nådde homogenitetskoefficienterna godkända nivåer för screeningstest (0,80-0,89). Alfakoefficienten för A-1 och A-2 var 0,73 och 0,80. Alfakoefficienten för B-1 och B-2 var 0,74 och 0,76. Alfakoefficienten för Hel-1 och Hel-2 var 0,87 och 0,84.

Tabell 3.

Korrelationskoefficienter mellan Klurisen Hel-1, Klurisen Hel-2, samt Iowa, Matte-Iowa och Info-Iowa.

Test	Klurisen Hel-1	Klurisen Hel -2	Iowa	Matte-Iowa	Info-Iowa
Klurisen Hel-1	1	0,74**	0,72**	0,71**	0,53**
Klurisen Hel -2		1	0,60**	0,58**	0,45**
Iowa			1	0,97**	0,76**
Matte-Iowa				1	0,59**
Info-Iowa					1

* $p < 0,01$; ** $p < 0,001$

Validitet. Begreppsvaliditeten för Klurisen undersöktes genom korrelationstester mot IOWA. Jämförelser mellan Klurisen Hel-1 och IOWA visade på starkt signifikant samband, $r(227) = 0,72$, $p < 0,001$, liknande jämförelse mellan Klurisen Hel-2 och IOWA visade på starkt signifikant samband, $r(234) = 0,60$, $p < 0,001$.

Då Iowa består av två deltest, exkluderades de delfrågor som inte rörde matematisk problemlösning. Testresultat delskalan Matte-Iowa jämfördes med testresultat för Klurisen. Jämförelser mellan Klurisen Hel-1 och Matte-IOWA visade på starkt signifikant samband, $r(227) = 0,71$, $p < 0,001$, liknande jämförelse mellan Klurisen Hel-2 och Matte-Iowa visade på ett medelstarkt signifikant samband, $r(234) = 0,58$, $p < 0,001$. När Info-Iowa, dvs delfrågor från IOWA som mäter informationsbearbetning, jämfördes med Klurisen Hel-1, fanns ett medelstarkt signifikant samband $r(227) = 0,53$, $p < 0,001$. När Info-Iowa jämfördes med Klurisen Hel-2 fanns ett medelstarkt samband $r(234) = 0,45$, $p < 0,001$.

Könsskillnader i prestation

IOWA. Pojkar uppvisade i genomsnitt 20,0 (SD = 5,5) i provresultat för IOWA och flickor hade i genomsnitt 19,4 i provresultat (SD=5,2). Ett oberoende t-test visade att det inte förelåg någon signifikant skillnad mellan pojkar och flickors testresultat, $t(251) = -0,81$, $p = 0,422$.

Matte-IOWA. Pojkar uppvisade i genomsnitt 14,4 (SD = 4,4) i provresultat för Matte-IOWA. Ett oberoende t-test visade att det inte förelåg någon signifikant skillnad mellan pojkar och flickors testresultat, $t(251) = -0,74$, $p = 0,414$.

Klurisen Hel-1. Pojkar uppvisade i genomsnitt 11,4 (SD= 6,0) i provresultat och flickor hade i genomsnitt 12,6 i provresultat (SD=4,8). Ett oberoende t-test visade att det inte förelåg någon signifikant skillnad mellan pojkar och flickors testresultat, $t(244) = -1,70$, $p = 0,09$.

Klurisen Hel-2.

Pojkar uppvisade i genomsnitt 11,0 (SD= 5,2) i provresultat och flickor hade i genomsnitt 11,8 (SD=4,3) i provresultat. Ett oberoende t-test visade att det inte förelåg någon signifikant skillnad mellan pojkar och flickors testresultat, $t(242) = -1,22, p = 0,225$.

Skillnader i prestation beroende av föräldrars utbildning och andel utlandsfödda

För att undersöka om andelen utlandsfödda elever eller andelen lågutbildade föräldrar i de olika skolorna, kunde predicerande testresultat på Klurisen och IOWA genomfördes linjär regressionsanalys av testresultat aggregerat per skola.

Jämförelse mellan skolor grupperat på ”andel lågutbildade föräldrar”. Andelen elever med lågutbildade föräldrar på skolan (låg SES) predicerade inte signifikant resultaten på något av de tre testen, även om en tendens fanns för Klurisen Hel-2 (se tabell 4). Tendensen för Klurisen Hel-2 visar att ju större andel elever på skolan med lågutbildade föräldrar, desto lägre sjunker medelvärdet på skolorna, $\beta = -0,650, t(6) = -2,09, p = 0,082, R^2 = 0,42$.

Tabell 4

Sammanfattning av linjära regressioner med andel lågutbildade elever per skola som oberoende variabel (N=7).

Test	B	SE(B)	β	t	p	R^2
IOWA	-0,08	0,04	-0,58	-1,76	0,129	0,34
Klurisen Hel-1	-0,07	0,05	-0,50	-1,39	0,213	0,25
Klurisen Hel-2	-0,09	0,04	-0,65	-2,09	0,082	0,42

Jämförelse mellan skolor grupperat på ”andel utlandsfödda elever”. Andelen utlandsfödda elever på skolan predicerade inte signifikant resultaten på något av de tre testen (se tabell 5), men en stark tendens fanns för Klurisen Hel-2. Tendensen för Klurisen Hel-2 visar att ju större andel utlandsfödda elever på skolan, desto lägre sjunker medelvärdet på skolorna, $\beta = -0,69, t(6) = -2,32, p = 0,059$.

Tabell 5

Sammanfattning av linjära regressioner med andel utlandsfödda elever per skola som oberoende variabel (N=7).

Test	B	SE(B)	β	t	p	R^2
IOWA	-0,06	0,03	-0,62	-1,96	0,098	0,39
Klurisen Hel-1	-0,05	0,03	-0,51	-1,45	0,197	0,26
Klurisen Hel-2	-0,07	0,03	-0,69	-2,32	0,059	0,47

Diskussion

En frågeställning för studien vara om Klurisen är ett tillförlitligt testverktyg. Det korta svaret på den frågan är ja, men bara om det ses som ett stort test istället för två mindre utbytbara test. Det långa svaret måste ta i beaktande Klurisens båda deltest och deras respektive tillförlitlighet. Det är otillfredsställande att de båda testdelarna, A- och B-delen inte når godkända nivåer av parallell-form reliabilitet (r -värden överstiger aldrig 0,80), trots att de är designade för detta syfte. Test-retest reliabiliteten för deltest A visar på starka samband mellan elevers prestationer mellan de båda testtillfällena, vilket motsvarar ett önskvärt resultat. Test-retest reliabiliteten för deltest B visar dock bara på medelstarka samband, vilket innebär att en större del av testresultatet för B-delen beror på slumpvisa mätfel än för A-delen. Tidigare utprovningar för urvalsgrupper i Lund och Uppsala har även de gett resultat som antyder att B-delen inte är lika tillförlitlig (personlig kommunikation med Martin Hassler, 2013). Tröskelvärden för homogenitetsreliabilitet styrs av vad användningsområdet för testet är, där screenings syften ökar kravnivån (Panter & Bracken, 2013). Klurisen är designad för kliniska sammanhang (att underlätta upptäckten av elever som behöver stödundervisning). När deltesterna undersöks vidare med avseende på Crohnbachs Alfa-värden, får båda deltesterna alltför låga värden för screenings syften. Om Klurisens A- och B-delars gemensamma homogenitetsreliabilitet undersöks (dvs Klurisen Hel-1 eller Hel-2) överstigs dock rekommenderade tröskelvärden. Klurisens A- och B-del är alltså tillsammans ett tillförlitligt testverktyg men de statistiska fynden antyder att, framför allt, B-delen behöver undersökas och förbättras på item-nivå för att öka sannolikheten att parallell-form reliabiliteten ökar vid kommande utprovningar. Om A- och B-delen skulle fås att fungera som utbytbara deltest, kan

de förutom att användas som screeningverktyg också användas som utfallsmått för att utvärdera hur väl elever svarar på stödinsatsen (där den ena delen görs innan stödinsatser och den andra delen görs efter stödinsatser).

En annan frågeställning för studien rörde begreppsvaliditet, huruvida Klurisen mäter matematisk problemlösning på samma sätt som Iowa mäter matematisk problemlösning. Det korta svaret på denna frågeställning är ja, Klurisen gör det. Det långa svaret behöver beakta Iowas egenskaper som test i denna studie, samt konvergent och divergent validitet. För urvalet som deltog i studien uppgår Iowas homogenitetskoefficient till likvärdigt goda nivåer som det ursprungliga deltestet rapporteras ha för deltagare i USA. Så, Klurisens begreppsvaliditet kan undersökas genom jämförelser med Iowa. Som nämnt tidigare, ett starkt samband mellan testresultaten för Klurisen och Iowa innebär att de mäter samma begrepp på liknande sätt. Men Iowa är designat att mäta en vidare form av problemlösning än Klurisen är designad för. För Iowa rör 22 frågor problemlösning (Matte-Iowa) och 8 frågor informationsbearbetning (Info-Iowa). Givet att båda delarna av Iowa når godkända nivåer för tillförlitlighet som mått, hade det varit önskvärt att Klurisen har ett starkare samband med Matte-Iowa, som den påminner mer om, än Info-Iowa som den skiljer sig från. En annan önskan är att sambandet mellan Klurisen och Matte-Iowa är likvärdigt eller starkare än Klurisen och Iowa som helhet. De statistiska analyserna visar att Info-Iowa ej når godkända nivåer för homogenitet, vilket enligt Wells och Wollack (2003) betyder att Info-Iowa ej bör undersökas vidare som separat konstrukt. Om detta ändå görs, mot bättre vetande, bör slutsatser dras med försiktighet. När Klurisen jämförs med Info-Iowa är sambandet medelstarkt, men när Klurisen jämförs med Matte-Iowa eller Iowa som helhet är sambandet starkt. Detta resultat gör att det går att dra slutsatsen att Klurisen mäter matematisk problemlösning på ett sätt som bättre motsvarar Matte-Iowa än Info-Iowa, men det är oklart om den också delvis motsvarar Info-Iowa som konstrukt.

En systematisk skillnad mellan testförhållandena när Klurisen gjordes jämfört med när Iowa utfördes, var att eleverna inte var lika distraherade av varandras prestationer när vid genomförande av Klurisen, jämfört med nära Iowa-testningen pågick (då eleverna hade hörlurar som störde ut andras sorl när Klurisen testades). Vidare gav Klurisen beröm under testningens gång och gjordes på surfplatta, vilket troligtvis ledde till att eleverna hade lite roligare när de gjorde Klurisen. Den här studien har inte undersökt eller tagit del av studier som undersöker under vilka omständigheter elever på lågstadiet gör sitt bästa, dvs presterar nära sin optimala förmåga, men det kan konstateras att testerna skiljer sig åt på väsentliga sätt,

i hur roliga de är att utföra för den enskilda eleven. Detta kan bidra till att testen har okända systematiska skillnader i testresultat (positiva eller negativa), som kommer av hur motivationsaspekter medierar prestationen. Då sambanden mellan Klurisen och Iowa är tillfredställande är detta hot mot tillförlitligheten möjligtvis försumbar, men vidare forskning bör ta hänsyn till motivationsaspekter. Ett hot mot den inre validiteten i studien, dvs. med vilken säkerhet studien kan dra slutsatser om samband, ex att Klurisen mäter matematisk problemlösning på det sätt Iowa gör, är att testtiden för Iowa översteg rekommenderad testtid vid samtliga testtillfälle, med en medeltesttid på 33 minuter (jämfört med 25 minuter som är rekommenderad testlängd). Detta kan ha skapat en confounder, där uthållighet medierar testresultatet i matematisk problemlösning. Då förmågan till att motstå uttröttnadseffekter inte undersökts i urvalet, är det svårt att uttala sig om i vilken utsträckning detta mätfel påverkat testresultatet. Homogenitetsreliabiliteten för Iowa och Matte-Iowa når godkända nivåer för urvalet och når liknande värden som deltestet ITBS-MPS gör i USA, vilket gör att slutsatsen kan dras att Iowa är ett tillförlitligt testinstrument att jämföra Klurisen med, samt att samband mellan testresultat beror på att de mäter samma begrepp.

Den sista frågan för studien rör huruvida delar av resultatet på Klurisen kan bero strukturella skillnader i urvalsgruppen, beroende på variabler såsom kön, SES eller etnicitet. När dessa variabler undersöktes upptäcktes inga signifikanta skillnader mellan pojkar och flickor, eller mellan skolorna när etnicitet eller SES undersöktes. En brist i studien var att data om utbildning och härkomst endast fanns på skolnivå, vilket resulterade i ett mindre n-värde för dessa jämförelsegrupper. Detta ökar risken att avfärda ett reelt samband (typ-II fel), som inte går att upptäcka med allt för grovkorniga variabler. När variabeln etnicitet undersöktes, uppgick signifikansnivån till 0,059 vid andra testtillfället av Klurisen. Detta resultat visar att det finns en tendens till att ju större andel elever som inte är födda i Sverige (etnicitet operationaliserades som om man var född i Sverige eller utomlands), desto lägre kommer skolans medelvärde vara på det enskilda testet. Steele (1997) utforskade hur stereotypot påverkar flickors testprestation i matematik negativt, och finner att med rätt testinstruktioner ökar flickors testprestation signifikant. Dessa instruktioner var inte en del av de som gavs inför Klurisen eller Iowa. Trots det, fanns inga signifikanta skillnader mellan pojkar och flickors resultat, men i ett kontrafaktiskt mättillfälle där "rätt" instruktioner givits, hade kanske flickor presterat bättre? Detta går inte att undersöka i efterhand, men kommande utprovningar av Klurisen bör förhålla sig till stereotypot som drabbar flickor negativt i matematik (ex. kan instruktörerna utan att ljuga säga att pojkar och flickor gör lika bra ifrån sig på det här testet). För att med större säkerhet uttala sig om hur strukturella faktorer påverkar

elevers testprestation, bör denna information insamlas på individnivå i den mån det är metodologiskt är möjligt och etiskt försvarbart.

Klinisk signifikans

Fynden från Duncan et al. (2007) studie visar att ett reliabelt testverktyg i matematik har användningsområden bortom ämnet matematik. Fyndet, som replikerats i flera andra studier visar att tidiga färdigheter i matematik är unikt prediktivt för kommande prestationer. Fuchs och Fuchs (2006) artikel om hur elever i riskzonen upptäcks, visar att det inte räcker att använda ett enskilt testresultat från ett test som metod att upptäcka vilka som kommer behöva hjälp under lång tid. Deras modell efterfrågar tidiga stödinsatser till elever med låga prestationer och att genom upprepade mätningar upptäcka de som visar på få eller uteblivna resultat.

Traditionell psykometrisk teori betonar vikten av att ha testverktyg som med hög träffsäkerhet uppräcker rätt personer. De som upptäcks av screeningverktyget (specificitet) ska vara samma som de som upptäcks av utfallsmåttet (sensitivitet), avvikelser sänker verktygens tillförlitlighet. För stora delar av elevgruppen bör detta stämma, men enligt Fuchs och Fuchs (2006) är det inte så enkelt för elever i riskzonen för skolmisslyckande. För denna grupp kommer det finnas oväntade avstånd mellan förväntad förmåga och uppvisad prestation. Än mer kännetecknande egenskap för denna grupp enligt Fuchs och Fuchs (2006), är att de är motståndskraftiga mot stödinsatser.

I ett senare skede, är det möjligt att Klurisen kan användas både vid screeningtillfälle och vid eftermätning och om parallell-form reliabiliteten ökar, kan halva testet göras vid screening och andra halvan av testet göras vid eftermätningar och testresultaten kommer fortfarande ha hög tillförlitlighet. Givet att stödinstituten har etablerat goda behandlingseffekter i andra sammanhang, kan skillnaden i testresultat före och efter behandling användas som sätt att upptäcka elever som följer en kompensatorisk utvecklingsmodell, från de som visar på en motståndskraft mot interventioner. Denna studie har inte fördjupat sig i vilka specialpedagogiska metoder som är verksamma för dessa elever, utan nöjt sig med att undersöka hur de upptäcks och hur ett test såsom Klurisen skulle kunna vara en del i denna processen.

Etik. Grundliga etiska ställningstaganden hade gjorts inom ramen för det större forskningsprojektet. Med hänsyn till detta samlades inga ytterligare samtycken in, inför dessa mättillfällen. Bifogat i slutet av uppsatsen (bilaga 2) är den informationsbroschyr och samtyckesblankett som föräldrar fick ta del av inför ett deltagande i det större projektet. Då

kommunen, skolorna och eleverna aidentifierats, inte tar del av resultaten annat än genom den här rapporten, torde den möjliga skadan av att delta för den enskilda eleven vara försumbar. Därför var det berättigat att använda en passiv samtyckesprocedur.

Avslutande summering

Som nämnt i teoridelen använder lärare mycket tid åt att utvärdera elevers inläring och lärare verkar vara i behov av tillförlitliga mått på elevens färdigheter. Testverktyg såsom Klurisen kunde användas för strukturerad utvärdering av elevers kunskapsutveckling eller oväntade brist på framsteg. Då skiftet i synen på inläringssvårigheter går i riktningen från att vara det oväntade avståndet mellan uppskattad förmåga och uppvisad prestation till att vara att en motståndskraft mot stödinsatser föreligger, kommer skolans behov av tillförlitliga test av akademiska färdigheter att öka. Utprövningen av Klurisen tyder på att den i nuläget är tillförlitlig nog att användas i screeningsyften och om vissa förändringar görs kan användningsområdet vidgas till att fylla fler syften.

Referenser

- Battin-Pearson, S., Newcomb, M. D., Abbott, R. D., Hill, K. G., Catalano, R. F., & Hawkins, J. D. (2000). Predictors of early high school dropout: A test of five theories. *Journal Of Educational Psychology*, 92(3), 568-582. doi:10.1037/0022-0663.92.3.568
- Bishop, P. A. (2008). *Measurement and evaluation in physical activity applications*. Scottsdale, AZ, Holcomb Hathaway.
- Boxer, P., DeLorenzo, T., Savoy, S., Mercado, I., & Goldstein, S. (2011). Educational aspiration-expectation discrepancies: Relation to socioeconomic and academic risk-related factors. *Journal Of Adolescence*, 34(4), 609-617. doi:10.1016/j.adolescence.2010.10.002
- Butterworth, B. (n.d). Dyscalculia: From brain to education (vol 332, pg 1049, 2011). *Science*, 334(6057), 761.
- Catts, H., Fey, M., Tomblin, J., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal Of Speech, Language, And Hearing Research*, 45(6), 1142-1157. doi:10.1044/1092-4388(2002/093)
- Cooter K. S., Cooter R. B. (2004). One size doesn't fit all: Slow learners in the reading classroom., *Reading Teacher*, 57(7), 680-684.

- DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., & ... Kostas-Polston, E. (2007). a psychometric toolbox for testing validity and reliability. *Journal Of Nursing Scholarship*, 39(2), 155-164. doi:10.1111/j.1547-5069.2007.00161.x
- Dunbar, S. B., Hoover, H. D., Frisbie, K. R., Ordman, D. A., Oberley, V. L., Naylor, R. J., & Bray, G. B., (2008). *The Iowa Tests*. Directions for administration. Iowa: Riverside Publishing.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., & ... Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, 43(6), 1428-1446. doi:10.1037/0012-1649.43.6.1428
- Fergusson, D., Swain-Campbell, N., & Horwood, J. (2004). How does childhood economic disadvantage lead to crime? *Journal Of Child Psychology And Psychiatry And Allied Disciplines*, 45(5), 956-966. doi:10.1111/j.1469-7610.2004.t01-1-00288.x
- Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly*, 25(1), 33-45.
- Fuchs, L. S., Fuchs, D., Prentice, K., Hamlett, C. L., Finelli, R., & Courey, S. J. (2004). Enhancing mathematical problem solving among third-grade students with schema-based instruction. *Journal Of Educational Psychology*, 96(4), 635-647. doi:10.1037/0022-0663.96.4.635
- Fuchs, D., & Fuchs, L. S. (2006). Introduction to Response to Intervention: What, Why, and How Valid Is It?. *Reading Research Quarterly*, (1). 93.
- Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., & ... Zumeta, R. O. (2009). Remediating number combination and word problem deficits among students with mathematics difficulties: A randomized control trial. *Journal Of Educational Psychology*, 101(3), 561-576. doi:10.1037/a0014701
- Garbarino, J., & Ganzel, B. (2006). The human ecology of early risk. In J. Shonkoff & S. Meisels (Eds.), *Handbook of early intervention* (2nd ed., pp. 76-93). New York, NY: Cambridge University Press.
- Gersten, R., Jordan, N., & Flojo, J. (2005). Early identification and interventions for students with mathematics difficulties. *Journal Of Learning Disabilities*, 38(4), 293-304.

- Gottfredson, L. S. (2008). Of what value is intelligence? I Prifitera, A., Saklofske, D.H., & Weiss, L.G. (Red), *WISC-IV clinical assessment and intervention* (2:a upplagan) (pp. 545-563). Amsterdam: Elsevier.
- Gregory, R. J. (1996). *Psychological testing : history, principles, and applications* / Robert J. Gregory. Boston : Allyn and Bacon, cop. 1996.
- Gustafsson, J. (2010). *School, learning and mental health : a systematic review* / Gustafsson, J.-E., Allodi M. Westling, Alin Åkerman, B., Eriksson, C., Eriksson, L. Fischbein, S., Granlund, M., Gustafsson, P. Ljungdahl, S., Ogden, T., Persson, R.S. Stockholm : Health Committee, Royal Swedish Academy of Sciences, cop. 2010.
- Gustafsson, J., & Erickson, G. (2013). To trust or not to trust?—Teacher marking versus external marking of national tests. *Educational Assessment, Evaluation And Accountability*, 25(1), 69-87. doi:10.1007/s11092-013-9158-x
- Gustafsson, J., Cliffordson, C., & Erickson, G., (2014). *Likvärdig kunskapsbedömning i och av den svenska skolan – problem och möjligheter*. Stockholm, SNS Förlag
- Hindman, A., Erhart, A., & Wasik, B. (2012). Reducing the matthew effect: lessons from the ExCELL head start intervention. *Early Education And Development*, 23(5), 781-806. doi:10.1080/10409289.2010.549443
- Hooper, S. R., Roberts, J., Sideris, J., Burchinal, M., & Zeisel, S. (2010). Longitudinal predictors of reading and math trajectories through middle school for African American versus Caucasian students across two samples. *Developmental Psychology*, 46(5), 1018-1029. doi:10.1037/a0018877
- Hoover, H. D., Dunbar, S. B., Frisbie, D. A., Oberley, K. R., Ordman, V. L., Naylor, R. J., Bray, G. B., Lewis, J. C., Qualls, A. L., Mengeling, M. A., Shannon, G. P. (2003). *The Iowa Tests. Guide to Research and Development*. Iowa: Riverside Publishing.
- Jitendra, A. K., (2008). *Solving Math Word Problems: Teaching Students with Learning Disabilities Using Schema-Based Instruction*. Austin, TX: Pro-Ed.
- Jensen, A. R. (1999). The G factor: The science of mental ability. *Psychology*, 10

- Jussim, L. (2012). *The Once raging and still smoldering pygmalion controversy*. in social perception and social reality: why accuracy dominates bias and self-fulfilling prophecy. Oxford University Press. 10.1093/acprof:oso/9780195366600.003.0013
- Kroesbergen, E., & Van Luit, J. (2003). Mathematics interventions for children with special educational needs: A meta-analysis. *Remedial And Special Education*, 24(2), 97-114.
- Kupiainen, S., Hautamäki, J., & Karjalainen, T. (2009). The Finnish school system and PISA. Helsinki: Ministry of Education Publications, Finland 2009:46
- Larheden, H., (2012). Barn till lågutbildade hamnar efter i skolan. *Välfärd*, 2012, 3, 8-9. Hämtad från http://www.scb.se/statistik/publikationer/LE0001_2012K03_TI_03_A05TI1203.pdf
- Leppanen, U., Niemi, P., Aunola, K., & Nurmi, J. (2004). Development of reading skills among preschool and primary school pupils. *Reading Research Quarterly*, (1). 72.
- Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreño, C., & Haas, K. (2010). Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology*, 46(5), 1062-1077. doi:10.1037/a0020066
- Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development*, (5). 1357. doi:10.1111/j.1467-8624.2004.00745.x.
- McCoy, A. R., & Reynolds, A. J. (1999). Grade retention and school performance: An extended investigation. *Journal Of School Psychology*, 37(3), 273-298. doi:10.1016/S0022-4405(99)00012-6
- McLeod, J. D., Uemura, R., & Rohrman, S. (2012). Adolescent mental health, behavior problems, and academic achievement. *Journal Of Health And Social Behavior*, 53(4), 482-497.
- Merton, R. K. (1968). The Matthew effect in science. *Science*, 159, 56-63.
- Morgan, P. L., Farkas, G., & Hibel, J. (2008). Matthew effects for whom? *Learning Disability Quarterly*, (4). 187.

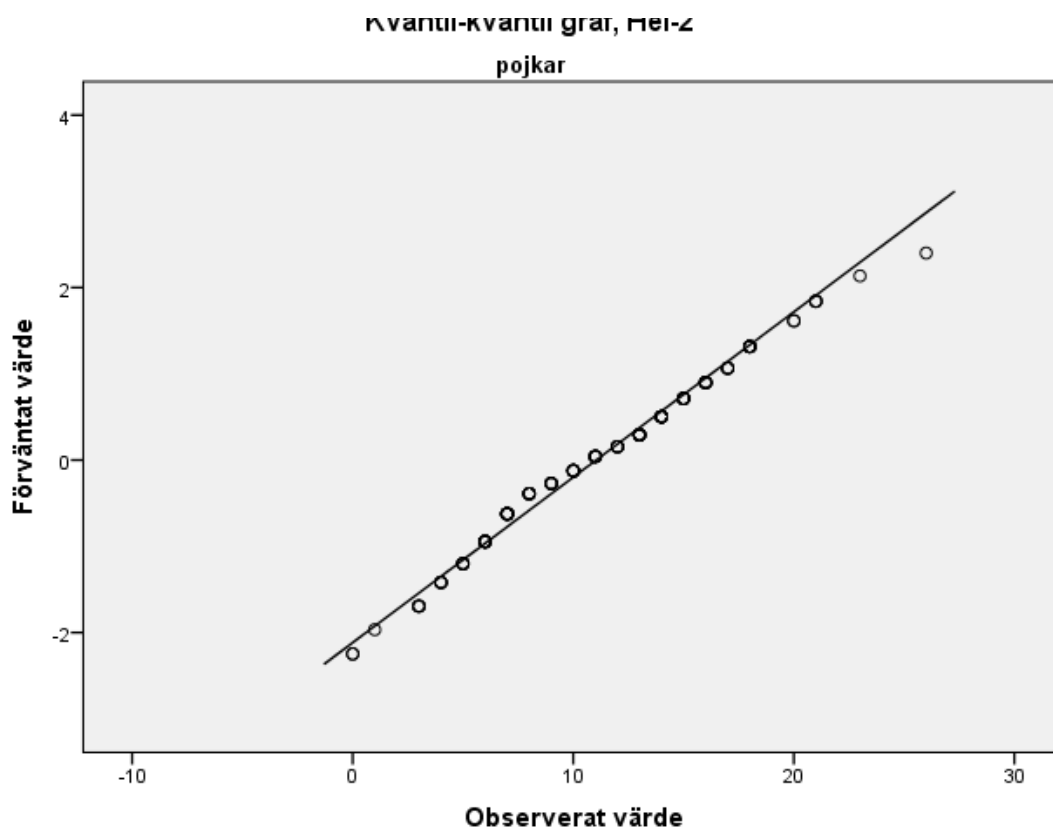
- Morgan, P.), Farkas, G.), & Wu, Q.). (2011). Kindergarten children's growth trajectories in reading and mathematics: Who falls increasingly behind? *Journal Of Learning Disabilities*, 44(5), 472-488. doi:10.1177/0022219411414010
- Naglieri, J. A. (2013). Psychological assessment by school psychologists: Opportunities and challenges of a changing landscape. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J. C. Hansen, N. R. Kuncel, S. P. Reise, M. C. Rodriguez (Eds.) , *APA handbook of testing and assessment in psychology, Vol. 3: Testing and assessment in school psychology and education* (pp. 3-19). Washington, DC, US: American Psychological Association. doi:10.1037/14049-001
- Niemi, P., Nurmi, J., Lyyra, A., Lerkkanen, M., Lepola, J., Poskiparta, E., & Poikkeus, A. (2011). Task avoidance, number skills and parental learning difficulties as predictors of poor response to instruction. *Journal Of Learning Disabilities*, 44(5), 459-471. doi:10.1177/0022219411410290
- Nunnally, J. C. (1979). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- Panter, J. E., & Bracken, B. A. (2013). Preschool assessment. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J. C. Hansen, N. R. Kuncel, S. P. Reise, M. C. Rodriguez (Eds.). *APA handbook of testing and assessment in psychology, Vol. 3: Testing and assessment in school psychology and education* (pp. 21-37). American Psychological Association. doi:10.1037/14049-002
- Parrila, R., Aunola, K., Leskinen, E., Nurmi, J., & Kirby, J. R. (2005). Development of individual differences in reading: results from longitudinal studies in English and Finnish. *Journal Of Educational Psychology*, 97(3), 299-319. doi:10.1037/0022-0663.97.3.299
- Passer, M. W., & Smith, R. E. (2007). *Psychology : the science of mind and behavior / Michael W. Passer, Ronald E. Smith*. Boston, Mass. : McGraw-Hill Education.
- Persson, L., Öhrvall, R., (2011) Januaribarn mer framgångsrika. *Välfärd*, 2011, 3, 14-17. Hämtad från http://www.scb.se/statistik/publikationer/LE0001_2011K03_TI_07_A05TI1103.pdf
- Pool, J., Carter, G., Johnson, E., & Carter, D. (2013). The Use and effectiveness of a targeted math intervention for third graders. *Intervention In School And Clinic*, 48(4), 210-217. doi:10.1177/1053451212462882
- Schad, E., (Red.). (2014). *Psykolog i skolan*. Lund: Studentlitteratur.

- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin, 2002.
- Siegler, R., Duncan, G., Davis-Kean, P., Susperreguy, M., Chen, M., Duckworth, K., & ... Engel, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, 23(7), 691-697. doi:10.1177/0956797612440101
- Sjöberg, A., (2012). Analys Mattetest 2012-10-08, www.psychometrics.se/? Efter personlig korrespondens med Martin Hassler Hallstedt per e-mail (2013-05-24).
- Snow, C. E., Burns, M., Griffin, P., & National Academy of Sciences - National Research Council, W. n. (1998). *Preventing reading difficulties in young children*.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6), 613-629. doi:10.1037/0003-066X.52.6.613
- Turton, A., & Green, S. (2013). Prevention strategies for students at-risk for learning disabilities. *Advances In Special Education*. 24, 109-128. doi:10.1108/S0270-4013(2013)0000024009
- VanDerHeyden, A., & Witt, J. (2005). Quantifying context in assessment: Capturing the effect of base rates on teacher referral and a problem-solving model of identification. *School Psychology Review*, 34(2), 161-183.
- Wells, C. S., Wollack, J. A. (2003) An instructor' guide to understanding test reliability. Testing & Evaluation Services, Wisconsin. Hämtad från <http://testing.wisc.edu/Reliability.pdf>
- Willis, J. O., Dumont, R., & Kaufman, A. S. (2013). Assessment of intellectual functioning in children. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J. C. Hansen, N. R. Kuncel, S. P. Reise, M. C. Rodriguez (Eds.) , *APA handbook of testing and assessment in psychology, Vol. 3: Testing and assessment in school psychology and education* (pp. 39-70). American Psychological Association. doi:10.1037/14049-003
- World Health Organisation, (2013) *Process of translation and adaptation of instruments*. Hämtad den 23 januari, 2013 från http://www.who.int/substance_abuse/research_tools/translation/en/

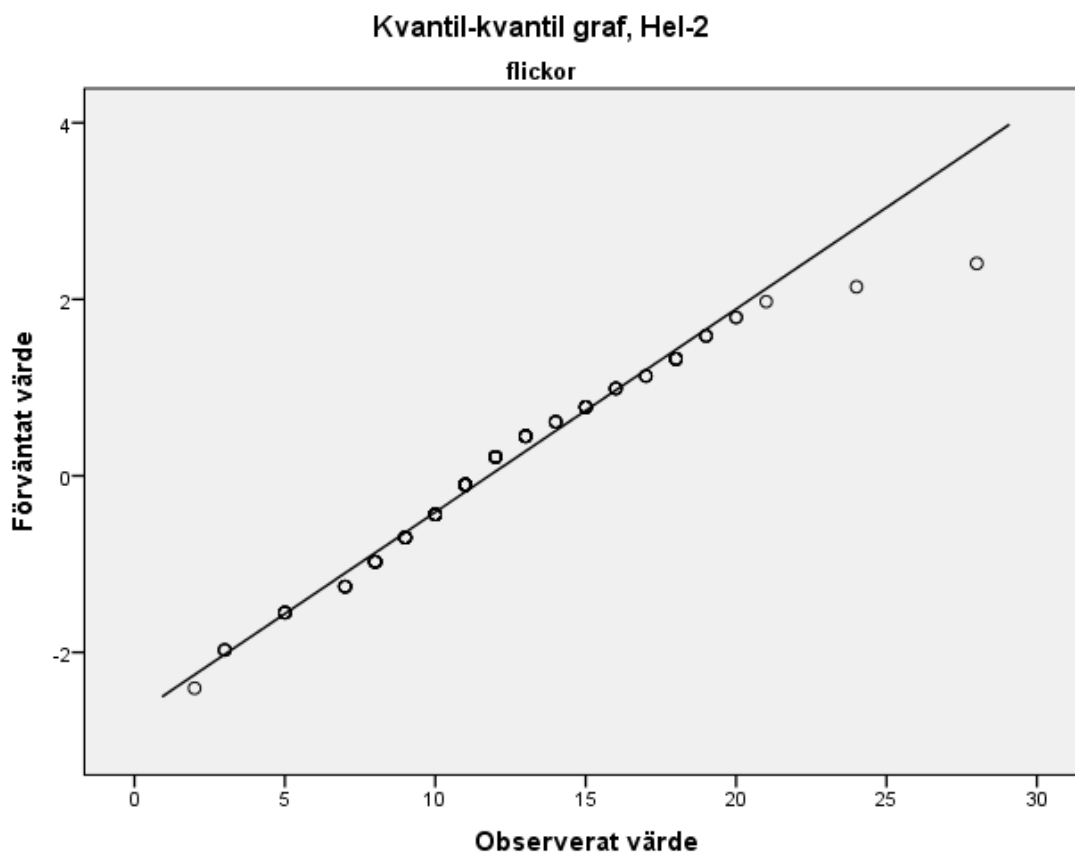
Bilaga 1

Kvantil-kvantil grafer för Klurisens helskala vid retest-tillfälle, grupperat på kön.

Figur 1.



Figur 2.



Information till föräldrar angående deltagande i studie

Information om forskningsprojektet

”Math Adventures”

Ditt/ert barn har genomfört ett matematiktest på surfplatta inom ramen för ett forskningsprojekt. Projektet utvärderar effekterna av ett matematikprogram på surfplatta riktat till elever med låga resultat i matematik i årskurs 2. Samtidigt är vi intresserade av sambandet mellan matematikförmåga och psykiska hälsa. Det finns studier som visar på ett samband mellan matematikfärdigheter och psykisk hälsa. Matematik är ett viktigt ämne. Syftet med matematiktestet var att undersöka ditt och andra barns färdigheter i matematik. Resultaten visar att ditt barn saknar en del viktiga färdigheter i matematik.

Vi vill undersöka i vilken utsträckning ett nytt matematikprogram på surfplatta som heter Math Adventures, ett lästräningsprogram på surfplatta eller om barnet fortsätter delta som vanligt i undervisningen kan hjälpa barn utveckla sin aritmetikförmåga och ifall Math Adventure även påverkar barns hälsa och välbefinnande i stort. Eftersom det är en forskningsstudie kommer slumpen att avgöra vilka barn som får ta del av Math adventure, lästräningsprogram, eller sedvanlig undervisning plus eventuella specialundervisningar som ges i skolan. Du och ditt barn tillfrågas härmed om deltagande i denna studie.

Hur går studien till?

Barn som tar del av matematikprogrammet eller lästräningsprogrammet tränar ca 15-30 min/dag under två terminer.

Som förälder besvarar man i början av första terminen ett frågeformulär på papper som består av 34 frågor. Frågorna handlar om hur barnet mår och fungerar samt några frågor som rör utbildningsnivå, sysselsättning och liknande. För att delta i projektet behöver man även ge sitt samtycke till att barnets lärare ska fylla i samma formulär för barnet samt frågeformulär där läraren bedömer barnets skolprestation. Även barnet får svara på frågor om hur barnet mår, men detta sker på surfplatta under skoltid. Barnet svarar totalt på 56 frågor som visas med text samtidigt som en röst läser texten. I skolan kommer barnet även att genomföra matematiktester, minnestester och ett begåvnings-test på surfplatta. Barnets lärare närvarar vid testerna och kan hjälpa till om något barn önskar det. Varje testtillfälle är som längst 20 min.

Effekterna av programmet utvärderas i början av första terminen, i slutet av första terminen, i slutet av den andra terminen då den aktiva delen avslutas, samt 6 och 18 månader efter att den aktiva delen avslutats. Det betyder att föräldrarna, lärarna och barnen ombeds besvara ett antal frågeformulär vid nästan varje utvärdering och att

barnen genomför matematik- och minnestester vid samtliga tillfällen. Vid varje utvärdering kan ni som föräldrar besluta er för att avstå från att fortsätta delta i projektet.

Ni kan även välja att medverka i en delstudie som undersöker hur barns minne påverkar matematikinläring. I så fall får barnet göra en undersökning med en så kallad magnetkamera, magnet resonanstomografi (MRI), som avbildar hjärnan. Föräldrarna sitter bredvid barnet under hela undersökningen. Om man bara är intresserad av matematikutvärderingen, men inte av minnets betydelse för matematikinläring, behöver man inte delta i denna undersökning. På det bifogade formuläret "Samtycke till deltagande i forskning" kan ni ange om ni är intresserade eller ej av mer information om delstudien av minne och matematikinläring.

Vilka är riskerna?

Inga risker med dessa tester, undersökningar eller deltagande i studien finns redovisade.

Finns det några fördelar?

Det är viktigt att tidigt uppmärksamma färdighetsbrister i matematik. Projektet som sker i mycket nära samarbete med skolorna skapar en ännu högre medvetenhet om betydelsen av matematik och sannolikheten för att alla barn som visar sig vara svagpresterande får en bra åtgärd. Hälften av alla barn som deltar kommer att få ta del av Math Adventures som på ett roligt och spännande sätt ska hjälpa barn lära sig grundläggande färdigheter i aritmetik. De barn som inte tar del av Math Adventures får utöver sin vanliga undervisning i matematik andra relevanta åtgärder då skolorna uppmanas att tillhandahålla de bästa specialprogrammen för matematik som en bra och relevant jämförelsegrupp för Math Adventures.

Sekretess, tystnadsplikt, databearbetning

Alla handlingar beträffande din medverkan i projektet kommer att behandlas konfidentiellt. Dina och ditt barns svar, samt det som rapporteras av lärare kommer att behandlas så att inga obehöriga kan ta del av dem. All personal du kommer i kontakt med i projektet har tystnadsplikt. Undantag i tystnadsplikten gäller enbart när ett barn far så pass illa att anmälningsplikt föreligger (t.ex. vid misshandel).

De frågeformulär du fyller i får ett kodnummer. Forskarna har en kodnyckel som berättar hur ett kodnummer kan paras ihop med en enskild individ. Utan kodnyckeln kan övrigt material inte spåras till enskild individ. Kodnyckeln förvaras separat och inlåst. Tillgång till materialet har endast de forskare som är knutna till projektet. Kodnycklarna förstörs så fort data är analyserad och rapporterad.

Undersökningsresultaten kommer enbart att presenteras på gruppnivå. Inga enskilda individer kommer att kunna identifieras.

Personuppgifter och information om studiens resultat

Hantering av dina personuppgifter regleras enligt Personuppgiftslagen (1998:204) och ansvarig för dina personuppgifter är Uppsala universitet. Uppgifterna lämnas inte vidare. Du har rätt att få ett registerutdrag en gång per år och att få rättelse av felaktiga uppgifter. Vänd dig i så fall till projektledaren Ata Ghaderi (adress och telefon på nästa sida).

Resultaten kommer att publiceras i form av vetenskapliga rapporter i internationella tidskrifter. Du som deltagare kan även få information om resultaten av studien genom att skriftligen kontakta projektledaren.

Försäkring och ersättning

Normal försäkringsskydd gäller. Ingen ersättning utgår till föräldrar.

Frivillighet

Deltagandet i studien är helt frivilligt. Du och ditt barn kan när som helst avbryta deltagandet utan att ange någon särskild förklaring. Detta påverkar inte ditt eller ditt barns sedvanliga omhändertagande och verksamhet i skolan eller andra instanser.

Om du har fler frågor så är du välkommen att ringa eller skriva till oss.

Vänligen fyll i om ni vill delta eller ej i det bifogade formuläret "Samtycke till deltagande i forskning". Posta svaret i det bifogade svarskuvertet. För att vi inte ska missa önskemål om att avstå från deltagande behöver vi ert svar inom två veckor.

Med vänliga hälsningar

Ata Ghaderi (ansvarig forskare)

Leg. psykolog, Leg. Psykoterapeut, Professor i psykologi

Institutionen för psykologi

Uppsala Universitet (forskningshuvudman).

Box 1225, 751 42 Uppsala

Tel: xxx-xxx xxxx, Fax: xxx-xxx xxxx

Samtycke till deltagande i forskning

Jag har informerats om studien och fått möjlighet att ställa frågor och få dem besvarade.

Jag är medveten om att deltagandet i studien är frivilligt och att jag när som helst kan avbryta mitt deltagande utan att ange någon särskild förklaring.

Jag samtycker härmed till att delta i denna studie och att mina och mitt barns personuppgifter får behandlas på det sätt som beskrivs i informationen om forskningsstudien. Jag ger också mitt samtycke att mitt barns lärare ska bidra med information om mitt barn från lärar- och skolperspektiv genom att besvara de frågeformulär som ingår i studien.

Jag är medveten om att resultatet från studien behandlas under sekretess, och inga obehöriga kommer att veta att jag eller mitt barn deltagit eller kunna se hur just vi svarade. Svaren kommer att sammanställas i anonymt grupp format, och presenteras så att enskilda personers svar inte kan spåras.

Ange om du vill ha mer information om delstudien av minnets betydelse för matematikinlärning med så kallad magnetkamera:



Ja, jag önskar mer information om delstudien av minnets påverkan på matematikinlärning.



Nej, jag önskar inte mer information om delstudien av minnets påverkan på matematikinlärning.

Observera att båda vårdnadshavares (i förekommande fall) godkännande och underskrift krävs för att bekräftelsen skall vara giltig.

Barnets namn

Målsmans underskrift och
Namnförtydligande

Ort och datum

Telefon och e-mail adress

Målsmans underskrift och
Namnförtydligande

Ort och datum

Telefon och e-mail adress

Information till barn angående deltagande i studie

Hej,

Vi är forskare som vill ta reda på om ett program på surfplatta eller iPad kan hjälpa barn att lära sig matematik. För att ta reda på det behöver vi barn som är intresserade av att vara med. Vi undrar därför om du skulle vilja vara med i vår forskningsstudie? Det är inget farligt. Du ska helt enkelt göra matematikuppgifter på surfplatta några gånger i skolan. Det är som ett spel på dator. Då kommer du också att få göra uppgifter där det gäller att komma ihåg saker och svara på frågor om hur du mår. Din lärare kommer att visa dig hur man använder surfplattan. Dessutom kommer du att vara med i en lottning som vi gör. Du kan lottas till att få träna mer matematik på surfplatta, eller träna läsning på surfplatta, eller att fortsätta med matematik som vanligt i skolan. Om du lottas till att träna på surfplatta gör du det ett litet tag varje dag under ett halvår eller ett år i skolan. Om du lottas till att fortsätta med matematik som vanligt får du fortfarande göra matematikuppgifter på surfplatta och svara på lite olika frågor några gånger.

Nu ska vi kort berätta lite om hur matematikuppgifterna som alla gör på surfplatta ser ut. Uppgifterna handlar om plus, minus och om att lösa problem med siffror. På surfplattan finns det en tecknad figur som lär dig hur man gör matematikuppgifterna, ungefär som en lärare. Först berättar figuren hur du ska göra uppgifterna och sen får du träna på att göra några uppgifter tillsammans med figuren. När du har tränat färdigt får du försöka göra så många uppgifter som du kan utan hjälp. Det tar ungefär 20 minuter att göra uppgifterna. Din lärare kommer att vara med hela tiden.

Om du vill vara med i studien och dina föräldrar vill att du är med, kommer din lärare hjälpa dig att komma igång. Det är helt frivilligt att vara med. Du kan när som helst sluta om du inte vill vara med längre.

Om du har frågor är du välkommen att ringa eller skriva till oss.

Med vänliga hälsningar

Ata Ghaderi

Professor i psykologi

Institutionen för psykologi

Box 1225, 751 42 Uppsala

Tel: xxx-xxx xxxx, Fax: xxx-xxx xxxx

xxx.xxxxxxx@xxx.xx.xx