



LUNDS
UNIVERSITET

Institutionen för psykologi
Kandidatuppsats

Kan du bli smartare?

En studie om förbättring vid upprepade försök på intelligenstester

Can you become smarter?

A study about improvement regarding repeated attempts on intelligence tests

Max Angelin, Hanna Digman och Oscar Forum

Kandidatuppsats HT 23

Handledare: Petri Kajonius

Examinator: Una Tellhed

Abstract

Intelligence tests are frequently used in recruitment contexts since these tests have shown to be a strong predictor when it comes to future work performance. Therefore, it is not uncommon for job seekers to take these tests on several occasions throughout their careers with scores fluctuating to some degree. The aim of the study was to investigate whether there is an improvement in the performance in intelligence tests after multiple exposures. The study was conducted with the use of three tests consisting of ten items of already existing items from cognitive tests which the participants ($n = 162$) filled out through an online survey, $n = 52$ completed all three tests. The results showed that there is a considerable improvement in performance between the first and second test as well as the first and third test. However, no significant difference was shown for the second and third tests. From this result we could conclude that we were able to observe a test-retest effect on intelligence tests, however, the rate of improvement decreases over time. This is in line with previous research conducted within the field. From these results we argue that there is a point of having candidates complete practice-tests before taking their final intelligence test in a recruitment process. This could improve the reliability of the use of intelligence tests, by taking the practice effect into consideration.

Keywords: intelligence test, test-retest effect, intelligence, recruitment, ICAR

Sammanfattning

Idag används intelligenstester frekvent i rekryteringssammanhang och då dessa tester har visat sig vara en stark prediktor för framtida arbetsinsats läggs vikt vid dessa när det kommer till olika urvalsprocesser. Av den anledningen är det inte ovanligt att arbetssökande får genomföra ett intelligenstest vid flera tillfällen i sin karriär vilket kan leda till olika resultat vid olika tillfällen. Studiens syfte var att undersöka om det kan finnas en förbättring av individens resultat på intelligenstester ju fler tester individen genomför. Studien bestod av tre stycken test där deltagarna ($n = 162$) fick svara på tio frågor bestående av items från redan existerande intelligenstester som deltagaren fick genomföra via ett webbformulär. $n = 52$ genomförde samtliga 3 test. Studiens resultat visar att det finns en betydande förbättring i prestationen mellan det första och andra testet samt det första och tredje testet. Däremot fann vi ingen signifikant skillnad mellan det andra och tredje testet. Vi kan konstatera att det existerar en så kallad test-retest effekt på intelligenstester då det sker en förbättring av resultatet efter det första testet. Vi kom även fram till att förbättringen avtog efter det andra testtillfället vilket stämmer överens med tidigare forskning. Utifrån resultatet anser vi att det finns en fördel i att låta kandidater genomföra övningstest inför sitt faktiska intelligenstest. På så sätt är det möjligt att öka testets reliabilitet i urvalssammanhang, genom att ta test-retest effekten i beaktning.

Nyckelord: intelligenstest, test-retest effekt, intelligens, rekrytering, ICAR

Tack!

Vi önskar att föra fram ett stort tack till alla som stöttat oss i arbetet med denna uppsats!

Stort tack till alla som valt att delta i studien och som på så sätt gjort hela projektet möjligt. Vi var redan på förhand medvetna om att det skulle krävas en hel del engagemang och hängivenhet från er och att vi lyckades få ihop så många deltagare glädjer oss mycket. Allas insats betyder så mycket för oss, och för det är vi så tacksamma.

Ett annat stort tack vill vi rikta till vår handledare, Petri Kajonius, som varit en otroligt viktig resurs för oss i dataanalys och för att bolla våra ibland spretiga idéer och funderingar med. Vi kan inte tänka oss en bättre handledare för detta som trott på denna studie från början, och alltid varit behjälplig och snabb med sin återkoppling.

Sist vill vi avsluta med ett tack riktat till varandra, för pepp och stöd oss emellan. Vi har haft stor respekt för varandra hela vägen igenom och gjort en uppsats som vi alla kan känna oss lika delaktiga i, och framförallt vara stolta över.

Kan du bli smartare?

En studie om förbättring vid upprepade försök på intelligenstester

Intelligenstester används frekvent i urvalssammanhang i syfte att mäta olika individers nivå av intelligens (Hausknecht et al., 2005). Eftersom testerna ofta förekommer i olika situationer är det vanligt att personer får göra flera intelligenstester vid olika tillfällen i karriären. Vid upprepade mätningar kan det uppstå en test-retest effekt vilket innebär att resultaten kan förbättras desto fler gånger test av samma typ genomförs. Denna studie undersöker om det finns en test-retest effekt i intelligenstest med hypotesen att en test-retest effekt existerar. Vidare diskuteras om en test-retest effekt påverkar reliabiliteten i intelligenstester samt om det är ett bra verktyg i urvalssammanhang.

Varför intressant att undersöka intelligens i rekryteringsprocesser

En av de viktigaste aspekterna för att hitta rätt kandidat i en rekryteringsprocess är att kunna förutsäga personens arbetsprestation. Schmidt och Hunter (1998) menar att intelligenstester är ett av de bästa sätten att predicera en persons framtida arbetsprestation, framförallt i kombination med andra metoder som arbetsprov, strukturerade arbetsintervjuer eller integritetstest. Resultatet i studien visar att intelligenstester för sig själva har en prediktiv validitet på $r=.51$, där 1 är perfekt prediktiv validitet. Kombinerar testerna med en strukturerad arbetsintervju stiger siffran till $R=.63$. I studien framkommer även att intelligens, General Mental Ability (GMA), är ett av de bästa sätten att predicera lärande i arbetet. Utifrån prediktionsgraden i GMA kopplat till arbetsprestation och lärande anser forskarna att intelligenstester bör ses som ett av de primära verktygen för att utvärdera kandidater i rekryteringsprocesser: "Because of its special status, GMA can be considered the primary personnel measure for hiring decisions, and one can consider the remaining 18 personnel measures as supplements to GMA measures" (Schmidt & Hunter, 1998, s. 266).

Varför tillåts kandidater göra om tester? Är det rimligt?

Det är vanligt att låta personer göra om tester i urvalsprocesser i både utbildningssammanhang och i arbetslivet (Hausknecht et al., 2005). Det finns två tydliga motiveringar till varför omtestning tillåts (Lievens et al., 2005). En anledning kopplas till

personliga förutsättningar vid testtillfället (exempelvis sjukdom) och den andra är kopplad till själva testet (exempelvis avvikelser från standarden för testadministration eller mätfel). Det är även vanligt med tester som skrivs vid specifika tillfällen som det svenska högskoleprovet. I Sverige har i snitt 40% av de som skriver högskoleprovet gjort det vid minst ett tillfälle tidigare (Cliffordson, 2004).

En tidigare metaanalys som undersökt test-retest effekt på intelligenstester för 134,436 individer visar att personer i snitt höjer sig med en kvarts standardavvikelse från det första testet till det andra (Hausknecht et al., 2005). Hur mycket en individ höjer sig från ett test till ett annat beror på vilken typ av intelligens som mäts, hur testet administreras och personens mentala förmåga eller g-faktor som det också kallas (Arendasy & Sommer, 2017). Förbättringar av resultat genom att göra om tester gör det rimligt att ifrågasätta rättvisan i användningen av intelligenstester i urvalssammanhang.

Intelligens

Enligt Sternberg et al., (2003) finns två olika tillvägagångssätt att se på fenomenet intelligens. Det ena är det psykometriska tillvägagångssättet, som handlar om hur vi mäter intelligens. Det andra är det kognitiva tillvägagångssättet som studerar vad intelligens är och hur det fungerar. Denna uppsats bygger på en studie som behandlar själva testningen, det vill säga den psykometriska delen av intelligens.

Intelligens utgår ifrån G-faktorn, ett mått skapat av Spearman som kan likställas med vår generella nivå av mental förmåga och hur väl en individ kan klara av olika uppgifter. G-faktorn delas upp i kristalliserad och flytande intelligens, där kristalliserad syftar till det vi lärt oss och tränat på tidigare och hur detta kan appliceras även på nya situationer. Flytande intelligens är istället hur vi kan resonera och tänka logiskt. G-faktorn kan ses som grunden till intelligens, av den anledningen kan en persons g-faktor användas för att predicera förväntad prestation (Holt et al., 2019, s. 335).

Utvecklingen av intelligenstester

Introduktionen av intelligenstestning går tillbaka till början av 1900-talet då Alfred Binet skapade vad som än idag antas ligga som grund för dessa tester (Lindelöw Danielsson, 2003, s. 107). Binet fick i uppgift att ta fram ett test för att kontrollera huruvida skolelever behövde skolstöd eller inte beroende på deras nivå av intelligens. Testerna byggde på antaganden om att kompetens utvecklas med ålder, och att det utvecklas kontinuerligt. Från

detta skapades en teori om "mental age" som också ligger som grund för fenomenet IQ (Holt et al., 2019, s. 340).

Utifrån Binets tester utvecklade amerikanen Lewis Terman ett nytt test, som kom att kallas The Stanford-Binet test. Som en konkurrent till detta skapade därefter Wechsler ytterligare ett test, Wechsler Adult Intelligence Scale (WAIS) som senare utvecklats i flera olika nivåer, då han menade att ens nivå av intelligens inte bara beror på verbala förmågor utan vi behöver även ta hänsyn till, och mäta, icke-verbala förmågor. WAIS- och Stanford-Binet-testerna är än idag två av de mest använda intelligenstesterna, och ligger till grund för flertalet vedertagna liknande tester som används idag. (Holt et al., 2019, s. 340).

Idag används dessa tester för att undersöka vilken nivå av intelligens en person besitter och vilka intellektuella förutsättningar hen har. Av den anledningen används dessa tester ofta i rekryteringssammanhang för att kontrollera hur pass analytisk en person är, hur skarp personen är när det kommer till problemlösning samt att processa ny kunskap, för att sedan kunna applicera detta på förväntad prestation i arbetet. (Lindelöw Danielsson, 2003, s. 120)

ICAR

I vår studie har vi använt oss av items från The International Cognitive Ability Resource (ICAR) vilket är en allmän domän med bestående items, som skapats med avsikten att förse forskning med ett effektivt verktyg för neuropsykologiska bedömningar, det vill säga items till intelligenstester. ICAR ville ta fram valida och beprövade items som skulle vara lätt-tillgängliga för forskare att använda, och samtidigt tillhandahålla tydlig information om dessa resurser. Genom att uppmuntra till användandet av ICAR items hoppas forskarna på att öka medvetenheten om intelligens och hur intelligensnivåer kan testas, samt hur det hör ihop med andra variabler. Projektet hoppades även på att kunna öka användandet av intelligensnivåer i forskningssyfte genom att skapa lätthanterliga items. (The Project - International Cognitive Ability Resource - the ICAR Project, 2014). The ICAR Project är bekostat av the National Science Foundation (NSF) i USA, the Deutsche Forschungsgemeinschaft (DFG) i Tyskland, och the Economic and Social Research Council (ESRC) i Storbritannien.

Test-retest effekt och intelligenstagster

Test-retest används för att kontrollera reliabilitet i tester, där förhoppningen är att testresultatet ska visa det samma oavsett om testerna utförs flera gånger. En låg test-retest effekt visar att ens mätinstrument är tillförlitligt över tid. (Hobbs, 2016). Det finns flertal studier som likt vår studie undersöker vad som sker med ens testresultat när någon gör liknande intelligenstagster vid upprepade tillfällen. Hausknecht et al., (2007) utförde en metastudie som kom fram till att det finns en så kallad test-retest effekt för intelligenstagster, även kallat kognitiva tester, när de görs fler än en gång. Det här innebär att det finns en sannolik ökning i prestation till följd av att någon genomför en viss typ av kognitivt test flera gånger. Studien visade att det fanns en tydlig förbättring mellan test 1-2, något mindre mellan 2-3 och att förbättringen därefter avtar. I en annan metaanalys har samma resultat framkommit; det finns en betydande förbättring i prestationen vid upprepade testtillfällen, men effekten avtar efter det tredje genomförda testet. Denna metaanalys såg att det fanns ett flertal variabler som påverkar hur stor test-retest effekten skulle bli. Bland annat handlar det då om intervallet mellan testerna, deltagarens ålder samt hur pass lika testerna som genomförts var (Scharfen et al., 2018).

Hausknecht et al., (2007) undersökte även vad det kan bero på att vi får en test-retest effekt, och kommer fram till flera olika aspekter. Dels kan det bero på att personerna faktiskt har övat på testerna mellan testtillfällena och därav får ett bättre resultat. Det kan också finnas andra yttre faktorer som kan spela in, bland annat testmiljön, till exempel om testet genomförts i forskningssammanhang eller i ett rekryteringssyfte, samt hur testerna utformats och med vilket intervall de gjorts.

Även Arendasy och Sommer (2017) har nått samma slutsats; Det finns en signifikant test-retest effekt vid upprepade testtillfällen. De har även undersökt vad vi kan göra för att minska test-retest effekten i den aspekt att det kan påverka utfallet av till exempel en rekrytering. Studien har kommit fram till att det i huvudsak finns två saker som påverkar storleken av test-retest effekten; hur testerna administreras samt deltagarnas generella mentala förmåga.

Samma analys visar på att administrering i form av Computerized Adaptive Tests (CAT), som innebär att testerna blir svårare eller lättare efterhand beroende på hur korrekt svar som angivits, kommer att ge en mindre test-retest effekt än helt identiska tester. Därav finns det en fördel med att använda sig av CAT-tester i syfte att minska test-retest effekten. Ytterligare resultat från studien visar på att personer med högre GMA kommer att få en större

effekt av test-retest. Analysen har även sett en skillnad i effektstorleken beroende på vilken typ av kognitivt test som administrerats. Exempelvis har forskare sett att det är skillnad om det varit rotationer, figurer eller verbala förmågor som mätts i testet. Där såg forskarna främst att verbala förmågor inte påverkas lika stort av test-retest effekten. (Arendasy & Sommer, 2017).

Syfte

Vi är intresserade av att studera om det finns en förbättring vid upprepade utföranden av intelligenstester. Syftet är således att undersöka om det finns ett samband mellan fler utförda tester och förbättring av resultatet på testet. Utifrån studiens resultat vill vi bland annat diskutera användningen av intelligenstester i rekryteringssammanhang. Studien fokuserar inte på hur väl deltagarna presterar på testen utan på om en person kan förbättra sitt resultat på dessa tester med hjälp av så kallad test-retest effekten.

Metod

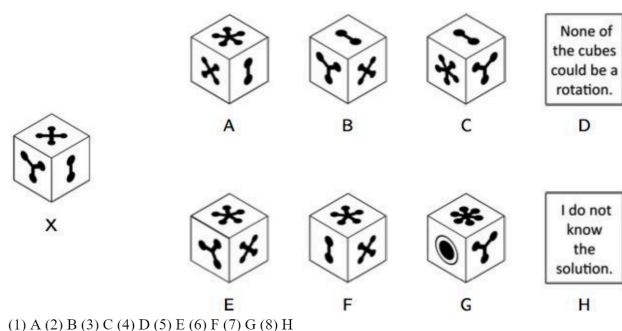
Urval

Forskningen är baserad på ett bekvämlighetsurval som har genomförts genom att författarna etablerat kontakt med sina befintliga sociala nätverk via olika sociala medieplattformar, inklusive Facebook, Instagram och LinkedIn. Utöver det gjordes e-postutskick till olika kontakter, såsom kollegor, professorer och vänner. Det tillämpades inga specifika restriktioner för deltagande i studien, eftersom det övergripande målet var att uppnå en så omfattande urvalsgrupp som möjligt, både vad gäller antal och demografisk mångfald. För att ytterligare diversifiera urvalet utformades enkäten och utskicken på engelska, vilket möjliggjorde en bredare nåbarhet gentemot olika nationaliteter. Sammanlagt deltog 162 personer i någon del av studien. Av dessa genomförde 92 deltagare den första testfasen, medan 52 personer slutförde samtliga fyra delar av studien.

Form och design

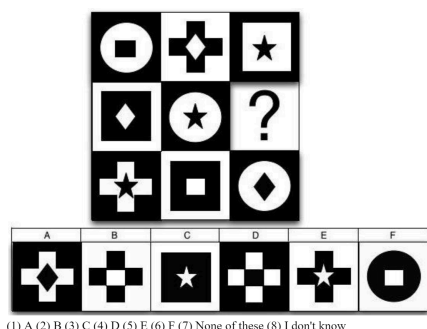
Till forskningen valdes en kvantitativ design med syfte att samla in en omfattande mängd datapunkter och därigenom möjliggöra generalisering av slutsatser kring fenomenet test-retest effekt, vilket innebär en förbättring av resultatet vid upprepade testtillfällen. För att genomföra studien konstruerades en internetbaserad enkät med hjälp av Google Forms. Enkäten bestod av fyra separata formulär, varav ett innehöll inledande frågor om deltagarnas demografi, inklusive ålder, utbildningsnivå och tidigare erfarenhet av intelligenstester. De tre återstående formulären innehöll separata intelligenstester. För att stärka validiteten inkluderades 21 befintliga frågor från ICAR 3D och 9 frågor från ICAR matrices i dessa tester, då dessa frågor tidigare har bevisats vara pålitliga (Kristjánsdóttir et al., 2023). Därefter konstruerades tre separata tester, vardera bestående av tio items. För en visuell illustration av dessa items hänvisar vi till figur 1 och 2.

Figur 1 - exempel på ICAR 3D



ICAR 3D items bygger på mentala rotationer där testdeltagaren får en bild på en tärning med olika symboler och sex olika alternativ på bilder på en annan tärning. Målet är att personen som genomför testet ska lista ut vilket alternativ som kan vara en rotation av den första tärningen. Utöver de 6 alternativen på tärning finns även alternativen "Ingen kuberna kan vara en rotation" samt "Jag vet inte vilket svar som stämmer".

Figur 2 - exempel på ICAR matrices



ICAR matrices består av figurer där deltagaren får se ett mönster i ett rutnät med 8 figurer och ska sedan räkna ut vilket av 6 givna alternativ som stämmer in som figur 9 i mönstret. Här finns alternativ för "Ingen av dessa stämmer" och "Jag vet inte vilken som stämmer".

Vardera test bestod av 7 items från ICAR 3D och 3 items från ICAR matrices. Varken rotationerna eller figurerna var progressiva utan följde ungefär samma svårighetsgrad genom hela enkäten och alla tre tester. Ordningen av items randomiserades för deltagarna i syfte att blanda mellan rotationer och figurer. Varje test inleddes med att deltagarna fick uppge sitt alias, och i slutet av varje test fanns även några avslutande frågor. De avslutande frågorna handlade om den upplevda svårigheten på testerna och hur pass dedikerad individen var till att prestera på testet. I de två senare testerna efterfrågades huruvida deltagaren upplevde skillnad på svårighetsgrad mellan de olika testerna. Varje test avslutades dessutom med en tacktext med externa länkar för att uppmuntra om vidare läsning kring intelligens och intelligenstestning. Här förtydligades även instruktioner och deltagarna fick länkar till nästkommande test (se i Appendix).

Insamling av data

Sju personer fick pilottesta undersökningen, varpå de återkom med feedback som korrigerades. Efter det inleddes insamlingen av materialet som ligger till grund för vår studie. Enkäterna låg ute i två och en halv vecka, och under denna perioden gjorde flera olika körningar på de olika kanalerna för att nå ut till så många personer som möjligt. I samband med det fick de som påbörjat testerna också en påminnelse om att slutföra de kvarstående testerna.

Deltagarna fick tillgång till länkarna till alla 4 tester direkt, men uppmuntrades att åtminstone vänta en timme mellan varje test. I övrigt fick deltagarna själva välja när de ville genomföra de olika testerna. Varje deltagare uppmuntrades att ta som mest 15 minuter på sig för att genomföra varje enskilt test. Den första, inledande delen, bestod framförallt av information kring studien och viss övergripande information om den deltagande. Det första formuläret hade också en del som handlade om personlighet (se appendix), där deltagarna fick skatta sig själva utifrån olika frågor baserat på personlighetsdrag från the Big Five i syfte att se en starkare eller svagare test-retest effekt baserat på olika personlighetsdrag. Detta valdes dock att inte tas med vidare i studien då det inte fanns signifikanta samband mellan personlighet och test-retest effekt kopplat till frågeställningen. I den inledande delen fick deltagarna dessutom chansen att uppge sin mailadress för att få en påminnelse om att slutföra samtliga tester. Till dessa personer skickades det ut påminnelser en vecka innan slutdatum samt när det var två dagar kvar.

Deltagandet i studien var helt anonymt. För att kunna se progressionen mellan de olika testerna och kunna koppla detta till vardera deltagare, samt för att se vilka som föll bort under tiden, fick deltagarna skapa ett alias som de skulle uppge i början av varje test. Genom detta alias kunde deltagarna behålla sin anonymitet men samtidigt kopplas ihop mellan vardera test.

Etik

Innan studien genomfördes gjordes en etik-deklaration som undertecknades av oss och vår handledare. Detta gjordes för att studien skulle följa lagen (2003:460) om etikprövning av forskning som avser människor.

Studien inleddes med en grundläggande förklaring av syftet samt instruktioner för hur deltagaren förväntas genomföra studien. Viss information om syftet som handlar om prestationsförbättring mellan testerna utelämnades dock för att undvika demand karakteristiska hos deltagarna, då det annars finns en risk i att deltagarna medvetet kunde prestera sämre på de tidigare testerna och bättre på de senare om de visste med sig från början att detta var syftet med studien. När samtliga tre tester var genomförda fick däremot deltagaren information om att detta var en del av syftet.

Deltagandet var anonymt, helt frivilligt och samtliga deltagare kunde när som helst välja att inte fortsätta med testerna. Detta resulterade i en del bortfall: 64% bortfall jämfört med den första och den sista enkäten, och ca 40% mellan intelligentest 1 och 3. För de

deltagarna som valde att uppge mailadress för att bli påmind om att genomföra hela studien förblev deltagandet inte anonymt för oss, men deras resultat är fortsatt anonymt i studien. Detta blev det också informerade om i samband med att de uppgav sin mailadress.

Analys av data

Första steget var att skapa kodning av datan för att lättare kunna skapa ett dataset till Jamovi. Fullständig kodning finns i appendixen. Här framtog det också vilka alias som genomfört vilka av testerna för att kunna få fram och analysera bortfallet. Därefter skapades ett data set som sedan infördes till Jamovi där analyserna genomförts. Analyserna gjordes utifrån statistiskt vedertagna metoder (Navarro & Foxcroft, 2022)

Vi skapade även en del egna variabler utifrån datan. Bland annat togs det fram en variabel för förbättringen mellan de olika testerna, vilka benämns imp1-2, imp1-3 samt imp2-3 (se Appendix för fullständig kodning). Dessa skapades genom att ta deltagarens resultat från test 1 och subtrahera med resultatet från test 2 för att skapa ett mått på hur mycket populationen samt den enskilda individen förbättrat eller försämrat sitt resultat jämfört med sina tidigare tester. Vidare delades rotationer och figurer upp i två olika delar, och vi samlade ihop medelvärden för dessa för att kunna se på hur prestationen har skiljt sig baserat på typen av item. Vi skapade descriptives för att få fram övergripande data över medelvärde, antal och uppdelningar mellan kön och ålder. Därefter gjordes ett paired sample t-test för att se skillnader i medelvärde, samt en repeated measure ANOVA för att analysera variansen mellan testerna tillsammans med det ett post hoc-test. För att analysera bortfall framställdes deskriptiv statistik över hur många som fallit bort under studien, och kopplades ihop med graden av upplevd svårighet som deltagarna fått självskatta.

Resultat

Bortfallsanalys

Totalt deltog 162 personer i studien och 52 personer slutförde samtliga enkäter. Innan studien hade vi räknat med ett stort bortfall men vi ville undersöka om det fanns en gemensam nämnare som kunde förklara vilka personer som föll bort. Samtliga intelligenstester inkluderade en självskattad fråga om upplevd svårighetsgrad på en skala 1-5.

Tabell 1 visar fördelningen på upplevd svårighetsgrad på det första testet samt hur bortfallet såg ut kopplat till vad deltagaren svarade på frågan. Det var få som svarade 1-2 ($n = 4$) på upplevd svårighetsgrad men samtliga av dem genomförde samtliga tester. Bland de som svarade 3-5 på upplevd svårighetsgrad ($n = 87$) ser vi ett bortfall på 39 individer, vilket motsvarar 45%.

Tabell 1
Deskriptiv statistik

	Totalt	Upplevd svårighetsgrad	Test 1	Test 2	Test 3
Antal deltagare	161		92	61	55
		1	1	1	1
		2	3	3	3
		3	26	16	15
		4	44	27	23
		5	17	11	10
Bortfall totalt			69	100	106
Bortfall från test 1		1	0	0	0
		2	0	0	0
		3	0	10	11
		4	0	17	21
		5	0	6	7
Medelvärde			4.37	6.21	5.93
Median			4.0	7.0	6.0
Standardavvikelse			2.74	3.04	3.23

Tabellen visar deskriptiv statistik över samtliga testtillfällen.

Notering: Test 1 är det första intelligenstestet. Test 2 är det andra intelligenstestet. Test 3 är det tredje intelligenstestet. Samtliga tester har poäng 1-10. Upplevd svårighetsgrad är en självskattnings skala 1-5 hur svårt deltagaren upplevde det första intelligenstestet.

Översikt resultat

Tabell 1 ger en översikt över studiens resultat och visar att medelvärdet för testresultat ökar från det första till andra och tredje intelligenstestet. Mellan det andra och tredje testet sänks medelvärdet från 6.21 till 5.93. Ökningen från det första till det andra är 1.84 vilket motsvarar 42% av medelvärdet i test 1. Att vi ser en ökning från det första testet kan indikera att det finns en test-retest effekt i intelligenstester.

Skillnad mellan testtillfällen

För att undersöka skillnaden mellan testomgångarna gjordes ett paired samples t-test (resultat i tabell 2) som undersöker skillnader i medelvärden. Resultatet av t-testet visar en signifikant höjning mellan från det första intelligenstestet till det andra ($t=7.28, p<.001$) och tredje ($t=3.97, p<.001$). Precis som i tabell 1 ser vi en att medelvärdet minskar mellan det andra och tredje testet men denna skillnaden är inte signifikant. Testet visar att skillnaden i medelvärdet mellan test 1-2 är större än mellan test 1-3. Det vill säga att förbättringen är som störst mellan test 1-2. Detta stärks av en högre effektstorlek mellan test 1 och 2 (Cohen's $D=.956$) som indikerar att deltagarna i snitt höjer sig med nästan en hel standardavvikelse. Mellan test 1 och 3 är effektstorleken medelstor (Cohen's $D=.55$) och indikerar en höjning på lite mer än en halv standardavvikelse.

I samband med t-testet undersöktes om antagandena för testet var uppfyllda. Antagandet om oberoende mätningar uppfylls eftersom varje persons svar är oberoende av vad någon annan svarat. Deltagarna har inte fått se eller påverkats av någon annans svar. Antagandet om varianshomogenitet uppfylls också eftersom urvalet är detsamma för samtliga testtillfällen. Antagandet om normalitet undersöktes genom ett Shapiro-Wilks test. Shapiro-Wilks testet visade att normalitet endast förekom mellan test 1 och test 3.

Tabell 2**Test-retest effekt över tre mätningar**

		T-värde	P-värde*	Medelvärdes- skillnad	Standardfel	Effektstorlek
Test 1	Test 2	-7.28	<.001	-1.724	0.237	-0.956
	Test 3	-3.97	<.001	-1.442	0.364	-0.550
Test 2		1.48	0.144	0.364	0.245	0.200

Testet visar skillnaden i medelvärdet mellan de olika testtillfällena. Värdet "statistik" visar hur stor skillnaden är.

Notering: Test 1 är deltagarnas poäng 1-10 på det första intelligenstestet. Test 2 är deltagarnas poäng 1-10 på det andra intelligenstestet. Test 3 är deltagarnas poäng 1-10 på det tredje testet.

Effektstorlek är Cohen's D.

* $p < .005$ = signifikant

Utöver t-tester genomfördes en repeated measures ANOVA samt post hoc jämförelser som ett sätt att undersöka alla skillnaderna i resultat mellan testtillfällena. Resultatet (se tabell 3) stärker hypotesen att det finns en test retest effekt i intelligenstester då vi ser en signifikant skillnad mellan testtillfällena ($F=20.3, p=.001$). Ur tabellen framgår det även en medelstor effektstorlek ($\eta^2=.06$). Post hoc jämförelsen ger ett liknande utfall som de initiala t-testerna men visar att resultatet fortfarande är signifikant (Tukey $<.001$) mellan test 1-2 samt 1-3 efter en Tukey-korrektion. I samband med testet undersöktes om antagandena för en repeated measures ANOVA var uppfyllda. Antagandet om oberoende mätningar uppfylls av samma anledning som för t-testet. Antagandena om normalitet och sfäricitet (se Sfäritetstest i Appendix) uppfylls också.

Tabell 3**En variansanalys av test-retest effekten**

		F-värde	P-värde*	Effektstorlek η^2
Testresultat		20.3	<.001	0.060
Post Hoc Jämförelse				
		Medelvärdes skillnad	Standardfel	Ptukey**
Test 1	Test 2	-1.769	0.251	<.001
	Test 3	-1.442	0.364	<.001
Test 2	Test 3	0.327	0.257	0.418

Tabellen visar resultatet från en repeated measures ANOVA samt ett Post Hoc-test.

Notering: Testresultat är baserat på samtliga intelligenstester. Test 1 är deltagarnas poäng 1-10 på det första testet. Test 2 är deltagarnas poäng 1-10 på det andra testet. Test 3 är deltagarnas poäng 1-10 på det tredje testet.

* $P < .005$ = signifikant

** $P_{tukey} < .005$ = signifikant

Takeffekt

För att undersöka om vi fått en takeffekt skapades en frekvensstabell som visar hur vanligt förekommande vardera poäng, 1-10, var på samtliga test. Typvärdet för test 1 är tre och visar att en majoritet av deltagarna (62%) fick fyra eller färre poäng vilket är lägre än medelvärdet på test 1 (4.37) som visas i tabell 1. I test 2 var spridningen större och typvärdet för testet var tio vilket kan indikera att vi fått en takeffekt då många deltagare fick full poäng och därför inte kunde höja sig till test 3.

Tabell 4
Frekvenser för testresultaten

Antal rätt	Test 1	Procent	Test 2	Procent	Test 3	Procent
0	1	1.1%	2	3.3%	3	5.5%
1	13	14.1%	1	1.6%	4	7.3%
2	11	12%	7	11.5%	5	9.1%
3	18	19.6%	4	6.6%	3	5.5%
4	14	15.2%	7	11.5%	3	5.5%
5	7	7.6%	3	4.9%	4	7.3%
6	8	8.7%	4	6.6%	8	14.5%
7	3	3.3%	10	16.4%	2	3.6%
8	5	5.4%	4	6.6%	5	9.1%
9	7	7.6%	7	11.5%	11	20%
10	5	5.4%	12	19.7%	7	12.7%

Tabellen visar uppdelningen av hur många antal rätt deltagarna haft på vardera test, samt en procentuell uppdelning jämfört med det totala antalet deltagare för varje test.

Notering: Test 1 visar fördelningen över hur många deltagare som fått ett visst totalpoäng på det första testet. Test 2 visar fördelningen över hur många deltagare som fått ett visst totalpoäng på det andra testet. Test 3 visar fördelningen över hur många deltagare som fått ett visst totalpoäng på det tredje testet.

Diskussion

Resultatet av studien visar att det finns en förbättring vid upprepade genomföranden av intelligenstester, och då framförallt mellan test 1 och 2. Störst skillnad är mellan test 1 och 2 men vi ser även att deltagarnas resultat förbättras signifikant mellan test 1 och 3. Resultatet går i linje med tidigare forskning (Calamia et al., 2012. Hausknecht et al., 2007. Scharfen et al., 2018), vilket styrker utfallet av studien. Till skillnad från samma tidigare forskning visar resultatet inte en signifikant skillnad mellan det andra och tredje testet.

Mellan första och andra testet får vi en hög effektstorlek i T-testet. Cohen's D på .96 antyder att deltagarna i snitt höjt sig med 96% av en standardavvikelse. Mellan första och

sista testet är Cohen's D .55 vilket innebär att deltagarna i snitt höjer sig med 55% av en standardavvikelse. Det är värt att nämna att *t*-testet mellan test 1 och 2 bröt mot antagandet om normalitet. Eftersom *t*-test är robusta kan vi fortfarande använda resultaten även om antagandet om normalitet inte är uppfyllt (Edgel & Noon, 1984) men det är viktigt att ha med sig att resultatet kan ha påverkats av att antagandet inte är helt uppfyllt. Trots att antagandet inte uppfylls i *t*-testet ser vi att resultatet blir signifikant. Dessutom är resultaten i samtliga *t*-tester i linje med Post Hoc testerna och antyder en signifikant skillnad mellan test 1 och 2 samt 1 och 3. De resultat vi fick från *t*-testerna om effektstorlek stämmer även överens med de resultat som framgick från tidigare presenterade metaanalyser (Calamia et al., 2012. Hausknecht et al., 2007. Scharfen et al., 2018).

Implikationer för urvalsprocesser

Resultatet av vår studie styrker att det finns en förbättring, det vill säga en test-retest effekt, som innebär att personer som gör liknande intelligenstester vid upprepade tillfällen generellt presterar bättre på det andra testet. Tidigare metaanalyser har visat att deltagarna även brukar förbättras mellan andra och tredje testet (Calamia et al., 2012. Hausknecht et al., 2007. Scharfen et al., 2018) men att förbättringen minskar desto fler tester individen gör. Det går alltså att förbättra sina resultat genom träning men att individen når en maxnivå efter ett fåtal tester och att det därefter är svårt att signifikant förbättra sitt resultat.

Utifrån resultaten i studien tillsammans med tidigare forskning kan vi argumentera för att det finns en fördel med att låta kandidater öva på intelligenstester innan det slutgiltiga testet som ska ligga till grund för en urvalsprocess. Att låta kandidater bekanta sig med testerna på förhand kan ge kandidaterna mer rättvisa förutsättningar att nå sin maxnivå vilket i sin tur kan generera en mer rättvis bedömning av deras faktiska nivå av intelligens och sedermera förväntad arbetsinsats. Eftersom skillnaden är störst mellan första och andra testet anser vi att ett träningstest kan vara ett sätt att ge kandidater mer rättvisa förutsättningar oavsett vilken tidigare erfarenhet kandidaterna har av intelligenstester.

En risk med att inkludera övningstester i urvalsprocesser är att företag eller andra institutioner riskerar att förlora kandidater som anser att denna typ av tester är för energikrävande. I avsnittet om bortfall diskuterade vi att många deltagare upplevde testerna som svåra och jobbiga. I en urvalsprocess kan det finnas en risk för ett liknande bortfall då kandidater som upplever processen som lång, svår eller jobbig, kanske drar sig från att söka en tjänst.

Begränsningar

Att vi i studien inte får ett signifikant resultat mellan test 2 och 3 kan ha flera olika orsaker. En anledning kan vara att vi får en takeffekt från test 2. En takeffekt uppstår om ett test är utformat på ett sådant sätt att det är för lätt att få full poäng (Liu & Wang, 2020). Typvärdet för resultatet på det andra intelligenstestet är 10 rätt av 10 och 20% av deltagarna fick full poäng på testet (se tabell 4). Detta innebär att en stor del av deltagarna inte har möjlighet att förbättra sig till det tredje testet.

En annan potentiell anledning till att vi inte nådde signifikans i resultat mellan test 2 och 3 kan ha att göra med urvalet och hur studien var utformad. Studiens designades på ett sätt som ställde krav på deltagarna att lägga ner cirka 50-60 minuter fördelat på ett inledande test och tre intelligenstester. Tiden som krävdes för att slutföra studien kan ha lett till att personer utmattades över testperioden och presterade sämre i sista testet. I slutet av studien fick deltagarna möjlighet att ge oss skriftlig feedback på upplevelsen av att delta i studien. I feedbacken nämnde en del av personerna att testerna upplevts som svåra och jobbiga samt tagit lång tid. Ett av våra alias skriver: "In the beginning I was very dedicated to get good results, but I found it very hard and lost some of my ambition by the second test." Och en annan skriver: "You get tired along the road. It does not take 15 min."

Bortfall

Vi ser ett relativt stort bortfall i studien, framförallt mellan test 1 och 2. Test 1 genomfördes av 92 individer medan 61 individer slutförde test 2. Det sista testet gjordes av 55 individer. Det var även några deltagare som missade första testet men gjorde andra eller tredje. Totalt genomförde 52 individer samtliga 3 intelligenstester.

Bortfallet mellan testerna kan bero på olika aspekter men en anledning skulle kunna vara tiden som krävdes för att genomföra testet. I avsnittet om begränsningar diskuteras att det krävdes 50-60 minuter per deltagare för att genomföra hela studien. I den skriftliga feedbacken nämner deltagare att studien tog tid, var svårt och att det var lätt att tappa motivationen under tiden (se citat i avsnittet om begränsningar).

En annan förklaring till bortfallet kan bero på hur svåra testerna upplevdes. Tabell 1 i resultatet visar att bortfallet var 0% bland de som svarat 1 eller 2 på upplevd svårighetsgrad på första testet. Bortfallet bland de som svarat 3-5 är betydligt högre (45%). Värt att nämna är att det var relativt få som svarade 1 eller 2 på upplevd svårighetsgrad. Vi ser alltså att de flesta deltagarna upplevde att testet var relativt svårt (3+), vilket kan förklara en del av

bortfallet. Detta bekräftas av den feedback vi fått från deltagare där en del skrivit att testet upplevdes som svårt och tog lång tid (se citat under avsnittet begränsningar i resultatet).

Förslag till vidare forskning

Som nämnt har vi fått signifikanta resultat från studien, vilka dessutom har stärkts från tidigare forskning. Resultatet blev som förväntat och syftet med studien gällande att undersöka om det finns en förbättring, en test-retest effekt, har således uppfyllts. Vid analysen av vår data uppkom en del andra intressanta områden som vi hade kunnat utforska vidare om det funnits utrymme för det. Bland annat hade det varit intressant att se hur test-retest effekten skiljer sig mellan kvinnor och män, men också generellt hur kvinnor och män skiljer sig prestationsmässigt på intelligenstester. Att djupare dyka ner i detta hade varit intressant för att undersöka hur aktiva inom rekryteringsbranschen kan optimera användandet av tester på ett så jämlikt sätt som möjligt mellan könen, eller för att se hur det faktiskt skiljer sig mellan kön vad gäller test-retest effekt. I vårt resultat kunde vi inte se att det fanns några signifikanta skillnader när vi gjorde uppdelningarna i kön, och därav kunde vi inte dra några slutsatser kring detta. I en studie med annat fokus finns det eventuellt intressanta slutsatser att dra kring detta.

För att bygga vidare på vår studie hade det varit önskvärt att utforma en studie med ett större spann av frågor med större kontroll över externa variabler såsom hur och vart deltagarna genomför testerna. Detta skulle kunna bidra till mer signifikanta resultat då en sådan studie skulle begränsa antalet felkällor. Således hade det varit möjligt att undvika den takeffekt vi förmodar att vi fått i vår studie då spannet för förbättring hade varit större. Utöver det här hade studiens chanser att finna signifikanta resultat samt validitet kunnat höjas genom att ta hjälp av fler deltagare och en större spridning på svårighetsgrad av items.

En annan aspekt som är intressant att undersöka i framtida forskning är huruvida det finns kopplingar mellan Big five personlighetsfaktorer och test-retest effekt vid intelligenstester. Det hade varit intressant att undersöka korrelationer mellan olika personlighetstyper och olika grad av förbättring mellan testtillfällen på intelligenstester. Detta skulle möjliggöra en ökad förståelse av om eller hur personlighet och intelligens samvarierar vilket är relevant då personlighets- och intelligenstester ofta förekommer tillsammans i urvalsprocesser.

Slutsats

Då vi nått slutsatsen om att det finns en test-retest effekt på intelligenstester kan vi observera en förbättring av deltagarnas resultat från det första testet till senare. Denna förbättring avtar efter att deltagaren genomför fler tester vilket resulterar i att en kandidat som aldrig tidigare gjort ett intelligenstest har en sämre förutsättning att prestera till sin potential på testet och av den anledningen kan minska sin chans att komma vidare i en urvalsprocess. Därav finns det ett syfte med att låta kandidater utföra intelligenstester flertalet gånger eller att förse kandidater med övningstester inför ett intelligenstest för att höja reliabiliteten av dessa tester. Detta skapar en mer rättvis bild av kandidaternas nivå av intelligens, vilken i sin tur kan användas för att predicera framtida arbetsinsats vilket är anledningen till att testet används.

Referenser

- Andersson, G., Hallén, N. & Smith, P. (2016). *Rekrytering och urval - Teori och praktik, 1e*. Studentlitteratur.
ISBN: 9789144107851
- Arendasy, M. E. & Sommer, M. (2017). Reducing the effect size of the retest effect: Examining different approaches. *Intelligence*, 62, 89–98.
<https://doi.org/10.1016/j.intell.2017.03.003>
- Calamia, M., Markon, K. & Tranel, D. (2012). Scoring Higher the Second Time Around: Meta-Analyses of Practice Effects in Neuropsychological Assessment. *The Clinical Neuropsychologist*, 26(4), 543–570. <https://doi.org/10.1080/13854046.2012.680913>
- Cliffordson, C. (2004). Effects of Practice and Intellectual Growth on Performance on the Swedish Scholastic Aptitude Test (SweSAT). *European Journal of Psychological Assessment*, 20(3), 192–204.
<https://doi.org/10.1027/1015-5759.20.3.192>
- Edgell, S. E. & Noon, S. M. (1984). Effect of violation of normality on the *t* test of the correlation coefficient. *Psychological Bulletin*, 95(3), 576–583.
<https://doi.org/10.1037/0033-2909.95.3.576>
- Hausknecht, J. P. Halpert, J. A., Di Paolo, N. T. & Moriarty Gerrard, M. O. (2007). Retesting in selection: A meta-analysis of coaching and practice effects for tests of cognitive ability. *Journal of Applied Psychology*, 92(2), 373–385.
<https://doi.org/10.1037/0021-9010.92.2.373>
- Hobbs, M. (2016, September 15). *What is test-retest reliability and why is it important?*
<https://cambridgecognition.com/what-is-test-retest-reliability-and-why-is-it-important/>
- Holt, N., Bremner, A., Sutherland, E., Vliek, M., Passer, M. & Smith, R. (2019). *Psychology: The Science of Mind and Behaviour, 4e*. McGraw-Hill Education.
ISBN: 9781526846983
- Kristjánsdóttir, D., Zaiter, A. & Kajonius, P. (2023). Public Domain Intelligence Tests: Psychometric properties of the Cog15 and ICAR16 cognitive ability scales. *In Press*
<https://lup.lub.lu.se/student-papers/record/9125503>

- Lievens, F., Buyse, T. & Sackett, P. R. (2005) Retest effects un operational selection settings: development and test of a framework. *Personnel Psychology*, 58(4), 981–1007.
<https://doi.org/10.1111/j.1744-6570.2005.00713.x>
- Lindelöw Danielsson, M. (2003). *Kompetensbaserad rekrytering, intervjuteknik och testning*. Natur & Kultur Akademisk.
ISBN: 9789127094437
- Liu, Q. & Wang, L. (2020). T-Test and ANOVA for data with ceiling and/or floor effects. *Behavior Research Methods*, 53.
<https://doi.org/10.3758/s13428-020-01407-2>
- Navarro, DJ. & Foxcroft, DR. (2022). Learning statistics with Jamovi: a tutorial for psychology students and other beginners. (Version 0.75).
<https://doi.org/10.24384/hgc3-7p15>
- Scharfen, J., Peters, J. M. & Holling, H. (2018). Retest effects in cognitive ability tests: A meta-analysis. *Intelligence*, 67, 44–66.
<https://doi.org/10.1016/j.intell.2018.01.003>
- Schmidt, F. & Hunter, J. (1998) The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings. *Psychological Bulletin*, 124, 262-274.
<https://doi.org/10.1037/0033-2909.124.2.262>
- Sternberg, R. J., Lautrey, J. & Lubart, T. I. (2003). *Models of intelligence: International perspectives*. Washington, DC: American Psychological Association.
ISBN: 9781557989710
- The Project - International Cognitive Ability Resource - The ICAR Project*. (2014).
Icar-Project.com. <https://icar-project.com/>

Appendix

Bilaga 1 - testning av antaganden

Tabell 5
Normalitetstest (Shapiro-Wilk)

		W	P-värde*
Test 1	Test 2	0.950	0.017
Test 1	Test 3	0.970	0.211
Test 2	Test 3	0.905	<.001

Notering: ett lågt p-värde indikerar att antagandet om normalitet är brutet

Tabell 6
Sfäritetstest

	Greenhouse-Geisser	P-värde*
Test resultat	0.788	<.001

Notering: Sfäritetstest för ANOVA uppfylls då greenhouse är mer än 0.75.

Bilaga 2 - Fullständig kodning

Rot - Rotation

Fig - Figur

Siffror #.# där första siffran är vilket Cog # och andra siffran är vilken fråga i ordningen.

Rot 1.1 = Första rotationsfrågan i Cog #1

Rot 1.3 = Tredje rotationsfrågan i Cog #1

Rot 2.1 = Första rotationsfrågan i Cog #2

Fig 1.1 = Första figuren i Cog #1

Fig 1.3 = Tredje Figuren i Cog #1

Fig 2.1 = Första Figuren i Cog #2

Score 1 - Resultat Cognitive #1

Score 2 - Resultat Cognitive #2

Score 3 - Resultat Cognitive #3

Specifika Scores

RotScore1 - Total poäng på rotationsfrågorna i Cog #1

FigScore1 - Total poäng på figurfrågorna i Cog #1

Avg över alla tester

RotAvg - Genomsnittligt poäng på rotationsfrågorna i genomförda Cog-Profiles.

FigAvg - Genomsnittligt poäng på figurfrågorna i genomförda Cog-Profiles.

ScoreAvg - Genomsnittlig totalpoäng på genomförda Cog-Profiles.

RotPercentScoreAvg - I snitt hur många procent av personens totala genomsnittliga poäng kommer från rotationsfrågor. **“Viktning 70% om man är lika bra på båda”**.

Imp - Improvement. Eventuell förbättring mellan tester.

Ex: **Imp1-2** Förbättring Från Cog #1-Cog #2. Negativt värde visar en försämring.

Uträkning ex: Score 2-Score 1 ger **Imp 1-2**

Difficulty # - To what extent did you feel the puzzles were hard?

DifAvg - Average of difficulty response

DifComp #- How would you compare the level of difficulty between this profile and the previous one?

Dedication # - How dedicated were you to get a good result on the puzzles?

DedAvg - Average of dedication response

Exp # - How was your overall experience of taking part in this study?

Efter dessa indikerar vilket quiz man har svarat på frågan i.

Feedback - Do you have any final thoughts or feedback that you want to give us, now that you have finished the study?

Personlighet

Neu# - Neuroticism och vilket nummer

Con# - Consciousness och vilket nummer

Ope# - Openness och vilket nummer

Agr# - Agreeableness och vilket nummer

Ext# - Extroversion och vilket nummer

Angående negativ Laddning

ER = Ej reversed. Gäller för de med negativ laddning.

Ex Neu1ER = Neurotism 1 utan reversead skala. Sen blir;

Neu1 Det faktiska reversade värdet som vi använder för att räkna ut medelvärde.

Avg - Average. Medelvärde för en faktor. Ex NeuAvg = Average för neurotism på en person.

Gender

1=Man

2=Kvinna

Age

0=<16

1=16-19

2=20-29

3=30-44

4=45-59

5=60+

Education - What is your highest level of completed education?

0=Primary education

1=Secondary education

2=Bachelors degree

3=Masters degree

4=Doctoral degree

5=No completed education

Previous - Have you previously completed any cognitive test/IQ-test, for example in a recruitment process? If so, roughly how many times?

0=Never

1=1-2 times

2=2-4 times

3=5+

4=I do not know

HPScore - Högsta uppnådda poäng om man har skrivit högskoleprovet.

0=I have not taken

1=0-0.4

2=0.45-0.85

3=0.9-1.3

4=1.35-1.75

5=1.8-2.0

6=I do not remember

EngProf - English proficiency. Hur bra man är på engelska

Bilaga 3 - Personlighetstest

Kod	No	Fråga	Stämmer inte alls	Stämmer helt	Faktor som mäts	Laddning
Neu1	1	I keep my emotions under control	5	1	Neuroticism	-
Neu2	2	I panic easily	1	5	Neuroticism	+
Neu3	3	I rarely get irritated	5	1	Neuroticism	-
Neu4	4	I remain calm under pressure	5	1	Neuroticism	-
Neu5	5	I often feel down (blue)*	1	5	Neuroticism	+
Neu6	6	I have frequent mood swings	1	5	Neuroticism	+
Neu7	7	I am afraid of many things	1	5	Neuroticism	+
Neu8	8	I am not easily frustrated	5	1	Neuroticism	-

Kod	No	Fråga	Stämmer inte alls	Stämmer helt	Faktor	Laddning
Con1	1	I pay attention to details	1	5	Conscientiousness	+
Con2	2	I often forget to put things back in their proper place	5	1	Conscientiousness	-
Con3	3	I am often late to work	5	1	Conscientiousness	-
Con4	4	I do not plan ahead	5	1	Conscientiousness	-
Con5	5	I am careful to avoid making mistakes	1	5	Conscientiousness	+
Con6	6	I love order and regularity	1	5	Conscientiousness	+
Con7	7	I accomplish my work on time	1	5	Conscientiousness	+
Con8	8	I need a push to get started	5	1	Conscientiousness	-

Kod	No	Fråga	Stämmer inte alls	Stämmer helt	Faktor	Laddning
Ope1	1	I have a rich vocabulary	1	5	Openness (Intellectual/experience)	+
Ope2	2	I have difficulty understanding abstract ideas	5	1	Openness	-
Ope3	3	I believe in the importance of art	1	5	Openness	+
Ope4	4	I like to solve complex problems	1	5	Openness	+
Ope5	5	I get excited by new ideas	1	1	Openness	+
Ope6	6	I rarely look for deeper meaning in things	5	1	Openness	-
Ope7	7	I am not interested in theoretical discussion	5	1	Openness	-
Ope8	8	I seldom get lost in thought	5	1	Openness	-

Kod	No.	Fråga	Stämmer inte alls	Stämmer helt	Faktor som mäts	Laddning
Agr1	1	I believe that i'm better than others	5	1	Agreeableness	-
Agr2	2	I love to help others	1	5	Agreeableness	+
Agr3	3	I believe that others have good intentions	1	5	Agreeableness	+
Agr4	4	I am hard to get to know	5	1	Agreeableness	-
Agr5	5	I trust what people say	1	5	Agreeableness	+
Agr6	6	I am annoyed by others mistakes	5	1	Agreeableness	-
Agr7	7	I am inclined to forgive others	1	5	Agreeableness	+
Agr8	8	I hold grudges	5	1	Agreeableness	-

Kod	No	Fråga	stämmer inte alls	stämmer helt	faktor som mäts	Laddning
Ext1	1	I have a natural talent for influencing people	1	5	Extraversion	+
Ext2	2	I have difficulty expressing my feelings	5	1	Extraversion	-
Ext3	3	I take charge	1	5	Extraversion	+
Ext4	4	I am not good at getting people to like me	5	1	Extraversion	-
Ext5	5	I am the life of the party	1	5	Extraversion	+
Ext6	6	I wait for others to lead the way	5	1	Extraversion	-
Ext7	7	I don't mind being the center of attention	1	5	Extraversion	+
Ext8	8	I don't like to draw attention to myself	5	1	Extraversion	-

Confirmation message

Good job completing the first form! This first profile consisted of IPIP-items that regards personality mapping, based on the Big Five Theory.

If you are interested in getting to know more about this, here is a clip from nyhetsmorgon about how you can find your love partner with the big five theory:

<https://www.tv4.se/klipp/va/13319253/hitta-ratt-karlekspartner-med-big-five-teorin>

And some more interesting facts about the theory:

<https://www.youtube.com/watch?v=5HEvVzjXA94>

The next part is the first of out the three cognitive tests. Below you will find a link to the first test. If you feel like it you can take it right away.

Carefully save the link to the cognitive profile #1:

<https://forms.gle/KVYFxsqwJsRTDhEL8>

Good luck!

Bilaga 4 - Intelligenstest

Kognitivt test 1

Welcome to the first cognitive test of our study. If you have any questions please contact max@angelin.se.

If you are interested in knowing your score, you will be able to see your scores ones you have submitted each profile.

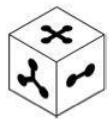

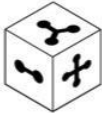


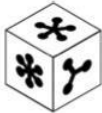

Thank you again for your participation!

The test consists of 10 questions with answer options A-H or A-F. There are two types of questions; figures and rotations. You will receive instructions on each question but the instructions will remain the same for all items of the same category.

Please notice that there is only one correct answer per item.

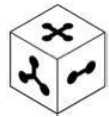
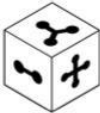




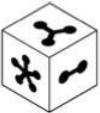
Good luck!

Rot 1.1 Which of these cubes below is a rotation of X?

 X	 A	 B	 C	None of the cubes could be a rotation.
	 E	 F	 G	I do not know the solution.

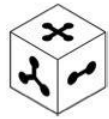

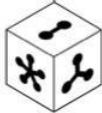




Rätt svar: F

Rot 1.2 Which of these cubes below is a rotation of X?

 X	 A	 B	 C	None of the cubes could be a rotation.
	 E	 F	 G	I do not know the solution.

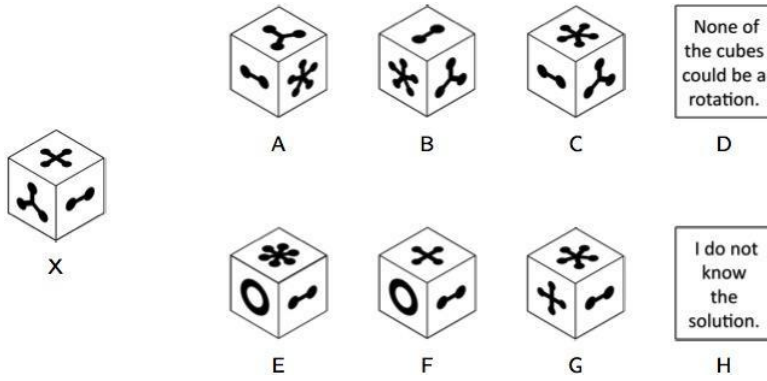
Rätt svar: F

Rot 1.3 Which of these cubes below is a rotation of X?

 X	 A	 B	 C	None of the cubes could be a rotation.
	 E	 F	 G	I do not know the solution.

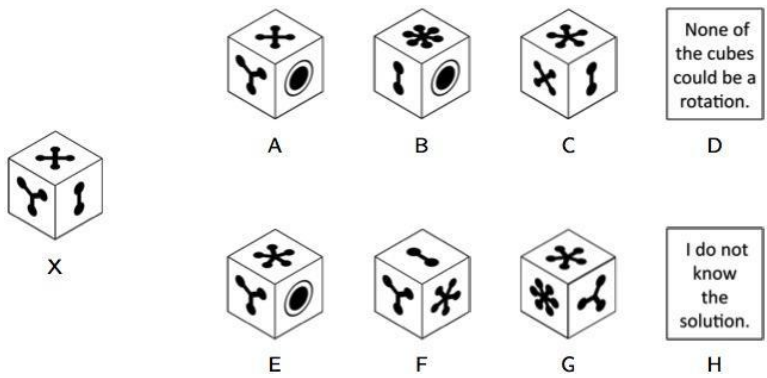
Rätt svar: A

Rot 1.4 Which of these cubes below is a rotation of X?



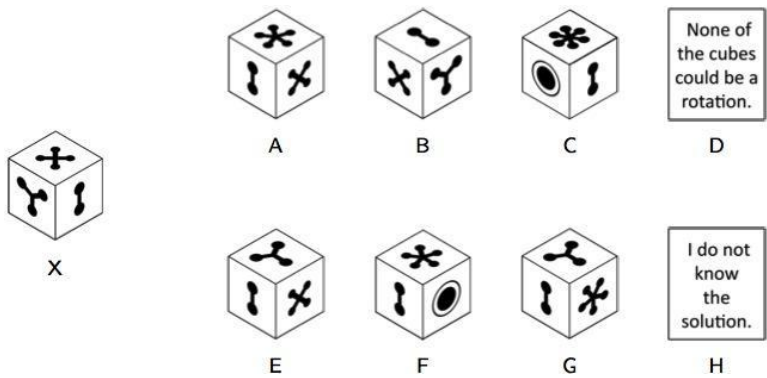
Rätt svar: E

Rot 1.5 Which of these cubes below is a rotation of X?



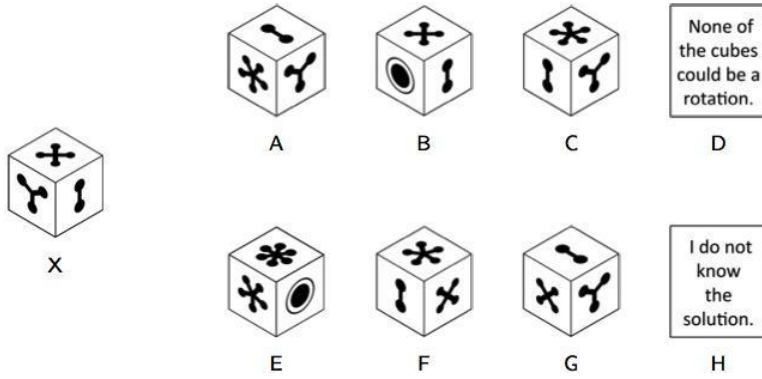
Rätt svar: G

Rot 1.6 Which of these cubes below is a rotation of X?



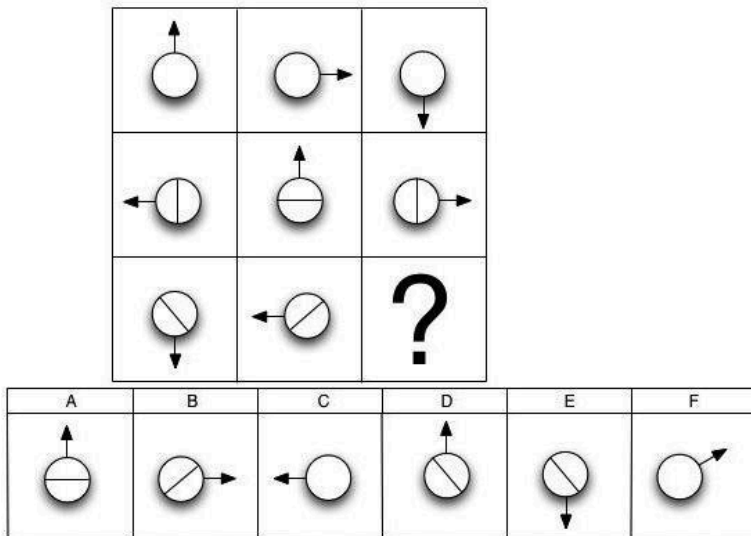
Rätt svar: C

Rot 1.7 Which of these cubes below is a rotation of X?



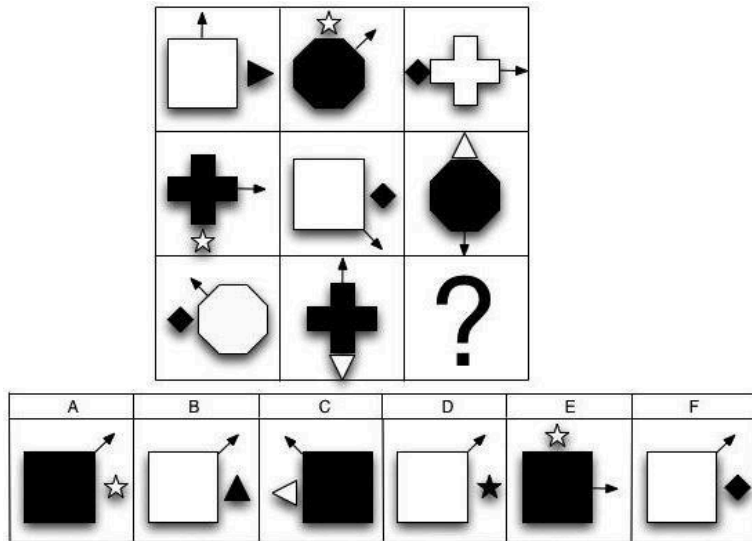
Rätt svar: E

Fig 1.1 Which alternative fits best in the question mark below (?)



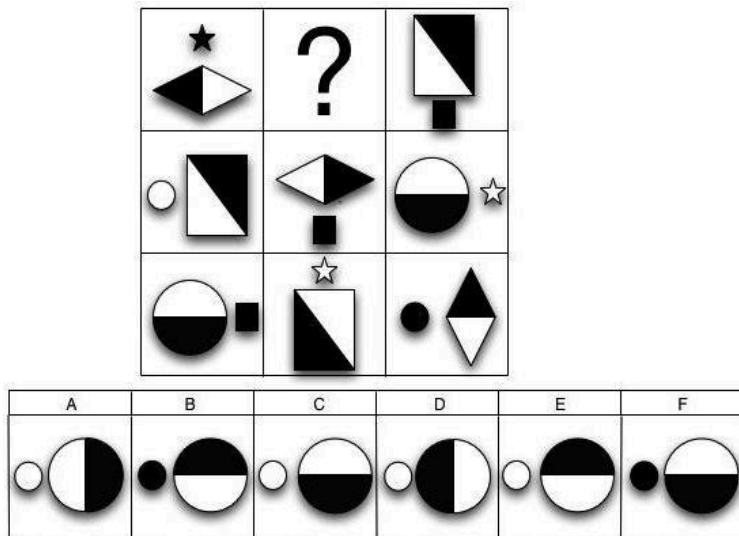
Rätt svar: D

Fig 1.2 Which alternative fits best in the question mark below (?)



Rätt svar: D

Fig 1.3 Which alternative fits best in the question mark below (?)



Rätt Svar: E

Closing questions

To what extent did you feel the puzzles were hard?

- Very easy
-
-
-
- Very hard

How dedicated were you to get a good result on the puzzles?

- Not at all dedicated
-
-
-
- Very dedicated

Confirmation message

Thank you for taking your time to complete the first profile. We hope it went well and that you enjoyed it.

Like we previously stated you can, when completed all three tests, receive your results from us. If you forgot to enter your email or changed your mind about wanting to know how you scored, you can always contact us afterwards to receive your result.

If you are interested in comparing your score to the population or want to find out what percentile your score would put you in you can read more about IQ and cognitive tests on Mensa's website.

<https://shorturl.at/gipFO>

Below you can find the link to cognitive test #2. Please take a break of at least an hour before completing this next test, (well deserved and better for the result). But remember to complete all three profiles before november 10th.

And once again, thank you!

Carefully save the link to Cognitive profile #2:

<https://forms.gle/2HFzYmnXtxsNAcCL8>

Kognitivt test 2

Welcome to the second cognitive test of our study.

If you are interested in knowing your score, you will be able to see your scores ones you have submitted each profile.

Thank you again for your participation!

Rot 2.1 Which of these cubes below is a rotation of X?

A 3x3 grid of cube options for Rot 2.1. The first column contains cube X and its label 'X'. The second and third columns contain cubes A, B, C, E, F, G and their labels. The fourth column contains two text boxes: 'None of the cubes could be a rotation.' (D) and 'I do not know the solution.' (H).

Rätt svar: E

Rot 2.2 Which of these cubes below is a rotation of X?

A 3x3 grid of cube options for Rot 2.2. The first column contains cube X and its label 'X'. The second and third columns contain cubes A, B, C, E, F, G and their labels. The fourth column contains two text boxes: 'None of the cubes could be a rotation.' (D) and 'I do not know the solution.' (H).

Rätt svar: B

Rot 2.3 Which of these cubes below is a rotation of X?

A 3x3 grid of cube options for Rot 2.3. The first column contains cube X and its label 'X'. The second and third columns contain cubes A, B, C, E, F, G and their labels. The fourth column contains two text boxes: 'None of the cubes could be a rotation.' (D) and 'I do not know the solution.' (H).

Rätt svar: G

Rot 2.4 Which of these cubes below is a rotation of X?

None of the cubes could be a rotation.

I do not know the solution.

Rätt svar: A

Rot 2.5 Which of these cubes below is a rotation of X?

None of the cubes could be a rotation.

I do not know the solution.

Rätt Svar: A

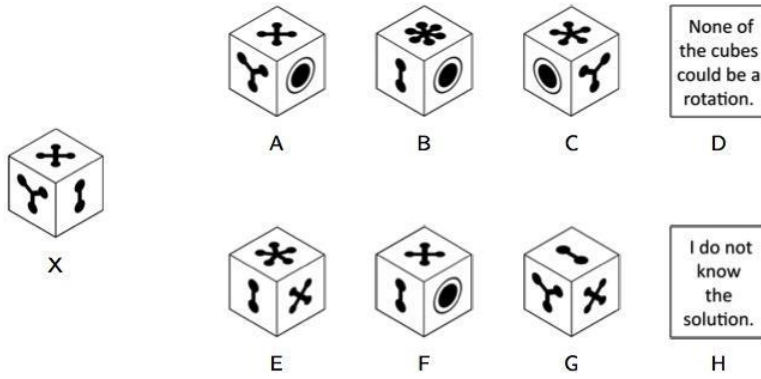
Rot 2.6 Which of these cubes below is a rotation of X?

None of the cubes could be a rotation.

I do not know the solution.

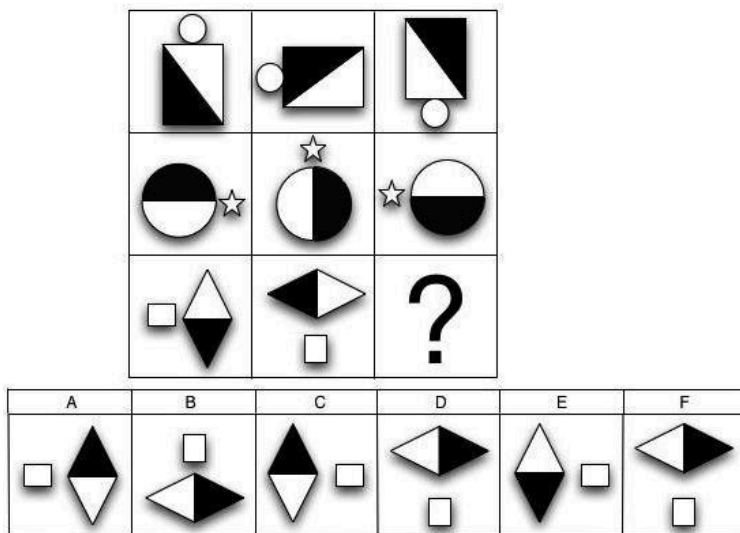
Rätt Svar: B

Rot 2.7 Which of these cubes below is a rotation of X?



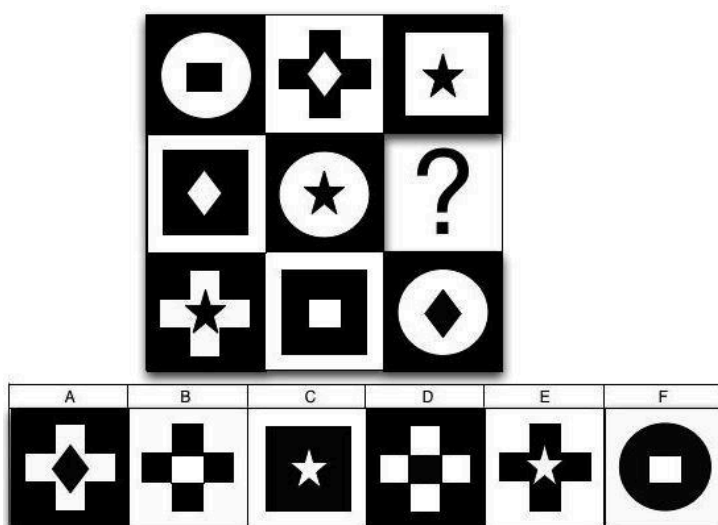
Rätt svar: F

Fig 2.1 Which alternative fits best in the question mark below (?)



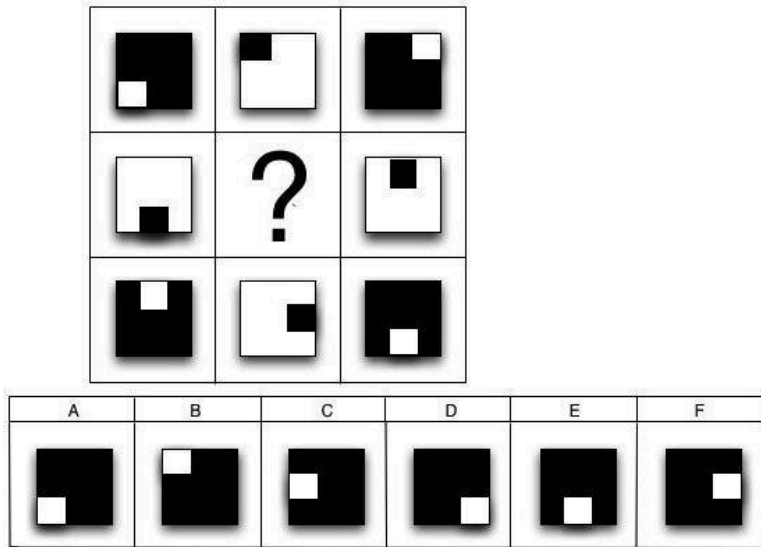
Rätt svar: C

Fig 2.2 Which alternative fits best in the question mark below (?)



Rätt svar: B

Fig 2.3 Which alternative fits best in the question mark below (?)



Rätt svar: C

Closing questions

To what extent did you feel the puzzles were hard?

- Very easy
-
-
-
- Very hard

How dedicated were you to get a good result on the puzzles?

- Not at all dedicated
-
-
-
- Very dedicated

How would you compare the level of difficulty between this profile and the previous one?

- A lot easier
-
-
-
- A lot harder

Confirmation message

Almost done! Great job so far!

Now there's just one more profile left for you to complete, we truly appreciate your time and effort. If you're like us and have a huge interest for cognitive testing for recruitment purposes you can find out more through this link.

<https://shorturl.at/sDGJT>

Below you can also find the link to the third and final cognitive profile. Just like before, we highly encourage you to take a one hour break before completing it.

If you want a reminder you can always contact us and we will give you a heads up before the final date (november 10th) to fill out the forms.

Carefully save the link to Cognitive profile #3

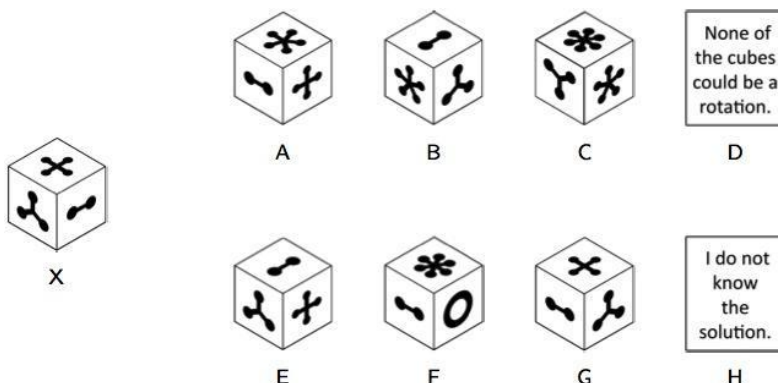
<https://shorturl.at/zGTX9>

Kognitivt test 3

Welcome to the third and final test. We are so grateful for your help. Make sure to use this final test as a practice for possible future job applications!

If you are interested in knowing your score, you will be able to see your scores once you have submitted each profile.

Rot 3.1 Which of these cubes below is a rotation of X?



Rätt svar: C

Rot 3.2 Which of these cubes below is a rotation of X?

X

A B C

E F G

D

H

Rätt svar: B

Rot 3.3 Which of these cubes below is a rotation of X?

X

A B C

E F G

D

H

Rätt svar: C

Rot 3.4 Which of these cubes below is a rotation of X?

X

A B C

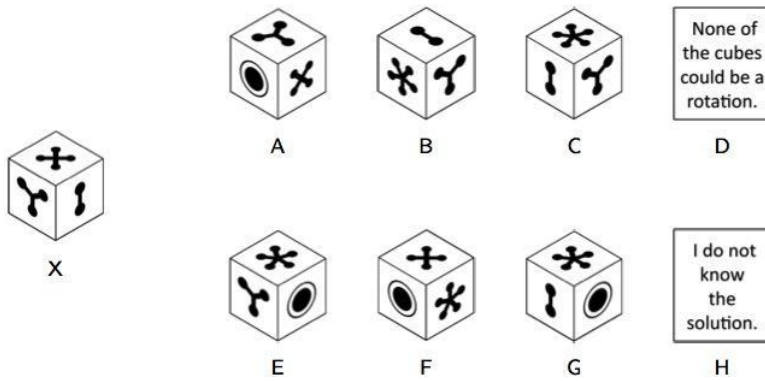
E F G

D

H

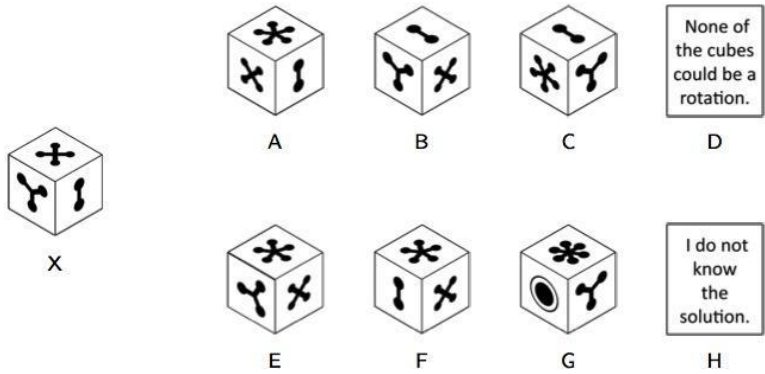
Rätt svar: G

Rot 3.5 Which of these cubes below is a rotation of X?



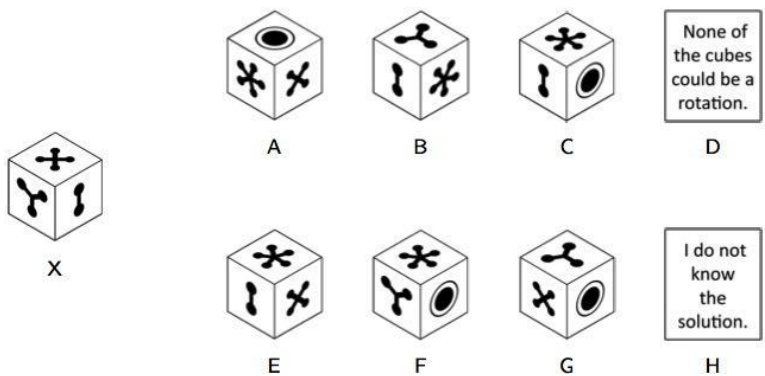
Rätt svar: F

Rot 3.6 Which of these cubes below is a rotation of X?



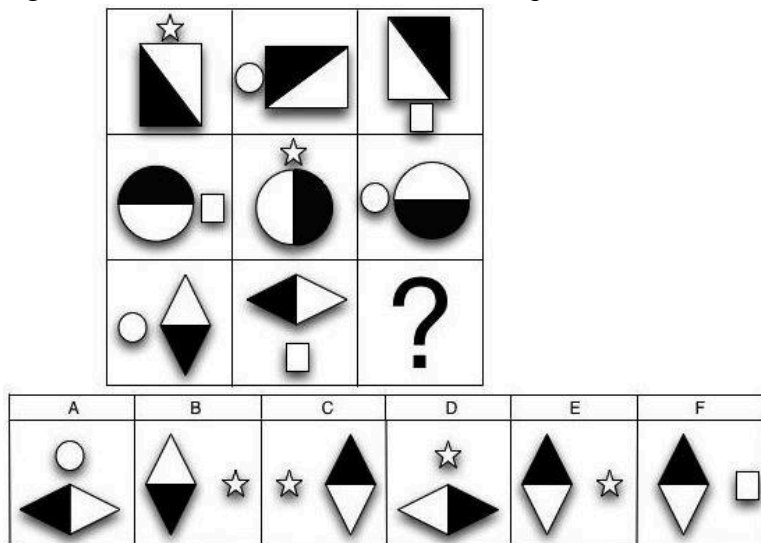
Rätt svar: E

Rot 3.7 Which of these cubes below is a rotation of X?



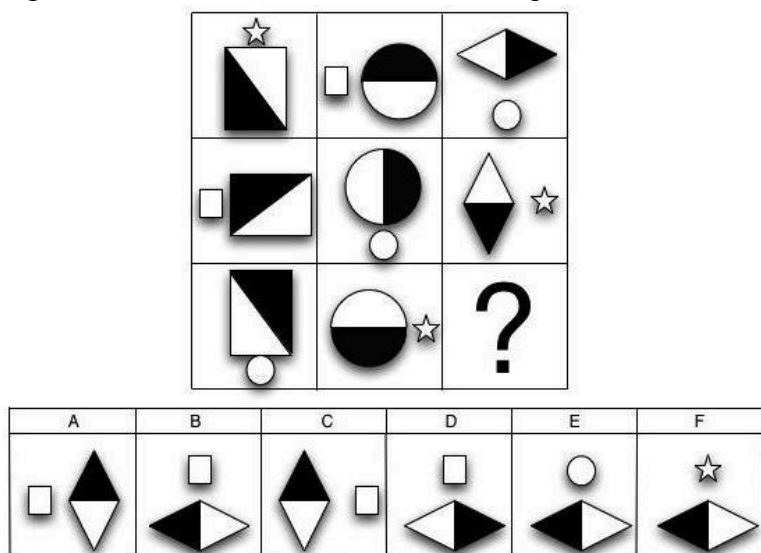
Rätt svar: A

Fig 3.1 Which alternative fits best in the question mark below (?)



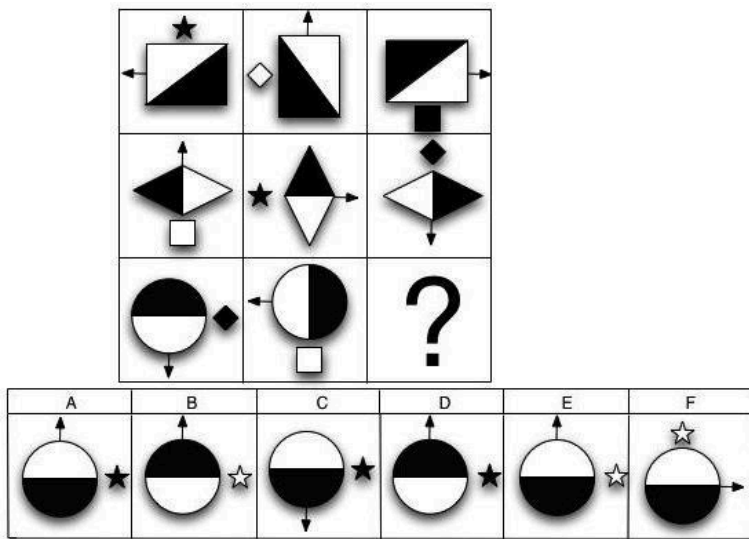
Rätt svar: E

Fig 3.2 Which alternative fits best in the question mark below (?)



Rätt svar: B

Fig 3.3 Which alternative fits best in the question mark below (?)



Rätt svar: A

Closing questions

To what extent did you feel the puzzles were hard?

- Very easy
-
-
-
- Very hard

How dedicated were you to get a good result on the puzzles?

- Not at all dedicated
-
-
-
- Very dedicated

How would you compare the level of difficulty between this profile and the previous one?

- A lot easier
-
-
-
- A lot harder

How was your overall experience of taking part in this study?

- I did not like it at all
-
-
-
- I loved it

Do you have any final thoughts or feedback that you want to give us, now that you have finished the study?

Confirmation message

All done, good job!!

We want to profusely thank you for your time and effort.

Now that you have completed your part of the study we can reveal a little more about the purpose of our research.

Like you've probably already figured out, we have taken a special interest in these cognitive tests which have become the norm when recruiting new personnel to companies and organizations. Another thing that has been a big question for us has been the presence of test-retest effect within these types of tests.

The aim of our study is to test if there's a correlation between having taken more of these tests and scoring better. From there we want to discuss the use of cognitive tests as a precursor of future work-performance and whether candidates can improve their chances at getting a certain job just from practicing cognitive tests.

If you feel like reading more about the test-retest effect there are a few interesting articles that we have taken inspiration from. Here is a link to one of these articles:

<https://shorturl.at/dnzN2>

Again, we are so grateful for your participation and hope you have enjoyed it. If you want to take part of your results, the results of the study once completed, or if you have any questions don't hesitate to contact us.

Best and kindest of regards,

Hanna Digman (hanna.digman@hotmail.com), Max Angelin (max@angelin.se) and Oscar Forum (oscarforum5@gmail.com)