

# Adaptiva och andra datorbaserade prov

KUNSKAPSÖVERSIKT



*Skolverket*



# **Adaptiva och andra datorbaserade prov**

En kunskapsöversikt

Beställningsadress:  
Fritzes kundservice  
106 47 Stockholm

Telefon: 08-690 95 76  
Telefax: 08-690 95 50  
E-post:skolverket@fritzes.se

[www.skolverket.se](http://www.skolverket.se)  
ISBN: 978-91-86529-17-8

Form: Ordförrådet AB  
Omslagsbild: MATTON

Stockholm 2010

## Förord

I takt med en ökande datortäthet och ett växande IT-kunnande inom skolans värld har man i många länder under en längre tid drivit projekt med adaptiva och andra datorbaserade prov. Adaptiv testning bygger på att uppgifterna väljs efter individens förmåga. Ett adaptivt prov känner fortlöpande av provtagarens kunskapsnivå och presenterar relevanta uppgifter tills provtagarens kunskapsnivå är fastställd. Då avbryts provet.

Projekt med datorbaserade uppgifter har i flera fall lett till nationella datorbaserade bedömningssystem till exempel i Norge och Danmark. Efter ett studiebesök i Danmark hösten 2007 väcktes tanken på att ta fram en kunskapsöversikt inför framtida beslut om datorbaserade bedömningsformer i Sverige.

Skolverket gav därför Institutionen för beteendevetenskapliga mätningar, Umeå universitet ett uppdrag att i samverkan med Göteborgs universitet ta fram en kunskapsöversikt över forskning om och användning av adaptiva prov inom och utom Norden, särskilt fokuserad på adaptiva prov i matematik, NO-ämnen och engelska.

Uppdraget innehöll två mindre pilotstudier med datorbaserade prov i engelska och matematik.

I en redovisning till regeringen våren 2009 framhöll Skolverket också möjligheten med datorbaserade bedömningar och vikten av att utveckla datorbaserade prov och andra bedömningsformer inom några pilotområden.

Parallellt med att kunskapsöversikten togs fram påbörjade Skolverket utvecklingen av en plattform för framtida datorbaserade bedömningar. Enligt ett regeringsuppdrag våren 2009 utvecklar Skolverket datorbaserade prov i svenska för invandrare. Dessa prov används numera av ett flertal utbildningsanordnare.

Föreliggande kunskapsöversikt avslutas med ett strategiskt discussionsunderlag för framtida beslut om adaptiva provs roll i Sverige.

Ett stort tack riktas till de elever och lärare som med stort engagemang deltog i pilotstudien.

Stockholm i september 2010

*Karin Hector-Stahre*  
Enhetschef

*Wolfgang Dietrich*  
Undervisningsråd



# Innehåll

Inledning.....	8
Sammanfattning.....	12
Adaptiva och andra datorbaserade prov .....	18
Datorer i skolan .....	20
Datorer, bedömning och prov .....	23
Adaptiva prov.....	28
Datorbaserade adaptiva prov .....	28
Datorbaserade prov i praktiken (exempel) .....	40
Datoriserade nationella prov.....	41
Prov utvecklade vid universitet och andra institutioner.....	45
Datoriserade och adaptiva prov i matematik och NO.....	48
Datoriserade och adaptiva prov i språk .....	55
Produktiva färdigheter.....	56
Studiebesök, demonstrationer och konferenser .....	62
Pilotstudier.....	74
Syfte med pilotstudierna .....	74
Pilotstudie i matematik .....	74
Pilotstudie i engelska.....	86
Diskussion.....	110
Några slutsatser utifrån kunskapsöversikten .....	110
Om Skolverkets redovisning angående datorbaserade prov .....	112
Datorbaserade prov och datorstödd undervisning.....	114
Rekommendation .....	117
Referenser .....	120
Litteratur .....	120
Andra källor .....	129
Länkar.....	133





# Inledning

## Inledning

Tanken att ta fram en kunskapsöversikt om adaptiva prov föddes på Skolverket efter att ha tagit del av utvecklingen av adaptiva prov i Danmark och även fått förslag på utvecklingsprojekt inriktade mot adaptiva prov inom Sverige. Uppdraget med kunskapsöversikten gick till Institutionen för beteendevetenskapliga mätningar,<sup>1</sup> Umeå universitet, att i samverkan med Göteborgs universitet, PRIM-gruppen och Skolverket ta fram ett bättre underlag för beslut om fortsatt utvecklingsarbete av i första hand datorbaserade adaptiva prov. Under arbetets gång har fokus i viss mån förskjutits mot datorbaserade prov i allmänhet.

I denna rapport presenteras resultatet av ett uppdrag från Skolverket som handlat om att

- ta fram en kunskapsöversikt över forskning om och användning av adaptiva prov inom och utom Norden, med särskilt fokus på adaptiva prov i matematik, NO-ämnena och engelska
- genomföra två pilotstudier med datorbaserade prov, en i matematik och en i engelska
- presentera en resultatredovisning av pilotstudien samt ett strategiskt diskussionsunderlag för framtida beslut om adaptiva provs roll i Sverige.

Uppdraget gällde alltså i första hand adaptiva prov och kunskapsöversikten skulle behöva göras mycket mer omfattande, för att också på ett heltäckande sätt täcka alla erfarenheter och kunskaper som byggts upp om datorbaserade prov i allmänhet.

Uppdraget har genomförts i ett samarbete mellan Institutionen för beteendevetenskapliga mätningar vid Umeå universitet,<sup>1</sup> Institutionen för pedagogik och didaktik, Enheten för språk och litteratur vid Göteborgs universitet<sup>2</sup> och Skolverket. Även PRIM-gruppen vid Stockholms universitet har deltagit i möten och diskussioner inom uppdraget.

---

<sup>1</sup> Sedan 1 januari 2010 ingår den tidigare Institutionen för beteendevetenskapliga mätningar i Institutionen för tillämpad utbildningsvetenskap vid Umeå universitet.

<sup>2</sup> Efter en institutionell omstrukturering, då indelningen i enheter upphörde, ingår sedan 1 juli 2010 den i detta projekt medverkande arbetsgruppen i Institutionen för pedagogik och specialpedagogik vid Göteborgs universitet.

Peter Nyström har varit huvudansvarig för projektet och är också ansvarig för utformningen av denna rapport. Underlag för rapporten kommer från många medverkande och mer fullständiga delrapporter har också publicerats.<sup>3</sup>

Den arbetsgrupp som tagit fram underlag för kunskapsöversikten, genomfört pilotstudien och medverkat i skrivandet av rapporten har förutom ovan nämnda projektledare bestått av Anna Lind Pantzare, Margareta Löfstedt, och Marie Wiberg från Umeå universitet, samt Hans Albinsson, Gudrun Erickson och AnnaKarin Lindqvist från Göteborgs universitet. Skolverkets kontaktperson har varit Wolfgang Dietrich.

Förutom denna kunskapsöversikt har utredningsarbetet redovisats vid två seminarier på Skolverket (den 17 december 2008 och den 15 juni 2009).

---

<sup>3</sup> Mer fullständiga rapporter från olika delar av projektet kommer att finnas tillgängliga på webbplatser tillhörande Göteborgs universitet och Umeå universitet.



# Sammanfattning

## Sammanfattning

Rapporten beskriver resultatet av ett projekt som främst handlar om adaptiva prov, det vill säga datorbaserade prov där datorn enligt någon regel väljer ut uppgifter som passar till den enskilde provtagaren. De adaptiva proven kan alltså ses som en delmängd av datorbaserade bedömningar. Internationellt används en rad olika termer och beteckningar för de bedömningar som använder sig av datorer eller annan modern teknik, och förslagsvis kan de svenska beteckningarna datorstödda prov (DSP), datorbaserade prov (DBP) och datorbaserade adaptiva prov (DBAP) fånga in de viktigaste aspekterna i en kategorisering av olika typer av datorstödd.

För att kunna förstå möjligheter och begränsningar med datoradaptiva prov (DAP), så är det angeläget att orientera sig om datorbaserade prov i allmänhet. Möjligheter och svårigheter med datorbaserade prov hänger i sin tur samman med generella frågor om IT och lärande samt hur datoranvändningen och datortillgången ser ut i svenska skolor.

Den relativt goda datortillgången i den svenska skolan och ett ökat fokus på datorn som verktyg i undervisningen borde ge goda förutsättningar för användningen av datorbaserade prov, men det är ytterst tveksamt om datorbaserade prov och bedömningar används i någon större utsträckning. Utvärderingar av datoranvändningen i svenska skolor tyder på en ganska omfattande användning i vissa ämnen, framför allt i språk, men nästan ingen alls i andra (till exempel matematik). Det finns inga undersökningar om motsvarande skillnader mellan ämnen även när det gäller användningen av datorbaserade prov. Möjligen skulle datorbaserade prov vara mest angelägna i ämnen med en större datoranvändning, eftersom det skulle öka samstämmigheten mellan vad som sker i undervisningen och vad som sker på proven.

Några av de fördelar och möjligheter som lyfts fram med datorbaserade prov är att de kan utformas så att de prövar kunskaper som inte låter sig prövas på traditionella prov, att de ökar provens autenticitet och samstämmighet med undervisningen samt att de är flexibla och kan möjliggöra en direkt återkoppling av resultat till eleven. När det gäller nackdelar och svårigheter så handlar det bland annat om tillgången på datorer och nätuppkopplingar av tillräcklig kvalitet, bristande datorvana och en begränsning av vad som bedöms.

Datorn gör det möjligt att skapa uppgiftstyper som mäter komplexa mål på innovativa sätt, men i praktiken gör bland annat utvecklingskostnader att de datorbaserade proven innehåller uppgiftsformat som är begränsande när det gäller att spegla kunskapsmål. Begränsningen blir särskilt tydlig när datorn ska automaträtta elevernas svar, vilket i sin tur är en förutsättning för att eleverna ska få en direkt återkoppling av resultaten. Detta är ett exempel på hur en av de möjligheter som framhålls med datorbaserade prov samtidigt förutsätter åtgärder som framhålls som risker med datorbaserade prov.

Det speciella med datorbaserade *adaptiva* prov är alltså att en dator används för att anpassa provet till provtagarens nivå. I stället för att ge varje deltagande individ samma uppgifter, så programmeras datorn att från en uppgiftsbank välja de uppgifter som mest effektivt ger information om provtagarens förmåga.

Det råder knappast något tvivel om att teori och teknik för datorbaserade adaptiva prov är väl utvecklade, men forskning och andra erfarenheter från datorbaserade adaptiva prov visar på både möjligheter och svårigheter i tillägg till vad som gäller för datorbaserade prov i allmänhet. Möjligheterna handlar bland annat om att de datorbaserade adaptiva proven i högre grad än andra prov kan ge information om provtagarnas förmåga på hela skalan med hög tillförlitlighet, att de kräver färre uppgifter och att fuskrisken kan minskas. Bland de specifika svårigheterna med datorbaserade adaptiva prov kan nämnas att de ställer stora krav på utprovning och kategorisering av uppgifter, att datorbaserade adaptiva prov inte är lämpliga för att mäta alla typer av kunskaper, samt att de har höga utvecklingskostnader.

Datorbaserade adaptiva prov finns i olika varianter, som bland annat handlar om hur uppgifterna grupperas innan de presenteras för provtagaren. I ett så kallat fullständigt datorbaserat adaptivt prov väljer datorn en uppgift i taget ur uppgiftsbanken medan ett begränsat adaptivt prov bygger på att uppgifterna grupperats i delprov som i sin helhet ges till provtagaren. I alla varianter av datorbaserade adaptiva prov behövs en uppgiftsbank med väl beskrivna uppgifter, en startpunkt som består av en eller flera uppgifter, möjligheter för datorn att bedöma svarets kvalitet, en algoritm som väljer vilken eller vilka uppgifter som provtagaren ska möta, samt ett avslutningskriterium som anger villkoren för att provet avslutas.

Att tekniken för datorbaserade adaptiva prov är väl utvecklad illustreras också av de exempel på sådana prov som redovisas i rapporten. Exemplet kommer från olika delar av världen och fokuserar i första hand på prov och uppgifter i matematik, naturvetenskap och språk. Som ett led i arbetet med kunskapsöversikten har kontakter tagits med miljöer i andra länder som utvecklat olika varianter av datorbaserade prov. Dessa kontakter har gett värdefulla erfarenheter. Studiebesök har genomförts i Danmark, Norge, Nederländerna och England. Projektmedlemmar har också deltagit i en konferens och seminarier om både danska och finska exempel på datorbaserade och adaptiva provmodeller.

I projektet har vidare två pilotstudier genomförts, en i matematik och en i engelska. I båda fallen användes datorbaserade prov från det norska nationella provsystemet. Proven var datorbaserade men inte adaptiva. I samband med att proven genomfördes i ett antal klasser undersöktes attityder, uppfattningar och reaktioner hos elever och lärare och skolledare. Pilotstudierna hade även fokus på beredskapen för användning av datormediet som sådant i relation till prov och bedömning.

Pilotstudien i matematik genomfördes i årskurs 3 och 6 med fyra klasser i varje årskurs, totalt 190 elever. Eleverna visade sig intresserade, nyfikna och allmänt positiva till att göra ett datorbaserat prov i matematik. Detta trots att uppgifterna i sig varken var särskilt nyskapande eller alternativa i förhållande till de uppgifter eleverna arbetade med på lektionstid. Vissa könsskillnader kunde observeras, framför allt att pojkarna hade en tendens till tävling, till exempel om hur många uppgifter de kunde hinna, medan flickorna ägnade sig åt uppgifterna med litet större noggrannhet. När det gäller lärarna så tycks datorvana lärare se större fördelar och vara mer positiva till datoriserade matematikprov än de som har mindre datorvana. Övergripande fungerade de tekniska lösningarna med bärbara datorer och trådlösa nät mycket väl.

Pilotstudien i engelska genomfördes i årskurs 5, 6 och 8 med totalt 250 elever. Redan innan eleverna sett och provat på de datorbaserade proven var de allra flesta mycket positiva till tanken och kände tillförsikt. Även efter provet uttryckte de allra flesta att de tyckte om att göra provet och att det kändes lättillgängligt. Ganska få elever upplevde att de för egen del hade behövt mer hjälp. Dock menar hälften av dem att det inte är lika lätt för alla att göra datorbaserade prov. Lärarna i engelska beskriver en förvånansvärt låg användning av datorer i undervisningen,



vilket troligen hänger ihop med relativt dålig tillgång till datorer och en del tekniska problem med bland annat Internetuppkopplingen på de deltagande skolorna. Lärarna, liksom skolledarna, är dock positiva till ökad användning av datorer och datorbaserade prov. Erfarenheterna från pilotstudierna tyder på att den attitydmässiga beredskapen för användning av datormediet som sådant i relation till prov och bedömning är mycket god. Bland de skolor som deltog i pilotstudien i engelska finns dock en oro att de tekniska förutsättningarna inte ska vara tillräckliga för att genomföra datorbaserade prov på ett bra sätt.

Sammanfattningsvis visar rapporten på att användningen av datorbaserade prov i Sverige måste betraktas som låg och outvecklad jämfört med flera europeiska länder, däribland Danmark och Norge. Datoranvändningen i undervisningen är också låg, i synnerhet i vissa ämnen. Pilotstudierna som genomförts i detta projekt har dock visat på en positiv attityd till datorbaserade prov från elever, lärare och skolledare, samt beredskap att använda dem trots att utrustning ibland saknas. Pilotstudierna har dessutom visat att eleverna i de yngre åldrarna inte förmår arbeta särskilt länge med ett datorbaserat prov. Det skulle tala för adaptiva datorbaserade prov, eftersom tekniken framför allt innebär kortare provtid för samma information. Eftersom målrelaterade prov torde vara mest intressanta i Sverige så bör i så fall ett begränsat adaptivt prov väljas där man kan bestämma om provtagaren når upp till en förutbestäm nivå av kunskap eller förmåga. Det är uppenbart att användningen av datorer i samband med pedagogisk bedömning öppnar för möjligheten att utveckla och använda intressanta uppgiftsformat som inte är möjliga i så kallade papper och penna-prov. Denna möjlighet tycks dock sällan ha utnyttjats i praktiken. En viktig begränsning för datorbaserade adaptiva prov är att de kräver datorrättning, vilket sätter relativt snäva gränser för hur uppgifterna i proven kan utformas.

I rapporten framhålls att Skolverket bör initiera och stödja ett långsiktigt arbete för att utveckla datorbaserade prov i Sverige, och då i första hand satsa på datorbaserade prov som inte är adaptiva. På sikt kan även datorbaserade adaptiva prov utvecklas, och primärt prov som syftar till att på ett tillförlitligt sätt identifiera kategorier av elever, till exempel vilka elever som är i behov av särskilt stöd. En satsning på datorbaserade prov kan vara mest angelägen inom områden där datorn används på ett naturligt sätt i skolan, till exempel i skriftlig produktion, eftersom

det skulle kunna öka samstämmigheten mellan prov och undervisning. Samtidigt bör det påpekas att det idag knappast finns någon kostnads-effektiv och tillräckligt bra modell för att låta datorn bedöma skriftlig produktion. Den skriftliga produktionen måste, även om den sker i en dator, bedömas av mänskliga bedömare, med eller utan datorstöd.

Det är avslutningsvis angeläget att det fortsatta arbetet med datorbaserade prov tar hänsyn till de teoretiska och praktiska begränsningarna för datorbaserade prov och datorbaserade adaptiva prov som redovisats i denna rapport. Det är även angeläget att i det fortsatta arbetet bygga upp kraftfulla och långvariga samarbeten med andra som kommit längre.

# **Adaptiva och andra datorbaserade prov**

## Adaptiva och andra datorbaserade prov

Datorbaserade prov, och särskilt datorbaserade adaptiva prov, är i fokus för denna kunskapsöversikt. För att kunna diskutera adaptiva prov är det inledningsvis angeläget att sätta in dem i ett sammanhang.

Kunskapsöversikten handlar om datorbaserade prov i undervisningssammanhang, vilket innebär att vi rör oss inom det fält som på senare år kommit att kallas pedagogisk bedömning. Med pedagogiska bedömningar avses här alla de aktiviteter i utbildningssammanhang som har ett mer eller mindre uttalat syfte att ge information och belägg som underlag för olika beslut.

Besluten kan vara små, som många av de avgöranden som sker i klassrummet när lärare till exempel tar ställning till om en elev behöver mer hjälp för att nå ett lärandemål, och de kan vara stora, som till exempel betygssättning. När en bedömningssituation kallas för ”prov” så avses i de allra flesta fall en formell bedömningssituation som är specifikt utformad för att kunna bedöma vissa saker och där spelreglerna är kända för elever och lärare. Motsatsen skulle då vara en informell bedömningssituation, där ett tillfälle grips i flykten och blir till ett bedömnings-tillfälle.

Denna rapport handlar vidare om adaptiva prov, vilket är ett begrepp som starkt kopplas till datorer. Mer allmänt kan man naturligtvis säga att ett prov i någon mening är adaptivt om det finns ett inslag av att olika elever gör olika uppgifter, och att valet av uppgifter hänger samman med elevens upplevda eller visade kunnande. Eleven kan själv vara delaktig att välja uppgifter i papper och penna-prov (vilket till exempel har provats i nationella prov av PRIM-gruppen, Stockholms universitet) och i datorbaserade prov,<sup>4</sup> eller läraren kan välja eller anpassa uppgiften för att den ska passa den enskilda eleven. Ett exempel på det senare finns i de nationella ämnesprov för NO-ämnena som presenterades våren 2009. Där ska läraren bedöma elevernas planering av en laborativ uppgift och om den inte är tillräckligt bra så kan lärare ge eleverna en färdig planering i stället. Begreppet adaptiva prov måste dock begränsas till anpassning av provuppgifter som sker inom ett prov, det handlar inte om att eleverna ges olika prov beroende på deras förutsättningar eller val. I allmänhet

---

<sup>4</sup> Rocklin & O'Donnell (1987)

avses dock med adaptiva prov att det är en dator som väljer och presenterar uppgifterna för eleven, vilket alltså mer fullständigt bör betecknas som datorbaserade adaptiva prov.

Kunskapsöversikten bygger på gängse metoder för att identifiera aktuell forskning och utveckling när det gäller adaptiva prov, framför allt litteratursökningar i databaser och kontakter med experter. En litteratursökning som gjordes i databasen ERIC med sökorden *adaptive test* och *adaptive assessment* gav 66 träffar i vetenskapliga tidskrifter mellan åren 2000 och 2008. En grov kategorisering av dessa publikationer visar att cirka två tredjedelar av dem handlar om tekniska/statistiska aspekter och att cirka en fjärdedel av den publicerade forskningen handlar om tillämpningar i praktiken eller generell argumentation om adaptiva provs möjligheter. Denna enkla jämförelse visar på var fokus i forskningen ligger.

När litteratursökningen i ERIC utvidgas bakåt i tiden framträder en bild av att det på senare år publicerats en avsevärt mindre mängd forskning om adaptiva prov. Detta kan inte tolkas som att utvecklingen avstannat (även om den största entusiasmen för vad som kan göras kanske har lagt sig) utan kan åtminstone delvis förklaras med att mycket utveckling på området sker inte i universitetsmiljön utan i kommersiella sammanhang, och publiceras därför inte i vanlig bemärkelse. Resultatet av detta utvecklingsarbete kan eventuellt spåras via webbplatser som beskriver olika produkter.

I allmänhet avses alltså med adaptiva prov att en dator väljer och presenterar uppgifterna för den enskilda eleven, beroende på hans eller hennes förmåga. För att kunna förstå möjligheter och begränsningar med datorbaserade adaptiva prov så är det angeläget att först kortfattat betrakta datorbaserade prov i allmänhet. Datorbaserade adaptiva prov kan anses vara en delmängd av datorbaserade prov. Möjligheter och svårigheter med datorbaserade prov hänger i sin tur samman med generella frågor om IT och lärande och hur datoranvändningen och dator-tillgången ser ut i svenska skolor. Kunskapsöversikten handlar primärt om datorbaserade adaptiva prov, och en verkligt seriös genomgång av kunskapsläget när det gäller datorbaserade prov i allmänhet ligger långt bortom vad som är möjligt i denna studie. Samtidigt är det angeläget att sätta in de datorbaserade adaptiva proven i sitt sammanhang och kortfattat diskutera några aspekter av datorbaserade prov.

## Datorer i skolan

Utvecklingen av datorbaserade prov hänger naturligtvis samman med tillgång och användning av datorer i skolan. I en aktuell utvärdering av IT-användningen i skolan skriver Skolverket att tillgången till IT-utrustning är olika i svenska förskolor, skolor och vuxenutbildningar.<sup>5</sup> Alla lärare och all förskolepersonal har i princip tillgång till dator på arbetsplatsen, men i grundskola och förskola delar man vanligtvis datorn med en eller flera kollegor. Bland grundskollärarna säger tre av fyra att de delar dator med kollegorna. Elevernas datortillgång är större i gymnasieskolan än i grundskolan, i genomsnitt går det 2,5 elever per dator i kommunala gymnasieskolor och 6,0 elever per dator i kommunala grundskolor. Enligt Skolverkets studie använder de flesta av lärarna IT för såväl pedagogiska som administrativa syften. Nio av tio lärare uppger att de använder dator på lektionstid åtminstone någon gång och drygt hälften av lärarna anger att de använder dator på lektionstid varje vecka. Omkring två tredjedelar av lärarna i gymnasieskolan och vuxenutbildningen bedömer att de är ganska bra eller mycket bra på IT, medan motsvarande andel är något lägre för personal i grund- och förskola. Många lärare vill ha kompetensutveckling inom IT-relaterade områden, främst när det gäller arbete med bild, ljud och video med hjälp av IT.

**Tabell 1** Andel (%) lärare som använder dator på lektionstid, varje dag eller mer sällan (förskolepersonal som använder dator i arbetet med barngruppen)<sup>6</sup>

	Förskolepersonal	Grundskolelärare	Gymnasie-lärare	Vuxenutbildning
Varje dag	11,9	20,4	27,5	43,2
Varje vecka	25,4	36,5	30,1	27,1
Varje månad	8,5	11,6	13,7	7,9
Mer sällan	24,7	20,2	19,5	15,1
Aldrig	28,3	10,7	9,1	6,8
Uppgift saknas	1,2	0,7	0,1	0,0

<sup>5</sup> Skolverket (2009)

<sup>6</sup> Skolverket (2009)

Om vi håller oss till grundskolan säger alltså Skolverkets utvärdering av IT-användning att alla lärare i princip har tillgång till dator, att uppåt 70 procent av lärarna använder datorn på lektionstid minst en gång i månaden, och att det i genomsnitt går 6 elever på varje elevdator i kommunala grundskolor.<sup>7</sup> Denna utvärdering säger dock ingenting om hur datoranvändningen ser ut i olika ämnen och om det skiljer sig åt mellan ämnen, men ett par av de internationella jämförande studier som Sverige deltar i kan bidra med mer ämnesspecifik information om datoranvändning. PIRLS undersöker läsförståelse hos elever i årskurs 4, och av den svenska rapporten från den senaste PIRLS-studien framgår att i Sverige går hälften av eleverna på skolor som har färre än fem elever per dator i årskurs 4.<sup>8</sup> I Norge är motsvarande andel cirka tre fjärdedelar och i Danmark och Island går flertalet av eleverna på skolor med så pass få elever per dator. Man konstaterar att de här grupperna har blivit mycket större sedan 2001 och att det har skett en markant och signifikant uppgång i Sverige, Norge och Island.

Enligt PIRLS 2006 har 35 procent av Sveriges elever i årskurs 4 lärare som minst någon gång per månad låter dem använda undervisningsprogram för att utveckla läsfärdigheter och lässtrategier.<sup>9</sup> I Danmark är motsvarande andel lägre (29 procent) medan den är högre i såväl Island (51 procent) som Norge (61 procent). I Sverige har denna typ av datoranvändning i språk blivit signifikant mindre till sin omfattning sedan den förra PIRLS-studien 2001. Hälften, eller något fler, av de nordiska eleverna får också använda datorn för att läsa berättelser eller andra texter, och den andelen har inte förändrats i Sverige sedan 2001.

Utifrån ett språkundervisningsperspektiv i årskurs 4 sammanfattar PIRLS 2006 att i stort sett alla elever i de nordiska länderna går på skolor där det finns datorer och att en tredjedel av de svenska eleverna har lärare som minst någon gång per månad låter eleverna använda undervisningsprogram på datorn för att utveckla läsfärdigheter och lässtrategier.<sup>10</sup>

När det gäller datortillgången i matematik och naturorienterande ämnen i årskurs 8 så har 40 procent av de svenska eleverna tillgång till dator under matematiklektionerna och 60 procent under NO-lektionerna,

---

<sup>7</sup> Skolverket (2009)

<sup>8</sup> Skolverket (2007)

<sup>9</sup> Skolverket (2007)

<sup>10</sup> Skolverket (2007)

enligt vad lärarna uppger i TIMSS 2007.<sup>11</sup> I TIMSS 2003, som alltså genomfördes fyra år tidigare, var motsvarande andelar 46 respektive 64 procent. Enligt lärarna har alltså cirka 4 av 10 elever i årskurs 8 tillgång till en dator i matematikundervisningen, och denna andel har till och med minskat något mellan 2003 och 2007. I NO-undervisningen är datortillgången något bättre, men även här visar TIMSS på en liten nedgång.

Förutom en bild av tillgången på datorer ger TIMSS vissa svar på frågan om hur datorer används i matematik och NO, och hur mycket de används. I TIMSS 2007 rapporterar lärarna från alla deltagande länder att det är mycket ovanligt med användning av datorer i matematikundervisningen så ofta som hälften av lektionerna, även i länder med relativt god tillgång till datorer. Enligt de svenska lärarna är det bara 1 procent av eleverna som vid minst hälften av lektioner använder datorer för att upptäcka principer och begrepp, öva färdigheter och procedurer, söka idéer och information, eller behandla och analysera data. Användningen av datorer är också mycket låg i NO, även om så mycket som 8 procent av eleverna använder datorn för att söka idéer och information vid minst hälften av NO-lektionerna. I den svenska rapporten om TIMSS 2003 skriver Skolverket att datorer inte används särskilt ofta i undervisningen i vare sig matematik eller NO i årskurs 8 i Sverige.<sup>12</sup> Då uppgav lärarna att den vanligaste matematikaktiviteten vid datorn var att öva färdigheter och procedurer, medan det var litet mer ovanligt att använda datorn för att upptäcka matematiska principer och begrepp eller söka efter idéer och information.

Sammanfattningsvis säger TIMSS-undersökningen att datoranvändningen i skolans matematik- och NO-undervisning har *minskat* marginellt mellan 2003 och 2007, enligt lärarnas uppfattning. Vidare får svenska lärare mindre fortbildning i att integrera IT i undervisningen jämfört med lärare i andra länder.

Trots relativt god datortillgång i skolorna tycks alltså datoranvändningen inte vara särskilt stor i NO-ämnena och nästan obefintlig i matematik. I språkundervisningen verkar dock datorn användas betydligt mer, åtminstone i årskurs 4.

---

<sup>11</sup> Skolverket (2008a)

<sup>12</sup> Skolverket (2004)



## Datorer, bedömning och prov

Tillgången till datorer och annan informations- och kommunikationsteknik (IKT) i skolan gör det möjligt att även använda modern teknik i samband med bedömning av vad elever vet och kan göra i olika ämnen. IKT och bedömning kan kopplas till varandra på en rad olika sätt. Det kan ske genom alltifrån användning av miniräknare och andra tekniska hjälpmedel i samband med prov och konstruktion och distribution av prov via dator, till att besvara frågorna i datorn och kanske också få dem bedömda direkt av datorn. Med en vid definition av vad som är datorer och om vi räknar med alla olika former av datoranvändning i samband med prov så är datorstödda prov inte så ovanliga i Sverige. En snävare form av datorbaserade prov, där datorn är bedömningssituationens primära medium, tycks dock vara ganska ovanlig.

Övergripande kan bedömning som i högre eller lägre grad använder sig av datorteknik beskrivas som alltifrån IKT-stödd till datorbaserad bedömning. IKT-stödd bedömning innefattar hela bredden från till exempel användningen av miniräknare på prov, eller deluppgifter där datorn används för att ta fram information och göra simuleringar men svaren ges på papper, till prov och hela bedömningssystem som görs direkt i datorn. Med datorbaserad bedömning och datorbaserade prov avses en snävare kategori där bedömningssituationen presenteras i datorn och eleverna ger sina svar i datorn. Bedömningen kan ske av datorn, med datorstöd, eller helt manuellt.

I den engelskspråkliga litteraturen om datorstödd bedömning används en rad olika termer för olika varianter av datorstöd. Här listas de viktigaste:

CBT – computer based testing  
CBA – computer based assessment  
CAA – computer aided assessment  
CAT – computer adaptive testing  
CMT – computer mastery testing  
E-assessment

Specifikt för språkbedömning förekommer även CALT – computer assisted language testing och WBLA – web-based language assessment.

Det är angeläget att hitta en lämplig svensk terminologi, och kategorierna *datorbaserade prov*, *datorbaserade adaptiva prov* och *datorstödda*

*prov* kan vara lämpliga för att beksriva de viktigaste särdragen hos prov som använder sig av IKT. Med datorbaserade prov (DBP) avses alla formella bedömningssituationer som använder datorn som det huvudsakliga mediet för att ställa frågor och ge svar. Med datorbaserade adaptiva prov (DBAP) avses sådana datorbaserade prov som har en adaptiv funktion. Datorstödda prov (DSP) syftar på sådana bedömningssituationer som använder sig av en dator (inkl. avancerade miniräknare) som en del av arbetet med provuppgifterna, men där svaren redovisas utanför datormediet.

## Fördelar och möjligheter med datorbaserade prov

Litteraturen om datorbaserade prov lyfter fram ett antal fördelar och möjligheter med att använda datorn och annan modern teknik i samband med pedagogiska bedömningar. Den lista på sådana fördelar som presenteras i tabell 2 gör inga anspråk på att vara heltäckande, men lyfter fram ett antal viktiga aspekter.

**Tabell 2** Fördelar och möjligheter med datorbaserade prov

Fördel/möjlighet	Beskrivning och källor
Datorn kan möjliggöra bedömning av andra kunskaper än prov med papper och penna	Det handlar om nya innovativa uppgifter som kan mäta andra aspekter av kunskap än traditionella uppgifter (Rudner, 1998; Sireci, 2004; Tian, Miao, Zhu & Gong 2007; Wise & Plake, 1990; Zenisky & Sireci, 2002), (Hamilton, Klein & Lorie, 2000). För en sammanställning av uppgiftstyper i datorbaserade prov, se (Zenisky & Sireci, 2002). Russell (2006) menar att "computers have the potential to alter the types of knowledge and skills that are assessed. Rather than focusing only on the end-product that students produce as they work on a test, computers make it possible to collect information about the processes and decisions students make as they solve problems" (sid. 8). Ett annat exempel är (Ridgway, McCusker & Pead, 2004) som skriver om det vidare begreppet "e-assessment" som de hävdar kan användas för att bedöma "nya" utbildningsmål: "Interactive displays which show changes in variables over time, microworlds and simulations, interfaces that present complex data in ways that are easy to control, all facilitate the assessment of problem-solving and process skills such as understanding and representing problems, controlling variables, generating and testing hypotheses, and finding rules and relationships. ICT facilitates new representations, which can be powerful aids to learning" (sid. 3).

<b>Fördel/möjlighet</b>	<b>Beskrivning och källor</b>
Ökad autenticitet och samstämmighet	Om datorer och annan modern teknik används i undervisningen och om de dessutom är centrala i användningen av kunskapen utanför skolan, så kan datoranvändning i samband med prov öka samstämmigheten mellan vad som bedöms och vad som är eftersträvansvärd kunskap. "Paper and pencil tests can be made more authentic by allowing students to word process essays, or to use spreadsheets, calculators or computer algebra systems in paper-based examinations". (Ridgway, McCusker & Pead, 2004, sid. 3) Studier har också visat att elever som är vana att skriva på datorn visar betydligt sämre resultat på skrivuppgifter som ska redovisas med papper och penna jämfört med vad de visar om samma frågor administreras i en dator (se till exempel Russell & Haney, 1997).
Omedelbar återkoppling	Datorbaserade prov kan möjliggöra att provtagaren får se resultatet på provet omedelbart efter avslutat prov. Provtagarna kan även få direkt feedback på inkorrekta svar och på deras provpoäng (Wise & Plake, 1990). När det gäller elevers upplevelser av datorbaserade prov rapporterar Pino-Silva (2008) om den studie där de två vanligaste fördelarna med datorbaserade prov var den omedelbara tillgången till provresultatet och det snabba och bekväma sättet på vilket provet besvarades och administrerades.
Eleverna positiva	Forskning har visat att elever kan uppfatta datorbaserade prov mycket positivt (Pino-Silva, 2008).
Flexibilitet	Datorbaserade prov är flexibla när det gäller konstruktion, upplägg och val av uppgifter (Weiss & Kingsbury, 1984; Sireci, 2004), samt genom att de kan ges när provtagarna vill och inte nödvändigtvis vid ett givet provtillfälle (Rudner, 1998; Tian, Miao, Zhu & Gong, 2007). Vidare kan provtiden enkelt anpassas till olika provtagare (Rudner, 1998; Tian, Miao, Zhu & Gong, 2007).
Säkerhet	Genom datorisering kan testsäkerheten ökas genom att uppgifterna kan ges i olika ordning till provtagarna (Linn, Graue & Sanders, 1990; Zenisky & Sireci, 2002), och genom en infrastruktur där inga papperskopior behöver cirkulera samt att det blir enklare att hålla reda på hur ofta en uppgift administrerats (Rudner, 1998; Sireci, 2004; Tian, Miao, Zhu & Gong (2007).

<b>Fördel/möjlighet</b>	<b>Beskrivning och källor</b>
Standardisering	När det är viktigt att alla provtagare har samma villkor för att göra ett prov, skapar datorbaserade prov större möjligheter att säkerställa att till exempel tid och hjälpmedel är desamma för alla provtagare. Datorbaserade prov kan ge en större standardisering av provgenomförandet och av proven (Wise & Plake, 1990).
Effektivitet	Det finns studier som visar att provtiden tenderar att reduceras med ett datorbaserat prov (English, Reckase, & Patience, 1977; Wise & Plake, 1990). Även Russell (2006) pekar på effektivitetsvinster med datorbaserade prov, och att detta har varit en viktig drivkraft bakom utvecklingen av DBP.

### Nackdelar och svårigheter med datorbaserade prov

De som skriver om datorbaserade prov tenderar att vara relativt positiva till möjligheterna som datorn skapar i förhållande till bedömning. Det kan bidra till att listan med fördelar och möjligheter är ganska lång, medan den lista med nackdelar och svårigheter som presenteras här är kortare (tabell 3). Listan på nackdelar gör liksom listan på fördelar inga anspråk på att vara heltäckande.

**Tabell 3** Nackdelar och svårigheter med datorbaserade prov

<b>Nackdel/svårighet</b>	<b>Beskrivning och källor</b>
Tillgång till datorer	Datortillgång samt datorernas prestanda kan begränsa möjligheten att genomföra datorbaserade prov och även begränsa vilka provuppgifter som kan användas. Exempelvis kan långa texter eller artistiska arbeten vara svåra att hantera i datorn (Rudner, 1998; Tian, Miao, Zhu & Gong, 2007). Vidare kan olika skärmtyper skapa svårigheter i standardiserade prov, till exempel genom att vissa elever kan se hela skärmbilden och andra måste skrolla för att hitta information.

<b>Nackdel/svårighet</b>	<b>Beskrivning och källor</b>
Datorvana	Datorvana kan spela roll för hur väl datorbaserade prov ger elevernas möjlighet att visa vad de vet och kan göra (Rudner, 1998; Tian, Miao, Zhu & Gong, 2007), och vana att skriva med tangentbord kan ha betydelse i uppsatsprov (Hamilton, Klein & Lorie, 2000). Russell (1999) noterade även att maskinskrivningshastighet var en god prediktor av provtagarnas resultat på öppna svar på språk- och naturvetenskapsprov på Massachusetts Comprehensive Assessment System samt National Assessment of Educational Progress (NAEP). Även Russell & Haneys (1997) studie visade att provtagare som var vana vid datorer presterade bättre på öppna frågor (till exempel uppsatser) om de fick använda en dator än på ett vanligt prov med papper och penna.
Testsäkerhet	Testsäkerhet kan bli ett problem om uppgifterna inte byts tillräckligt ofta, speciellt om det är ett prov som betyder mycket för provtagarna, till exempel påverkar deras betyg i hög grad (Linn, 2000).
Vana	Elever behöver anpassa sig till en annan form för bedömning (Pino-Silva, 2008). Om eleverna är vana vid att arbeta med papper och penna kan det vara en stor omställning att i stället jobba i datorn.
Läsning på skärmen	En visuell trötthet kan infinna sig som en följd av alltför mycket läsning från dataskärmen (Pino-Silva, 2008).
Begränsning av vad som bedöms	Datorbaserade prov riskerar att begränsa vad som kan bedömas om uppgiftsformaten är begränsade. Datorisering kan till exempel innebära att komplexa mål inte utvärderas, eftersom de inte så lätt låter sig fångas i de uppgiftsformat som används i datorn. Problemet är särskilt stort om datorn ska rätta provet, eftersom det inte finns några riktigt väl fungerande och ekonomiskt försvarbara system för automatisk rättning av mer komplexa uppgifter, till exempel essäer.
Kostnadseffektivt men inte kvalitetshöjande	Det är inte ovanligt att datorbaserade prov endast är direkta överföringar av prov för papper och penna till datorskärmen, och att de därmed inte erbjuder någon utveckling av elevernas möjligheter att visa vad de vet och kan göra. En sådan datorisering kan spara pengar i form av bland annat tryckkostnader, men riskerar att lida av datoriseringens problem utan att ge några positiva effekter. Dragow, Luecht och Bennett (2006) konstaterar att erfarenheterna visar att en övergång till datoriserade prov inte minskar kostnaderna.

## Adaptiva prov

Idén om prov som på olika sätt anpassas till den som genomför provet är inte ny. Redan Alfred Binets prov från 1908, som anses vara starten på det mer systematiska arbetet med prov, hade en adaptiv aspekt. Syftet med Binets skalor var att identifiera elever som inte drog fördel från ett typiskt utbildningssystem. Idén att elever som inte kan besvara en enkel fråga troligtvis inte kan besvara en svår fråga ledde till att Binet sorterade uppgifterna i stigande svårighetsgrad. Utifrån elevernas svar så avbröts provet när han hade tillräckligt med information om en elev. Först när datorerna utvecklades kunde adaptiva prov användas i storskaliga mätningar.<sup>13</sup> Forskning om datorbaserade adaptiva prov har därför en relativt kort historia, från 70-talet och framåt.<sup>14</sup>

## Datorbaserade adaptiva prov

Datorbaserade adaptiva prov bygger alltså på att en dator används för att anpassa provet till provtagarens förmåga. I stället för att ge varje deltagande individ samma uppgifter, så programmeras datorn att från en uppgiftsbank välja de uppgifter som mest effektivt ger information om varje individs förmåga. På grundval av de svar som provtagaren lämnat beräknar datorn ett värde på individens förmåga och väljer sedan nästa uppgift utifrån den informationen.<sup>15</sup> I praktiken innebär detta att en provtagare som besvarar en uppgift korrekt får en svårare uppgift medan en provtagare som besvarar en uppgift inkorrekt får en enklare uppgift. Utifrån svaren på uppgifterna uppdateras ett mått på provtagarens förmåga allteftersom. Datorn fortsätter att välja ut uppgifter till dess ett så kallat slutkriterium har uppnåtts.<sup>16</sup> Detta slutkriterium handlar ofta om noggrannheten i skattningen av individens förmåga, det vill säga när individens förmåga kan uppskattas med tillräcklig noggrannhet så avbryts provet, men det kan också innehålla andra villkor som gör att datorn avbryter proceduren att välja uppgifter. Beroende på provtagarnas svar

---

<sup>13</sup> Sireci (2004)

<sup>14</sup> En startpunkt anses vara den första officiella konferensen om CAT (Clark, 1975), den första boken om CAT (Weiss, 1983) eller när CAT började användas i större skala (ex. Graduate Record Examination i USA). Det har även skrivits en del böcker om CAT (Eggen, 2004; Sands, Waters, & McBride, 1997; van der Linden & Glas, 2000; Wainer, 2000; Wainer, Dorans, Flaugher, Mislevy, Green, Steinberg, & Thissen, 2000) samt olika vetenskapliga sammanfattningar (se ex. Meijer & Nering, 1999; Tian, Miao, Zhu, & Gong, 2007)

<sup>15</sup> Weiss & Kingsbury, 1984 och van der Linden & Glas (2000)

<sup>16</sup> Thissen & Mislevy (2000)

kan färre uppgifter behövas i ett datorbaserat adaptivt prov än om ett traditionellt prov med fast antal uppgifter används.<sup>17</sup>

### Fördelar med datorbaserade adaptiva prov

Utöver de fördelar som nämndes vid datorbaserade prov så har även användandet av datorbaserade adaptiva prov i sig ett antal specifika fördelar som sammanfattas här.

- (1) Information om provtagare kan ges på hela skalan av förmåga, med god mätprecision. I prov med ett givet antal uppgifter sätter oftast provtiden hinder ivägen för detta, eftersom det är omöjligt att ha tillräckligt antal uppgifter på alla förmågenivåer i ett och samma prov. Genom att datorn väljer ut de uppgifter som är lämpligast för varje provtagares förmåga kan man åstadkomma mätprecision på alla nivåer.
- (2) Det behövs vanligen färre uppgifter i ett datorbaserat adaptivt prov än i ett motsvarande prov med ett fast antal uppgifter. Med bibehållen tillförlitlighet i förmågemätningen kan ett datorbaserat adaptivt prov vara upp till 50 procent kortare.<sup>18</sup> Detta medför att provtagarna inte slösar tid på för svåra eller för enkla uppgifter<sup>19</sup> och kan minimera provtagarnas frustration, uttråkning och provångest.<sup>20</sup> Dessutom gör det provtagarna mindre uttrötade, vilket är en faktor som påverkar deras prestationer på prov.<sup>21</sup>
- (3) Fuskrisken kan minskas, eftersom två provtagare som sitter bredvid varandra vid provtillfället inte får samma prov, vilket ökar test-säkerheten.<sup>22</sup>
- (4) Problemet med att provuppgifter blir kända, trots att de är tänkta att hemlighållas för att till exempel kunna återanvändas, kan minska med hjälp datorbaserade adaptiva prov. Detta beror på att olika provtagare får olika uppgifter samt att de får färre uppgifter, vilket minskar antalet exponerade uppgifter.<sup>23</sup>

---

<sup>17</sup> Weiss & Kingsbury (1984)

<sup>18</sup> Meijer & Nering (1999), Wang & Kolen (2001), Weiss & Kingsbury (1984), Weiss (1983)

<sup>19</sup> Bunderson, Inouye, & Olsen (1989), Sireci (2004)

<sup>20</sup> Hamilton, Klein & Lorie (2000)

<sup>21</sup> Rudner (1998), Tian, Miao, Zhu & Gong (2007)

<sup>22</sup> Hamilton, Klein & Lorie (2000) och Rudner (1998), i Tian, Miao, Zhu & Gong (2007)

<sup>23</sup> Notera dock att det finns en risk att uppgifter av medelsvårighetsgrad kan exponeras mer om alla provtagare får en sådan typ av uppgift i inledningen av ett prov (Thissen & Mislevy, 2000).

- (5) Datorbaserade adaptiva prov skapar goda möjligheter att utvärdera provtagarnas kunskapsutveckling över tid, eftersom provet kan administreras vid flera tillfällen till samma provtagare utan att provtagaren behöver möta samma uppgifter ur uppgiftsbanken.<sup>24</sup>

#### Nackdelar med datorbaserade adaptiva prov

Även om datorbaserade adaptiva prov har många fördelar finns det även ett antal nackdelar. Förutom de nackdelar som delas av alla datorbaserade prov har de adaptiva varianterna sina egna problem.

- (1) I datorbaserade adaptiva prov tenderar provtagare med samma förmåga att få samma provuppgifter och dessutom tenderar provtagare att minnas provuppgifter och dela med sig av informationen till potentiella provtagare.<sup>25</sup> Det säkraste sättet skulle vara att ge slumpmässigt valda uppgifter, så att vissa uppgifter inte överutnyttjas, men då blir det problem med att tillämpa den adaptiva grundidén. Testsäkerheten kan även ökas i datorbaserade adaptiva prov om fler uppgifter ges och om man utformar uppgifterna så att de är svårare att minnas och återberättas för potentiella provtagare.<sup>26</sup>
- (2) Provtagare vill gärna ha en möjlighet att revidera sina givna svar,<sup>27</sup> men att i ett adaptiv prov skapa möjlighet för provtagare att gå tillbaka och revidera svar på givna uppgifter är komplicerat.<sup>28</sup> Orsaken är att när en provtagare besvarar en uppgift fel så ges oftast en enklare uppgift. Om en provtagare medvetet skulle välja att besvara en mängd uppgifter felaktigt skulle uppgifterna bli enklare och enklare. Om samma provtagare då skulle ha möjlighet att gå tillbaka och ändra sina svar, och svara korrekt på ett antal frågor som är enklare än de frågor som provtagaren egentligen skulle ha fått, skulle förmågan överskattas.<sup>29</sup>
- (3) För att kunna skapa ett bra datorbaserat adaptivt prov krävs att uppgifternas egenskaper (svårighetsgrad, diskrimineringsförmåga etc.) är kända. Detta innebär i sin tur att alla uppgifter som ska

---

<sup>24</sup> Hamilton, Klein & Lorie (2000)

<sup>25</sup> Guo (2007); Thissen & Mislevy (2000), van der Linden & Glas (2000)

<sup>26</sup> Guo (2007) ger exempel på uppgifter som är svårare att dela med sig, eftersom de innehåller både mer text och mer förståelse i uppgifterna.

<sup>27</sup> Vispoel, Rocklin & Wang (1994)

<sup>28</sup> Sireci (2004)

<sup>29</sup> Rudner (1998), Tian, Miao, Zhu, & Gong (2007), Wainer (1993), Weiss & Kingsbury (1984)



användas måste prövas ut. Detta bör göras inom befintliga prov, det vill säga utprovningssuppgifter och skarpa uppgifter blandas i provet men provtagarna får bara sitt resultat utifrån de skarpa uppgifterna.<sup>30</sup> Vi kan dock inte enbart använda oprövade uppgifter i ett adaptivt prov, eftersom vi då inte kan skatta provtagarnas förmåga. I stället kan en del uppgifter vara nya men merparten bör vara uppgifter som vi har kunskap om. Vi behöver ofta uppemot tusen provtagare för att kunna skatta uppgifternas karakteristika med lämplig precision.<sup>31</sup> Vidare är det inte säkert att uppgifternas egenskaper är desamma i ett datorbaserat adaptivt prov som när de används i prov som besvaras med papper och penna.<sup>32</sup>

- (4) Datorbaserade adaptiva prov är inte lämpliga för att mäta alla typer av kunskap och förmågor. Vanligen baseras de adaptiva proven på modern testteori, men det finns inte alltid någon lämplig modell för att beskriva den förmåga eller kunskap som är av intresse.<sup>33</sup>
- (5) Eftersom olika provtagare får olika provuppgifter så kan provtagare uppfatta det som att proven är orättvisa om inte tillräckligt många uppgifter ges för att ringa in provtagarnas förmåga.<sup>34</sup>
- (6) Jämfört med andra prov innebär datorbaserade adaptiva prov en ökad utvecklings- och administrationskostnad i och med att data-program måste skapas och att provet måste administreras. Vidare måste en stor uppgiftsbank skapas för att det adaptiva provet ska bli säkert och fungera bra, och den måste underhållas på alla nivåer av provtagarnas förmåga. Detta kräver både tid och pengar.<sup>35</sup> Testsäkerheten kan bli problematisk om inte uppgiftsbanken uppdateras kontinuerligt.<sup>36</sup> Eftersom välfungerande uppgifter tenderar att användas mycket så krävs ständiga uppdateringar av uppgiftsbanken för att garantera testsäkerheten. I den bästa av världar skapar provkonstruktörerna enbart bra uppgifter, men eftersom vi inte vet det förrän uppgifterna är utprovade så måste en mängd uppgifter produceras. Det finns även en kostnad i själva provgenomförandet,

---

<sup>30</sup> Thissen & Mislevy (2000)

<sup>31</sup> Wainer & Mislevy (2000)

<sup>32</sup> Rudner (1998), Tian, Miao, Zhu & Gong (2007)

<sup>33</sup> Rudner (1998), Tian, Miao, Zhu & Gong (2007)

<sup>34</sup> Rudner (1998), Tian, Miao, Zhu & Gong (2007)

<sup>35</sup> Guo (2007), Sireci (2004)

<sup>36</sup> Meijer & Nering (1999)

eftersom det handlar om hur många datorer som finns tillgängliga. Om fler datorer finns tillgängliga så behövs färre provtillfällen, vilket leder till en ökad testsäkerhet.<sup>37</sup>

- (7) Även om varje uppgift kan granskas innan den ges till en provtagare så kan inte varje provversion utvärderas av experter i förväg, eftersom vi inte vet hur den kommer se ut i och med att vi inte vet hur provtagarna besvarar provet i förväg.<sup>38</sup>

### Uppbyggnad av datorbaserade adaptiva prov

En första förutsättning för ett fungerande datorbaserat prov är en uppgiftsbank med välkalibrerade uppgifter för hela spektrat av den förmåga provet vill mäta.<sup>39</sup> Uppgifternas egenskaper beskrivs vanligen med hjälp av så kallad modern testteori (Item Response Theory, IRT) utifrån svårighetsgrad, diskrimineringsförmåga och möjlighet att gissa det korrekta svaret.<sup>40</sup> och denna information används för att bestämma vilka uppgifter som väljs ut och ges till provtagaren.

För att kunna bestämma uppgifternas egenskaper krävs resultat från ett ganska stort antal elever, för att det över huvud taget ska vara möjligt att göra de statistiska beräkningarna, och ju mer avancerad statistisk modell som används desto fler elevresultat krävs. Det behövs vidare en stor uppgiftsbank för att den adaptiva processen ska fungera, men det är svårt att säga exakt hur stor den behöver vara. Kravet på uppgiftsbankens storlek påverkas bland annat av provets längd, hur många gånger uppgiften får användas innan den anses för exponerad, och den optimala fördelningen av uppgifternas svårighetsgrad.<sup>41</sup> Vidare måste innehållet i uppgiftsbanken matcha provspecifikationerna, det måste finnas tillräckligt många uppgifter inom varje område, samt en lämplig fördelning av uppgifternas svårighetsgrad.

Utvecklingen av nya uppgifter tar mycket tid och medför en stor kostnad.<sup>42</sup> Dessutom måste uppgiftsbanken underhållas genom tillägg av nya uppgifter, eftersom uppgifter som används exponeras så mycket att de blir kända och därför måste utgå ur uppgiftsbanken. Det behövs

---

<sup>37</sup> Guo (2007)

<sup>38</sup> Luecht & Nungester (1998), Sireci (2004)

<sup>39</sup> Wright & Bell (1984)

<sup>40</sup> Weiss & Kingsbury (1984), Wise (1997)

<sup>41</sup> Hamilton, Klein & Lorie (2000)

<sup>42</sup> Hamilton, Klein & Lorie (2000)

därför strategier både för att ta bort och lägga till uppgifter i banken.<sup>43</sup> Ett annat underhåll av uppgiftsbanken handlar om att provuppgifternas karakteristika kan förändras över tid, vilket innebär att uppgifterna kan behöva kalibreras igen.<sup>44</sup>

Genomförandet av ett datorbaserat adaptivt prov behöver en startpunkt i form av någon eller några uppgifter. Det finns möjlighet att använda olika strategier vid val av startpunkt i provet, beroende på vilket syfte man har. När datorn efter startuppgifterna ska välja uppgifter som passar provtagarens förmåga finns en mängd olika metoder att välja bland, så kallade algoritmer. Flera av dessa algoritmer använder sig av modern testteori, eftersom uppgifterna och provtagarnas förmåga kan läggas på en gemensam skala.<sup>45</sup> Den grundläggande principen för alla algoritmer är att datorn väljer en uppgift som ger maximal information om provtagarens förmåga. Valet av uppgifter bygger dels på den beräkning av provtagarens förmåga som datorn gör efter varje besvarad uppgift och dels på de egenskaper som beskrivits för varje uppgift i uppgiftsbanken.

Eftersom vilka provuppgifter som används i ett datorbaserat adaptivt prov varierar, så måste man kontrollera uppgifternas exponering för att undvika att en enskild uppgift används för ofta och därmed blir känd för potentiella provtagare. För att minska uppgiftsexponeringen kan man kontrollera vilken uppgift som administreras, hur den väljs ut i uppgiftsbanken<sup>46</sup> samt notera märkliga svarsmönster efter att den har administrerats.

Bedömningen av provtagarens svar måste kunna göras direkt av datorn i ett datorbaserat adaptivt prov, och hur man väljer att sätta poäng på provtagarnas svar har betydelse för hur rättningen går till och hur resultatet tolkas. Efter varje uppgift som en provtagare besvarat så rättas den och provtagarens skattade förmåga uppdateras, vilket kan ske med olika metoder. Det slutgiltiga provresultat som ges till en provtagare är skapat utifrån den slutgiltiga skattningen av provtagarens förmåga. Generellt bör processen fortsätta så länge som informationen om provtagarens förmåga anses otillräckliga, men för att provet inte ska pågå i oändlighet eller tills alla uppgifter i uppgiftsbanken har använts, så

---

<sup>43</sup> Wise (1997)

<sup>44</sup> Wise (1997).

<sup>45</sup> Thissen & Mislevy (2000), van der Linden & Glas (2000), Weiss & Kingsbury (1984)

<sup>46</sup> McBride och Martin (1983)

behövs ett tydligt avslutningskriterium. Det finns en rad olika varianter av avslutningskriterier och valet av kriterium bör grunda sig på provets syfte. Ett målrelaterat prov<sup>47</sup> har till exempel syftet att undersöka om en provtagare uppnått en viss nivå av kunskap. Provets avslutas då vi med säkerhet vet att provtagaren har uppnått en given nivå eller inte.<sup>48</sup> Notera dock att vissa datorbaserade adaptiva prov alltid ges med ett förutbestämt antal uppgifter oavsett hur provtagaren besvarat uppgifterna.<sup>49</sup> En orsak till det kan vara att man vill undvika stora skillnader i antalet besvarade provuppgifter mellan provtagarna.

### Uppgifter i datorbaserade adaptiva prov

Att använda datorbaserade adaptiva prov öppnar möjligheter till användandet av olika former av innovativa provuppgifter. Denna möjlighet har dock inte alltid utnyttjats, trots att mer komplicerade uppgifter bland annat minskar risken att provtagaren kan gissa det korrekta svaret.<sup>50</sup> Uppgiftstyper kan delas upp i kategorierna *valda svar* samt *konstruerade svar*.<sup>51</sup> Inom den första kategorin väljer provtagarna sitt svar från en lista av presenterade svar, till exempel flervalssuppgifter, medan inom den senare kategorin skapar provtagaren ett svar, till exempel en uppsats. Beror på vilka uppgiftstyper som väljs så behöver man fundera över hur de ska kunna rättas maskinellt. Uppgifter med valda svar samt kortare konstruerade svar är ofta relativt enkla att maskinrätta om tydliga rättningssmallar skapas. Om man däremot använder konstruerade svar var det inledningsvis svårt att maskinrätta dem men detta har förändrats i takt med att bättre verktyg har skapats.<sup>52</sup> Vad gäller att bedöma vad som är korrekt eller inkorrekt svar vid en maskinrättning där provet använde ett automatiskt expertsystem har det i en studie visat sig överensstämma relativt bra med manuell rättning av algebraiska ordproblem.<sup>53</sup> I tabell 4 och 5 ges exempel på uppgifter inom de två kategorierna med referens till studier om dem.<sup>54</sup> Notera att en del uppgifter används i skarpa prov medan andra enbart är föreslagna som möjliga att använda.

---

<sup>47</sup> För en introduktion till denna typ av adaptiva prov se Parshall, Spray, Kalohn & Davey (2006) eller Thompson (2007)

<sup>48</sup> Lewis & Sheehan (1990)

<sup>49</sup> Sireci (2004)

<sup>50</sup> Parshall, Davey & Pashley (2000)

<sup>51</sup> Hamilton, Klein & Lorie (2000)

<sup>52</sup> Hamilton, Klein & Lorie (2000), Parshall, Davey & Pashley (2000), Zenisky & Sireci (2002)

<sup>53</sup> Bennett & Sebrecths (1997)

<sup>54</sup> Fler studier om innovativa uppgifter finns i Zenisky och Sireci (2002)

**Tabell 4** Uppgiftsformat inom kategorin "valda svar", kort beskrivning samt exempel på studier.

Uppgiftsformat	Beskrivning	Exempel på studie(r)
Mångvalsfrågor	Välj alla korrekta svarsalternativ.	Koch (1993)
Välj och placera	Provtagarna väljer ett objekt och drar det till korrekt svarsfält.	Luecht (2001)
Dra eller samman-koppla förhållanden	Provtagarna identifierar par-förhållanden mellan givna objekt.	Luecht (2001)
Flytta figurer eller symboler i bilder	Provtagarna ska flytta in till exempel kurvor i en figur eller graf utifrån en given situation.	Martinez (1991)
Sorteringsuppgift	Utifrån givna kategorier eller prototyper ska provtagarna matcha uppgifter som har likheter med dem.	Bennett & Sebrechts (1997)
Ordna information	Ordna information ex. från minst till störst.	Luecht (2001)
Sätta in text	Provtagarna drar och släpper text i sektioner utifrån något kriterium (ex. var den passar bäst in).	Taylor, Jamieson, Eignor, & Kirsch (1998)
Concept mapping	Provtagare visar sin kunskap om förhållanden mellan datapunkter genom att grafiskt representera bilder och texter på skärmen genom länkar och noder.	Chung, O'Neil & Herl (1999)
Markera texter	Provtagarna väljer ut specifika meningar i en sektion utifrån syfte eller speciell information.	Taylor, Jamieson, Eignor, & Kirsch (1998)
Multipel selektion	Provtagarna presenteras för stimuli (text, ljud, bild) och väljer svar från en lista. (Svaren kan användas flera gånger i en serie av uppgifter.)	Ackerman, Evans, Park, Tamassia & Turner (1999)
Rangordna korrekthet	Rangordna delar utifrån grad av korrekthet.	Crocker & Algina (1986)
Editering av en sektion	Provtagarna editerar en kort sektion text. Inbegriper val av om att förändra texten utifrån olika alternativ eller låta den stå som den presenteras.	Davey, Godwin & Mittelholtz (1997)

**Tabell 5** Uppgiftsformat inom kategorin "konstruerade svar", kort beskrivning samt exempel på studier.

Uppgiftsformat	Beskrivning	Exempel på studie(r)
Fånga eller välja ramar	Datormusen används för att välja delar av en bild, karta eller graf utifrån direktiv.	Hambleton (1997)
Grafisk modellering	Provtagare skissar med kurvor och linjer på ett ruttmönster utifrån en given situation.	Bennett, Morley, Quardt, & Rock (2000)
Interagera med figurer eller symboler i bilder	Provtagarna ska manipulera eller interagera med en figur eller bild i en given situation.	Martinez (1993)
Matematiska uttryck	Provtagarna skriver in ett unikt uttryck för att uttrycka ett förhållande.	Martinez & Bennett (1992)
Numeriska ekvationer	Provtagarna slutför numeriska meningar med nummer eller matematiska symboler.	Hambleton (1997)
Multipla numeriska svar	Provtagarna skriver in flera numeriska svar, ex. i en tabell.	Hambleton (1997)
Analysera situationer	Utifrån korta texter samt ljud/visuella klipp ska provtagarna fatta ett beslut.	Ackerman, Evans, Park, Tamassia & Turner (1999)
Generering av multipla lösningar	Utifrån en given situation ska provtagarna leverera möjliga lösningar eller förklaringar.	Bennett & Rock (1995)
Generering av exempel	Utifrån givna situationer så skapar provtagarna exempel eller restriktioner.	Bennett, Quardt, Rock, Singley, Katz, & Nhouyvanisvong (1999)
Uppsats/korta svar	Med eller utan restriktion på längden.	Burstein, Kukich, Wolff, Lu & Chodorow (1998)
Problemlösningssituationer	Problemlösningssituationer som betygsätts utifrån skapad lösning.	Luecht (2001)
Sekventiell problemlösning/rollspel	Provtagarna ges en serie med respons utifrån en dynamisk situation. Både process och produkt poängsätts.	Clauser, Harik & Clyman (200)

Vid val av uppgiftstyp bör man förutom att beakta syftet med uppgiften även beakta vilka hjälpmedel provtagarna har att besvara uppgiften med, till exempel datormus, tangentbord, joystick, pekskärm, samt papper och penna. Om man till exempel använder en mus så kan provtagaren uppmanas kryssa för rätt svar eller dra ner olika menyer.<sup>55</sup> På liknande sätt används olika hjälpmedel om provtagaren ska skriva ett svar, navigera på en karta, peka på korrekt svar eller markera ett stycke i en text. Provtagaren kan även uppmanas att markera ett felaktigt stycke i en text och korrigerar det.<sup>56</sup> Notera att pekskrämar kräver mindre datorvana än användandet av till exempel mus.<sup>57</sup>

Olika medier i provet kan hjälpa provtagarna att visualisera det efterfrågade problemet bättre. Med medier åsyftas bland annat ljud, grafik, film och animationer i uppgifterna. Ljud används framför allt vid musik eller språkprov, som TOEFL.<sup>58</sup> Provtagarna kan även visa på olika grader av interaktivitet. Exempel inkluderar att provtagaren pekar på rätt svar, markerar en text, flyttar en del av ett svar, förändrar bilder och så vidare.<sup>59</sup> Notera dock att det är ovanligt med mycket interaktivitet i ett prov. Ett exempel på ett undantag är dock det prov som administreras från National Council of Architectural Registration Boards. Provet visar en bild och ett arkitektproblem som provtagaren ska lösa med hjälp av datorns ritverktyg.<sup>60</sup> En orsak till att det sällan används är att det kan vara svårt att poängsätta prov med mycket interaktivitet. En fråga som man måste ta ställning till är exempelvis om en provtagare ska få lika många poäng för att identifiera ett fel som om provtagaren lyckas korrigera felet korrekt.

### Varianter av datorbaserade adaptiva prov

I den mest utpräglade formen av datorbaserade adaptiva prov får provtagaren en uppgift i taget och beroende på svaret väljer datorn nästa uppgift. Det finns dock varianter av datorbaserade adaptiva prov där den adaptiva aspekten är mer begränsad av till exempel innehållsmässiga skäl eller på grund av uppgiftsexponering eller säkerhetskrav.<sup>61</sup>

---

<sup>55</sup> Baker & O'Neil (1995)

<sup>56</sup> Davey, Godwin & Mittelholtz (1997), Martinez (1993)

<sup>57</sup> Parshall, Davey & Parshley (2000)

<sup>58</sup> Parshall, Davey & Parshley (2000)

<sup>59</sup> Zenisky & Sireci (2002)

<sup>60</sup> Parshall, Davey & Pashley (2000)

<sup>61</sup> Thissen & Mislevy (2000), van der Linden (2000)

En variant går ut på att grupper av uppgifter väljs ut i stället för enskilda uppgifter, vilket leder till att provet blir uppdelat i provdelar eller paneler som kan variera i antal.<sup>62</sup> Denna typ av datorbaserade adaptiva prov kallas *multistage testing* (MST) eller *computerized adaptive sequential test* (CAST).<sup>63</sup>

I en annan variant ges uppgifterna i så kallade *testlets* eller delprov. Varje delprov är i sig ett slags minitest,<sup>64</sup> och kan till exempel bestå av en grupp uppgifter som hör till samma text, samma figur eller samma tabell. Provtagarens förmåga beräknas efter att den besvarat ett delprov i stället för en enda uppgifter. Utifrån provtagarens skattade förmåga väljs sedan ett nytt delprov ut som administreras till provtagaren. Detta innebär att provet är adaptivt på delprovsnivå och inte på uppgiftsnivå.<sup>65</sup> Tre orsaker till att använda delprov är att de kan minska bekymren med den atomistiska naturen som enskilda uppgifter har, minska effekten av att uppgifternas sammanhang varierar (varje delprov bär med sig sitt eget sammanhang), samt öka provets effektivitet.<sup>66</sup>

---

<sup>62</sup> Luecht & Nungester (1998), van der Linden (2000), Sireci (2004)

<sup>63</sup> För en introduktion se Mead (2006) samt Luecht, Brumfield och Breithaupt (2006)

<sup>64</sup> Wainer & Kiley (1987), Wainer & Lewis (1990)

<sup>65</sup> Sireci (2004)

<sup>66</sup> Wainer, Bradlow & Du (2000)



# **Datorbaserade prov i praktiken (exempel)**

## Datorbaserade prov i praktiken (exempel)

Runt om i världen finns det exempel på datorbaserade adaptiva prov i användning. I USA finns till exempel Test of English as a Foreign Language (TOEFL), Graduate Record Examinations (GRE) och Graduate Management Admission Test (GMAT). Vidare har Northwest Evaluation Association (NWEA) flera datorbaserade adaptiva prov inom olika ämnen med tillhörande uppgiftsbanks som är anpassade för elever i grundskolan och gymnasiet.<sup>67</sup> I Europa har Nederländernas CITO utvecklat ett antal adaptiva prov inom olika ämnen och Danmark och Storbritannien kan erbjuda prov som är datorbaserade och adaptiva. I detta avsnitt kommer ett mindre urval av datorbaserade adaptiva prov att beskrivas, med särskilt fokus på prov i matematik, naturvetenskap och språk.

Ett argument för datoriserade och adaptiva prov – förutom ekonomi, kortare provtider och rättningseffektivitet – är möjligheten att skapa andra typer av uppgifter än vad som går att ha via prov med papper och penna.<sup>68</sup> Datormediet innebär att det går att inkludera filmer, ljud och animeringar om man vill. Dessutom öppnar datoriseringen för möjligheten att redovisa svar på andra sätt. Flervalsfrågor behöver inte bara vara att markera korrekt svarsalternativ utan det kan även vara olika former av interaktiva aktiviteter. Zenisky och Sireci (2002) konstaterar att endast ett fåtal av de 21 svarstyper de listat används i praktiken. Vissa av dem finns till och med bara på forskningsstadiet. De har bara tagit med svarstyper där provtagarna ska använda mus och tangentbord, eftersom andra inmatningsmetoder till exempel tryckkänsliga skärmar, programvara som tolkar tal inte är speciellt vanliga. Teknikutvecklingen gör dock att till exempel tryckkänsliga skärmar blir vanligare och även billigare, vilket kan förändra läget på sikt. I de uppgiftsexempel från datoriserade prov som går att finna är det också bara ett fåtal av dessa olika svarstyper som finns representerade.

Idag har många internationella testföretag och organisationer datoriserade eller adaptiva prov i sina testprogram. Vad är det då för typer av uppgifter som ingår i de prov som finns tillgängliga? Det är inte speciellt svårt att finna enstaka exempel på uppgifter som skulle kunna ingå i

---

<sup>67</sup> Hamilton, Klein & Lorie (2000)

<sup>68</sup> Zenisky & Sireci (2002)

proven. Däremot finns inga exempel på något adaptivt system där man kan få möjlighet att testa hur det fungerar. Att ett komplett adaptivt system inte är möjligt att finna beror troligtvis på ekonomiska orsaker. Det kostar mycket pengar och arbete att bygga upp en välkalibrerad uppgiftsbank och enda anledningen till att släppa den vore för att provet inte ska användas längre. Å andra sidan blir provet i och med det ointressant att ha som exempel.

En begränsande faktor för vilken typ av uppgifter som går att använda i adaptiva prov, men även i många linjära datoriserade prov, är kravet på automatisk rättning.<sup>69</sup> Det är en nödvändighet i proven, eftersom valet av nästkommande uppgift hänger på om provtagaren svarade rätt eller fel på den aktuella uppgiften. Då krävs det att datorn kan rätta uppgiften utan mänsklig assistans om testsystemet ska vara effektivt.

### Datoriserade nationella prov

PISA-undersökningen 2000 kom som en obehaglig överraskning för många i Norge. Det visade sig att resultaten, mot allmän förväntan, enbart låg strax över medel; landet var klart efter Sverige och långt bakom Finland. För att komma tillrätta med problemet lanserades ett nationellt provsystem, där färdigheterna norska (läsa och skriva), matematik och engelska (läsa och skriva) ingick. Proven är *low stakes*<sup>70</sup> för provtagaren och syftet är att ge elever, lärare och skolledare nödvändig information för att befrämja pedagogisk utveckling, samt att förse lokala och nationella myndigheter och allmänheten med information som kan stimulera till dialog och utveckling av utbildningsstandarden. Det pedagogiska syftet togs dock bort i den förändring 2006 som beskrivs nedan.

Provutvecklingen började 2003. Under 2004 genomförde man storskaliga utprövningar över hela landet och under 2004 och 2005 gjordes proven av elever i 4, 7, 10 och 11 trinn, vilket motsvarar årskurs 3, 6, 9 och första året på gymnasiet i Sverige. Provet i engelska var datorbaserat, de övriga var baserade på papper och penna. År 2006 låg provverksamheten nere och vissa förändringar gjordes. Färdigheten skriva utslöts ur provet, eftersom interbedömarreliabiliteten ansågs för låg. Dessutom plockades gymnasieproven bort, delvis på grund av kraftiga protester mot centraliserade prov från gymnasieelevernas sida. Det var alltså inte

---

<sup>69</sup> Russel (2006)

<sup>70</sup> Med *low stakes* avses att provet inte har en avgörande betydelse för individen.

protester mot datorbaserade prov som gjorde att de försvann. Övriga prov flyttades från vår (4, 7 trinn) till höst (5, 8 trinn) men prövar målen för 4 trinn och 7 trinn (en sommar emellan). Anledningen till denna förändring var att lärarna kände att även deras arbete utvärderades och eftersom eleverna byter stadium och får nya lärare efter dessa årskurser blev detta mindre känsligt.

Något som också väckte starka reaktioner var att provresultaten offentliggjordes. Proven som sådana hade och har fortfarande stor acceptans och numera anses publiceringen positiv. I förlängningen kan ifrågasättandet initiera en behövlig förändringsprocess.

Lärarna var mycket nöjda med att de, tack vare den datorbaserade rättningen, inte fick en ökad arbetsbörda i och med provens införande och eleverna uppskattade snabb feedback (de får nämligen resultatet direkt). Vissa initialsvårigheter med föråldrad datorpark och dålig bandbredd fanns, men de är nu i huvudsak åtgärdade och skolorna anser i allmänhet att det tekniska fungerar smärtfritt. Relativt stora skillnader i resultaten mellan olika skolor och delar av landet har noterats, däremot inte mellan pojkar och flickor. I de enkäter som gjorts upplever dock fler pojkar än flickor att proven är lätta. Vissa elever tycker att textmassan är stor.

Provkonstruktörernas reflektioner är att

- det varken är lätt eller billigt att framställa datorbaserade prov
- det är en utmaning att få administratörerna i utbildningsväsendet, politikerna och testkonstruktörerna att samverka
- det är nödvändigt att ha ett nära samarbete mellan lärare, konstruktörer av datasystemen och statistiker<sup>71</sup>
- skolorna är positiva till att delta i utprovningar
- det är en välsignelse att data genereras automatiskt.<sup>72</sup>

Övningsmaterial finns tillgängligt på Internet och alla testtagare uppmanas att bekanta sig med detta för att undvika att ovana vid distributionssätt, uppgiftstyper och instruktioner ska påverka resultatet.

---

<sup>71</sup> Hittills har såväl provkonstruktörer som datateknisk personal för engelskans del funnits vid universitetet i Bergen och för matematikens del vid universitetet i Trondheim, vilket har ansetts vara en stor fördel. Den tekniska sidan skall dock nu överföras till Utdanningsdirektoratet i Oslo.

<sup>72</sup> Moe (2008)

Danmark har nationella prov, bland annat i engelska i årskurs 7, i matematik i årskurs 3 och 6 och i NO-ämnena i årskurs 8. Proven har ingen avgörande betydelse för provtagaren och när utbyggnaden är klar kommer de att vara obligatoriska för skolorna att genomföra. Systemet kommer att finnas tillgängligt under en tidsperiod från februari till april och kapaciteten är cirka 6 000 samtidigt inloggade elever. Den enskilda läraren avgör när proven ska genomföras och det är inte nödvändigt att hela klassen gör provet samtidigt. Utöver detta erbjuds eleverna att göra två frivilliga prov under den övriga läsårstiden. Systemet kan på så vis mäta elevens utveckling.

Syftet är att ge information till elever och föräldrar om kursmålen har uppnåtts och att kartlägga i vad mån skolorna lyckas med sitt uppdrag. Det betonas dock att endast en del av elevens färdigheter mäts i dessa prov och att det är lärarens uppgift att kontinuerligt bedöma de övriga, till exempel testas inte egen produktion i engelska. I alla prov prövas dock tre så kallade profilmråden, vilket innebär att tre separata adaptiva prov pågår samtidigt.

De danska proven utgör ett fullständigt adaptivt prov, det vill säga de bygger på att enskilda uppgifter distribueras till provtagarna. Intressant är dock att provet inte avbryts när systemet har gjort en statistiskt säkerställd bedömning, utan eleverna har 45 minuter på sig att göra så många uppgifter som de hinner. Denna tid kan förlängas för elever som eventuellt inte har hunnit göra nödvändigt antal uppgifter.

Starten skedde i maj-juni 2007 med en reducerad uppgiftsbank. En revision hösten 2007 visade att banken inte var tillräckligt stor och att kvaliteten på en hel del uppgifter var för låg. Under 2008 utvecklades och utprövades nya uppgifter och inga prov genomfördes. Enligt planen skulle de nya provuppgifterna pilottestas under våren 2009 och systemet vara färdigt och i drift under våren 2010. Full effekt anser man sig uppnå efter några år då man kan följa elevers och gruppers utveckling över tid på ett sätt som är unikt i världen.

Resultaten på alla nivåer (individ, grupp, skola, kommun) är konfidentiella. Tanken med detta är delvis att undvika att undervisningen ändras så att man studerar för provet. De som berörs av resultaten har dock rätt att se dem. Det är alltså möjligt för till exempel en kommun att göra jämförelser mellan skolor, och för en rektor att se den egna skolans alla resultat. Läraren får tillgång till en detaljerad rapport om

varje elevs prestation, där det till exempel framgår vilka uppgifter varje elev har fått och det är på så vis möjligt för lärare och elever att gå igenom uppgifterna tillsammans. Föräldrarna ska också informeras och datorn genererar för detta ändamål automatiskt en utförlig redogörelse.<sup>73</sup> Kritik har riktats mot valet av ingenjörsfirman COWI som ansvarig för konstruktion av proven, samt mot de stora summor som lagts ner på dessa datorbaserade prov. Enligt kritikerna håller provuppgifterna inte måttet.<sup>74</sup>

Island håller på att utveckla adaptiva prov. I första omgången (2009) kombineras prov med papper och penna och datorbaserade prov för årskurs 10 i ämnena matematik, isländska och engelska. År 2010 kommer datorbaserade adaptiva prov att genomföras i samma årskurs och om försöket faller väl ut, kommer även årskurs 4 och 7 att omfattas. Syftet är att informera elever, föräldrar och lärare om hur väl eleverna uppnår målen, samt att se hur skolor och nationen som helhet uppfyller ställda krav.

*CITO* (Nederländerna) konstruerar nationella prov i bland annat språk och matematik för olika stadier. Proven kan fås antingen i en datorbaserad eller i en variant med papper och penna, där den datorbaserade används av cirka 90 procent. Tidigare distribuerades proven via Internet, men detta övergavs av säkerhetsskäl och numera tillhandahålls den datorbaserade versionen på cd. Ett Cito-prov används inför gymnasievalet. Detta är dock omdebatterat och cirka 20 procent av skolorna väljer alternativa test. Utveckling av nya prov för gymnasiets teoretiska linjer pågår med ett pilotprov 2010.

I USA har Virginia Department of Education, Texas Education Agency och Idaho State Board of Education sina prov online, med prov med papper och penna som alternativ. Skoldistriktet avgör vilken distributionsmetod som ska användas. Det vanligaste skälet att man väljer provet med papper och penna är begränsad tillgång till datorer. Språkproven är inte adaptiva.

Flera universitet och stora privata institut i USA har utvecklat adaptiva prov. De listas i det följande kapitlet ”Prov utvecklade vid universitet och andra institutioner”.

---

<sup>73</sup> Wandall (2008)

<sup>74</sup> Richter (2009)

För Skottlands del har Scottish Qualifications Authority (SQA) inlett en breitt upplagd satsning på ”e-assessment” för såväl formativt som summativt inriktade syften.

Även Nya Zeeland har ett slags nationella prov, asTTle. Dessa finns online på Internet från och med 2008, tidigare levererades de på cd-rom.

I Alberta (Kanada) har delstatsproven, i ett mycket ambitiöst projekt, flyttats från prov med papper och penna till CAA (Computer Adaptive Assessment). Ett privat företag, Castle Rock Research, fick överta provbanken. Resultatet har blivit ett ramaskri från lärare och lärarfack, som anser att kostnaden blir för hög och att de flervalsfrågor som distribueras motverkar läroplanens tankar om att utveckla kreativitet. Vidare kritiserats att lärarfacken inte fick vara med då besluten fattades. Av 150 000 elever som skulle ha omfattats av proven hade cirka 4 000 genomfört dem i februari 2007. De tidigare proven med papper och penna hade däremot stor acceptans.<sup>75</sup>

### Prov utvecklade vid universitet och andra institutioner

GRE (the Graduate Record Examination) testar engelska och matematik och utvecklas av ETS, Educational Testing Service, i USA. GRE innehåller ett allmänt prov och ett ämnesprov, som framför allt genomförs av studenter som är intresserad av en master eller en forskarutbildning i Nordamerika, speciellt inom ingenjörsprogrammen. GRE ges i cirka 150 länder och ungefär en halv miljon studenter genomför GRE varje år. Det allmänna provet syftar till att mäta provtagarnas förmåga inom matematik och engelska. Mer specifikt syftar det till att mäta analytisk skrivning samt verbala och kvantitativa färdigheter som har erhållits under en längre tid och som inte är relaterade till något visst område eller ämne. De flesta uppgifterna inom den verbala och kvantitativa delen är flervalsfrågor, medan den skriftliga delen består av uppsatsskrivning baserat på ett slumpmässigt valt ämne (från en stor bank av uppsatsämnen). GRE datoriseras 1992 av ETS och 1999 tog man bort möjligheten till att genomföra det som ett prov med papper och penna, och 2003 ströks möjligheten att handskriva uppsatsen. GRE ges nu enbart som ett datorbaserat prov, där de verbala och kvantitativa delarna även är adaptiva. För att kunna ge ut GRE så krävs en stor uppgiftsbank där ständigt nya uppgifter adderas efter att de har utprovats. Efter avslutat prov får prov-

---

<sup>75</sup> Teachers Alberta (2007)

tagaren tre skalade provpoäng på de olika delarna: verbal (200–800), kvantitativ (200–800), samt analytisk skrivning (0–6). Resultatet är giltigt i fem år och provet kan tas en gång i månaden upp till fem gånger inom samma tolv månadersperiod.

TOEFL (Test of English as a Foreign Language) syftar till att mäta förmågan att använda och förstå nordamerikansk engelska (som den används inom högskolan och akademien) hos personer som inte har det som modersmål. De flesta universitet och högskolor i USA kräver att utländska studenter genomför ett TOEFL innan de antas till en utbildning, med undantag för studenter från Storbritannien, Irland samt Australien. TOEFL används även inom andra engelskspråkiga länder samt en del statliga myndigheter, stipendieprogram och dylikt. Ungefär en miljon studenter genomför TOEFL varje år. TOEFL datoriserades i juli 1998 men erbjuds fortfarande som ett prov med papper och penna. Det datorbaserade TOEFL består av fyra delar: hörförståelse, läsförståelse, uppsatsskrivning (om ett slumpmässigt givet ämne) samt struktur (känna igen fel samt slutföra meningar). Ett gratis linjärt exempelprov finns på Internet. Proven var tidigare adaptiva, men vid en grundlig översyn 2005/2006 ändrades detta och de är nu linjära. De fyra basfärdigheterna testas med hjälp av dator. Ingen speciell instruktion ges för hanteringen av dator eller programvara. Färdigheterna *tala* och *skriva* bedöms efter senaste revisionen inte maskinellt. Testtagaren tillåts numera föra anteckningar under provets gång.

Anledningen till omarbetningen som gjordes var att validiteten ifrågasattes. Provet mäter nu färdigheter mer integrerat och ska bättre återspegla akademiska språkkrav. Syftet är att testa det språk man behöver som student i USA ”från klassrummet till bokhandeln”. Förmågan att kombinera och sammanställa information från olika källor är viktig, till exempel testas *läsa* och *höra* genom en skriftlig eller muntlig sammanställning.

Skillnaden mellan prov med papper och penna och de datorbaserade versionerna av TOEFL är att inom hörförståelsen och läsförståelsen används en del nya innovativa provuppgifter som matchning av objekt eller text till kategorier samt olika visuella uppgifter. De flesta uppgifter är dock flervalsuppgifter. Provpoängen baseras på lika många uppgifter för de olika provtagarna. Provtiden är cirka 3,5–4 timmar.



GMAT är ett storskaligt prov som syftar till att mäta individers möjligheter att lyckas akademiskt inom företagsekonomiska högre studier. Ungefär 200 000 genomför provet varje år. Provet används som urvalsinstrument till mastersprogram på handelshögskolor i USA. GMAT består av fyra delar; två delar består av analytiska skrivuppgifter (0–6 poäng) och de övriga består av en verbal och en kvantitativ del i form av flervalsfrågor i ett datoradaptivt format. Den verbala delen (0–60 poäng) innehåller 41 flervalsuppgifter som handlar om korrigerig av meningar, logiskt tänkande och läsförståelse. Korrigerig av meningar innebär att delar i en mening är understrukna och fem möjliga svarsalternativ ges. Logiskt tänkande innebär att provtagaren ska dra slutsatser, identifiera antaganden eller hitta svagheter eller styrkor i en argumentation. Den kvantitativa delen (0–60 poäng) består av 37 flervalsfrågor om problemlösning och datatillräcklighet. Problemlösningen innehåller aritmetik, algebra och elementär geometri. Fem svarsalternativ ges till varje uppgift. I uppgifterna med datatillräcklighet får provtagaren en uppgift och två påståenden som skulle kunna användas för att lösa uppgiften. Provtagarens uppgift är att bedöma om uppgiften går att lösa med hjälp av någon, båda eller inget av påståendena, det vill säga dessa är flervalsuppgifter med fem svarsalternativ. Denna typ av uppgiftsformat är densamma som tillämpas inom högskoleprovet på provdelen NOG. En del består av analytiskt skrivande, vilket först bedöms av en maskin med ett program som kallas Intellimetric, sedan av mänskliga bedömare (GMAT). Den totala provtiden är 3,5 timmar. Resultaten är giltiga i fem år och maxpoängen är 800.<sup>76</sup>

ACT (Compass College Placement Test) är ett test som ska hjälpa colleges och studenter att hitta styrkor och kompetensområden som behöver speciell uppmärksamhet och ge hjälp vid kursval, alltså en typ av formativt/diagnostiskt material. Beträffande ESL (English as a Second Language) testas grammatik/språkfärdighet, läsa och höra. Provet använder sig av delprov.<sup>77</sup>

MELAB (Michigan English Language Assessment Battery) är ett språkligt inträdesprov till universitet i USA och Kanada som alternativ till TOEFL.

---

<sup>76</sup> Rudner (2007)

<sup>77</sup> Chapelle & Douglas (2006)

WebLAS (Language Assessment System) vid University of California har prov knutna till undervisning i engelska som andraspråk och främmande språk (japanska och koreanska) på universitetet. Flera institutioner samarbetar för att utveckla prov och bedriva forskning om datorbaserade prov. Målsättningen är att en hög grad av autenticitet ska uppnås. Ett intressant inslag i proven är att man använder video för hörförståelse. En undervisningssituation med bland annat whiteboard och Powerpoint-bilder visas. Proven är så flexibelt utformade att de lätt ska kunna anpassas till andra språk.<sup>78</sup>

IELTS är ett samarbete, påbörjat 1989, mellan University of Cambridge ESOL Examinations, the British Council och IDP Education Australia. Två varianter finns, en som testar akademisk nivå för dem som vill studera på universitet, och en mer allmän för icke-akademiska, yrkesutbildningsrelaterade behov.

BULATS (Business Language Testing Service) gör språkprov för yrkeslivet. Det är ett resultat av ett samarbete mellan University of Cambridge ESOL Examinations, Alliance Française, Goethe-Institut och Universidad de Salamanca.

Cambridge ESOL har provcentra över hela världen för olika certifikat: Key English Test (KET), Preliminary English Test (PET), Business English Certificate (BEC), Preliminary level, Business English Certificate (BEC), Vantage level, Teaching Knowledge Test.

## Datoriserade och adaptiva prov i matematik och NO

De flesta exemplen på uppgifter i matematik och NO som går att finna via testföretagens webbplatser är flervalsfrågor. Det finns även en del exempel på provuppgifter med öppna svarsformat där provtagaren ska skriva in sitt svar i en ruta. Alternativt ska man konstruera formler eller beräkningar med hjälp av givna knappar eller delar av svar som ska sättas samman till en helhet. Denna typ av uppgifter är att betecknas som kortsvarsuppgifter där endast ett svar är korrekt. Svaret ska skrivas på en given form och är därmed även möjligt att automaträtta. Uppgifter med mer komplicerade svar så som kompletta beräkningar eller egna förklaringar finns inte som exempeluppgifter till datoriserade prov. Det finns

---

<sup>78</sup> Chapelle & Douglas (2006)

försök med datoriserade uppgifter som genererar svar i form av längre texter men det är mestadels på forskningsstadiet ännu.<sup>79</sup>

### Exempel på publicerade uppgifter i matematik

De flesta uppgifter som finns i datoriserade prov är vanliga flervalsuppgifter där provtagaren ska markera rätt svar, uppgifter som lika gärna hade kunnat ingå i ett prov med papper och penna. Exempelen nedan kommer från datoriserade prov, som inte nödvändigtvis är adaptiva.

**Figur 1** Uppgift från ACT:s matematikdel

What is the greatest common factor of 42, 126, and 210?

- F. 2
- G. 6
- H. 14
- J. 21
- K. 42

**Figur 2** Uppgift från GRE:s matematikdel

$$\int_{-3}^3 |x + 1| dx =$$

- (A) 0      (B) 5      (C) 10      (D) 15      (E) 20

<sup>79</sup> Drasgow et al. (2006), Thomas, Smith, & Waugh (2008)

**Figur 3** Uppgift från Algebra end-of-course assessment, ETS

Eduardo's bowling scores for his first 3 games were 145, 136, and 156. If he wants to have an average score of  $x$  after 4 games, which equation describes  $s$ , the score he needs for his fourth game?

- F  $x = \frac{145 + 136 + 156}{s}$
- G  $x = \frac{145 + 136 + 156}{3} + s$
- H  $x = \frac{145 + 136 + 156 + s}{4}$
- J  $x = \frac{145 + 136 + 156 + s}{3}$
- K  $x = \frac{145 + 136 + 156}{4} + s$

Även om de tre uppgifterna ovan bara kräver att provtagaren klickar på svaret, så kräver åtminstone de två övre uppgifterna att en beräkning görs innan svaret kan ges.

Uppgifter med öppna svarsformat men som ska automaträttas är oftast av en typ där provtagarna ska skriva in rätt svar, ett ord eller en siffra, i en ruta eller på en given rad.

**Figur 4** Uppgift från NAEP:s matematikdel. Från Danmark och Norge kommer följande uppgifter

A club needs to sell 625 tickets. If it has already sold 184 tickets to adults and 80 tickets to children, how many more does it need to sell?

Answer: \_\_\_\_\_

**Figur 5** Uppgift från Danmark, matematik årskurs 3

Find trekanter

Klik i de 3 trekanter.

Klik på svar



The image shows a collection of ten different polygons scattered on a white background. There are three triangles: one right-angled triangle, one acute-angled triangle, and one obtuse-angled triangle. There are also a square, a rectangle, a parallelogram, a trapezoid, a pentagon, and a hexagon.

I uppgiften ovan är svarsformatet fast men av lite annan typ. Provtagaren ska klicka på de figurer som föreställer trianglar. När en figur är vald så byter den färg för att markera för provtagaren att den ingår i svaret.

**Figur 6** Uppgift från Norges nasjonale prøver, matematik 8 trinn

Per kjøper 3 sjokolader.  
En sjokolade koster 12 kroner.  
Per betaler med en 50-kroneseddel.

Hvor mye skal han få tilbake?

Svar:  kr



The image shows three chocolate bars stacked together. The top one is white chocolate, the middle one is milk chocolate, and the bottom one is dark chocolate. Each bar is divided into a grid of squares.

Det finns även exempel på uppgifter där provtagarna ska komplettera diagram genom att dra i staplar eller rita en given linje, uppgifter där de ska sortera givna svar i rätt ordning och uppgifter där provtagaren ska dra och släppa bilder och ord på rätt plats. Det som dessa olika typer av uppgifter har gemensamt är att de är lätta att automaträtta och att de bedöms dikotomt, det vill säga antingen är svaret rätt eller så är det fel.

#### Exempel på publicerade uppgifter i naturvetenskapliga ämnen

Då det gäller datoriserat provmaterial i de naturvetenskapliga ämnena så som biologi, fysik och kemi, finns inte lika många exempel som i matematiken. Grundläggande matematik prövas som en del i många storskaliga datoriserade mätningar, det är däremot inte lika vanligt i de naturvetenskapliga ämnena. Att finna adaptiva prov är ännu svårare. Det material som finns att tillgå för de naturvetenskapliga ämnena visar även de på uppgifter som till största del är uppgifter med fasta svarsförslag. Proven i NAEP:s naturvetenskapliga del är delvis datoriserade prov som innehåller uppgifter där provtagaren ska ge fullständiga svar till ett antal frågor. Dessa svar är möjliga att skriva in via datorn men de rättas av en mänsklig bedömare. Det är därmed ingen automatisk rättning av fritextsvaren.

Även om det finns de som hävdar att flervalsuppgifter inte har möjlighet att mäta alla typer av förmågor och då kanske speciellt i de naturvetenskapliga ämnena, så är det ändå denna typ av uppgifter som är dominerande på de datoriserade proven. Se några exempel nedan.

**Figur 7** Uppgift från Danmark, biologi

### Polarræven

Temperaturen kan have stor indflydelse på livsbetingelserne for dyr og planter. I polare områder er kulden et problem for f.eks. polarræven. Hvis man sammenligner med ørkenræve, har polarræven et anderledes udseende.

Vælg de rigtige ord.

#### Klik på svar

I forhold til ørkenræven skal polarræven kunne holde på varmen. Derfor har den  ører. Man kan også se på polarræven, at den er tilpasset et koldt klima, fordi den har  ben og en , varmeisolerende pels.

Denna uppgift är av flervalstyp där eleven på tre ställen ska välja ett korrekt ord av fyra möjliga. Provtagaren svarar genom att klicka på det ord som passar in i meningen.

Sedan finns det en hel del exempel på flervalsuppgifter som varierar i svårighet från rena faktafrågor till relativt avancerade förståelsefrågor.

**Figur 8** Uppgift från MAP:s naturvetenskapliga del, NWEA

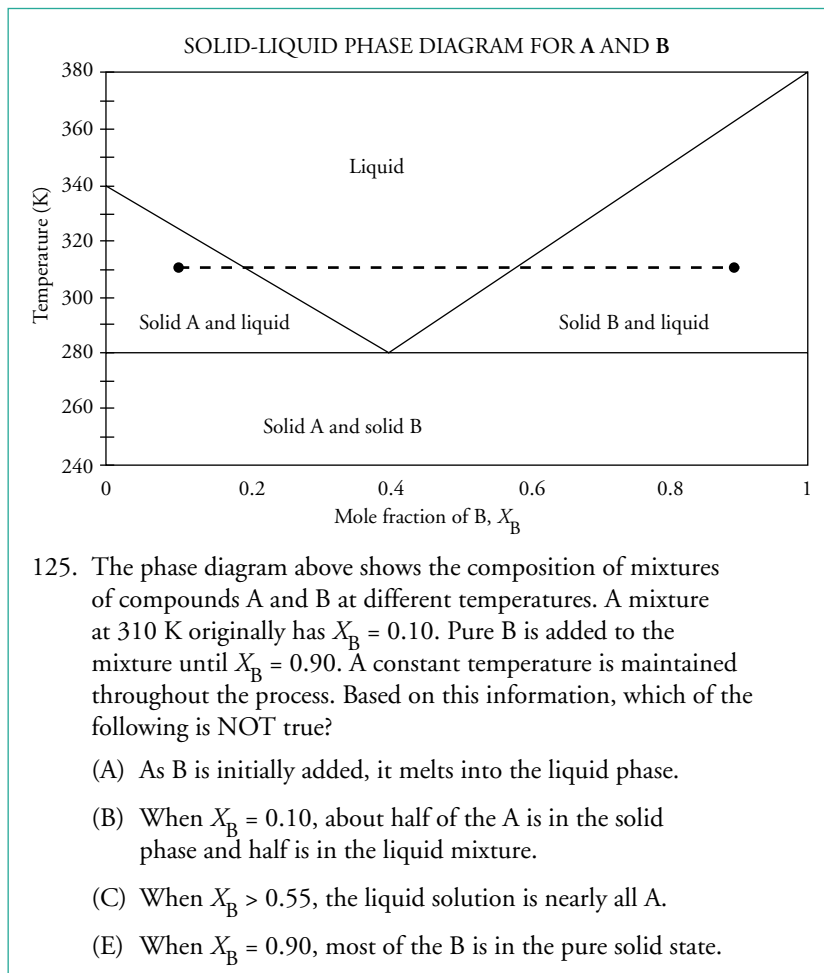
In guinea pigs, fur color is dependent on only one pair of genes, and black is dominant over white.

If no mutations occur, what will happen if a purebred black guinea pig is crossed with a white guinea pig?

- A. 1/2 of the offspring will be black; 1/2 will be white.
- B. 3/4 of the offspring will be black; 1/4 will be white.
- C. 9/16 of the offspring will be black; 7/16 will be white.
- ✓ D. All of the offspring will be black.
- E. All of the offspring will be white.

**Figur 9**

Uppgift från GRE:s kemidel



125. The phase diagram above shows the composition of mixtures of compounds A and B at different temperatures. A mixture at 310 K originally has  $X_B = 0.10$ . Pure B is added to the mixture until  $X_B = 0.90$ . A constant temperature is maintained throughout the process. Based on this information, which of the following is NOT true?

- (A) As B is initially added, it melts into the liquid phase.
- (B) When  $X_B = 0.10$ , about half of the A is in the solid phase and half is in the liquid mixture.
- (C) When  $X_B > 0.55$ , the liquid solution is nearly all A.
- (E) When  $X_B = 0.90$ , most of the B is in the pure solid state.

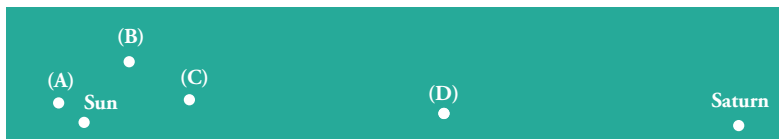


**Figur 10** Uppgift från NO-prov, EAA, Australien

The table shows the average distance of each of the planets in our solar system from the Sun.

Planet	Average distance (millions of kilometres)
Mercury	58
Venus	108
Earth	150
Mars	229
Jupiter	779
Saturn	1 427
Uranus	2 871
Neptune	4 496
Pluto	5 913

The distance between the Sun and Saturn is shown to scale below. Four other planets labelled A, B, C and D are also shown to the same scale. Which of these is Mars?



## Datoriserade och adaptiva prov i språk

### Exempel på språkprov

Dialang och Surveylang är två europeiska projekt som handlar om datorstödd bedömning i språk. Dialang-projektet finansieras av EU under det så kallade Sokratesprogrammet, med syfte att ta fram diagnostiska språkprov i 14 olika europeiska språk. Öppet för alla och ett nedladdningsbart gratisprogram finns, dock endast för Windows. Många intressanta exempel på provtyper finns på webbplatsen. Surveylang (European Survey on Language Competences) omfattar 32 länder och ska på EU:s

uppdrag testa fem språk och primärt tre färdigheter (läsa, höra, skriva). Provet kommer att vara datorbaserat för dem som vill, men prov med papper och penna kommer att finnas som alternativ. Ursprungligen var tanken att proven enbart skulle vara datorbaserade, men det visade sig att flera länder helt enkelt inte hade den IT-infrastruktur som behövdes. Cirka 1 500 elever per land och språk ska genomgå provet. År 2011 ska proven genomföras i hela Europa och slutrapporten är planerad till 2012. Tanken är att systemet ska använda den hårdvara skolorna har och det utvecklas som open-source-projekt, det vill säga det är fritt att använda för all icke-kommersiell verksamhet. Alla allmänna instruktioner för provets genomförande kommer att vara på provtagarnas modersmål, medan instruktioner kopplade direkt till provuppgifterna blir på målspråket. Klientdatorerna ska startas (bootas) från en usb-sticka som har ett Linux-system, en lösning som ger unika möjligheter att styra vad som kan och får göras i systemet.<sup>80</sup>

Några andra aktörer som utvecklat färdighetsprov i olika språk är Alliance Française, Goethe-institutet och Universidad de Salamanc. Dessa prov distribueras av ett antal aktörer på marknaden. Ett annat exempel är CARLA, Center for Advanced Research on Language Acquisition, University of Minnesota, som är finansierat av USA:s utbildningsdepartement och arbetar bland annat för att förbättra språkundervisning och språkinläring. Som ett led i detta utvärderar man kvaliteten i existerande amerikanska färdighetsprov i franska, tyska och spanska.

Inom språkbedömning finns särskilda utmaningar i förhållande till de olika typer av färdigheter som bygger upp språkförmågan. Ytterligare exempel på datorbaserade prov har sorterats efter olika språkliga färdigheter.

## Produktiva färdigheter

### Skriva

När det gäller skriftlig produktion är det överlägset vanligaste att bedömningen sker med mänskliga bedömare, men det finns även system som kan bedöma skriftliga arbeten maskinellt. Det mest använda är e-rater. E-rater är baserad på en teknik kallad NLP (Natural Language

---

<sup>80</sup> Ryssevik (2008), Bjerkestrand (2008)

Processing), och kan analysera strukturen i det skrivna (inledning, avhandling och avslutning), hur avancerat ordförrådet är, ordförråd som är relevant för ämnet, syntax, stavning, satsstruktur, variation i meningsbyggnad och interpunktion. Hög korrelation med mänskliga bedömare har kunnat konstateras på de flesta punkter,<sup>81</sup> men den kanske viktigaste aspekten, nämligen hur en text uppfattas av en läsare, kan datorn inte bedöma fullt ut.<sup>82</sup> Texter som har skrivits för att få systemet att misslyckas har demonstrerat detta.<sup>83</sup> Manuell medbedömning sker kontinuerligt på ett slumpurval av de datorbedömda arbetena och på grundval av de diskrepanser man upptäcker förbättras algoritmen fortlöpande. Ett program som liknar e-rater till sin funktion är Intellimetric från Vantage Learning. Beträffande problemet som nämndes ovan med e-rater och hur programmet hanterar nonsens med rätt ord och grammatik, sägs Intellimetric klara 95 procent av sådana försök. Det används bland annat i GMAT.

Somliga menar att vi med stor sannolikhet kommer att se stora framsteg inom detta område de närmaste åren. Genom en förväntad utveckling inom datortekniken i allmänhet och artificiell intelligens i synnerhet, kan man i framtiden räkna med en mer ”intelligent bedömning”.<sup>84</sup>

## Tala

Ofta kopplas ”tala” ihop med ”samtala” och då vidgas begreppet till att gälla interaktion mellan två eller flera parter. Detta samspel kan datorn idag, så vitt vi vet, inte bedöma alls. Inte heller kan en maskin tolka och ”förstå” mänskligt tal. De problem som diskuteras ovan under färdigheten skriva och nedan kring bedömningen av öppna svar i läs- och hörförståelseprov gäller naturligtvis även vid talat språk och till detta kommer svårigheten att överföra talet till ett för datorn begripligt format. Den största delen av forskningen och utvecklingen på detta område ägnas åt just detta steg. Vad kan då datorn göra?

Molholt och Presler konstaterade i en pilotstudie att datorer är bra på att bedöma uttal. Hög korrelation med mänskliga bedömare uppmättes vad beträffar saknade fonem (språkljud), fel fonem, partiella fonem,

---

<sup>81</sup> Lee, Gentile & Kantor (2008)

<sup>82</sup> Bridgeman (2008)

<sup>83</sup> Chapelle & Douglas (2006)

<sup>84</sup> Bennet (1999) refererad av Chapelle & Douglas (2006)

extra fonem och betoningsfel.<sup>85</sup> Andra studier uppvisade inte samma positiva resultat.<sup>86</sup> Gemensamt för de maskinbedömda modellerna är att yttre form bedöms, men hög korrelation med mänskliga bedömare har i de flesta fall uppmätts.<sup>87</sup>

Versant test är ett prov konstruerat för färdigheter i främmande språk. Korrekthet vid repetition av ord bedöms, liksom uttal, flyt vid uppläsning av text, upprepning av ord och uppgifter av typen ”välj alternativ A eller B”. Ytterligare en del som inte bedöms utan sparas, så att behöriga personer kan lyssna, innehåller frågor av typen ”What qualities do you look for in a friend?” Provtagaren har sedan 20 sekunder på sig att formulera ett svar. Programmet använder sig av röstigenkänningsprogramvara som innehåller en algoritm baserad på ett stort bibliotek av röster tillhörande infödda talare med olika regional och social tillhörighet.<sup>88</sup> Chun riktar kraftig kritik mot detta test, som han menar saknar autenticitet (”it fails by any reasonable measure of authenticity”) och därmed validitet.<sup>89</sup>

Ett annat exempel är SpeechRater från ETS (Educational Testing Services), ett relativt nytt program som sägs kunna tolka fritt tal. TOEFL använde sig tidigare av denna teknik men har nu gått tillbaka till mänskliga bedömare. Det hävdas att uttal, ”fluency”, grammatisk korrekthet och vokabulär kan bedömas. Systemet klarar i dagsläget inte att bedöma innehållet i det som sägs och går att lura genom att man läser innantill i något som inte har det minsta med det aktuella ämnet att göra.

Ett system som förutsätter att programmet lär sig varje talare är Dragon Naturally Speaking från Dragon Systems. En text bestående av cirka 3 800 ord läses in för att programmet ska lära sig talaren och sedan läses den egentliga provtexten upp, cirka 1 050 ord. Försöket har gett resultat som är jämna och konsekventa. Vidare har POET/OEPT Purdue University utvecklat ett eget prov för färdigheten ”tala”. Systemet arbetar med mänskliga bedömare och benchmarks som presenteras på skärmen.<sup>90</sup> Ytterligare ett exempel är COPI (Computerized Oral Proficiency Instrument) från Center for Applied Linguistics. Det använder

---

<sup>85</sup> Molholt och Presler (1986)

<sup>86</sup> Reid (1986), refererad av Chapelle & Douglas (2006)

<sup>87</sup> Chapelle & Douglas (2006)

<sup>88</sup> Chapelle & Douglas (2006)

<sup>89</sup> Chun (2006)

<sup>90</sup> Chapelle & Douglas (2006)

sig av mänskliga bedömare som lyssnar på testtagarna via Internet, och resultatet sparas och finns tillgängligt.

Framtidsperspektivet är intressant. Utvecklingen går framåt, både vad gäller röstigenkänning och talsyntes. Kanske blir det en dag möjligt att koppla samman röstigenkänning, analys genom artificiell intelligens och talsyntes och på så vis skapa interaktiva talsituationer med hög autenticitet, och kanske vågar man till och med hoppas att datorn då förmår bedöma detta på ett relevant sätt. Men vägen dit är fortfarande oerhört lång.

## Receptiva färdigheter

I datorbaserade språkprov som prövar receptiva färdigheter används huvudsakligen de uppgiftstyper som listats i tabell 4. Datorns inboende begränsningar styr alltså i mycket högre grad provens utformning än frågan om vilka konkreta språksituationer som är meningsfulla att testa och om dessa låter sig prövas med datorbaserade prov. Valet av uppgiftstyper, utformningen av dem och det sätt på vilket vi utvärderar svaren i relation till det vi önskar bedöma, är naturligtvis avgörande för kvaliteten i bedömningen.<sup>91</sup>

Multimedia i hörförståelse förtjänar särskild uppmärksamhet eftersom detta oftast nämns som en fördel med datorbaserade och adaptiva prov, främst därför att det anses skapa större autenticitet. Dialang har ett enkelt exempel på video i hörförståelse, och WebLAS har en mer avancerad modell som innebär att en akademisk psykologiföreläsning om minnet spelas upp och följs av frågor. Coniam menar att om hörförståelse ska bedömas kan videoinslag riskera validiteten hos provet.<sup>92</sup> Han fann å andra sidan inga noterbara skillnader då han jämförde ren lyssning med videostödd lyssning. Ockey noterar också att introduktion av video i hörförståelse ökar autenticiteten, men samtidigt förändrar vad som bedöms.<sup>93</sup> Provdeltagarna tolkade gester och ansiktsuttryck och läste på läpparna. Förmågan att dra nytta av detta varierade från en testtagare till en annan. Personerna i undersökningen fick också ta ställning till om videostödet var till nytta för förståelsen eller distraherande. Även där varierade svaren mellan testtagarna. Ockey's testunderlag var mycket litet (sex personer), men hans slutsatser kan ändå anses relevanta. Wilhelm &

---

<sup>91</sup> Chapelle & Douglas (2006), Messick (1989), Kane (2006)

<sup>92</sup> Coniam (2001), refererad av Chapelle & Douglas (2006)

<sup>93</sup> Ockey (2007)

Schroeders fann att lyssningsprov med bildstöd fungerade mycket bra i ett försök med engelska som främmande språk.<sup>94</sup>

I förhållande till autenticitetsbegreppet är det värt att poängtera att ren hörförståelse ("receptive listening") sällan förekommer i verkligheten; man ser oftast den som talar eller också har man bildstöd som illustrerar det som sägs, till exempel vid en nyhetssändning på tv.<sup>95</sup> Unga människor lyssnar inte i större omfattning på tal-radio, vid telefonsamtal kan man be om repetition.

---

<sup>94</sup> Wilhelm & Schroeders (2008)

<sup>95</sup> Buck (2001)

**Studiebesök,  
demonstrationer  
och konferenser**

## Studiebesök, demonstrationer och konferenser

I detta projekt har flera studiebesök och seminarier genomförts, för att samla erfarenheter från andra länder och viktiga forskningsmiljöer. Närmare kontakter har på detta sätt tagits med universitetet i Bergen, Danmarks Skolestyrelse och @ventures som ansvarar för utvecklingen av adaptiva nationella prov i Danmark, CITO i Nederländerna, Durham University i Storbritannien och det finska företaget Renet Oy.

### Studiebesök Bergen 22 juni 2008

Den 22 juni 2008 besöktes den del av universitetet i Bergen, som ansvarar för utvecklingen av de nationella proven i engelska för det norska skolsystemet.<sup>96</sup> Närvarande från det norska projektteamet var projektledaren, professor Angela Hasselgreen, projektsamordnaren Eli Moe, samt olika medarbetare som arbetar med provutveckling, bilder, grafik och dataanalys. Från Sverige deltog Tommy Lagergren och Wolfgang Dietrich (Skolverket), samt Hans Albinsson, AnnaKarin Lindqvist och Gudrun Erickson (Göteborgs universitet).

Under besöket gavs information om bakgrunden till och utvecklingen av projektet, med fokus delvis på den allmänna och skolpolitiska situationen, men med huvudsaklig koncentration på proven i engelska. Vi gavs också möjlighet att träffa personer som arbetar med provutveckling (inklusive bilder och grafik), implementeringsaspekter och frågor om dataanalys.

De norska proven i engelska är utvecklade på basis av, och nivåbestämda i relation till, den gemensamma europeiska referensram för språk som tagits fram av Europarådet, och som används i stor omfattning i Europa. Provens funktion, utveckling och innehåll beskrivs närmare i samband med pilotstudien i engelska. Sammanfattningsvis kan dock sägas, att uppdraget ursprungligen avsåg adaptiva prov, som utvecklades under stor tidspress och användes under en kort period och med resultat offentliggjorda på ett sätt som inte var varslat initialt. I enlighet med centrala beslut, övergick dock systemet tämligen snart

---

<sup>96</sup> Verksamheten bedrivs inom "forskningsselskapet" Uni Research vid en avdelning som heter Uni Digital (tidigare AKSIS). Universitetet i Bergen är huvudägare till Uni Research.



till att avse linjära prov, som numera tillhandahålls i tre parallella versioner, och vars domäntäckning reducerats jämfört med vad som först var avsikten. Det senare motiverades från centralt håll med otillräcklig interbedömarreliabilitet, en uppfattning som uppenbarligen inte delades av provutvecklarna, som i sin analys fokuserade det didaktiska värdet av skrivproven, de norska språklärarnas ovana vid den europeiska referensramen samt deras positiva utveckling under den mycket korta förberedelsestid som fanns att tillgå. I provgruppen finns en anställd konstnär, som i samråd med provkonstruktörerna producerar bildmaterial och grafik till proven. Vid kravgränssättning medverkar reguljärt internationellt välrenommerade och av Europarådet ofta anlitate experter.

Under besöket gavs tillfälle till samtal om såväl principiella som mera praktiskt inriktade frågor. Till exempel diskuterades olika aspekter av validitet, reliabilitet och implementering. Vidare framkom att våra norska kolleger, förutsatt formellt medgivande av den norska nationella skolmyndigheten, skulle vara både villiga att och intresserade av att 'låna ut' norska prov till det svenska pilotförsöket i engelska.

För vidare information om utvecklingen av de norska nationella proven i engelska hänvisas till webbplatsen [digital.uni.no](http://digital.uni.no).

### Studiebesök Köpenhamn 24 juni 2008

Den 24 juni 2008 besöktes Danmarks Skolestyrelse i Köpenhamn. Närvarande från Skolestyrelsen var Jakob Wandall och Margit Holm Nielsen. De svenska deltagarna var Wolfgang Dietrich och Klas Unger från Skolverket; Gudrun Erickson, AnnaKarin Lindqvist och Hans Albinsson från Göteborgs universitet; Katarina Kjellström från Stockholms universitet; Anna Lind Pantzare och Peter Nyström från Umeå universitet.

Syftet med besöket var att få närmare information om de nationella prov i Danmark som är datorbaserade och adaptiva. Danmark har som enda land i Norden utvecklat adaptiva prov i ett antal ämnen.

Olika typer av avgångsprov har använts under lång tid i Danmark. Examensproven har varit frivilliga men nästan alla har genomfört dem. Bakgrunden till att Danmark dessutom beslutat införa obligatoriska nationella prov finns i PISA-studierna från 2000 och 2003 samt i OECD:s översikt från 2004 som visade på brister i det danska utvärderingssystemet. Föräldrar fick för lite information och lärarna ansågs inte kunna

bedöma. Efter en upphandling fick det danska företaget COWI A/S i juni 2006 ett uppdrag att konstruera de nationella adaptiva proven som omfattade utveckling av såväl uppgifter som en plattform för att leverera dem. Uppdraget var att färdigställa tre prov på mindre än ett år, för att börja användas i maj 2007. Fokus i dessa prov var utvärdering och utvärderingskultur, inte att testa enskilda elever. Tanken var att skolan skulle använda de nationella proven som en del i den egna kvalitetsdiskussionen. Proven är obligatoriska för eleverna, men det är upp till skolan att använda resultaten för att utveckla pedagogiken. Enligt företrädarna för Skolestyrelsen finns ingen egentlig relation mellan examensproven och de nationella proven, och en viktig skillnad är att examensproven prövar slutmålen medan de nationella proven inte prövar allt. I klartext sades att de nationella proven bara prövar sådant som är möjligt att pröva med ett adaptivt prov.

Kännetecken för de danska nationella proven är att de ska vara ett verktyg för lärande, det är ett adaptivt system, proven distribueras via Internet, de är självvärderande, systemet innehåller ett rapporteringssystem med feedback till elever, föräldrar och lärare, de är low-stake för den enskilde eleven och resultaten inte är tillåtna att offentliggöras. Det är bara eleven, föräldern och läraren som kan se den enskilda elevens resultat och mer detaljerad information om hur provet genomförts. Skolledarna får information på skolnivå och övergripande klassnivå. För mer detaljerad information måste de kontakta den enskilda läraren. Kommunen får information på skolnivå. Den enda information som får offentliggöras är resultatet på landsnivå, det vill säga om resultatet över landet var bättre eller sämre än tidigare år. Systemet innehåller många möjligheter att få olika typer av rapporter, däremot vet man inte hur mycket de nyttjas av lärarna. Vid den första omgången blev rapporterna försenade vilket innebar att de inte fanns tillgängliga förrän under sommaren.

Det framkom mycket kritik mot de nationella proven efter den första omgången. Det fanns tveksamheter om kvaliteten på uppgifterna var tillräcklig och om uppgiftsbankerna var nog stora. Relativt snabbt kom ett beslut om att göra ett omtag i processen. Inga nya prov skulle levereras förrän våren 2009, Skolestyrelsen skulle vara mer inblandad i kvalitets-säkringen av uppgifterna och uppgifterna skulle arbetas om. Den kanske viktigaste erfarenheten som representanterna från Skolestyrelsen ville dela med sig av var att skynda långsamt. Det tar tid att bygga upp ett

nytt provsystem från grunden. I efterhand hade det varit bättre att vänta lite längre med att genomföra första omgången och därmed ha tid att säkerställa materialets kvalitet.

Studiebesöket följdes upp med ett seminarium på Skolverket den 13 november 2008. Då presenterade Michael Lund-Larsen (chef) och Kaj Haupt (konsulent på Evalueringskultur-projektet) från @ventures/eVidenCenter, *Kompetence og videncenter for e-læring*. Detta center har ansvaret för provutvecklingen för de danska adaptiva proven. Vid detta tillfälle demonstrerades processen från bokning av prov till rapportering i de nationella proven. Dessutom diskuterades specifika frågor om de adaptiva proven och möjligheten att samarbeta i samband med pilotprojekten. Detta samarbete diskuterades vidare under de närmaste månaderna, men kunde i slutändan inte förverkligas.

För vidare information om de danska adaptiva nationella proven hänvisas till projektets webbplats [www.evaluering.uvm.dk](http://www.evaluering.uvm.dk).

### Seminarium med Renet OY 20 augusti 2008

Vid ett seminarium på Skolverket den 20 augusti 2008 berättade representanter för det finska företaget Renet OY om vad de utvecklat när det gäller datorbaserade prov. Renet arbetar med planering, utveckling och tillverkning av digitala multimediala språkundervisningssystem och erbjuder språkexamina i en digital miljö. Digitala språktest och språkexamina har genomförts med Renets system sedan 1998.

Vid seminariet presenterades Renets datorbaserade prov i finska för immigranter från det forna Sovjetunionen. Proven genomförs på särskilda datorer som finns på ett par olika testcentra. Provtagarna använder en dator och ett headset, och proven simulerar bland annat kommunikativa situationer genom att datorn visar en person som ställer frågor som provtagaren ska svara muntligt på. Demonstrationen av det datorbaserade provet visade en ganska övertygande simulering som definitivt kan skapa en känsla av närvaro och kommunikation. Lika intressant var presentationen av det system för bedömning av provtagarnas svar som utvecklats av Renet. Varje svar från provtagaren registrerades, och för de muntliga delarna sparades varje svar som en ljudfil. I bedömningssystemet kunde så bedömarna lyssna på varje svar och ange sin bedömning. En fråga kunde bedömas för alla provtagarna och varje svar kunde lyssnas på igen

och igen om det skulle behövas. Systemet borde underlätta bedömning-  
en avsevärt och även stärka kvaliteten i bedömningen.

Företrädarna för Renet demonstrerade även ett prov de utvecklat för  
att pröva språkförmåga i tyska. Provet hade liknande funktioner som det  
som beskrivits ovan, men det var ännu mer sofistikerat. Hela provet var  
organiserat utifrån ett tema om en beställning av en resa. Provtagaren  
togs med i en historia där han eller hon ska beställa resa åt en god vän. I  
samband med detta får provtagaren uppgifter att till exempel besöka en  
resebyrå. I denna del av provet visar datorn en resebyråmiljö med en säl-  
jare som talar tyska och som provtagaren måste beskriva sitt ärende för.  
I andra delar ska provtagaren skriva ett mail och ett sms på tyska. Hela  
konceptet skapar en närvarokänsla och erbjuder ett sammanhang för att  
visa sina kunskaper som inte kan åstadkommas med varken skriftliga  
eller muntliga prov av konventionellt slag. Även för dessa prov användes  
ett system för bedömning där varje elevsvar registrerades och sorterades i  
datorn, vilket möjliggjorde enkel åtkomst och registrering av omdömen.

Renets presentation visade på datorns möjligheter att skapa uppgifts-  
och provformat som är innovativa och skapar nya möjligheter att be-  
döma väsentliga förmågor i språk. Renets produkter är goda exempel  
på när datorbaserade prov kan bredda vad som bedöms. Bedömnings-  
modulen är också synnerligen intressant med dess mycket strukturerade  
stöd för bedömning av uppgifter med komplexa och svårhanterliga  
svarsformat, till exempel muntliga svar.

Mer information om Renet och dess produkter finns på deras  
webbplats [www.renet.fi](http://www.renet.fi).

### Internationell konferens om datorbaserad bedömning, september 2008

Som ett led i en paneuropeisk strävan att belysa frågan om datorbaserad  
bedömning anordnades i september 2008 en internationell konferens  
på Island, samordnad av EU-organet CRELL (Centre for Research on  
Lifelong Learning) och det isländska utbildningsministeriet. Konferen-  
sen hade tre huvudteman:

1. Comparison between paper and pencil tests and computer based  
assessment
2. Electronic tests and gender differences

### 3. Adaptive vs. linear computerbased assessment.

Ett tydligt fokus var på naturorienterande ämnen, vilket förklaras av konferensens fullständiga titel: ”The Transition to Computer-Based Assessment – Lessons learned from the PISA 2006 Computer Based Assessment of Science (CBAS) and implications for large scale testing”. PISA:s datorbaserade linjära instrument prövades ut i ett relativt stort antal länder, men bara tre valde att genomföra provet i det skarpa läget 2006 (Island, Danmark och Korea). På basis av dessa tre länders data-insamlingar, som också innefattade ett prov med papper och penna (dock, enligt uppgift, utan gemensamma uppgifter), presenterades ett antal analyser och resultat. Förutom CBAS- och NO-perspektivet var mycket få presentationer relaterade till traditionella skolämnesområden. Ett av inläggen handlade om språk, och då i ett huvudsakligt systemiskt perspektiv (Eli Moe om de nya nationella proven i Norge, som vi i detta projekt fått använda i pilotstudien i engelska). Däremot var språk i allmän bemärkelse ofta diskuterat, nämligen som avsedd eller oavsedd faktor i bedömningen av olika ämnen på papper och i dator. Generellt bör också nämnas att fokus låg på prov snarare än på bedömning i vidare mening. Vare sig linjära eller adaptiva ansatser i kontinuerliga, formativa ansatser berördes följaktligen.

Som ett resultat från konferensen publicerades i mars 2009 en rapport, som omfattar drygt 200 sidor och 27 bidrag och som är indelad i fem sektioner med följande rubriker:

- I Assessment needs and European approaches
- II General issues of computer-based testing
- III Transition from paper-and-pencil to computer-based testing
- IV Methodologies of computer-based testing
- V The PISA 2006 computer-based assessment of science (CBAS)

Bidragen speglar den stora spännvidd och de ytterst skilda perspektiv som finns och kan anläggas i diskussionen om datorbaserad bedömning. Exempelvis fokuseras såväl mera principiella frågor som syfte, potential för såväl breddning och fördjupning som insnävning och förtunning av det som ska bedömas och bias, som frågor om tekniskt tillhandahållande och kostnader. Likaså finns rapporter från närliggande skolsystem. De nationella proven i Danmark har exempelvis fokus på datorbase-

rade adaptiva prov som pedagogiskt redskap. Norge har erfarenheter av införandet av datorbaserad, nationell bedömning av engelska. Island och Danmark har studerat eventuella systematiska könsskillnader mellan datorbaserade prov och prov med papper och penna, med vitt skilda metodologiska ansatser.

Sammanfattningsvis kan konstateras att användning av datorer vid bedömning av kunskap, belystes från ett antal skilda synvinklar under konferensen. Det generella intryck som gavs var positivt, utan att därför potentiella eller reella problem negligerades.

Material från konferensen (program och presentationer) finns tillgängligt på [crell.jrc.ec.europa.eu/WP/workshoptransition.htm](http://crell.jrc.ec.europa.eu/WP/workshoptransition.htm). Rapporten *The Transition to Computer-Based Assessment – New Approaches to Skill Assessment and Implications for Large-Scale Testing* (Friedrich Scheuermann & Julius Björnsson, Eds.) finns tillgänglig i fulltext (pdf) på CRELL:s webbplats: [crell.jrc.ec.europa.eu/](http://crell.jrc.ec.europa.eu/) och går även att beställa i pappersform.

#### Studiebesök i Durham 25 november 2008

Den 25 november 2008 besöktes Centre for Evaluation and Monitoring (CEM), en självständig del av universitetet i Durham i Storbritannien. Centret utvecklar datorbaserade/adaptiva indikatorsystem inom olika domäner (*cross curricular* såväl som ämnesbaserade). Systemen erbjuds skolor i Storbritannien som komplement till de nationella proven, men även intressenter i övriga delar av världen. De har i huvudsak formativa och diagnostiska syften, används inte för *accountability purposes* och karakteriseras på grund av detta som *low stakes*. Material produceras inte bara för reguljära förskole- och skolsystem utan även för högre utbildning. Enligt projektets webbplats är CEM, som grundades 1983, den största utbildningsvetenskapliga forskningsenheten vid ett brittiskt universitet. Deras verksamhet är självbärande, det vill säga finansierad av försäljning av egna produkter samt uppdrag i anslutning till detta.

Besöket inleddes med en kort ömsesidig presentation, som i det svenska fallet även inkluderade det nationella styr- och provsystemet. Närvarande från CEM var projektledaren, professor Peter Tymms samt medarbetare (bland andra Christine Merrell, Paul Jones och David Thomson) med skilda funktioner och ansvar för olika system och prov. Från Sverige deltog Wolfgang Dietrich och Klas Unger från Skolverket;

AnnaKarin Lindqvist och Gudrun Erickson från Göteborgs universitet; Katarina Kjellström från PRIM-gruppen vid Stockholms universitet.

Under besöket presenterades och diskuterades så kallade *baseline systems*, med avsikt att fastställa utgångspunkter för analys av *value added*-karaktär. Ett system av detta slag är bland annat ASPECTS (Assessment Profile on Entry for Children and Toddlers), med syfte att bedöma och mäta framsteg hos mycket små barn inom områdena ”personal, social and emotional development, language and mathematics development”. Metoden för detta är en kombination av vad som rubriceras som objektiva uppgifter och lärarobservationer. För *primary school* presenterades bland annat PIPS (Performance Indicators in Primary Schools), som används i flera länder och vars informationsmaterial finns nedladdningsbart från CEM:s webbplats och InCAS (Interactive Computerised Assessment System), som är ett helt adaptivt, diagnostiskt bedömningssystem med fokus bland annat på icke-verbal förmåga, attityder, läsning och allmän matematik. Vidare berördes SOSCA (Secondary On Screen Curriculum Assessments), som är ett datorbaserat provsystem för elever i 14-årsåldern, med syfte att ”measure pupil progress in Maths, Science and Reading”. Samtliga dessa material relaterar till ”the national curriculum” och ger skolor en möjlighet att bedöma sina elevers framsteg och måluppfyllelse. I de fall produkterna marknadsförs utanför Storbritannien görs anpassningar till de aktuella ländernas styrdokument.

De presenterade, datorbaserade och till viss del adaptiva materialen diskuterades även utifrån teknisk och metodologisk synvinkel. Till exempel berördes aspekter av uppgiftsbanker, underliggande analysmetoder (i regel enparametrisk IRT/Rasch) och tillhandahållande (LAN eller webbaserat), samt möjligheter respektive svårigheter i relation till detta. Bland annat underströk representanterna från CEM vikten av att materialen är *low stakes* till sin karaktär och därför inte inbjuder till missbruk på det sätt som inte sällan är en potentiell eller reell risk med andra typer av prov.

För utförlig information om de olika systemen producerade vid CEM hänvisas till projektets omfattande webbplats ([www.cemcentre.org/](http://www.cemcentre.org/)).

## Studiebesök CITO 8 december 2008

CITO är det nationella institutet för beteendevetenskapliga mätningar i Nederländerna. CITO har utvecklat ett antal olika datorbaserade adaptiva prov inom utbildning. Det gäller till exempel prov som handlar om matematiska och aritmetiska färdigheter inom vuxenutbildningen, prov som handlar om holländska som första och andra språk samt prov i geografi. Verschoor och Straetmans<sup>97</sup> beskrev bland annat ett placeringsprov för vuxna inom matematik. Syftet med det provet var att placera in en person på en av tre nivåer. Deras uppgiftsbank innehöll 478 uppgifter. Provtagaren fick först två enklare uppgifter för att göra oerfarna provtagare bekväma och uppmuntra dem, därefter valdes uppgifter ut utifrån maximal information i provtagarens skattade förmåga. Varje provtagare får sedan mellan 12 och 25 uppgifter och provet avbryts när man är 90 procent säker att man placerat provtagaren i rätt kategori. Utöver de tidigare nämnda proven har CITO även utvecklat datorbaserade adaptiva prov för åldrarna 4–6 år som ska testa grundläggande aritmetiska färdigheter (ex. ordning), verbala färdigheter och tid. Dessutom ges adaptiva prov för åldrarna 6–12 inom stavning, vokabulär samt inom till exempel läsförståelse och aritmetisk förmåga för åldrarna 7–12 år.<sup>98</sup>

Vid studiebesöket på CITO den 8 december 2008 presenterade Theo Eggen psykometriska aspekter på datorbaserade adaptiva prov som ges av CITO. Innan proven ges till provtagarna har de prövats, för att uppgifternas egenskaper ska vara kända. Ofta sker utprövningen genom prov med papper och penna även om de ska användas i datorbaserade adaptiva prov. Utprövningen är därmed ett *low-stakes*-prov, medan när provet ges kan det vara ett *high-stakes*-prov. I normalfallet kräver de minst 400 observationer per uppgift, men kan ibland nöja sig med cirka 250 observationer. Uppgiftsbanken som skapas för varje prov är i storleksordningen 200–800 uppgifter, men kan ibland innehålla mer än en innehållskategori vilket leder till att varje område kanske inte innehåller fler än 200–300 uppgifter. I inledningen av ett datorbaserat adaptivt prov ges ett antal startuppgifter utifrån innehållsliga orsaker och dessa är ofta lite enklare. Att uppgifterna är lite enklare gäller speciellt om de ges till yngre provtagare, eftersom uppgifterna i ett normalt datorbaserat adaptivt prov väljs så att varje provtagare svarar rätt på ungefär hälften av

---

<sup>97</sup> Verschoor och Straetmans (2000)

<sup>98</sup> För mer detaljerad information se Eggen (2008)



frågorna och det kan vara psykologiskt ansträngande att besvara många uppgifter felaktigt. I proven går själva uppgiftsselektionen till på följande sätt. Först väljs en selektion av uppgifter och sedan väljs optimala uppgifter ur den selektionen. När man framför allt är intresserad av att klassificera provtagare så används SPRT,<sup>99</sup> och för att skatta provtagarens förmåga används WLE.<sup>100</sup> För att välja optimala uppgifter används så kallad Fisher-information på den nuvarande skattade provtagarens förmåga. Proven har ett antal praktiska restriktioner som handlar om uppgifternas innehåll och svårighetsgrad samt kontroll av uppgiftsexponeringen. Uppgiftsexponeringen kontrolleras dels med Simpson-Hetter,<sup>101</sup> dels med att max 25–30 procent får samma uppgift. Som slutkriterium av ett datorbaserat adaptivt prov används olika valda kriterier. Exempel på kriterier som de använder vid CITO är maximalt och minimalt antal givna uppgifter samt tillräckligt litet standardfel.

Mia van Boxel demonstrerade ett datorbaserat adaptivt prov som används för att testa grundläggande matematikfärdigheter för grundskollärare i Nederländerna, det så kallade WISCAT\_pabo. Alla som ska bli grundskollärare i Nederländerna måste bli godkända på detta prov, annars måste de sluta på utbildningen. Den underliggande uppgiftsbanken innehåller cirka tusen uppgifter. Provet ges över Internet, men varje provtagare måste ladda ner och installera ett program för att kunna genomföra provet. Programmet ser bland annat till att provtagaren inte kan köra andra program eller öppna flera webbläsare medan provet genomförs. Reflektioner från provtagarna var bland annat att de inte är vana att besvara ungefär 50 procent av uppgifterna felaktigt, vilket inträffar om datorbaserade adaptiva prov fungerar enligt teorin. Detta bör provtagarna informeras tydligt om när ett datorbaserat adaptivt prov införs. Vidare kommenterade de att de inte kunde gå tillbaka till tidigare besvarade uppgifter. Slutligen så skiljer sig poängsättningen på provet från ett prov med papper och penna eller datorbaserat prov, vilket gör det svårt för provtagarna att tolka resultaten. Angela Verschoor visade även hur ett datorbaserat adaptivt prov för dyslexi bland 5–7 åringar fungerade. I det provet fanns 242 uppgifter i uppgiftsbanken och en simuleringsmodul hade byggts in, så att man skulle kunna simulera vilka

---

<sup>99</sup> Wald (1947)

<sup>100</sup> Warm (1989)

<sup>101</sup> Simpson & Hetter (1985) och Hetter & Simpson (1987)

uppgifter provtagare med olika förmåga kan få. Utifrån hur provtagarna besvarade de två första delarna gavs de sedan någon av två slutdelar (en mycket enkel och en mycket svårare). Detta var ett sätt att skatta förmågan som har visat sig fungera effektivt.

Mark Molenaar demonstrerade och diskuterade tidigare och framtida datasystem och mjukvara som används i datorbaserade adaptiva prov. CITO har skapat sina egna program som används för att administrera och genomföra prov. Egna programvaror togs fram framför allt för att få flexibilitet och en hög säkerhet. När ett prov ska genomföras laddas hela provet ner och sparas krypterat lokalt. När provet är genomfört tas det bort från datorn. En viktig del är att CITO kan kontrollera miljön som provet ges i. Alla provtagare får korrekta prov med lika mycket tid på varje uppgift, och de kan dessutom användas även om Internetanslutningen bryts. Begränsningar med denna mjukvara är dock att den kräver användandet av en Windows plattform om man vill vara säker på att provet laddas upp korrekt, vara oberoende av Internet samt för att kunna låsa datorn förutom den del som rör provet. Slutligen beskrev Henny Tax (8 december, 2008) olika aspekter med datorbaserade provningar med det så kallade CitoTester. Tax berörde aspekter som hur vi skapar provmiljöer som är generella trots att datortillgången på skolor ser otroligt olika ut. Vidare berördes fördelar med datorbaserade prov jämfört med prov med papper och penna (för exempel se tidigare avsnitt i rapporten). Tax diskuterade olika situationer när datorbaserade prov användes, till exempel vid praktiska bedömningar, konstnärliga prov, kompetensexamination samt generella examinationer. För tillfället undersöker de möjligheter med Internetbaserade prov. Slutligen berörde han de olika problemen som skapas när datorbaserade prov används, som olika datormiljöer och olika datorer inklusive olika plattformar.

# Pilotstudier

# Pilotstudier

## Syfte med pilotstudierna

För att stärka den empiriska grunden för ställningstaganden om fortsatt utveckling av adaptiva prov i Sverige skulle även två pilotstudier genomföras, en i matematik och en i engelska. I pilotstudien skulle tyngdpunkten ligga på attityder, uppfattningar och reaktioner hos elever och lärare och eventuellt skolledare. Vidare skulle pilotstudien ha fokus på beredskapen för användning av datormediet som sådant i relation till prov och bedömning. Pilotstudierna skulle alltså inte handla så mycket om elevernas resultat, som om förutsättningarna för att kunna genomföra datorbaserade prov i svenska skolor. Pilotstudien var aldrig tänkt att vara särskilt omfattande och det fanns aldrig någon ambition att göra ett representativt urval av skolor.

För detta syfte planerades och genomfördes en pilotstudie i matematik av Umeå universitet och en pilotstudie i engelska av Göteborgs universitet.

## Pilotstudie i matematik

### Övergripande utformning

Pilotstudien i matematik använde sig av datorbaserade prov som utvecklats och används i Norge. För de norska *nasjonale prøver* har Utdanningsdirektoratet tillsammans med en projektgrupp vid Matematikksenteret i Trondheim utvecklat linjära datorbaserade prov, det vill säga prov där alla elever gör samma uppgifter. Förhoppningen var länge att kunna hitta en samarbetspartner som kunde erbjuda ett adaptivt prov, men när den möjligheten var uttömd så var möjligheten att använda de norska proven en mycket bra lösning för att kunna genomföra en pilotstudie.

Pilotstudien i matematik genomfördes på några skolor inom en och samma kommun. Den utformningen medför såväl fördelar som begränsningar. En begränsning i förhållande till pilotstudiens syften är att alla skolorna hade samma tekniska lösning för sina datorsystem, såväl hårdvara som mjukvara. Å ena sidan innebär det att pilotstudien inte ger en särskilt bra grund för slutsatser om hur olika tekniska lösningar kan påverka möjligheten att genomföra datorbaserade prov i stor skala. Å andra

sidan innebär den tekniska homogeniteten att den lösning som används i denna kommun utvärderas bättre, och att stabilitet och annan möjlig variation inom ett och samma system kan studeras.

### **Planering**

Efter att information om pilotstudien gick ut till alla rektorer i kommunen kontaktades skolor i januari 2009, och relativt snabbt upprättades kontakter med intresserade lärare. Fem olika skolor kom att ingå i försöket, med totalt fyra klasser i årskurs 3 (88 elever) och fyra klasser i årskurs 6 (102 elever). Bland annat på grund av den tidspress som uppkom av olika förseningar i projektet så informerades lärarna om att någon från projektgruppen skulle komma till skolan och hålla i informationen om försöket (mer om detta nedan).

För att skapa bästa möjliga förutsättningar för planering och genomförande åkte två av projektgruppens medlemmar (Margareta Löfstedt och Peter Nyström) till Oslo den 25 mars för att träffa ansvariga vid Utdanningsdirektoratet. Efter detta möte gjorde projektgruppen översättningar av de norska matematikuppgifterna och skickade även listor över de elever som skulle ingå i pilotstudien. Utdanningsdirektoratet erbjöd tjänsten att föra in alla elever i systemet och lägga upp användare för dem, ett arbete som i det norska systemet normalt görs av lärare ute på skolorna.

De datorbaserade prov som användes i pilotstudien används i Norge tidigt på höstterminen i femte och åttonde trinn, vilket motsvarar årskurs 4 och 7 i Sverige. Vid planeringen av pilotstudien valdes därför att låta elever i årskurs 3 och 6 göra proven, för att i möjligaste mån efterlikna mognadsgrad och antal år i skolan hos de norska eleverna. Dessa årskurser valdes också därför att de fortfarande representerar två delvis olika kulturer ("lågstadiet" respektive "mellanstadiet") och har olika lärare.

Provet som genomfördes i årskurs 3 består av 48 uppgifter. Cirka 70 procent av uppgifterna är av flervalstyp och de övriga av en typ där eleverna ska skriva in ett kort svar. De besvaras en efter en, men det finns även möjlighet att hoppa mellan uppgifterna och gå tillbaka till en uppgift man arbetat med tidigare. I Norge får eleverna 90 minuter för att arbeta med provet, men i pilotstudien fick de aldrig så mycket. I årskurs 6 innehöll provet på motsvarande sätt 58 uppgifter som skulle besvaras på max 120 minuter, men de svenska eleverna fick inte så mycket tid.

Förutom själva provet genomfördes en elevenkät innan provtillfället och elevintervjuer i direkt anslutning till att de genomfört provet. Dessutom har berörda lärare och skolledare på de deltagande skolorna intervjuats efter att proven har genomförts.

Eftersom pilotstudien skulle genomföras med omyndiga barn, så ombads vårdnadshavare att ge sitt medgivande. Institutionen tillhandahöll en information till vårdnadshavaren tillsammans med en blankett för underskrift. Lärarna i respektive klass ansvarade för att dela ut informationen och samla in underskrifterna. Det var endast några få föräldrar som nekade till att deras barn skulle delta.

## Genomförande

Pilotstudien i matematik genomfördes på skolorna under vecka 18 och 19, det vill säga mellan 27 april och 8 maj och provtillfällena var förlagda till ordinarie lektionstid. I vissa fall delades klasserna i två grupper, och klassens lärare tog hand om halva klassen medan någon från projektgruppen genomförde det datorbaserade provet med den andra halvan. Tidsåtgången vid proven anpassades efter klassens ordinarie schema, vilket innebär att provtiden varierade mellan de olika grupperna. Vissa grupper fick cirka 40 minuter och andra uppemot 80 minuter för att arbeta med de datorbaserade proven.

Det norska systemet är Internetbaserat, vilket innebär att eleverna ska logga in på en webbplats och genomföra provet när de är uppkopplade mot Internet. Det behöver inte laddas in någon speciell programvara i datorerna för att det ska fungera, utan det enda som krävs är en fungerande nätuppkoppling och en webbläsare.

## Resultat

### Tekniska aspekter

Den kommun där pilotstudien genomfördes har ett relativt väl utvecklat datasystem och de flesta skolorna har ett trådlöst nätverk. Alla skolor som deltog i pilotstudien har en eller flera så kallade nomader, det vill säga rullbara skåp som innehåller 16 bärbara datorer. Skåpet rullas in i klassrummet, ansluts till eluttag och om inte trådlöst nätverk finns tillgängligt, även till nätverksuttag. Datorerna går på batteri, vilket kräver att läraren varit förutseende och laddat dem innan provtillfället. Klass-

rummen, speciellt i äldre skolor, har normalt inte speciellt många eluttag vilket innebär ett problem om datorerna måste anslutas till ett eluttag.

Efter att varje elev hämtat sin dator måste hon eller han använda sitt eget lösenord för att logga in på datorn. I pilotstudien skulle sedan eleverna gå till en namngiven webbsida, skriva in ett användarnamn och ett lösenord som de fått på en utdelad lapp. Som en säkerhetsåtgärd skulle de i nästa steg skriva in ytterligare ett lösenord som var gemensamt för klassen och som skrevs på tavlan av den som var ansvarig för genomförandet. Därefter kunde de börja med första provuppgiften.

Elevernas uppkoppling fungerade överlag mycket bra. Någon gång hade eleverna problem vid inloggningen på datorn, eftersom de inte kände till eller kommit ihåg sina egna lösenord för att logga in på Windows. Inloggningen till det webbaserade norska systemet, med användarnamnen, lösenorden och dagslösenorden, fungerade dock pricksfritt. Hela inloggningsprocessen var i allmänhet betydligt längre för årskurs 3 än för årskurs 6. De äldre elevernas datorvana syns tydligt.

Den nätverkslösning som används i de skolor som ingår i studien har fungerat bra. Kontakten med servern bröts vid några få tillfällen för någon elev. Vid något tillfälle berodde det på att eleven klickat på tillbakaknappen i webbläsaren. Detta blev ett störningsmoment, men det ordnade sig relativt snabbt. Svaren på de uppgifter som redan var gjorda fanns kvar när inloggning gjordes på nytt. Däremot var eleven tvungen att gå igenom alla uppgifter för att svaren skulle registreras.

Vid ett tillfälle var det flera elever som inte kunde komma in på Internet, vilket ställde till med problem. Det berodde troligtvis på att det var något fel med datorerna. Vid ett annat tillfälle var tre av datorerna i nomaden utan ström. Detta löstes med att vissa elever fick sitta vid stationära datorer i ett angränsande rum. Problem har också uppstått när eleverna efter att ha gjort uppgifterna ska skicka in sina svar. Vissa gånger var de tvungna att logga in på nytt för att skicka in proven.

Uppkopplingen mot Internet har dock fungerat över förväntan. Tidigare var det vanligt att uppkopplingen bröts med jämna mellanrum. Därför fanns det farhågor innan proven genomfördes att detta skulle hända igen, men så blev det inte. Det är viktigt att tekniken fungerar eftersom barn, speciellt de som är lite mer datorovana hinner klicka flera gånger om inte datorn reagerar direkt och då kan de riskera att missa uppgifter när systemet startar igen.

Den norska webblösningen har fungerat mycket bra. Systemet är snabbt och förhållandevis driftssäkert utifrån det vi har sett. Alla lösenord krävde mycket hantering, men det löste sig smidigt.

Sammanfattningsvis gör lösningen med bärbara datorer till varje elev att provgenomförandet har gått mycket smidigt. Några datorer kunde inte startas på grund av att batterierna inte var laddade, men de var ett marginellt problem som i de flesta fall kunde lösas genom att datorn nätanslöts. Användningen av bärbara datorer innebär att eleverna måste använda datorns tangentbord och styrplatta, och att de inte har tillgång till en mus. En hel del elever är ovana vid detta och det orsakar vissa problem. Till uppgifter där de ska klicka och dra i något så blir hanteringen betydligt besvärligare när de ska göra det med en styrplatta i stället för mus.

### **Förstår eleverna vad de ska göra?**

Eleverna har haft lätt att förstå uppgifterna, även om de inte alltid vetat hur de ska räkna ut dem. De flesta uppgifterna är av flervals- eller kortsvarstyp, vilket inte kräver någon speciell instruktion. Det var endast ett fåtal uppgifter där eleverna hade problem med att förstå vad de skulle göra. En uppgift i det prov som årskurs 3 gjorde innehöll ett moment där eleverna skulle svara genom att dra upp staplar till rätt höjd. Det var inte uppenbart, och vissa elever frågade hur de skulle kunna svara. Ett annat problem orsakades av en miss i översättningen av en uppgift som innehöll ett divisionstecken. Den norska beteckningen (ett kolon) är okänd som divisionstecken för svenska elever.

Det datorbaserade provets gränssnitt innebar att det längst ned på sidan visades vilka uppgifter som eleven gjort och vilka som återstod. Detta skapade vissa problem eftersom eleverna var tvungna att skrolla ned för att se informationen på de datorskärmar som användes (15 tum). Vissa elever såg inte sammanställningen förrän de fick den visad för sig. Att olika skärmtyper skapar problem vid användning av datorbaserade prov över Internet är känt från andra sammanhang.<sup>102</sup>

---

<sup>102</sup> John Hattie, personlig kommunikation.



### **Hur används tiden – före, under och efter provtillfället?**

Koncentrationen före provtillfället har varit olika i grupperna. Många har varit fokuserade och lyssnat ordentligt, medan andra har varit mer oroliga. Även här märks det stor skillnad på årskurserna. Eleverna i årskurs 3 har generellt haft svårare att lyssna koncentrerat på genomgången inför provtillfället. Vid det tillfället när datorerna krånglade blev koncentrationen sämre.

Under provtillfället har eleverna genomgående varit mycket koncentrerade. Vi kan dock konstatera att eleverna i årskurs 3 inte orkar arbeta koncentrerat så mycket längre än 30–45 minuter och för årskurs 6 är motsvarande tid cirka 60 minuter. Det är någonting att tänka på vid konstruktionen av proven. Just detta talar för att adaptiva prov ska användas för att på så sätt minska provtiden för den enskilde eleven. Många av eleverna tyckte att uppgifterna var svåra, men de försökte verkligen efter bästa förmåga.

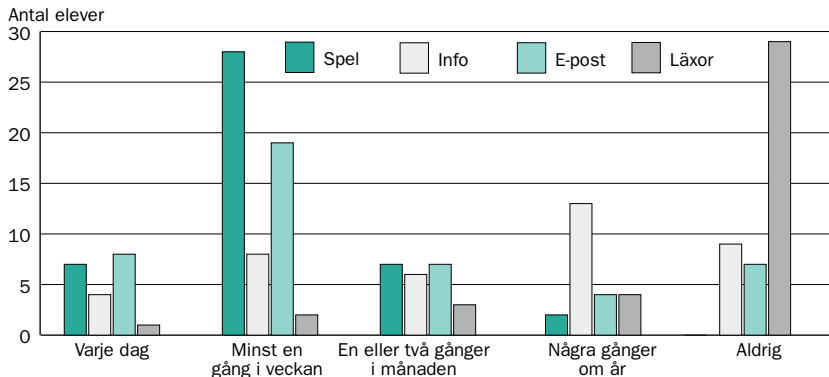
Tiden efter provtillfället har använts på lite olika sätt. Vissa klasser har haft med sig bänkbok för att använda när de har varit klara. Andra hade inte så mycket tid på sig, så att det blev någon tid över efter provgenomförandet.

### **Elevenkäten**

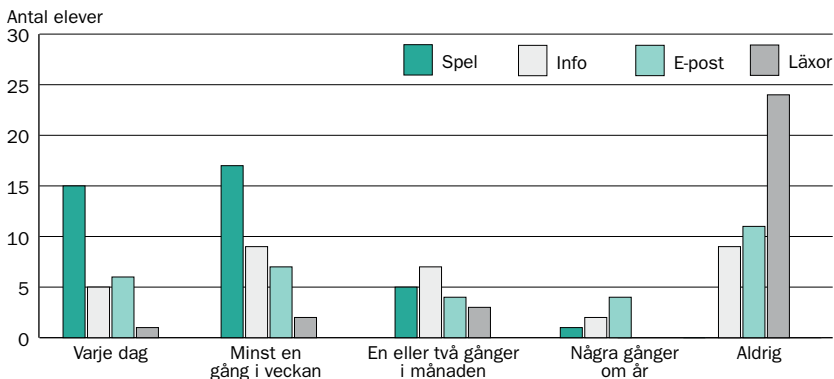
För att få veta lite mer om elevernas datorvana och om de arbetar med datorer hemma och på skolan genomfördes en elevenkät innan provtillfället.

Av de 190 elever som deltog i försöket var det endast två elever i årskurs 3 som angav att de inte använder dator hemma. De elever som angett att de använder dator hemma fick svara på ett antal frågor om vad de använder datorn till och hur mycket. Svaren fördelade sig enligt nedan.

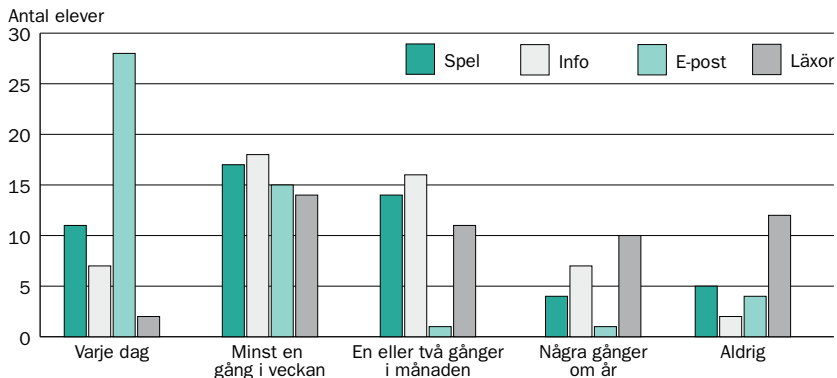
**Figur 11** Användning av dator hemma, flickor årskurs 3.



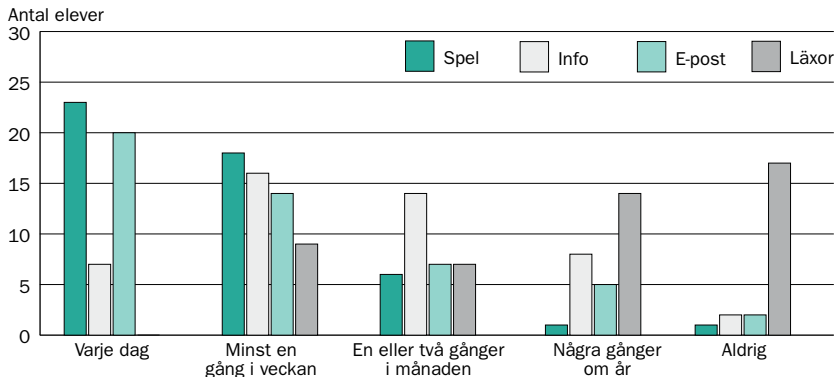
**Figur 12** Användning av dator hemma, pojkar årskurs 3.



**Figur 13** Användning av dator hemma, flickor årskurs 6.



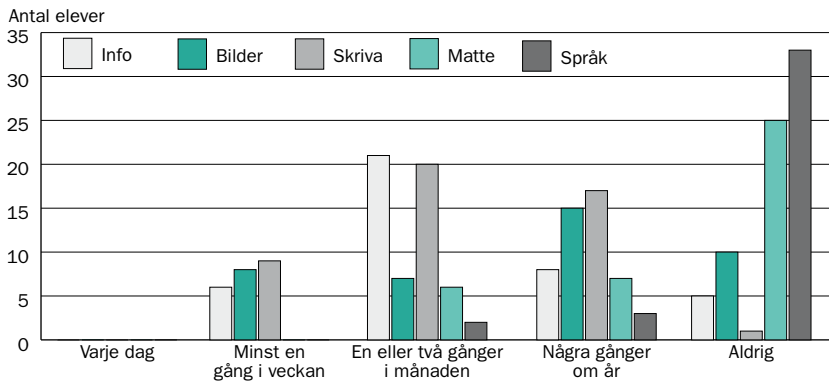
**Figur 14** Användning av dator hemma, pojkar årskurs 6.



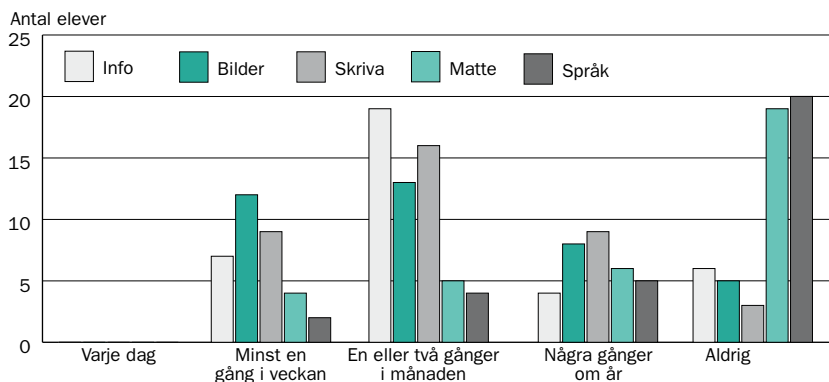
Utifrån figurerna ovan går det att konstatera att datoranvändningen ökar med ökad ålder. Eleverna i årskurs 6 använder dator hemma betydligt mycket mer än vad eleverna i årskurs 3 gör. Utifrån figurerna kan man också konstatera att pojkarna i större utsträckning spelar spel. Det gäller även om man jämför flickor och pojkar i årskurs 3. Användningen av olika kommunikationsmedel som e-post och chatt ökar med stigande ålder, men där är flickor och pojkar lika aktiva. Användning av datorn för att göra läxorna är i det närmaste obefintlig i årskurs 3 medan eleverna i årskurs 6 använder datorn regelbundet även vid läsläsning.

Då det gäller datoranvändningen i skolan är det totalt sju elever, två i årskurs 3 och 5 i årskurs sex som har angivit att de inte använder dator i skolan.

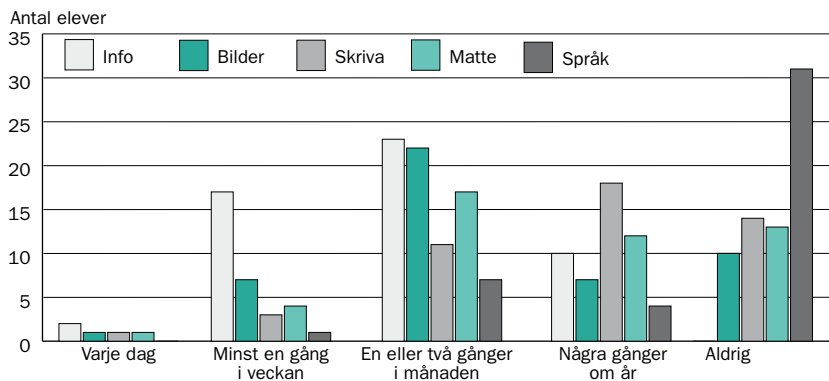
**Figur 15** Användning av dator i skolan, flickor årskurs 3.



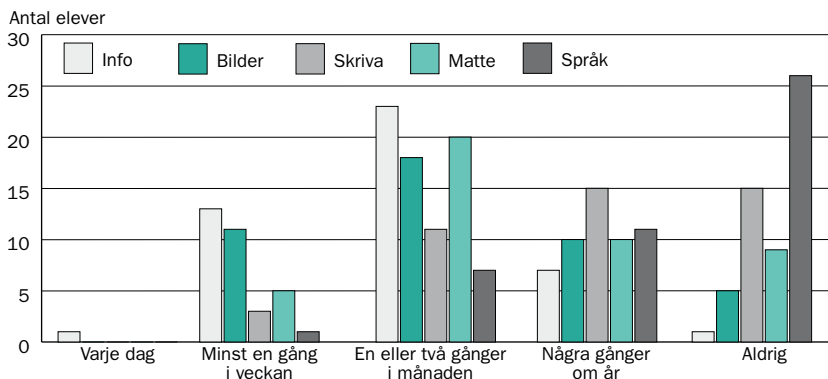
**Figur 16** Användning av dator i skolan, pojkar årskurs 3.



**Figur 17** Användning av dator i skolan, flickor årskurs 6.



**Figur 18** Användning av dator i skolan, pojkar årskurs 6.



Utifrån figurerna ovan kan man konstatera att användningen av datorer i skolan ökar något med stigande ålder. Eleverna använder datorerna till största del till att leta information och bilder samt att skriva text. Datorerna används dock även för att arbeta med matematik och språk.

På frågan om eleverna tidigare genomfört något matematikprov på datorn så svarar 181 av eleverna att de inte gjort det.

På den sista frågan om det skulle bli spännande att göra ett matematikprov via datorn så svarar närmare 83 procent av eleverna i årskurs 3 att de tycker att det ska bli spännande. Det är ingen skillnad på svaren från flickor och pojkar. I årskurs 6 är resultatet ett annat. Nästan 90 procent av flickorna men bara 64 procent av pojkarna tycker att det ska bli spännande med matematikprov på datorn.

### Elevintervjuer

Efter de flesta av provtillfällena har en grupp elever intervjuats, för att vi skulle få en uppfattning om vad de tyckt om de datoriserade matematikproven. Många elever som intervjuades uttryckte att det var roligt att göra de datorbaserade proven, och det märktes en något större entusiasm hos eleverna i årskurs 3 jämfört med årskurs 6. I årskurs 3 var det nog i högre grad att de fick arbeta med datorer som lockade och inte bara provet i sig. Det märktes även på elevintervjuerna att eleverna i årskurs 6 var något mindre entusiastiska och där fanns fler elever som tyckte att provet inte var speciellt roligt. Någon uttryckte det som att det var som i matteboken, men man fick skriva in svaret på datorn i stället för på ett

papper. Eleverna uttryckte även att det var spännande att se vad de hade för resultat. Många tyckte också att uppgifterna var svåra och att de inte kände igen vissa typer av uppgifter.

Av intervjuerna framgick att eleverna uppfattade de flesta uppgifterna i provet som lättbegripliga. I en uppgift skulle staplar i ett stapeldiagram dras upp till rätt höjd genom att ta tag i stapeln, klicka fast den och dra upp den med musen eller styrplattan. Detta uppfattades som svårbegripligt av en del elever.

I elevintervjuerna ställdes frågan om det var någon uppgift som var särskilt svår att lösa. I årskurs 3 tyckte eleverna att divisionsuppgifterna var svåra, men även decimaltalen ställde till problem. I årskurs 6 nämnde eleverna frågor om euro, statistik och diagram som svåra. Någon tyckte också att det var lite för mycket frågor som hade anknytning till Norge. Vilka frågor som var lätta var inte så enkelt för eleverna att svara på.

När eleverna fick chansen att ta upp något annat de tänkt på, så var eleverna i årskurs 3 mer talföra än eleverna i årskurs 6. Eleverna i årskurs 3 tyckte det var jätteroligt att arbeta med uppgifterna i datorn, mycket roligare än matteboken. ”Kanske man kan ha matteboken på datorn”, var det några som funderade. En flicka i sexan tyckte det kändes bra att göra prov på datorn, att hon kände sig som en del av det moderna samhället, ”uppdaterad helt enkelt”.

### Intervjuer med lärare och rektorer

Lärarna i de klasser som deltog i pilotstudien intervjuades också. När det gällde datorvana så svarade lärarna alltifrån ”inte speciellt van” till ”mycket stor vana”. Även svaren på frågan om hur ofta lärarna använder datorn i undervisningen, och specifikt i matematik, visade stor spännvidd. Lärarna hade svårt att kvantifiera sin datoranvändning, någon sa tre gånger per månad, och vi kan konstatera att datorn nästan inte alls används i matematik.<sup>103</sup> Någon nämnde att den används till att göra diagram vid något tillfälle och vissa skolor använder datorn för att spela matematikspel på datorn, till exempel ”Chefrens pyramid”.

Enligt lärarna på de deltagande skolorna går det 6–8 elever på varje dator. Lärarna är nöjda med tekniken och tycker att det har börjat fungera bättre efter att ”nomaderna” börjat användas. Det problem som lärarna framför allt haft erfarenhet av är att uppladdningen av de bär-

---

<sup>103</sup> Den låga datoranvändningen i matematik har även konstaterats i TIMSS (se till exempel Skolverket, 2008a).

bara datorerna krånglar. Det som också kan krångla är den så kallade multimedianoleden. Detta är ett liknande skåp som nomaden, men denna innehåller en dator och en projektor. Projektorn ger möjlighet att visa digitala bilder på storbild.

I intervjuerna tillfrågades lärarna om vilka fördelar och nackdelar de kan se med datoriserade prov. En fördel som framkommer i lärarintervjuerna är att rättningsarbetet med proven borde minska markant. Lärarna uttrycker också att det är en fördel att barnen tycker att det är spännande med datorer och att datorbaserade prov kan vara bra för skrivsvaga barn. En nackdel med datoriserade prov som lärarna tar upp är att svarsformatet innebär att man inte kan följa elevernas tankegångar, något som de menar prioriteras i kursplanerna.

Även rektorerna vid de deltagande skolorna intervjuades. På frågan om hur mycket skolans datorer används svarar alla rektorer att de används ofta och att användningen har ökat efter införandet av de så kallade nomaderna. Alla rektorer är också överens om att datorerna nästan aldrig används i ämnet matematik. Enligt rektorerna går det 6–8 elever på varje dator i de deltagande skolorna. När det gäller hur väl tekniken fungerar var vissa rektorer inte alls nöjda. De menar bland annat att uppkopplingen till Internet strular eller är trög. En rektor saknade teknisk support på skolan och en annan tyckte att underhållet hade minskat väsentligt efter nomadinförandet. Som fördelar med datoriserade prov nämnde rektorerna mindre pappershantering, mer lästrättat för lärarna och stöd för dyslexieleverna som positiva saker med datoriserade prov. Negativa synpunkter på datoriserade prov var att allt blir ”fyrkantigt” och att eleverna missar att göra saker för hand.

### Slutsatser av pilotstudien i matematik

Övergripande fungerade de tekniska lösningarna mycket väl. Det norska webbaserade systemet fungerade problemfritt och de datorer och nätverklösningar som fanns på de skolor som deltog fungerade också väl. Det måste dock påpekas att alla deltagande skolor hade samma tekniska lösningar, och studien ger därför ingen information om hur olika system skulle kunna fungera. Genom att eleverna hade tillgång till var sin bärbar dator så gavs mycket goda förutsättningar för att arbeta med de datorbaserade proven. Samtidigt kan bärbara datorer innebära att eleverna placerar sig i klassrummet på ett sätt som inte är helt optimalt för

ett prov som bygger på individuellt arbete. Om de får sätta sig var de vill eller som de sitter i vanliga fall finns en uppenbar risk att de också samarbetar, vilket inte var tanken med de prov som användes i denna pilotstudie. Det är angeläget att även elevernas placering i rummet planeras av läraren innan genomförandet. Trots att eleverna hade använt dessa datorer tidigare så var flera ovana vid datorernas styrplattor. Det är viktigt att uppgifter i datorbaserade prov inte utformas så att de är beroende av en viss teknisk utrustning som en datormus eller en skärm av en viss storlek.

Eleverna visade sig intresserade, nyfikna och allmänt positiva till att göra ett datorbaserat prov i matematik. Detta trots att uppgifterna i sig varken var särskilt nyskapande eller alternativa i förhållande till de uppgifter eleverna arbetar med på lektionstid. I något fall var också den positiva inställningen i något fall större före genomförandet än efter. Observationer i klassrummet tydde på att eleverna i årskurs 3 var ännu mer positiva till att arbeta med det datorbaserade provet än eleverna i årskurs 6. Dessutom fanns tendenser till att elever som inte använder dator så mycket blev mer stimulerade av att göra uppgifter i datorn än de som var mer erfarna av datoranvändning. Koncentrationen var på en mycket hög nivå under genomförandet, trots att uppgifterna var svåra för många. Vissa könsskillnader kunde observeras, framför allt att pojkar hade en tendens till tävling, till exempel om hur många uppgifter de kunde hinna, medan flickorna ägnade sig åt uppgifterna med litet större noggrannhet. När det gäller lärarna så tycks datorvana lärare se större fördelar och vara mer positiva till datoriserade matematikprov än de som har mindre datorvana.

## Pilotstudie i engelska

### Övergripande utformning

Som ett praktiskt och empiriskt komplement inom ramen för denna kunskapsöversikt ansvarade NAFS-projektet (Nationella prov i främmande språk) vid Enheten för språk och litteratur, Göteborgs universitet, bland annat för en pilotstudie med datorbaserade provuppgifter i engelska vid sex skolor i Västra Götaland under vårterminen 2009. I ett samarbete med universitetet i Bergen öppnades möjligheten att använda 2008 års norska datorbaserade nationella prov i engelsk läsning i den svenska pilotstudien i engelska. Datorbaserade prov i engelsk läsförståelse



infördes i Norge 2004 på uppdrag av stortinget. Proven avser att mäta måluppfyllelse av kunskapsmålen för läsning i engelska för fjärde (9–10 års ålder) och sjunde (12–13 års ålder) trinn/årskurserna<sup>104</sup> och bygger på beskrivningarna för läsförståelse i ”Common European Framework of Reference for Languages”.<sup>105</sup>

Genom avtalet med universitetet i Bergen förmedlades också samarbete med det norska företaget ENOVATE AS (<http://enovate.no>) som utvecklar programvara och nätbaserade lösningar för utbildnings-sammanhang. ENOVATE har skött tekniken för de norska nationella proven och bistod också vår västsvenska pilotstudie med all distribution och löpande support.

För att kunna skapa en bred och god bild av attityder, uppfattningar och reaktioner användes olika metoder för insamling av underlag. Den mesta informationen samlades in i form av enkätsvar, men också i samband med deltagande observation samt vid informella samtal och kontakter via e-post. I studien ingick så småningom sex grundskolor och cirka 250 såväl yngre som äldre elever i årskurserna 5, 6 och 8, samt deras 13 lärare och 8 skolledare.

Flera överväganden i pilotstudien har sin grund i erfarenheter gjorda under det inledande och inventerande planeringsarbetet under hösten. Längre ambitionen att i enlighet med Skolverkets uppdrag genomföra en pilotstudie med en adaptiv provutformning. Det visade sig dock mycket svårt att åstadkomma med den tid som stod till buds. När möjligheten att använda de norska datorbaserade proven öppnade sig så ansågs detta vara ett intressant alternativ, trots att de inte är adaptiva. Vidare beslutades att i pilotstudien använda uppgifter med fokus på receptiva färdigheter – läsning med slutna svarsformat, som rättas löpande på maskinell väg. Detta är den uppgiftstyp som hittills visat sig fungera bäst i datorbaserade prov.

Då pilotstudien bland annat riktar sig till omyndiga elever och det möjligen kan förekomma en viss misstänksamhet inför registrering av resultat och liknande, beslutades att rekommendera skolorna att ta in vårdnadshavarnas medgivanden. För att bespara lärarna tid, och för att det skulle bli enhetligt, tillhandahöll vi en sådan blankett. Dock överlämnade vi ansvaret till skolan att distribuera och samla in densamma.

---

<sup>104</sup> Den norska läroplanstexten LK06 återfinns på [www.utdanningsdirektoratet.no/grep/Lareplan/?laereplanid=122422](http://www.utdanningsdirektoratet.no/grep/Lareplan/?laereplanid=122422).

<sup>105</sup> Council of Europe (2001)

När det gäller urval av skolor och elevgrupper fanns en strävan att inom det begränsade antalet ändå få viss representation av något olika skolor till exempel avseende storlek och geografiskt läge. Eftersom flera besök och ett relativt nära samarbete var en förutsättning för att få en god bild av pilotstudiens genomförande, riktades urvalet till sex kommunala skolor i Västra Götaland, tre i vardera två olika kommuner – en storstadskommun och en landsortskommun. Skolorna har mellan 100 och 500 elever. Två av skolorna har åldersblandade klasser, en skola har grupper som undervisas enligt Montessoripedagogik och en skola har enligt sin webbplats en tydlig IT-satsning. En skola deltar i pilotstudien med både yngre och äldre elever.

Uppdraget var att genomföra en pilotstudie med elever i såväl yngre som äldre åldrar i grundskolan. Eftersom det under rekryteringsskedet verkade som vi skulle låna danska provuppgifter för motsvarande sjunde årskursen i Danmark, engagerades årskurs 8 och 6, för att sedan, när avtalet med Bergen stod klart, även ta med elever i årskurs 5 i de fall det gick att ordna på skolorna. Anledningen är att de norska proven riktar sig till yngre barn, men det föreföll också rimligt att ta hänsyn till att årskurs 5, just vid tiden för pilotstudien, genomförde de nationella ämnesproven. Totalt engagerades 276 elever från de olika skolorna.

De tretton lärare, tre män och tio kvinnor, som involverades i pilotstudien är alla lärare i engelska eller klasslärare som undervisar i engelska. Av dessa tretton lärare kom elva att genomföra prov med elevgrupper. De sex skolorna representeras också av åtta skolledare. En skola har två rektorer och på en skola har rektorsuppgiften för yngre respektive äldre åldrar uppdelad på två tjänster.

### **Provuppgifter**

De nationella proven för engelsk läsning i Norge genomförs under en fjortondagarsperiod alldeles i början av femte respektive åttonde trinn/årskurs. Proven är linjära (det vill säga att alla elever gör samma prov) och varje år ges tre likvärdiga versioner som slumpas ut, för att de ska kunna genomföras vid något olika tidpunkter på skolorna.

I båda proven förekommer rikligt med bilder och olika typer av uppgiftsformat, men endast slutna svarsformat, det vill säga olika typer av flervalsuppgifter som rättas löpande. Ämnesmålen för läsförståelse i engelska i de norska respektive svenska styrdokumenterna är i stort mycket lika varandra, men de norska proven har också ett uttalat syfte att testa

elevens förmåga att använda datormediet, IKT. Provet för årskurs 5 består av 38 uppgifter som prövar elevens förmåga att finna information och förstå huvudinnehållet i texter, som till exempel kunskapsmålet att förstå engelska ord och uttryck knutna till vardagsliv, fritid och intressen. De 40 uppgifterna för årskurs 8 prövar förmåga att finna information, förstå huvudinnehållet i texter samt reflektera över innehållet i texter. Några uppgifter handlar således om att lokalisera detaljer och andra om att dra slutsatser utifrån given information. Uppgifterna i de båda proven är när det gäller format och tillvägagångssätt nästan identiska, men de yngre eleverna behöver aldrig skrolla i längre texter. Eleven svarar till exempel genom att flytta bildobjekt, klicka på ord eller bilder, färglägga föremål eller välja bland givna namn, ord eller fraser i flervalsalternativ.<sup>106</sup>

Provuppgifterna rättas maskinellt, och normalt sett får de norska eleverna sitt resultat omedelbart efter avslutat prov. För pilotstudiens del beslutades dock att meddela skolorna resultaten först efter studiens slut, då eventuellt framför allt elevernas utvärdering av utprövningen annars skulle riskera att färgas av deras individuella prestationer.

## Enkäter

Samtliga enkäter utformades när det gäller frågeformat och även en del frågeställningar enligt modeller som under många år använts inom NAFS-projektet för så kallad Test Taker Feedback (TTF) och Teacher Feedback (TF), till exempel vid utprövningar av de nationella proven i engelska.<sup>107</sup> De innehåller slutna format med fem-punkts så kallade Likert-skolor, i vilka eleverna ombeds ta ställning till ett påstående genom att markera något av svaralternativen från ”Ja, absolut” till ”Nej, absolut inte” med tre steg däremellan, samt möjlighet att lämna öppna kommentarer. Eleverna gjorde en enkät före (Elevenkät I) och en enkät efter (Elevenkät II) provtillfället. Lärarna och skolledarna besvarade digitala enkäter efter periodens slut, vilka utväxlades per e-post. Efter som planeringsbesöken med lärare och skolledare inför pilotstudien erbjöd möjlighet att få en uppfattning om deras attityder, datorvanor och liknande, bedömdes det när det gäller dessa vara tillräckligt med en enkät efter avslutandet av studien.

---

<sup>106</sup> Moe, Carlsen & Hasselgren (2006)

<sup>107</sup> Erickson (1998)

## Genomförande

### Förberedelsearbete, planering och arbetsgång

Genomförandet av provuppgifterna ute på skolorna föregicks av en relativt omfattande planering och information. Kontakter inleddes med intresserade lärare och skolledare under oktober månad, och samtliga tillfrågade skolor visade sig intresserade av att delta. Kontakten togs först muntligt eller per e-post och följdes av ett beskrivande brev till rektor och berörda lärare.

Under januari och februari 2009 genomfördes planeringsmöten på samtliga skolor med vidtalade lärare och på tre skolor även med skolledare. Då diskuterades det praktiska genomförandet, skolans förutsättningar vad gäller teknisk utrustning samt till exempel de schemamässiga anpassningar som skulle behöva göras i de fall elevgruppen skulle behöva ta flera lektionstillfällen i anspråk för att alla skulle kunna få tillgång till datorer. Planeringsbesöken berörde även pedagogiska frågor och specifika frågor kring den egna skolans förutsättningar, lärarens datorvana eller den aktuella elevgruppen.

Vid planeringsmötet fick lärarna en presentation av studien, med bifogad blankett för godkännande, att distribuera till vårdnadshavare. De fick också skriftlig information att delge eleverna och den inledande attitydenkäten (Elevenkät I) att genomföra innan eleverna tagit del av exempeluppgifter eller gjort provet. Kontakten med och godkännande av vårdnadshavare lämnades därmed över till skolan att ansvara för och följa upp.

Genomförandet planerades i samtliga grupper till veckorna 11–14 före påsklovet, och ett besök för deltagande observation vid ett provtillfälle per skola planerades in. I de flesta fall föredrog lärarna att besöket förlades till det första provtillfället. Elevenkät II genomfördes direkt efter provet för de flesta elever och kunde distribueras vid observationstillfället då också den inledande enkäten samlades in.

## Resultat

### Klassrumsobservationer

Vid ett provtillfälle per skola genomfördes en deltagande observation. Sammanfattningsvis var eleverna mycket intresserade och engagerade, samt allvarliga och måna om att göra bra ifrån sig och bidra till pilot-

studien genom att kommentera proven i enkäterna. Flera påpekade spontant att de tyckte det var bättre med datorprov än ”vanliga” prov, och många kröp djupt in i skärmen när de gjorde provuppgifterna. Det var en del elever som ville förstora texten på skärmen och några, framför allt de som inte gjort de Internetbaserade exempeluppgifterna, behövde viss stöttning. Lärarens förkunskaper om provet, till exempel genom demo-versionen, verkade viktig. Det första provtillfället innebar förmodligen de största tekniska utmaningarna och utan tvekan krävdes också en del organisatoriska förberedelser för att kunna dela grupper etc. Samtliga besökta skolor har datasalar eller kombinerade studie- och datasalar med väggfasta datorer, vid vilka eleverna sitter relativt tätt bredvid varandra. Generellt föreföll den tekniska standarden relativt låg och vid varje skola var någon dator ur funktion. Helhetsintrycket blev dock att såväl lärare som elever välkomnade studien och att de generellt är både positiva och optimistiska när det gäller datoranvändning.

### **Elevenkäter**

De två elevenkäterna har som syfte att fånga upp elevernas förväntningar och tankar innan de sett proven, att jämföra med deras erfarenheter och slutsatser efter det att de genomfört ett datorbaserat prov. Efter ett visst bortfall var det totalt 248 elever som besvarade någon av enkäterna och av dem gjorde 209 elever båda enkäterna. 237 elever besvarade Enkät I och 220 Elevenkät II. Det var i stort sett lika många pojkar som flickor, samt ungefär lika många elever fördelat över äldre (årskurs 8) och yngre åldrar (årskurs 5 och 6) som gjorde någon av enkäterna. Däremot var det nästan dubbelt så många elever från de tre storstadsskolorna: 165 elever jämfört med 83 från de tre landsbygdsskolorna.

### **Elevenkät I**

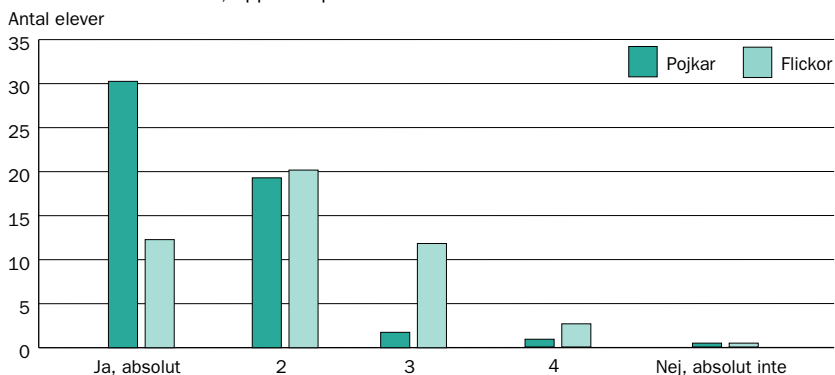
Elevenkät I genomfördes alltså tidigt under pilotstudiens gång, innan elevernas uppfattningar fångats av det datorbaserade provet. Enkätens inleds med påståenden om elevernas inställning till och kunskaper om datorer, samt deras datorvanor i skolan och hemmet. En stor andel av eleverna menar att påståendena stämmer mycket väl in på dem och markerar något av de två positiva alternativen på Likert-skalan, dock med en markant lägre överensstämmelse när det gäller datoranvändning i skolan:

**Tabell 6** Påståenden och svarsfrekvenser från Enkät I till pilotstudien i Engelska.

Påstående	Andel elever som valt ett av de två mest instämmande svarsalternativen
Jag gillar datorer	93 %
Jag är bra på (att använda) datorer	82 %
Jag använder ofta dator i skolan	35 %
Jag använder ofta dator hemma	90 %

Vid en uppdelning av elevsvaren på pojkar och flickor framträder relativt stora skillnader. Pojkarna har större självförtroende när det gäller att använda datorer.

**Figur 19** Elevernas grad av instämmande i påståendet "Jag är bra på (att använda) datorer", uppdelat på kön.



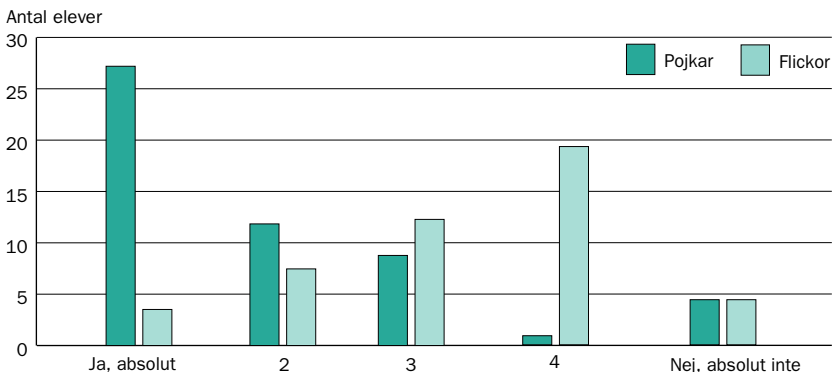
Vid en analys av enkätsvaren visar det sig att de elever som deltog i pilotundersökningen tycker om ämnet engelska, att de har ett gott självförtroende i ämnet samt att ganska exakt hälften av dem har erfarenhet av att använda engelska vid datorn, även när det gäller pedagogiska sammanhang:

**Tabell 7** Påståenden och svarsfrekvenser från Enkät I till pilotstudien i Engelska.

Påstående	Andel elever som valt ett av de två mest instämmande svarsalternativen
Jag gillar engelska	77 %
Jag är bra i engelska	71 %
Jag använder ofta engelska vid datorn	51 %
Jag har gjort språkövningar på datorn förut	48 %

En ytterligare granskning av svaren på frågan om eleverna använder engelska vid datorn, visar inte några stora skillnader mellan åldersgrupperna men däremot markanta könsskillnader (se diagram nedan). Ett mycket snarligt påstående förekommer i elevenkäten till de nationella proven i engelska för årskurs 5 ("I use English when I do things on a computer") och visar där liknande mönster enligt Skolverkets ämnesrapport 2008:<sup>108</sup> "Mötet med engelska i samband med datoranvändning ökar för varje år. I år uppger 65 procent av eleverna att de använder engelska i sådana sammanhang. Bland dessa elever finns dubbelt så många pojkar som flickor."

**Figur 20** Elevernas grad av instämmande i påståendet "Jag använder ofta engelska vid datorn", uppdelat på kön.



<sup>108</sup> Skolverket (2008b)

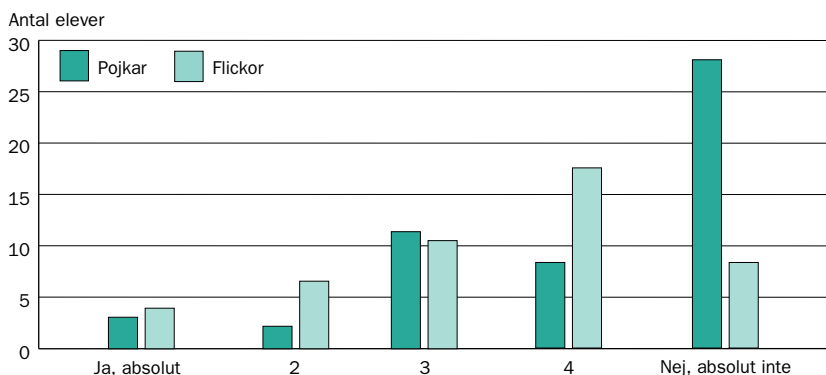
Redan innan eleverna sett och provat på de datorbaserade proven är de allra flesta mycket positiva till tanken och känner tillförsikt. Några av dem skulle dock föredra ett traditionellt pappersprov och ungefär lika många uttrycker genom sina svar att de känner en viss osäkerhet (se tabell 8).

**Tabell 8** Påståenden och svarsfrekvenser från Enkät I till pilotstudien i Engelska.

Påstående	Andel elever som valt ett av de två mest instämmande svarsalternativen
Jag tycker det verkar bra att ha engelskprov på dator	86 %
Jag skulle hellre göra ett vanligt prov med papper och penna	16 %
Jag behöver hjälp för att kunna göra prov på dator	15 %
Jag känner mig beredd på att göra det här provet	83 %

Skillnaden är åter ganska stor mellan flickors och pojkars inställning till att hellre göra vanligt prov med papper och penna:

**Figur 21** Elevernas grad av instämmande i påståendet "Jag skulle hellre göra ett vanligt prov med papper och penna", uppdelat på kön.





Knappt hälften av eleverna utnyttjar möjligheten att ge skriftliga kommentarer i enkäten, och bland dessa kommentarer är nästan hälften odelat positiva (endast en elev ger uttryck för en odelat negativ känsla inför datorbaserade prov). De allra flesta skriver kort och gott att det verkar bra eller mycket bra, och många uttrycker också att det ska bli roligt eller roligare att göra ett sådant prov. En del elever motiverar sig något och gör då en uttalad eller uttalad jämförelse med pappersprov, och några elever menar att det är lättare eller går snabbare och är smidigare att göra prov på datorn.

Några elever lyfter fram hur skivarbetet påverkas av datoriseringen av proven. De menar bland annat att det är odelat bättre att få skriva på dator, medan några andra problematiserar kring att de/man eventuellt behöver skriva för hand – för att träna till exempel sin handstil eller för att det är viktigt för ämnets skull. En elev skriver att ”Det kanske skulle vara bättre i och med att man slipper vässa pennor, sudda m.m.”, och en annan att ”Det är nog bra bara för vissa kanske är dåliga på att stava och får till orden bättre då. Det tycker jag är en stor möjlighet.” Enstaka elever tycker att det vore bättre att få skriva för hand, till exempel om man inte är så snabb vid tangentbordet.

Förutom de elever som skriver om datorn som skrivverktyg, är det ytterligare några som tar upp datoranvändning, och som är mer tveksamma till andra aspekter. Några elever tar upp en skillnad i förutsättningar beroende på datorvana och några elever efterlyser IT-utbildning för elever, men eventuellt också för lärare. Det förekommer också ett par andra reflektioner kring datoranvändning, till exempel när det gäller skärm- och migränkänslighet.

En grupp elever väljer att skriva något om att de ser nackdelar när det gäller möjligheterna att fuska vid datorbaserade prov. De verkar ha vissa farhågor om att det kan bli svårt att genomföra proven under sådana förhållanden som de förväntar sig vid prov. Här återfinns kommentarer som ”Vid en dator kan du oftast komma ut på internet eller nå filer vilket kan leda till fusk” och ”Vad som inte är så bra är att om alla gör provet i en datasal är det lätt att tjuvkika på grannens prov.”

## Elevenkät II

Elevenkät II besvarades av 220 elever, i de flesta elevgrupperna direkt efter genomförandet av det datorbaserade provet. Enkäten tar upp 16 påståenden, varav några har direkt anknytning till eller är likadana som påståenden på den första enkäten. Generellt syftar denna elevenkät till att fånga upp elevernas utvärdering – deras nyvunna erfarenhet av, upplevelse av och attityder till att göra ett datorbaserat prov i engelska samt möjliggöra vissa jämförelser med deras tankar innan de gjort utprövningen.

De fyra första påståendena på Elevenkät II vill fånga upp den omedelbara upplevelsen av och attityden till att göra proven. Resultaten visar att de allra flesta tyckte om att göra provet, att det kändes lättillgängligt och ganska få elever upplevde att de för egen del hade behövt mer hjälp. Dock menar hälften av dem att det inte är lika lätt för alla att göra datorbaserade prov. Påståendet kring rättvisa kan naturligtvis tolkas på en mängd sätt, men kan i viss mån ändå indikera hur elever problematiserar kring begreppet och då komplettera frågor kring tillgänglighet. Eleverna verkar å ena sidan inte tycka att datorbaserade prov innebär någon stor risk för ojämlikhet eller särbehandling ur rättvisesynpunkt, men å andra sidan kan i detta sammanhang andelen som markerar Nej-alternativ eller skalans mittalternativ anses vara anmärkningsvärd – 27 respektive 51 procent.

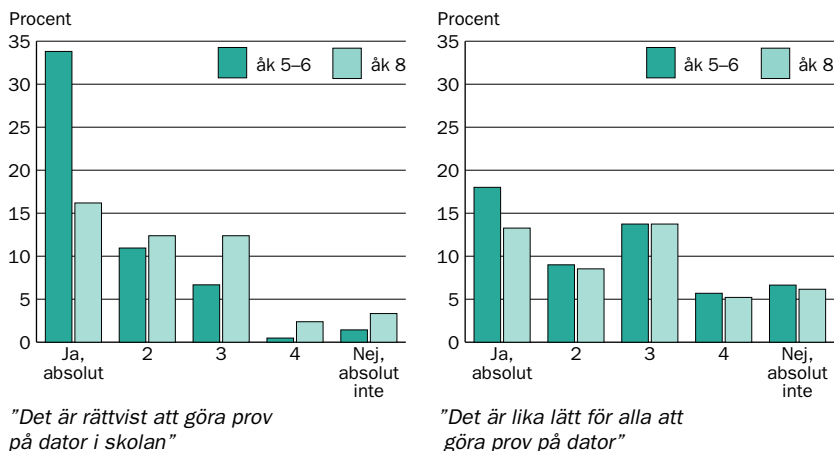
**Tabell 9** Påståenden och svarsfrekvenser från Enkät I till pilostudien i Engelska.

<b>Påstående</b>	<b>Andel elever som valt ett av de två mest instämmande svarsalternativen</b>
Jag gillade att göra engelskprov på dator	71 %
Det var lätt att förstå hur man skulle göra	75 %
Det var till hjälp att göra exempeluppgifterna innan*	56 %
Jag skulle ha behövt mer hjälp för att kunna göra prov på dator	10 %
Det är rättvist att göra prov på dator i skolan	73 %
Det är lika lätt för alla att göra prov på dator	49 %

\* De exempeluppgifter som fanns att tillgå före provet låg på norska Utdanningsdirektoratets webbplats, men fungerade inte för alla grupper och 78 elever har inte kunnat besvara nr 3. Förmodligen påverkar detta resultatet. För 2009 års exempeluppgifter se [www.utdanningsdirektoratet.no/Artikler/\\_Nasjonale-prover/Eksempeloppgaver-til-nasjonale-prover-2009/](http://www.utdanningsdirektoratet.no/Artikler/_Nasjonale-prover/Eksempeloppgaver-til-nasjonale-prover-2009/).

Fördelningen av elevsvaren på hela den femgradiga skalan framgår av figuren nedan.

**Figur 22** Elevernas grad av instämmande i påståendena "Det är rättvist att göra prov på dator i skolan" och "Det är lika lätt för alla att göra prov på dator", uppdelat på åldersgrupp.



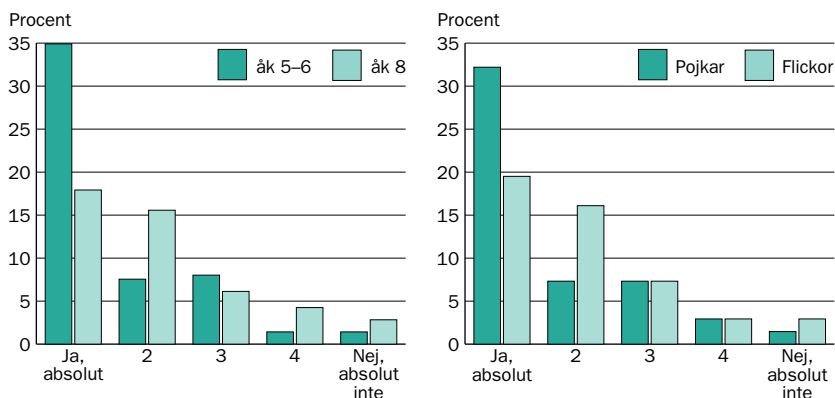
Fyra frågor handlar om elevernas helhetsintryck och sammanfattande åsikter om de datorbaserade provens vara eller icke vara, på deras hemmaplan och i deras framtid. Tre fjärdedelar av eleverna är positiva eller mycket positiva till datorbaserade prov i engelska och detsamma gäller genomförandet på den egna skolan. 46 elever av 220 föredrar att göra vanliga pappersprov. Vid en närmare granskning av den sista frågan framkommer att pojkar är mer positiva än flickor, och att årskurs 8 är lite mer tveksamma än årskurs 5 och 6 inför tanken att införa fler datorbaserade prov i framtiden (se figur 23 nedan).

**Tabell 10** Påståenden och svarsfrekvenser från Enkät I till pilotstudien i Engelska.

Påstående	Andel elever som valt ett av de två mest instämmande svarsalternativen
Det fungerade bra på min skola att göra prov på dator	74 %
Jag tycker det är bra att ha engelskprov på dator	78 %
Jag gör hellre vanliga prov med papper och penna	22 %
Jag tycker skolor i Sverige ska ha fler datorprov i framtiden	76 %

Spridningen över hela den femgradiga skalan för påståendet ”Jag tycker skolor i Sverige ska ha fler datorprov i framtiden” visar några skillnader mellan flickor och pojkar samt mellan åldersgrupper:

**Figur 23** Elevernas grad av instämmande i påståendet ”Jag tycker skolor i Sverige ska ha fler datorprov i framtiden”, uppdelat på åldersgrupp respektive kön.



Drygt en tredjedel av elevernas öppna kommentarer i Elevenkät II visar att de upplever en ökad motivation när de föreställer sig möjligheten att göra datorbaserade prov. Till övervägande del skriver de att de tycker

det är roligt ("Det var jättekul och jag skulle kunna göra det igen!"), bra och att de vill ha fler liknande prov, men det finns också en relativt stor grupp som utvecklar sina omdömen ytterligare och menar till exempel att det blir intressantare, lättare att koncentrera sig, inte så nervöst och skönt att slippa skriva. Endast två elever av 124 i hela urvalet öppna kommentarer är enbart negativa, utan att på ett konstruktivt sätt motivera eller precisera sig ("Kunde inte koncentrera mig, det tog tid och det var tråkigt.")

Drygt 20 elever tar upp problem som relaterar till datoranvändning i allmänhet och miljön i skolornas datasalar i synnerhet. De flesta tar upp att man sitter tätt bredvid varandra och att man därför lätt blir störd eller ser varandras skärmar och kan fuska. Ungefär tio elever skriver specifikt om att det är tröttsamt att läsa på datorskärmen och ytterligare andra om obekväma stolar eller att program kan haka upp sig. De allra flesta elever flikar dock in positiva kommentarer, till exempel den elev som skriver "Jag tycker att det är lite mer intressant att göra proven i datorn och det är lättare. Men i så fall borde skolan byta datorskärmarna." En elev uttrycker också att datorbaserade prov är mer miljövänliga i vidare mening: "Det är miljövänligt och inget slöseri med papper!"

Cirka en femtedel av elevernas kommentarer belyser deras funderingar kring sina egna och andras olika förutsättningar att genomföra datorbaserade prov, i både positiva och negativa hänseenden. Några verkar inspirerade av enkätfrågorna och anknyter till formuleringar i dem, men flera tar också upp nya infallsvinklar. De resonerar kring datorvana, ämneskunskaper och läs- och skrivsvårigheter, och menar att det finns olika beredskap och förmåga att göra proven. En elev skriver till exempel: "Det kan både vara bättre och sämre att göra det på dator, för att man kanske kan utvärdera mer när man har skriftligt." En annan lyfter fram datorvanan: "Frågan: "Det är lika lätt för alla att göra prov på dator" det beror på vad man har för data kunskaper. När man gör det här provet kan man ju inte visa hur bra man är på att läsa och uttala allt på engelska." Några tar upp provens domäntäckning och variation. De menar till exempel att de inte får visa om de kan annat inom ämnet än det provet testas och att skolan bör variera med skriftliga prov med papper och penna. Någon elev berör bedömning och tycker att det verkar lättare på papper. En elev tar upp att det är bättre om man kan förbereda sig inför provet.

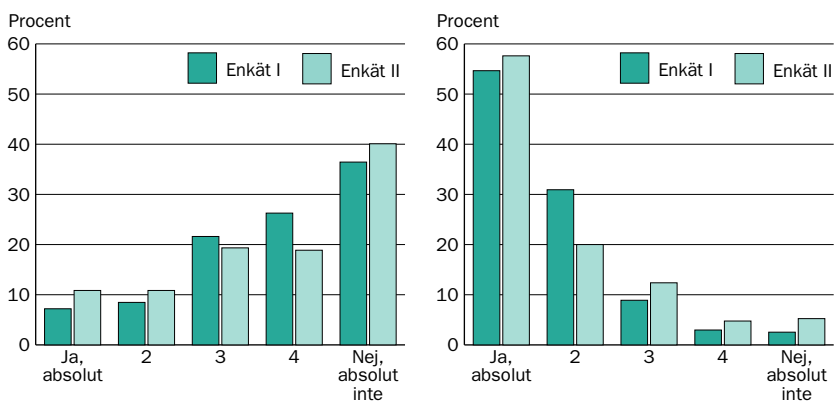
Ett tiotal elever kritiserar funktioner i provet och har också konstruktiva förslag på förbättringar. Flera av dem har råkat ut för att inte kunna gå tillbaka i provet och rätta till eller kontrollera sina svar, vilket bör ha fungerat inom ramen för en uppgift men inte mellan uppgifter. En elev skriver: ”Jag tycker att det är värdelöst att man inte kan gå tillbaka och ändra svar som man skrev till exempel i början. Det kan man med papper och penna!! Annars var det bra.”

### Sammanfattande kommentar till de båda elevenkäterna

Eftersom flera påståenden är helt eller i stort sett identiska på de båda enkäterna möjliggörs jämförelser mellan hur eleverna uppfattade något innan de sett eller gjort proven samt hur de utvärderar (i stort sett) samma frågeställning efteråt.

Två par utsagor som har direkt anknytning till pilotstudiens syfte att ta reda på attityder till datorbaserade prov jämfört med traditionella prov är ”Jag tycker det verkar/är bra att ha engelskprov på datorn” respektive ”Jag skulle hellre göra/gör hellre vanliga prov med papper och penna”. Eleverna är som grupp något mer positiva till datorbaserade prov i jämförelse med vanliga prov med papper och penna efter genomförandet av proven jämfört med före (se figur 24 nedan).

**Figur 24** Elevernas grad av instämmande i påståendena ”Jag gör hellre vanliga prov med papper och penna” respektive ”Jag tycker det verkar/är bra att ha engelskprov på datorn”, uppdelat på enkättillfälle. Enkät I gavs före, och Enkät II efter, genomförandet av det datorbaserade provet.



*"Jag gör hellre vanliga prov med papper och penna"*

*"Jag tycker det verkar/är bra att ha engelskprov på datorn"*

Eleverna har lagt ner stort arbete i sina enkäter på att ge konstruktiv feedback, samt på att i de öppna kommentarerna klä sina reflektioner i ord och tillföra ytterligare värdefulla aspekter som inte tas upp i de givna påståendena. Att sammanfatta i korthet är en grannliga uppgift, men anmärkningsvärt är att även om de har vissa farhågor och invändningar, avslutar de allra flesta sina kommentarer med ett positivt omdöme om utökad datorbaserad av provmaterial i skolan. De tycker generellt att det fungerade bra att göra proven, att det var roligt, ungdomsanpassat, motivationshöjande och lättare (att inte skriva), men de funderar också insiktsfullt kring såväl lärande- och likvärdighetsaspekter som tekniska och lokalmässiga förutsättningar på skolorna. Deras synpunkter på proven som sådana handlar ofta om svårighetsgrad och textmängd. Många elever tyckte proven var lätta, men flera menade också att det var mycket text och att det kan vara jobbigt att läsa på skärmen. Dessa och flera av deras synpunkter på proven sammanfaller med norska resultat från utvärderingar med elever.<sup>109</sup>

### Lärarenkäter

De tretton lärare som involverades i pilotstudien är alla lärare i engelska eller klasslärare som undervisar i engelska. Av dessa tretton lärare besvarade tolv den enkät som skickades ut elektroniskt efter genomförandet. En lärare valde att inte besvara den eftersom hon endast var med under planeringsstadiet.

I den enkät som besvarades digitalt av lärarna efter pilotstudiens slut fanns sex påståenden som lärarna skulle ta ställning till genom att markera graden av instämmande på en femgradig Likert-skala, från ”Ja, absolut” till ”Nej, absolut inte”. De sex påståendena var:

1. Jag använder ofta datorer i min engelskundervisning.
2. Jag kommer att använda fler datorbaserade övningar och prov i min undervisning.
3. De datorbaserade proven var lätta att förstå sig på och använda.
4. På min skola finns tillräcklig beredskap att genomföra datorbaserade prov.

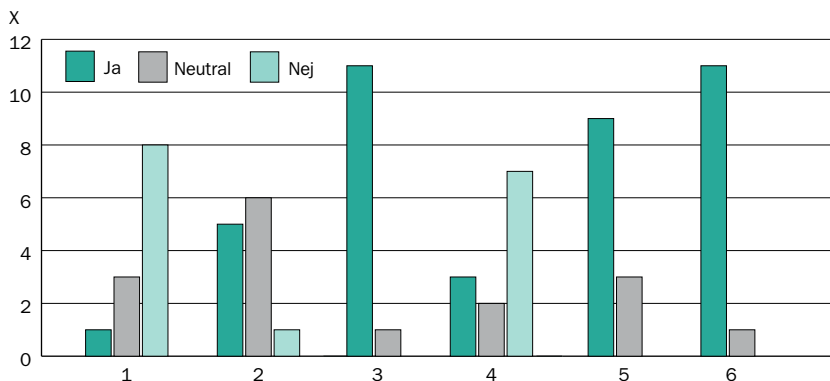
---

<sup>109</sup> Moe (2009)

5. Jag tycker det finns pedagogiska fördelar med datorbaserade prov.
6. Jag skulle ställa mig positiv till datorbaserade nationella prov.

Diagrammet nedan visar hur de tolv lärarna svarat. För tydlighetens skull har de två svarsalternativen i vardera ytterkanten av Likert-skalan slagits ihop till ett ja- respektive ett nej-alternativ.

**Figur 25** Frekvens av lärarsvar på de sex påståendena i lärarenkäten.



Som framgår av lärarnas individuella värdering av påståendena är det en majoritet som inte använder datorer i sin engelskundervisning, medan fem lärare menar att de kommer att göra det oftare efter pilotstudien. Med ett undantag menar lärarna att de datorbaserade proven var lätta att förstå sig på och använda. Något mer än hälften av lärarna anser inte att deras skola har tillräcklig beredskap att genomföra datorbaserade prov, men de flesta är efter studien positiva till, samt ser pedagogiska fördelar med, datorbaserade prov i engelska.

Tio av tolv lärare har skrivit i enkätens fält för positiva kommentarer. Vid en kategorisering av deras reflektioner framträder fyra innehållsliga grupper. För det första handlar de flesta kommentarerna om att eleverna är positiva till de datorbaserade proven, eller möjligen till själva pilotstudien. Två lärare beskriver vad de anser vara möjliga förklaringar till elevernas inställning: ”Eleverna positivt inställda. Jag tror eleverna uppfattar datorprov som mer ’rättvisa’ kanske” och ”Det var uppskattat av såväl eleverna som mig. Roligt! Jag tror att man kan få med sig fler elever som



i vanliga fall inte gillar att skriva för hand.” För det andra berör åtta av de tolv svarande läraren sin egen situation och pekar på arbets- och tidsbesparing. En lärare skriver helt enkelt att det är ”administrativt smidigt” och menar möjligen ett vidare perspektiv än bara lärarens. Fyra lärare nämner emellertid specifikt lättningen i rättningsarbete som en fördel, men även andra aspekter förekommer. Två lärare talar om själva genomförandet och då möjligen om pilotstudien i sig. För det tredje förekommer kommentarer som lyfter fram bedömningsaspekten och uttrycker att datorbaserad prov erbjuder goda möjligheter till snabb feedback till eleverna och en ”objektiv bedömning”. För det fjärde skriver nio lärare kommentarer som direkt eller indirekt relaterar till det genomförda provet som sådant. En lärare betonar att det var bra att få exempeluppgifter att träna på innan och två lärare talar om uppgifterna som elevnära respektive roliga. Några kommentarer har mera karaktären av reflektioner kring framtiden. Två av dem handlar om att de genomförda proven inte täcker allt som bör testas vid ett nationellt prov och två om möjligheten att ha adaptiva prov. Till exempel skriver en lärare: ”Mycket bra DELVIS vid nationella prov – den skriftliga biten måste göras för hand. Jag tror att det kommer vara ett bra komplement till andra test.”

De synpunkter lärarna valt att skriva i fältet för negativa kommentarer är nästan dubbelt så omfattande som de positiva. Även här har tio av tolv lärare lämnat reflektioner. Vid en kategorisering av dessa framträder ytterligare fyra innehållsliga grupper och även här kan de i någon mån överlappa varandra.

Den första kategorin av negativa kommentarer handlar om dator-tekniska problem på skolan. Lärargruppen uttrycker farhågor om att den egna skolans otillräckliga och opålitliga datorpark kan bli ett problem. Åtta av de tio lärarna som lämnat kommentarer skriver om detta. Tre lärare påpekar att en del av datorerna (ibland) är ur funktion eller att man helt enkelt har för få för att kunna genomföra med en undervisningsgrupp samtidigt. De menar också att det förekommer att datorerna låser sig, att Internetuppkopplingar och programåtkomst går mycket långsamt eller att elevernas egna inloggningsprogram i det lokala nätverket krånglar. Några av lärarna belyser också att de råkadet ut för liknande bekymmer under genomförandet av proven i pilotstudien. En lärare skriver att ”Vissa datorer fungerade mycket långsamt. En fjärdedel av våra datorer gick inte att använda.” En annan lyfter fram att två elever

inte kunde genomföra provet då datorn låste sig mitt i programmet. En tredje lärare påpekar begränsningen i antalet datorer och deras funktionsduglighet: ”Vi har enbart 8 datorer vilket försvårar genomförandet, (men fungerar i grupper). Inloggning och internet krånglar ofta.”

Den andra kategorin av negativa lärarkommentarer handlar om svårigheten att skapa en tillfredsställande provsituation främst med tanke på möbleringen i datasalarna. En lärare följer en liknande tanke, men syftar på den känslighet som möjligen kan finnas med Internetbaserade material. Den tredje kategorin av farhågor som lärarna uttryckte handlade om pedagogiska och ämnesrelaterade, samt bedömningsrelaterade farhågor avseende de datorbaserade proven. Ett par lärare menar att elever hade svårt att läsa, eller blev trötta av att läsa på skärmen: ”Flera elever i klassen tyckte att det var svårt att läsa så mycket text på datorn. Det tog längre tid än de hade väntat sig. De såg trötta ut när de lämnade datasalen.” Flera påpekar att pilotstudien endast erbjöd prov i receptiv läsförmåga och att detta bara utgör en del av funktionell språkfärdighet. Ytterligare några lyfter upp svarsformaten, och saknar elevproducerade svar. En lärare skriver att ”Man kan nog använda datorbaserade prov, men jag anser att de måste kompletteras. Språkförmåga är svår att mäta med enbart flervalfrågor.”

Den fjärde kategorin av negativa lärarkommentarer handlar om det aktuella provets funktion och innehåll. Kommentarererna handlar om att lösenord vid inloggningen till provet inte fungerade för någon elev, att det inte gick att ångra givna svar i provet samt om att man vid något tillfälle upplevde att instruktioner var otydliga eller att frågorna innehöll svåra ord i relation till nivån. En lärare påpekar också att de Internetbaserade övningsexempel, som var tänkta som hjälp före provet, inte var tillgängliga vid vissa tillfällen.

Sammanfattningsvis kan vi konstatera att åtta lärare av tolv markerar ”Nej, absolut inte” vid påståendet om huruvida de ofta använder datorer i sin undervisning, vilket kan uppfattas som en förvånande hög siffra. Förmodligen har den samband med de tekniska hinder lärarna belyser, samt inte minst med att man på flera skolor inte har datorer så det räcker till en hel undervisningsgrupp, vilket innebär att läraren behöver dela grupper och gå emellan. Några lärare ställer sig positiva till att hädanefter oftare använda datorer. Kanske har de påverkats av elevernas entusiasm, men de förväntningar de här tolv lärarna har på datorprovets

smidighet verkar grusas av en oro över respektive skolors datorkapacitet. Att så många lärare ändå ställer sig tveklöst positiva till datorbaserade prov kan kanske ses som en förhoppning om framtida förbättringar på IT-området. De antagna pedagogiska fördelarna med datorbaserade prov skulle vara intressant att följa upp och få mer specificerade, men kanske kan dessa tolkas som ett utslag av en positiv inställning till pilotstudien som sådan och en visst tillmötesgående, så kallad compliance-effekt det vill säga att lärarna svarar det de tror förväntas.

### Skolledarenkät

Rektorerna vid de sex deltagande skolorna fick inledningsvis ett informerande brev om att deras skola skulle delta i studien och genomföra datorbaserade prov i ett antal elevgrupper. De ombads också vara med vid planeringssamtalen som föregick studien. Vid tre skolor deltog rektor i dessa och vid de övriga ombads lärarna informera sina skolledare om planeringen av genomförandet och om det fortlöpande arbetet. Efter det att studien avslutats skickades en digital enkät till skolornas åtta rektorer. Fem skolledare som arbetar vid fyra av skolorna har besvarat enkäten. Skolledarenkätens upplägg var identiskt med lärarenkätens och innehöll dels påståenden att ta ställning till på en femgradig Likert-skala, dels utrymme att lämna öppna reflektioner om nationella datorbaserade prov

Samtliga svarande rektorer instämmer i att de på skolorna har tillräcklig beredskap att genomföra datorbaserade prov, att det finns pedagogiska fördelar med datorbaserade prov och att de skulle ställa sig positiva till datorbaserade nationella prov. Fyra av de fem rektorerna skriver kommentarer i enkäten. De menar att datorbaserade prov skulle innebära fördelar för såväl elever som lärarpersonal och ett par påpekar att datorer idag är ett modernt och naturligt verktyg. En rektor lyfter också fram ”fördelar med statistik osv” och refererar då förmodligen till huvudmannanivån. De fördelar man ser för elever är främst ökad motivation och omedelbar feedback i form av resultat på provet. För lärarnas del betonar rektorerna den arbetsbesparing när det gäller rättningsarbete de datorbaserade proven skulle innebära. Samma fyra rektorer har också reflekterat över eventuella nackdelar med en datorbasering av nationella prov. Tre av dem har farhågor när det gäller tekniken och menar till exempel att en satsning på datoriserade prov ställer stora krav på att den egna datorparken är tillräcklig och fungerar tillfredsställande. En rektor svarar:

”Tillgängligheten och allt krångel när det gäller tekniken. Fungerar datorerna idag eller just när jag ska använda dem? Vi har datasalar så att vi kan samla en klass med elever som gör provet samtidigt utan problem men alla elever i en årskurs kan inte göra det vid samma tillfälle. Det måste alltså vara prov som inte går att samtala om.”

Ett par kommentarer handlar om provens begränsade omfattning när det gäller att spegla elevens förmåga och utgöra underlag för bedömning. En rektor menar också att det kan förekomma olika attityder till datorer bland lärarpersonalen: ”Lärarens inställning till datorer. Är det ett medium för möjligheter eller är det något skrämmande?”

Sammanfattningsvis är de fem skolledarna som besvarat enkäten, liksom lärarna, generellt positiva till datorbaserade prov och ser flera fördelar, även pedagogiskt. Även om de markerat endast ja-alternativen i det slutna formatet, verkar de dock i de öppna kommentarerna hysa vissa tvivel på den egna skolans förutsättningar att klara ett verkligt genomförande och har också några pedagogiska invändningar. Detta är dock en för liten studie för att några generella slutsatser ska kunna dras.

### **Elevernas resultat på proven**

Av de till pilotstudien engagerade 276 eleverna genomförde, på grund av ett visst bortfall, totalt 219 elever proven; sammanlagt 113 elever i årskurs 5 och 6 gjorde provet för yngre åldrar och 106 elever i årskurs 8 det för äldre åldrar.

Av lärarenkäten framgår att lärarna undantagit en handfull elever på grund av att de går i särskilda undervisningsgrupper. För övrigt deltar elever med såväl vissa funktionshinder som elever med åtgärdsprogram i engelska på grund av att de ännu ej når målen för ämnet, samt elever som har undervisning i nybörjarengelska. En lärare i årskurs 6 var tvungen att undanta en relativt stor grupp av sina elever på grund av att de inte lämnade in medgivanden från sina vårdnadshavare.

Resultaten för deltagarna i pilotstudien visas nedan som lösningsfrekvens och i jämförelse med utfallet av testen när de gick som nationella prov i Norge 2008:<sup>110</sup>

---

<sup>110</sup> Material som erhållits vid personlig kontakt med Eli Moe.

**Tabell 11** Sammanställning av elevernas resultat på de datorbaserade proven i engelska.

Årskurs	Antal elever	Maxpoäng	Pilotstudien		Norge 2008
			Range min-max	Lösningsfrekvens	Lösningsfrekvens
5	29	38	17–37	0,75	0,67
6	84	38	22–38	0,87	(–)
8	106	40	1–40	0,68	0,58

Gruppen elever som deltog i pilotstudien fick alltså jämförelsevis hög lösningsfrekvens på de datorbaserade proven för såväl yngre som äldre åldrar och därmed goda resultat. Det bör dock noteras att de norska resultat som redovisas i tabell 11 kommer från elever som gjorde proven i början av höstterminen i femte och åttonde trinn, vilket motsvarar årskurs 4 och 7 i den svenska grundskolan. De norska eleverna är yngre och har gått färre år i skolan.

Sammanfattningsvis kan provresultaten för eleverna i pilotstudiegruppen sägas vara mycket goda. Emellertid bör beaktas att en direkt jämförelse med de norska resultaten egentligen inte är adekvat. Å ena sidan är alla de svenska eleverna i pilotstudien äldre än de norska som går i motsvarande årskurs – men å andra sidan börjar många eller de flesta svenska elever med engelska i skolan något år senare. Dessutom, eftersom föreliggande studie främst syftar till att undersöka attityder och förutsättningar att genomföra datorbaserade prov, lades inte större vikt vid att skolorna skulle genomföra proven under provlika former. För pilotstudiens del kan resultaten dock sägas bekräfta att svenska elever i regel har goda kunskaper i engelska och att datormediet i sig i detta fall inte tycks ha hindrat dem.

### Sammanfattning och slutsatser av pilotstudien i engelska

Pilotstudien i engelska visar att involverade skolläda-re, lärare och elever generellt är mycket positiva till den datorbaserade formen för prövning. Deras invändningar har snarare för avsikt att stödja än att hindra kommande arbete och bör beaktas inför beslut när det gäller datorbaserade material för den svenska skolan. Kärnan i både lärarnas och elevernas reaktioner verkar handla om förutsättningar, på ett konkret plan tekniska

och maskinella/materiella sådana för skolornas del, men också om att datormediet tillför en ytterligare och ny aspekt när det gäller elevers olika beredskap och möjligheter att visa vad de kan i engelska.

De norska prov, och system för distribution av datorbaserade prov, som användes i pilotstudien har fungerat bra. En viktig förutsättning för genomförandet av pilotstudien har varit det goda samarbetet med kollegerna vid Universitetet i Bergen. Eli Moe m.fl. har, med erfarenhet från utveckling av den norska versionen av Dialang, arbetat med att utveckla digitaliserad prövning i Norge sedan ett uppdrag från stortinget 2002. Man har gått igenom flera faser och sedan 2007 har proven enbart en rapportrande funktion. För femte och åttonde trinn genomförs vardera tre likvärdiga linjära prov i engelsk läsning som slumpas ut till skolorna under en fjortondagarsperiod. Under 2009 skulle man också pröva vokabulär och grammatik. Erfarenheter i Norge, efter de fyra år man nu genomfört datoriserade prov, är att "Once the tests are in place, and the students know the test formats, schools, teachers and pupils seem to like this form of testing."<sup>111</sup>

Till sist kan konstateras att resultaten och erfarenheterna från denna ytterst begränsade pilotstudie tyder på att den attitydmässiga beredskapen för användning av datormediet som sådant i relation till prov och bedömning är mycket god. Dock framskymtar också en oro för de rent tekniska förutsättningarna att på ett bra sätt genomföra proven vid de deltagande skolorna. Det mest övergripande intrycket är emellertid att såväl skolledare och lärare som elever i olika åldrar på ett insiktsfullt och konstruktivt sätt delat med sig av sina erfarenheter och synpunkter. Detta understryker att kvaliteten i vidare utveckling av datorbaserade bedömningsformer för den svenska skolan har allt att vinna på en medveten och nära samverkan mellan berörda parter på den politiska, akademiska, tekniska och lokala nivån. Detta gäller inte minst de bidrag som lämnas av dem det allra mest berör, nämligen eleverna.

---

<sup>111</sup> Moe (2009)

# Diskussion

## Diskussion

Denna rapport handlar om kunskapsläget när det gäller datoradaptiva prov, men den har också berört datorbaserade prov i allmänhet och presenterat erfarenheter från två pilotstudier med datorbaserade prov i grundskolan. När det gäller praktiska exempel på användning av datorer i provsammanhang fokuserar rapporten på matematik, naturvetenskap och språkämnen. I detta avslutande kapitel sammanfattas några viktiga aspekter inför en fortsatt diskussion om framtiden för datorbaserade och adaptiva prov i Sverige.

### Några slutsatser utifrån kunskapsöversikten

Utifrån kunskapsöversikten och de genomförda pilotstudierna kan vi formulera några viktiga slutsatser som kan utgöra utgångspunkter i den fortsatta diskussionen om eventuella satsningar på adaptiva prov inom det nationella provsystemet i Sverige:

- Sverige ligger efter våra nordiska grannländer och även andra länder i Europa i utvecklingen av datorbaserade prov i stor skala.
- Pilotstudierna visar på att många skolor har en beredskap för datorbaserade prov, även om inte alla har den utrustning som krävs. Erfarenheterna från Norge är att införandet av datorbaserade nationella prov stimulerar skolorna att uppdatera sin datorutrustning.
- Pilotstudierna har visat på en positiv förväntan och attityd till datorbaserade prov från elever, lärare och skolledare.
- Pilotstudien i matematik visar på att eleverna i de yngre åldrarna inte förmår arbeta särskilt länge med ett datorbaserat prov. Samtidigt innebär linjära datorbaserade prov som även ska kunna användas för formativa syften att det behövs en lång provtid, sannolikt för lång för yngre barn. Detta är ett argument för att vidare undersöka möjligheterna med adaptiva prov.
- Om man utgår från prov som görs med papper och penna så är ett första steg att göra det datorbaserat. I nästa steg kan man gå vidare och skapa ett datorbaserat adaptivt prov.



- För målrelaterade prov bör ett begränsat datorbaserat adaptivt prov väljas där man kan bestämma om provtagaren når upp till en förutbestämd nivå av kunskap eller förmåga. Utifrån syftet och den målgrupp man vänder sig till kan man sedan bestämma hur provet ska vara uppbyggt. Utifrån resurser och möjlighet att skapa en välkalibrerad uppgiftsbank, så kan man bestämma sig för ett fullskaligt eller ett mer begränsat datorbaserat adaptivt prov.
- Erfarenheterna från våra grannländer och från institutioner som arbetat länge med olika datorbaserade prov säger att det är klokt att skynda långsamt. Utvecklingsarbetet, och även implementeringsarbetet i skolorna, är en komplex och mödosam process.
- Datorbaserade prov skapar möjligheter men även risker när det gäller säkerhet. Säkerhetsaspekten är mindre viktig ju yngre barnen är, och ju mindre proven används för viktiga beslut på individnivå, till exempel i betygssättning.
- Datorn ger potential för intressanta uppgiftsformat. Denna möjlighet tycks dock sällan ha utnyttjats i praktiken.<sup>112</sup>
- Datorbaserade adaptiva prov kräver datorrättning, vilket innebär begränsningar i hur frågor kan ställas och besvaras. Detta kommer att begränsa möjligheterna att pröva alla mål och kunskapskvaliteter som beskrivs i kursplaner och andra styrdokument. Syftet med proven är dock avgörande för hur problematiskt detta kan anses vara.
- Trots relativt god datortillgång är datoranvändningen ganska låg i vissa ämnen, och i synnerhet i matematik. Detta innebär att datorbaserade prov i matematik i dagsläget inte bidrar till en ökad samstämmighet med undervisningen. Datorn används däremot för att skriva texter, vilket kan vara ett argument för att prioritera utvecklingen av datorbaserade prov i skriftspråkliga producerande sammanhang.
- Teori och teknik i samband med datorbaserade adaptiva prov är väl utvecklade, medan det tycks finnas avsevärt mindre av studier som behandlar de pedagogiska aspekterna av dessa prov.

Punkterna ovan utgör några angelägna utgångspunkter för nationella överväganden i förhållande till datorbaserade prov, och datorbaserade

---

<sup>112</sup> Parshall, Davey, & Pashley (2000)

adaptiva prov. En annan viktig utgångspunkt är de ställningstaganden som Skolverket redan gjort.

## Om Skolverkets redovisning angående datorbaserade prov

I sin redovisning av ett regeringsuppdrag om hur det nationella provsystemet bör utformas och utvecklas skriver Skolverket<sup>113</sup> en del om datorbaserade och även adaptiva prov.

Skolverket konstaterar att datorbaserade bedömningar bör vara viktiga inslag i ett framtida kunskapsbedömningssystem, och att det då kan vara både linjära prov med uppgifter som eleven kan besvara mer eller mindre interaktivt vid datorn och adaptiva prov där datorn anpassar uppgifternas svårighetsgrad fortlöpande efter elevens svar på uppgifterna.

Skolverkets text uttrycker en positiv förväntan på vad som kan åstadkommas med datorbaserade och datoradaptiva prov. De möjligheter som Skolverket lyfter fram stämmer väl med vad som framkommit i denna kunskapsöversikt. Skolverkets text är mycket sparsam när det gäller nackdelar och svårigheter, och det finns definitivt fler problem än begränsningen av vilka mål och färdigheter som kan provas.

I sitt svar på regeringsuppdraget skriver Skolverket ”När datorbaserade prov ska genomföras måste huvudmännen se till att skolornas datorer har nödvändig prestanda, bandbredd och att det finns support för genomförandet.” I pilotstudien i engelska som redovisats i den här rapporten ingick en skola som inte kunde tillhandahålla tillräcklig bandbredd för genomförandet av ett webbaserat prov med relativt mycket bilder. Om möjligheterna med datormediet ska utnyttjas kommer inslaget av mer interaktiva och multimediala uppgiftsformat att bli ännu större och därmed ställa ytterligare krav på skolornas datorutrustning och nätverk. Samtidigt säger erfarenheterna från Norge att införandet av datorbaserade nationella prov har varit ett incitament för skolhuvudmännen att se till att skolorna håller en god teknisk standard när det gäller IT-frågorna. Angående provsäkerhet skriver Skolverket i sitt svar till departementet att proven måste genomföras med tillräcklig säkerhet och utan att provtagaren har möjlighet att använda datorn till annat än att genomföra provet, till exempel chatta, söka och hämta information på Internet eller från provtagarens egen dator. Tekniska lösningar för att förhindra att

---

<sup>113</sup> Dnr: 2009:00337

otillåtna hjälpmedel används innebär att datorerna måste startas om med hjälp av ett provprogram som tillfälligt begränsar datorns funktionalitet.

Här lyfter alltså Skolverket fram säkerhetsfrågorna och betonar vikten av att kontrollera hur eleverna använder datorn i samband med datorbaserade prov. Detta är med största sannolikhet ett svårlost praktiskt problem. Det är intressant att notera att de norska nationella prov som använts i pilotstudierna inte har någon sådan spärr och att de norska myndigheterna inte anser att detta är ett stort problem. De norska proven syftar inte till att utvärdera den enskilde eleven, och det är troligt att problemen med icke önskvärd användning av datorns funktioner är större för prov som används för individbedömning. För prov som inte har en direkt koppling till till exempel elevens betyg och för de yngre åldrarna torde inte dessa säkerhetsfrågor vara ett stort problem.

Skolverket hänvisar till regeringsuppdraget om IT-uppföljning<sup>114</sup> där man beräknat att det går cirka 6 elever per dator i grundskolan, och denna relativt låga datortäthet antas få konsekvenser för möjligheten att genomföra IT-baserade prov om alla elever ska göra proven samtidigt. Skolverket konstaterar att det i dagsläget får anses ”uteslutet att välja lösningar som kräver att alla elever gör datorbaserade prov samtidigt”.

Det nationella provsystemet hade tidigare provperioder som gav lärare och skolor möjlighet att välja när proven skulle genomföras. Denna frihet fick stark kritik från lärarhåll och skapade även misstro mot resultatens hos dem som ville använda resultaten för utvärderande syften, och successivt har de nationella proven begränsats till en enda provdag. En möjlighet att kunna genomföra datorbaserade prov vid olika tillfällen är genom en adaptiv utformning, eftersom olika elever inte gör samma prov, och Skolverkets ställningstagande sätter fokus på datorbaserade adaptiva prov som en tänkbar lösning för framtida datorbaserade prov. De pilotstudier som genomförts inom ramen för det uppdrag som redovisas i den här rapporten har också visat att det går att genomföra datorbaserade prov med halva eller till och med hela klasser samtidigt. Skolverket problematiserar i sin redovisning möjligheten att föreskriva skolhuvudmännen att genomföra datorbaserade prov. Skolverket har bemyndiganden att meddela föreskrifter om nationella prov. För att kunna föreskriva att skolhuvudmännen ska genomföra datorbaserade prov kan föreskrifterna behöva omfatta krav på viss prestanda hos skolans datorer

---

<sup>114</sup> Dnr 75-2007:3775

och hastighet på bredbandsuppkoppling såväl som krav på teknisk kompetens hos skolan eller anordnaren. Vidare kan huvudmannens IT-policy komma att påverkas till exempel om ett prov innehåller video och om man har begränsningar för strömmande media i brandväggen. Detta är en viktig aspekt av problemen med utvecklingen av datorbaserade prov. Det är inte uteslutet att de krav som måste ställas för att kunna genomföra datorbaserade prov för vissa huvudmän kommer att innebära kostnader som inte är oväsentliga.

## Datorbaserade prov och datorstödd undervisning

Russell<sup>115</sup> menar att såväl modern teknologi som bedömningar och prov har varit populära vägar för att förbättra undervisning och lärande, men att de har utvecklats och implementerats separat. Han menar att detta har resulterat i en klyfta mellan prov och bedömning å ena sidan och teknologi för lärande å den andra. Till exempel har mycket av skrivandet även i skolan gått från papper och penna till dator, medan eleverna ofta i provsammanhang förväntas redovisa vad de lärt sig genom att skriva på papper. Russell menar att denna klyfta i första hand handlar om föreställningar och inte om praktiker, och att klyftan representerar mer än dålig samstämmighet mellan sätt att lära sig och sätt att bedöma vad man lärt sig. De som förordar en utveckling av IKT och lärande för att reformera utbildningen fokuserar mer på *hur* elever lär sig och arbetar än på att föreskriva *vad* de lär sig. De som utvecklar nationella system för bedömning av kunskaper och utvärdering av skolan föreställer sig att prov är ett effektivt verktyg för att påverka kunskapsutveckling. De har en föreställning om möjligheterna med dessa verktyg som primärt fokuserar *vad* eleverna lär sig, mer än *hur* de lär sig.

I fortsatta diskussioner om datorbaserade prov i Sverige är det angeläget att fundera över hur denna klyfta ser ut i vårt land. Kunskapsöversikten har pekat på relativt god datortillgång på svenska skolor, men mycket varierande datoranvändning mellan ämnen. I matematik kan det kanske inte vara fråga om så mycket av en klyfta, eftersom det är så låg datoranvändning i matematikklassrummet. I språkundervisningen tycks dock datorn ha en större plats i undervisningen och det är där mer angeläget att fråga sig i vilken utsträckning detta motsvaras av datoranvändning i samband med bedömning. Datorbaserad bedöm-

---

<sup>115</sup> Russell (2006)

ning bör därför stärka samstämmigheten mellan prov och undervisning, åtminstone i språkämnen. Det råder dessutom ingen tvekan om att datoradaptiva och allmänt datorbaserade prov skapar möjligheter som inte finns i prov som bygger på användning av papper och penna. Till exempel kan datorn i viss utsträckning tänkas simulera det som kan åstadkommas i mötet mellan en lärare och enstaka elever, och när det gäller till exempel multimedia kan datorn vidga det som kan göras i andra sammanhang. Det är därför angeläget att stödja en fortsatt utveckling som tar vara på möjligheterna i datoradaptiva och datorbaserade eller datorstödda prov. Ytterligare en anledning till en sådan satsning är att datorer och annan informationsteknik används i allt fler sammanhang. Datoriseringen av prov och andra pedagogiska bedömningar är därför närmast oundviklig, och en satsning på central nivå kan bidra till att denna datorisering sker på ett genomtänkt, systematiskt och högkvalitativt sätt.

Forskningen om datorbaserade adaptiva prov är väldigt inriktad på tekniska och statistiska problem, i första hand med målet att åstadkomma reliabla mått på elevers förmåga i något avseende. Det finns dock i litteraturen exempel på utvecklingsarbeten som siktar mot att åstadkomma adaptiva prov som kan vara diagnostiska verktyg, med möjlighet att användas formativt. Till exempel beskriver en grupp forskare hur de adaptiva prov de utvecklat hittills designats för att endast indikera en övergripande kompetensnivå,<sup>116</sup> vilket har inspirerat dem att påbörja ett arbete som syftar till att inte enbart ge resultat på en aggregerad nivå utan och detaljerad återkoppling om elevens prestationer inom varje delområde. De konstaterar att det behövs ett fyratimmars diagnostiskt prov för att få tillförlitlig information om måluppfyllelse i förhållande till det relativt breda måldokument som de använder, vilket naturligtvis är alldeles för långt även om det i deras fall handlar om vuxna studerande. Författarna menar att ett adaptivt diagnostiskt prov kan klara detta på mycket kortare tid genom bra algoritmer, en hierarkisk struktur som innebär att den som visar sig kunna de högre nivåerna också kan antas kunna de lägre, samt att eleven går in i provet på en nivå som de själva eller deras lärare bedömer är rimlig.

Ett exempel på när adaptiva prov åtminstone i teorin borde kunna uppfylla de mångfasetterade krav som ställs på nationella prov gäller års-

---

<sup>116</sup> Winkley, Rainbow, & Baki (2008)

kurs 3. Skolverket har definierat ett antal olika syften med de nationella proven i årskurs 3. Provet ska kunna

- användas av lärare som stöd för bedömning av elevers målpuppfyllelse,
- fylla såväl en summativ som en formativ funktion,
- identifiera elever i behov av särskilt stöd,
- ge underlag för en kunskapsprofil i syfte att stödja kunskapsutvecklingen mot de mål som följer senare i grundskolan, och
- användas i uppföljande och utvärderande syfte såväl på skolnivå, kommunal nivå och nationell nivå.

Med ett adaptivt datorbaserat prov kan eventuellt alla dessa syften uppnås. Den adaptiva utformningen skapar en unik möjlighet att ge relevant information för olika syften och utifrån varje individs kunskapsprofil och behov. Lärare kommer att kunna få stöd för bedömningen av elevers målpuppfyllelse i stora drag, men även ge en reliabel värdering av elevernas kunskapskvaliteter inom till exempel olika innehållsliga områden. Utan att förlänga provtiden så kan detta endast åstadkommas med en adaptiv design. Provet kan få en formativ roll genom en trovärdig kunskapsprofil kopplad till en kommentarfunktion där varje elevs resultat kommenteras för att bland annat föreslå åtgärder för att arbeta med elevens svårigheter.

Inför framtagandet av nationella prov i årskurs 3 betonade också Skolverket vikten av att

- provet inte blir ett stressmoment för eleverna, och att det därför bör konstrueras så att det kan infogas på ett så naturligt sätt som möjligt i den ordinarie undervisningen,
- formerna för provuppgifterna bör kännas bekanta för eleverna,
- provuppgifterna ska tala till flera sinnen och eleverna ska tillåtas använda olika uttrycksformer för att visa sin kunskap,
- eleverna ska ges tillfälle att visa de förmågor som beskrivs i målen, och
- provets utformning och innehåll i fråga om tidsaspekt och provtillfällen ska ta hänsyn till elevernas ålder och varierande mognad.

Även dessa förväntningar på provet kan infråas med ett adaptivt datorbaserat prov, och en av de möjligheter med datorbaserade prov som

kommer fram i kunskapsöversikten är möjligheten till spännande och intressanta uppgiftsformat. En väl utformad datormiljö kan erbjuda eleverna en lekfull och avdramatiserad möjlighet att visa vad de kan i matematik. Utformningen av datamiljön kommer att hämta sin inspiration från dataspelsvärlden. Provet kan infogas på ett naturligt sätt i den ordinarie undervisningen genom att någon eller några elever i taget får möjlighet att sätta sig vid datorn och arbeta med uppgifterna. Datorn ger nya, och i många fall unika, möjligheter att skapa uppgiftsformat som vänder sig till andra uttrycksformer än skriftliga svar. Eleverna får använda sin fantasi och arbeta med interaktiva uppgifter som blir avsevärt mindre beroende av det svenska skriftspråket. I synnerhet kan den adaptiva utformningen, mer än andra former för prov, göra att provets utformning och innehåll tar hänsyn till elevernas olikheter i en rad olika hänseenden.

Men vägen till ett datorbaserat adaptivt prov är mycket lång och det finns många problem att övervinna längs vägen.

## Rekommendation

Avslutningsvis redovisas några rekommendationer för det fortsatta arbetet med datorbaserade prov, och datorbaserade adaptiva prov, i Sverige.

För det första bör Skolverket initiera och stödja ett långsiktigt arbete för att utveckla datorbaserade prov i Sverige. Det är angeläget att inte förvänta sig snabba resultat utan att göra ett gediget arbete med såväl utveckling och utprövning, som implementering och tekniska förutsättningar.

För det andra bör utveckling av datorbaserade prov som inte är adaptiva prioriteras. Skolverkets ställningstagande att skolornas datortäthet inte möjliggör att eleverna kan göra provet vid samma tillfälle kan innebära att satsningen i första hand måste göras på bedömningar med stödjande, formativa syften. En sådan satsning innebär att vissa säkerhetsfrågor, de som handlar om att provtagaren har möjlighet att använda datorn till annat än att genomföra provet, inte blir särskilt framträdande eller viktiga.

För det tredje kan eventuellt datorbaserade adaptiva prov utvecklas på sikt, och primärt med tanke på indikerande prov som kan ge ett reliabelt summativt omdöme om till exempel vilka elever som är i behov av särskilt stöd. Skolverket bör även stödja ett mer långsiktigt forsknings- och

utvecklingsarbete som syftar till att utveckla ett adaptivt provsystem som kan ge värdefull diagnostisk information inom några kunskapsdomäner i ett ämne.

För det fjärde kan en strävan efter samstämmighet mellan prov och undervisning innebära att Skolverket i första hand satsar på områden där datorn används på ett naturligt sätt i skolan, till exempel i skriftlig produktion.

För det femte är det mycket angeläget att bygga upp kraftfulla och långvariga samarbeten med andra. I några av våra nordiska grannländer finns viktiga erfarenheter och troligen även en vilja att samarbeta kring gemensamma problem. Det finns även universitet och andra institutioner inom och utom landet som har expertis på området.

För det sjätte är det viktigt att i de fortsatta diskussionerna vara medveten om några angelägna problem som gäller datorbaserade prov och datorbaserade adaptiva prov, i såväl teori som praktik. En viktig begränsning är att datorbaserade adaptiva prov förutsätter (och bygger på) att datorn kan avgöra kvaliteten i en elevs svar. Detta påverkar i hög grad hur uppgifterna kan se ut, uppgiftsformatet. Det är troligt att man i första hand måste använda sig av ett svarsformat som kan bedömas som godtagbart eller inte, det vill säga uppgiften måste utformas så att den bestämmer svarets kvalitet (strukturerad, eller sluten, uppgift). Ett annat uppmärksammat problem handlar om säkerhets- och distributionsfrågorna, vilka måste anses kritiska i utvecklingen av datorbaserade prov. Vidare har samtliga kända tillämpningar av datorbaserade adaptiva prov normrelaterade utgångspunkter, vilket är ett problem i förhållande till det svenska systemet. Slutligen kan återkopplingen av resultat till elever och lärare nämnas som en särskild utmaning som kräver särskild uppmärksamhet.



# Referenser

# Referenser

## Litteratur

Ackerman, T. A., Evans, J., Park, K., Tamassia, C., & Turner, R. (1999). Computer assessment using visual stimuli: a test of dermatological skin disorders. I F. Drasgow & J. B. Olson-Buchanan (Eds.), *Innovations in computerized assessment* (137-150). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Baker, F. B. & O'Neil, H. F. (1995). Computer technology futures for the improvement of assessment. *Journal of Science Education and Technology*, 4, 37-45.

Bennett, R. E., Morley, M., Quardt, D., & Rock, D. A. (2000). Graphical modeling: a new response type for measuring the qualitative component of mathematical reasoning. *Applied Measurement in Education*, 13, 303-322.

Bennett, R. E., Morley, M., Quardt, D., Rock, D. A., Singley, M. K., Katz, I. R., & Nhouyvanisvong, A. (1999). Psychometric and cognitive functioning of an under-determined computer-based response type for quantitative reasoning. *Journal of Educational Measurement*, 36, 233-252.

Bennett, R. E. & Rock, D. A. (1995). Generalizability, validity, and examinee perceptions of a computer-delivered formulating-hypotheses test. *Journal of Educational Measurement*, 32, 19-36.

Bennett, R. E. & Sebrechts, M. M. (1997). A computer-based task for measuring the representational component of quantitative proficiency. *Journal of Educational Measurement*, 34, 1, 64-77.

Bjerkestrand, Ø. (2008) [http://crell.jrc.it/Presentations\\_Iceland%202008/Bjerkestrand.pdf](http://crell.jrc.it/Presentations_Iceland%202008/Bjerkestrand.pdf) (Hämtad 27/10, 2008)

Bridgeman, B. (2008) *Experiences from Large-Scale Computer-Based Testing in the USA*, [http://crell.jrc.it/Presentations\\_Iceland%202008/Bridgeman.pdf](http://crell.jrc.it/Presentations_Iceland%202008/Bridgeman.pdf) (Hämtad 30/10, 2008)

Buck, G. (2001). *Assessing Listening*. Cambridge: Cambridge University Press.

- Bunderson, C. V., Inouye, D. K., & Olsen, J. B. (1989). The four generations of computerized educational measurement. I R. L. Linn (Ed.), *Educational Measurement* (3<sup>rd</sup> ed.) (367-408). New York: Macmillan.
- Burstein, J. C., Kaplan, R. M., Rohen-Wolff, S., Zuckerman, D. I., Lu, C. (2000) *A Review of Computer-Based Speech Technology for TOEFL*, ETS, Princeton, NJ RM-99-5 <http://www.ets.org/Media/Research/pdf/rm-99-05.pdf> (Hämtad 12/4, 2008)
- Burstein, J., Kukich, K., Wolff, S., Lu, C., & Chodorow, M. (1998). *Computer analysis of essays*. Paper presented at the NCME symposium on Automated Scoring, San Diego, CA.
- Chapelle, C. A. & Douglas, D. (2006). *Assessing Language through Computer Technology*. Cambridge: Cambridge University Press.
- Chun, W. C. (2006). An Analysis of a Language Test for Employment: The Authenticity of the PhonePass Test, *Language Assessment Quarterly*, 3(3), 295-306.
- Chung, G. K. W. K., O'Neill, H. F. & Herl, H. E. (1999). The use of computer-based collaborative knowledge mapping to measure team processes and outcomes. *Computers in Human Behavior*, 15, 463-494.
- Clark, C. (ed.) (1975). *Proceedings of the first conference on computerized adaptive testing*. Washington D.C., June 12-13, 1975.
- Clauser, B. E., Harik, P., & Clyman, S. G. (2000). The generalizability of scores for a performance assessment scored with a computer-automated scoring system. *Journal of Educational Measurement*, 37, 245-261.
- Crocker, L. & Algina, J. (1996). *Introduction to classical and modern test theory*. New York: Holt, Thinehart & Winston.
- Davey, T., Godwin, J., & Mittelholtz, D. (1997). Developing and scoring an innovative computerized writing assessment. *Journal of Educational Measurement*, 34, 21-41.
- Drasgow, F., Luecht, R. M., & Bennett, R. M. (2006). Technology and testing. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 471-515). Westport: ACE/Praeger.

Eggen, T. (2008). Testing programs that have been implemented CATs developed at Cito, the national institute for educational measurement in the Netherlands. Tillgänglig 25 november 2008 på <http://www.psych.umn.edu/psylabs/catcentral/pdf files/citocats.pdf>

Eggen, T. J. H. M. (2004). *Contributions to the theory and practice of computerized adaptive testing*. Doktorsavhandling, Twente university, Enschede.

English, R. A., Reckase, M. D., Patience, W. M. (1977). Application of tailored testing to achievement measurement. *Behavior Research Methods and Instrumentation*, 9, 158-161.

Erickson, G. (1998): Test-taker feedback – Elevers synpunkter som del av arbetet med nationella prov i engelska. I Berit Ljung & Astrid Pettersson, red: *Perspektiv på bedömning av kunskap*, 79-100. Stockholm: Lärarhögskolan i Stockholm, Institutionen för pedagogik

Erickson, G. (2006): Bedömning av och för lärande. En kollaborativ ansats i arbetet med nationella prov i språk. I Ulrika Tornberg, red: *Mångkulturella aspekter på språkundervisningens kommunikativa praktiker*, 187-207. Örebro: Pedagogiska institutionen, Örebro universitet.

Guo, F. (2007). CAT Security: A practitioner's perspective. In D. J. Weiss (Ed.), *Proceedings of the 2007 GMAC Conference on Computerized Adaptive Testing*.

Hambleton, R. K. (1997). *Promising GMAT item formats for the 21<sup>st</sup> century*. Invited presentation at the international workshop on the GMAT, Paris, France.

Hamilton, L. S., Klein, S. P., & Lorie, W. (2000). *Using web-based testing for large-scale assessment*. Arlington, VA: Rand education.

Hetter, R. D. & Sympson, J. B. (1977). Item exposure control in CAT-ASVAB. I W. A. Sands, B. K. Waters, & J. R. McBride (Eds.), *Computerized adaptive testing: From inquiry to operation* (141-144). Washington, DC: American Psychological Association.

Kane, M. (2006). Validation. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed.). Westport: American Council on Education/ Praeger.

- Koch, D. A. (1993). Testing goes graphical. *Journal of interactive Instruction Development*, 5, 14-21.
- Lee, Y.-W. (2007) (with M. Chodorow and Claudia Gentile) Seoul National University, Seoul, Korea Abstract presenterad vid TESL/ Applied Linguistics: Conference on Technology for Second Language Learning, Friday, September 21 and Saturday, September 22, 2007 Fel! Ogiltig hyperlänkreferens. (Hämtad 20/10, 2008)
- Lee, Y.-W., Gentile C., Kantor, R. (2008) *Analytic Scoring of TOEFL® CBT Essays: Scores From Humans and E-rater*, ETS, Princeton, NJ, 2008.
- Lewis, C. & Sheehan, K. (1990). Using Bayesian decision theory to design a computer mastery test. *Applied Psychological Measurement*, 14, 367-386.
- Linn, R. L. (2000). Assessments and accountability. *Educational Researcher*, 29, 2, 4-16.
- Linn, R. L., Graue, M. E., & Sanders, N. M. (1990). Comparing state and district test results to national norms: The validity of claims that "everyone is above average". *Educational Measurement: Issues and Practice*, 9, 5-14.
- Luecht, R. M. (2001). *Capturing, codifying, and scoring complex data for innovative, computer-based items*. Paper presented at the annual meeting of the National Council on Measurement in Education, Seattle, WA.
- Luecht, R. M. & Nungester, R. J. (1998). Some practical examples of computer adaptive sequential testing. *Journal of Educational Measurement*, 35, 229-249.
- Luecht; R. M. & Brumfield, T., & Breithaupt, K. (2006). A testlet assembly design for adaptive multistage tests. *Applied Measurement in Education*, 19, 189-202.
- Martinez, M. E. (1991). A comparison of multiple-choice and constructed figural response items. *Journal of Educational Measurement*, 28, 131-145.
- Martinez, M. E. (1993). Item formats and mental abilities in biology assessment. *Journal of Computers in Mathematics and Science Teaching*, 12, 289-301.

- Martinez, M. E., & Bennett, R. E. (1992). A review of automatically scorable constructed-response item types for large-scale assessment. *Applied Measurement in Education*, 5, 151-169.
- McBride, J. R. & Martin, J. T. (1983). Reliability and validity of adaptive ability tests in a military setting. I D. J. Weiss (Ed.), *New horizons in testing: Latent trait test theory and computerized adaptive testing* (223-236). New York: Academic Press.
- Mead, A. D. (2006). An introduction to multistage testing. *Applied Measurement in Education*, 19, 185-187.
- Meijer, R. R. & Nering, M. L. (1999). Computerized adaptive testing: overview and introduction. *Applied Psychological Measurement*, 23, 187-194.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (Vol. 3, pp. 13-103). New York: American Council on Education.
- Mills, C. N. & Steffen, M. (2000). The GRE computer adaptive test: operational issues. I W. J. van der Linden & C. A. W. Glas (Eds.), *Computerized adaptive testing: theory and practice* (75-99). Dordrecht: Kluwer Academic.
- Moe E., Carlsen C., & Hasselgren A. (2006). Digitale leseprøver i engelsk– utfordringer og muligheter. *Norsk Pedagogisk Tidsskrift* 5(90), 391-403
- Moe, E. (2009). Introducing Large-scale Computerised Assessment, Lessons Learned and Future Challenges. I Scheuermann, F. & Björnsson, J. (Eds.): *The Transition to Computer-Based Assessment, New Approaches to Skills Assessment and Implications for Large-scale Testing*. European Commission, Joint Research Centre <http://crell.jrc.it/RP/reporttransition.pdf> (<http://crell.jrc.it/RP/reporttransition.pdf>)
- Molholt, G., & Presler, A. (1986). Correlation between human and machine ratings of Test of Spoken English reading passages. In C. Stansfield (Ed.), *Technology and language testing* (pp. 111-128). Washington, DC: TESOL Publications.

- Ockey, G. J. (2007). Construct implications of including still image or video in computer-based listening tests. 2007; 24; 517 *Language Testing*. DOI: 10.1177/0265532207080771. Online-version: <http://ltj.sagepub.com/cgi/reprint/24/4/517> (Hämtad 24/11, 2008)
- Parshall, C. G., Davey, T., & Pashley, P. (2000). Innovative item types for computerized testing. I W. J. van der Linden & C. A. W. Glas (Eds.), *Computerized adaptive testing: theory and practice* (129-148). Dordrecht: Kluwer Academic.
- Parshall, C. G., Davey, T., & Pashley, P. J. (2000). Innovative item types for computerized testing. In W. J. van der Linden & C. A. W. Glas (Eds.), *Computerized adaptive testing: theory and practice*. Dordrecht: Kluwer Academic Publishers.
- Parshall, C. G., Spray, J. A., Kalohn, J. C., & Davey, T. (2006). *Practical considerations in computer-based testing*. New York: Springer.
- Pino-Silva, J. (2008). Student perceptions of computerized tests. *ELT Journal*, 62(2), 148-156.
- Richter, L. (2009). Eksperter: Nationale test er spild af penge, *Dagbladet Information*, 6/4, 2009 <http://www.information.dk/187374> (Hämtad 24/4, 2009).
- Rocklin, T., & O'Donnell, A. (1987). Self-adapted testing: A performance-improving variant of computerized adaptive testing. *Journal of Educational Psychology*, 79(3), 315-319.
- Rudner, L. (1998). An on-line interactive, computerized adaptive mini-tutorial. ERIC *Clearinghouse on Assessment and Evaluation*.
- Rudner, L. M. (2007). Implementing the Graduate Management Admission Test® computerized adaptive test. In D. J. Weiss (Ed.), *Proceedings of the 2007 GMAC Conference on Computerized Adaptive Testing*.
- Russel, M. (2006). *Technology and assessment. The tale of two interpretations*. Greenwich: Information Age Publishing Inc.

- Russell, M. & Haney, W. (1997). Testing writing on computers: an experiment comparing student performance on tests conducted via computer and via paper-and-pencil. *Education Policy Analysis Archives*, 8, 19. Tillgänglig på <http://olam.ed.asu.edu/epaal/>.
- Russell, M. (1999). Testing on computers: A follow-up study comparing performance on computer and on paper. *Education Policy Analysis Archives*, 7, 20. Tillgänglig på <http://olam.ed.asu.edu/epaal/>.
- Ryssevik, J. (2008). [http://crell.jrc.it/Presentations\\_Iceland%202008/Ryssevik.pdf](http://crell.jrc.it/Presentations_Iceland%202008/Ryssevik.pdf) (Hämtad 27/10, 2008)
- Sands, W. A., Waters, B. K., McBride, J. R. (eds.) (1997). *Computerized adaptive testing: from inquiry to operation*. Philadelphia: American Psychological Association.
- Sireci, S. G. (2004). Computerized-adaptive testing: An introduction. I J. Wall and G. Walz (Eds.). *Measuring up: Assessment issues for teachers, counselors, and administrators* (pp. 685-694). Greensboro, NC: CAPS Press.
- Skolverket. (2004). TIMSS 2003. Svenska elevers kunskaper i matematik och naturvetenskap i skolor 8 i ett nationellt och internationellt perspektiv (Rapport 255). Stockholm: Skolverket.
- Skolverket. (2007). PIRLS 2006. Läsförmågan hos elever i årskurs 4 – i Sverige och i världen (Rapport 305). Stockholm: Skolverket.
- Skolverket. (2008a). TIMSS 2007. Svenska grundskoleelevers kunskaper i matematik och naturvetenskap i ett internationellt perspektiv (Rapport 323). Stockholm: Skolverket.
- Skolverket. (2008b). *Ämnesproven 2008 i grundskolans årskurs 5*. <http://www.skolverket.se/sb/d/306/a/14246>
- Skolverket. (2009). *Redovisning av uppdrag om uppföljning av IT-användning och IT-kompetens i förskola, skola och vuxenutbildning*. Stockholm: Skolverket.
- Sympson, J. B. & Hetter, R. D. (1985). Controlling item exposure rates in computerized adaptive testing. *Proceedings of the 27<sup>th</sup> annual meeting of the military testing association*, 973-977.



Taylor, C., Jamieson, J., Eignor, d., & Kirsch, I. (1998). *The relationship between computer familiarity and performance on computer-based TOEFL test tasks*. ETS research resport Nr. 98-08. Princeton, NJ: Educational Testing Services.

Thissen, D., & Mislevy, R.J. (2000). Testing Algorithms. I Wainer, H. (Ed.) *Computerized Adaptive Testing: A Primer*. Mahwah, NJ: Lawrence Erlbaum Associates.

Thomas, P., Smith, N., & Waugh, K. (2008). Automatically assessing graph-based diagrams. *Learning, Media and Technology*, 33(3), 249-267.

Thompson, N. A. (2007). A practioner's guide for variable-length computerized classification testing. *Practical Assessment Research & Evaluation*, 12, 1.

Tian, J-Q, Miao, D-M, Zhu, X, & Gong, J-J (2007). An introduction to the computerized adaptive testing. *US-China Education Review*, 4, 72-80.

Wainer, H. & Kiley, g. L. (1987). Item clusters and computerized adaptive testing: A case for testlets. *Journal of Educational Measurement*, 24, 185-201.

Wainer, H. & Lewis, C. (1990). Toward a psychometrics for testlets. *Journal of Educational Measurement*, 27, 1-14.

Wainer, H. (1993). Some practical considerations when converting a linearly administered test to an adaptive format. *Educational Measurement, Issues and Practice*, 12, 15-20.

Wainer, H. (2000). *Computerized adaptive testing: a primer* (2<sup>nd</sup> ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Wainer, H., & Mislevy, R.J. (2000). Item response theory, calibration, and estimation. I Wainer, H. (Ed.) *Computerized Adaptive Testing: A Primer*. Mahwah, NJ: Lawrence Erlbaum Associates.

Wainer, H., Bradlow, E. T., & Du, Z. (2000). Testlet response theory: an analog for the 3PL model useful in testlet-based adaptive testing. I W. J. van der Linden & C. A. W. Glas *Computerized adaptive testing: theory and practice*, 245-269. Dordrecht: Kluwer Academic Publishers.

- Wainer, H., Dorans, N. J., Flaugher, R., Mislevy, R. J., Green, B. F., Steinberg, L., & Thissen, D. (eds). (2000). *Computerized adaptive testing: a primer* (2<sup>nd</sup> ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Wald, A. (1947). *Sequential analysis*. New York: Wiley.
- van der Linden, W. J. & Glas, C. A. W. (2000). *Computerized Adaptive Testing: theory and practice*. Dordrecht: Kluwer Academic Publishers.
- van der Linden, W. J. (2000). Constrained adaptive testing with shadow tests. I van der Linden, W. J. & Glas, C. A. W. (Eds.) *Computerized Adaptive Testing: theory and practice* (27-52). Dordrecht: Kluwer Academic Publishers.
- Wandall, J. (2008). [http://crell.jrc.it/Presentations\\_Iceland%202008/Wandall.pdf](http://crell.jrc.it/Presentations_Iceland%202008/Wandall.pdf) (Hämtad 27/10, 2008).
- Wang, T. & Kolen, M. J. (2001). Evaluating comparability in computerized adaptive testing: issues, criteria and an example. *Journal of Educational Measurement*, 38, 1, 19-49.
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54, 427-450.
- Weiss, D. J. (Ed.) (1983). *New horizons in testing: latent trait theory and computerized adaptive testing*. New York: Academic Press.
- Weiss, D. J., & Kingsbury, G. G. (1984). Application of computerized adaptive testing to educational problems. *Journal of Educational Measurement*, 21, 361-375.
- Verschoor, A., & Straetmans, G. (2000). MATHCAT: A flexible testing system for adult mathematics education. I W. J. van der Linden & C. A. W. Glas (Eds.), *Computerized adaptive testing: theory and practice* (101-116). Dordrecht: Kluwer Academic Publishers.
- Wilhelm, O., & Schroeders, U. (2008). Traditional and Computerized Ability Measurement: Stressing Equivalence vs. Exploiting Opportunities The Transition to Computer-Based Assessment Reykjavik, Iceland 30.09.2008 [http://crell.jrc.ec.europa.eu/Presentations\\_Iceland%202008/Wilhelm.pdf](http://crell.jrc.ec.europa.eu/Presentations_Iceland%202008/Wilhelm.pdf) (Hämtad 20/10, 2008).

Winkley, J., Rainbow, B., & Baki, H. (2008, 8-9 juli). *Designing and using adaptive tests for large scale formative assessment 1999 to 2008*. Paper presented at the The 12th International Computer Assisted Assessment Conference, Loughborough.

Wise, S. L., & Plake, B. S. (1990). Computer-based testing in higher education. *Measurement and Evaluation in Counseling and Development*, 23, 3-10.

Wise, S. L. (1997). Overview of practical issues in a CAT program. *Paper presented at the 1997 annual meeting of the National Council on Measurement in Education*, Chicago, IL.

Vispoel, W. P., Rocklin, T. R., & Wang, T. (1994). Individual differences and test administration procedures: a comparison of fixed-item, computerized-adaptive, and self-adapted testing. *Applied Measurement in Education*, 53, 53-79.

Wright, B. D., & Bell, S. R. (1984). Item banks: what, why, how. *Journal of Educational Measurement*, 21, 331-345.

Zenisky, A. L., & Sireci, S. G. (2002). Technological innovations in large-scale assessment. *Applied Measurement in Education*, 15, 337-362.

## Andra källor

ACT <http://www.act.org/compass> (Hämtad 28/10, 2008).

@ventures <http://www.ventures.dk> (Hämtad 28/10, 2008).

@venx <http://uddannelsesforum2006.emu.dk/udstillere/oversigt/aventures.html> (Hämtad 21/10, 2008).

Assessing Learning in Australian Universities: Ideas, strategies and resources for quality in student assessment (2002) <http://www.cshe.unimelb.edu.au/assessinglearning/03/online.html> (Hämtad 29/10, 2008) asTTle <http://www.tki.org.nz/r/asttle/> (Hämtad 27/10, 2008).

Cambridge ESOL <http://www.cambridgeesol.org/exams/exams-info/computer-based-testing.html> (Hämtad 30/10, 2008), om item-banken: [http://www.cambridgeesol.org/rs\\_notes/offprints/pdfs/RN23p3-5.pdf](http://www.cambridgeesol.org/rs_notes/offprints/pdfs/RN23p3-5.pdf) (Hämtad 30/10, 2008).

CARLA <http://www.carla.umn.edu/about/lrc> (Hämtad 29/10, 2008).

Castle Rock <http://www.castlerockresearch.com/caa/>,  
<http://www.castlerockresearch.com/caa/FAQ.aspx>  
(Hämtad 27/10, 2008).

CatGlobal [http://www.catglobal.com/CATGlobal8/pdf/support/catswsys\\_8\\_2\\_rnotes.pdf](http://www.catglobal.com/CATGlobal8/pdf/support/catswsys_8_2_rnotes.pdf) (Hämtad 20/10, 2008).

CITO <http://www.cito.com> (Hämtad 29/10, 2008).

CitoTester [http://www.cito.nl/com\\_comprod/citotester/eind\\_fr.htm](http://www.cito.nl/com_comprod/citotester/eind_fr.htm)  
(Hämtad 27/10, 2008).

COPI <http://www.cal.org/topics/ta/copi.html> (Hämtad 18/10, 2008).

Danmark [http://evaluering.uvm.dk/templates/velkomst\\_layout.jsf](http://evaluering.uvm.dk/templates/velkomst_layout.jsf)  
(Hämtad 31/10, 2008).

Dialang <http://www.dialang.org/swedish/index.htm>, <http://dialang.org/project/english/index.html> (Hämtad 25/10, 2008).

Dialang *insertion*

<http://www.lancs.ac.uk/fss/projects/linguistics/experimental/new/item7/item7.htm> (Hämtad 18/10, 2008).

Dialang indirekt tal med ljudalternativ

<http://www.lancs.ac.uk/fss/projects/linguistics/experimental/new/item11/item11.htm> (Hämtad 18/10, 2008).

Dialang *Benchmarking* i färdigheten ”skriva”

<http://www.lancs.ac.uk/fss/projects/linguistics/experimental/new/item12/item12.htm> (Hämtad 18/10 2008).

Dialang *Benchmarking* i färdigheten ”tala”

<http://www.lancs.ac.uk/fss/projects/linguistics/experimental/new/item13/item13.htm> (Hämtad 18/10 2008).

Dialang *Confidence in chosen response* (slags självbedömning)

<http://www.lancs.ac.uk/fss/projects/linguistics/experimental/new/item8/item8.htm> (Hämtad 20/10, 2008).

Dialang *Deletion*

<http://www.lancs.ac.uk/fss/projects/linguistics/experimental/new/item8/item8a.htm> (Hämtad 20/10, 2008).

Dialang *Hot spot* Markera mening i textmassa samt lexikon  
<http://www.lancs.ac.uk/fss/projects/linguistics/experimental/new/item6/item6.htm> (Hämtad 20/10, 2008).

Dialang *Re-organisation*  
<http://www.lancs.ac.uk/fss/projects/linguistics/experimental/new/item5/item5.htm> (Hämtad 20/10, 2008).

Dialang Tematisk gruppering  
<http://www.lancs.ac.uk/fss/projects/linguistics/experimental/new/item9/item9.htm> (Hämtad 20/10, 2008).

Dialang Videoclip i hörförståelse  
<http://www.lancs.ac.uk/fss/projects/linguistics/experimental/new/item3/item3.htm> (Hämtad 20/10, 2008).

e-rater April 2005 issue of R&D Connections  
<http://www.ets.org/portal/site/ets/menuitem.1488512ecfd5b8849a77b13bc3921509/?vgnextoid=ed562d3631df4010VgnVCM10000022f95190RCRD&cvgnextchannel=a330253b164f4010VgnVCM10000022f95190RCRD> (Hämtad 17/10, 2008).

GMAT <http://www.gmac.com/gmac/thegmat/> (Hämtad 30/10, 2008).

GRE <http://gre-exams.com/gre-tests-mons.html?gclid=CNjHhYGbZJYCFQ2Y1QodZSPmyg> (Hämtad 29/10, 2008).

IELTS <http://www.ielts.org> (Hämtad 29/10, 2008).

Intellimetric (2008) <http://www.vantagelearning.com/school/products/intellimetric> (Hämtad 30/10, 2008).

Island [http://crell.jrc.it/Presentations\\_Iceland%202008/Björnsson%20-%20proposal.pdf](http://crell.jrc.it/Presentations_Iceland%202008/Björnsson%20-%20proposal.pdf) [http://crell.jrc.it/Presentations\\_Iceland%202008/Skulason.pdf](http://crell.jrc.it/Presentations_Iceland%202008/Skulason.pdf) (Hämtad 28/10, 2008).

Longman Market Leader <http://www.market-leader.net/> (Hämtad 23/10, 2008).

MELAB <http://www.lsa.umich.edu/eli/testing/melab> (Hämtad 29/10, 2008).

Norge, Norska nationella prov [http://udir.no/templates/udir/TM\\_Tema.aspx?id=3418](http://udir.no/templates/udir/TM_Tema.aspx?id=3418) (Hämtad 27/10, 2008).

NWEA <http://www.nwea.org/about/index.asp> (Hämtad 20/3, 2009).

PARES <http://www2.computer.org/portal/web/csdl/doi/10.1109/ICALT.2004.1357652> (Hämtad 29/10, 2008).

Pearson VUE <http://www.pearsonvue.com/sponsors/dev/> (Hämtad 21/10, 2008).

POET/OEPT Demo och information finns på <http://web.ics.purdue.edu/~aginther> (Hämtad 29/10, 2008).

QuestionMark <http://www.questionmark.com/us/perception/index.aspx> (Hämtad 29/10, 2008).

Quiz Center <http://school.discoveryeducation.com/quizcenter/quizcenter.html> (Hämtad 29/10, 2008).

Renet Oy <http://www.renet.fi/> (Hämtad 18/10, 2008).

Renstar <http://www.renlearn.com/sr/> (Hämtad 6/2, 2009).

Respondus <http://www.respondus.com> (Hämtad 21/10, 2008).

Skottland [www.sqa.org.uk/e-assessment](http://www.sqa.org.uk/e-assessment) (Hämtad 5/5, 2009).

SpeechRater [http://toefl.startpractice.com/programs/toefl/toefl\\_faq.htm](http://toefl.startpractice.com/programs/toefl/toefl_faq.htm) (Hämtad 30/10, 2008) <http://www.mkzchner.net/SLaTE07.pdf> (Hämtad 30/10, 2008).

Surveylang <http://www.surveylang.org/en/index.html> (Hämtad 27/10, 2008).

Teachers Alberta (2007)

<http://www.teachers.ab.ca/Issues%20In%20Education/Real%20Learning%20First/RLF%20Library/Pages/Computer%20Adaptive%20Assessment%20FAQs.aspx> (Hämtad 27/10, 2008).

TOEFL

<http://www.ets.org/portal/site/ets/menuitem.fab2360b1645a1de9b3a0779f1751509?vgnextoid=69c0197a484f4010VgnVCM10000022f95190RCRD> (Hämtad 30/10, 2008); [http://www.ets.org/Media/Tests/TOEFL/pdf/TOEFL\\_Tips.pdf](http://www.ets.org/Media/Tests/TOEFL/pdf/TOEFL_Tips.pdf) (Hämtad 27/2, 2009).

WebLAS UCLA, WebLAS demonstration av hf med multimedia  
[http://www.humnet.ucla.edu/web/departments/alt/weblas\\_esl\\_demo/  
demo\\_listen\\_psych1\\_vid.htm](http://www.humnet.ucla.edu/web/departments/alt/weblas_esl_demo/demo_listen_psych1_vid.htm) (Hämtad 20/10, 2008), öppna svar:  
<http://www.weblas.ucla.edu/> för exempel (klicka på am. knappen)  
(Hämtad 20/10, 2008).

Versant (PhonePass) <http://www.ordinate.com/versant/versant.jsp>  
(Hämtad 18/10, 2008).

## **Länkar**

<http://www.computerbasedtest.com>


<http://www.eric.ed.gov/>

<http://www.ets.org/gre/>

<http://www.toefl.org/>

<http://www.psych.umu.edu/psylabs/catcentral>

<http://www.evaluering.uvm.dk>



Skolverket gav 2008 Institutionen för beteendevetenskapliga mätningar, Umeå universitet ett uppdrag att i samverkan med Göteborgs universitet ta fram en kunskapsöversikt över forskning om och användning av adaptiva prov inom och utom Norden, särskilt fokuserad på adaptiva prov i matematik, NO-ämnen och engelska.

Uppdraget innehöll två mindre pilotstudier med datorbaserade prov i engelska och matematik.

Föreliggande kunskapsöversikt avslutas med ett strategiskt diskussionsunderlag för framtida beslut om adaptiva provs roll i Sverige. Författarna rekommenderar bland annat att datorbaserade prov som inte är adaptiva bör prioriteras och att sådana framtida satsningar i första hand måste göras på bedömningar med stödjande, formativa syften. Skolverket bör även stödja ett mer långsiktigt forsknings- och utvecklingsarbete i samarbete med universitet och andra institutioner inom och utom landet.

*Skolverket*

[www.skolverket.se](http://www.skolverket.se)