

Utprovning av uppgifter till högskoleprovet

Utvärdering av försöksverksamheten med en ny utprovningsmodell

Gunilla Ögren

Pm nr 140, 1998



ISSN 1100-696X
ISRN UM-PED-PM--140--SE

ABSTRACT

The Swedish Scholastic Aptitude Test (SweSAT) has been used since 1977 for selection to higher education. Up to the spring of 1996 new items were pretested in upper secondary school. As from the spring of 1996 the pretesting of new items is performed in connection with the regular administrations of the test. The aim of this report was to evaluate the new pretesting model, i.e. to study how well the new pretest results serve as a basis for predicting regular test results. A survey and analysis of items pretested in accordance with the new model and in the four regular tests administered from the autumn of 1996 to the spring of 1998 show that the new model is clearly superior to the old one. The new model exhibits reduced differences between pretest and regular test and also reduced gender differences.

INNEHÅLLSFÖRTECKNING	sid
SAMMANFATTNING	
INLEDNING	1
SYFTE	2
DEN NYA UTPRÖVNINGSMODELLEN	2
RESULTATREDOVISNING; reguljärt - utprovning	8
DELPROVET ORD (ordkunskap)	9
ORD 96B	9
ORD 97A	12
ORD 97B	15
ORD 98A	18
En jämförelse av ej reviderade och reviderade uppgifter i ORD 96B-98A	21
DELPROVET NOG (kvantitativa resonemang)	23
NOG 97A	23
NOG 97B	24
NOG 98A	26
DELPROVET LÄS (läsförståelse)	28
LÄS 96B	29
LÄS 97A	30
LÄS 97B	33
LÄS 98A	36

	sid
En jämförelse av ej reviderade och reviderade uppgifter i LÄS 96B-98A	39
DELPROVET DTK (tolkning av diagram, tabeller och kartor)	40
DTK 96B	40
DTK 97A	42
DTK 97B	43
DTK 98A	45
En jämförelse av ej reviderade och reviderade uppgifter i DTK-provet 96B-98A	48
AVSLUTANDE REFLEKTIONER	49
REFERENSER	
BILAGA	

TABELLFÖRTECKNING

Tabell	sid
1 Exempel på en provdag	3
2 Delprovets placering i reguljärproven 96A-98A	3
3 Antal uppgifter som prövats ut enligt den nya utprövningsmodellen och ingått i något av proven 96B, 97A, 97B eller 98A. Uppgifterna är fördelade på delprov och utprövningsprov tillfälle.	6
4 Reguljära värden och utprövningsvärden för de ORD-uppgifter i 96B som prövats ut enligt den nya utprövningsmodellen.	10
5 Reguljära värden och utprövningsvärden för ORD 96B uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.	11
6 Reguljära värden och utprövningsvärden för de ORD-uppgifter i 97A som prövats ut enligt den nya utprövningsmodellen.	13
7 Reguljära värden och utprövningsvärden för ORD 97A uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.	14
8 Reguljära värden och utprövningsvärden för de ORD-uppgifter i 97B som prövats ut enligt den nya utprövningsmodellen.	16
9 Reguljära värden och utprövningsvärden för ORD 97B uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.	17
10 Reguljära värden och utprövningsvärden för de ORD-uppgifter i 98A som prövats ut enligt den nya utprövningsmodellen.	19
11 Reguljära värden och utprövningsvärden för ORD 98A uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.	20
12 Medelvärde för ej reviderade och reviderade uppgifter som prövats ut enligt den nya modellen och ingick i ORD 96B, 97A, 97B och 98A. Inom parentes anges antal uppgifter.	22
13 Reguljärt värde och utprövningsvärde för den NOG-uppgift i 97A som prövats ut enligt den nya utprövningsmodellen.	23

14	Reguljära värden och utprövningsvärden för NOG 97A uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.	24
15	Reguljära värden och utprövningsvärden för de NOG-uppgifter i 97B som prövats ut enligt den nya utprövningsmodellen.	25
16	Reguljära värden och utprövningsvärden för NOG 97B uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.	25
17	Reguljära värden och utprövningsvärden för de NOG-uppgifter i 98A som prövats ut enligt den nya utprövningsmodellen.	27
18	Reguljära värden och utprövningsvärden för NOG 98A uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.	27
19	Reguljära värden och utprövningsvärden för de LÄS-uppgifter i 96B som prövats ut enligt den nya utprövningsmodellen.	29
20	Reguljära värden och utprövningsvärden för LÄS 96B uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.	30
21	Reguljära värden och utprövningsvärden för de LÄS-uppgifter i 97A som prövats ut enligt den nya utprövningsmodellen.	31
22	Reguljära värden och utprövningsvärden för LÄS 97A uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.	33
23	Reguljära värden och utprövningsvärden för de LÄS-uppgifter i 97B som prövats ut enligt den nya utprövningsmodellen.	35
24	Reguljära värden och utprövningsvärden för de LÄS-uppgifter i 98A som prövats ut enligt den nya utprövningsmodellen.	37
25	Medelvärde för ej reviderade och reviderade uppgifter i 96B, 97A, 97B, och 98A. Inom parentes anges antal uppgifter.	39
26	Reguljära värden och utprövningsvärden för de DTK-uppgifter i 96B som prövats ut enligt den nya utprövningsmodellen.	40

27	Reguljära värden och utprövningsvärden för DTK 96B uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.	41
28	Reguljära värden och utprövningsvärden för de DTK-uppgifter i 97A som prövats ut enligt den nya utprövningsmodellen.	42
29	Reguljära värden och utprövningsvärden för DTK 97A uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Antal uppgifter anges inom parentes.	43
30	Reguljära värden och utprövningsvärden för de DTK-uppgifter i 97B som prövats ut enligt den nya utprövningsmodellen.	44
31	Reguljära värden och utprövningsvärden för DTK 97B uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Antal uppgifter anges inom parentes.	45
32	Reguljära värden och utprövningsvärden för de DTK-uppgifter i 98A som prövats ut enligt den nya utprövningsmodellen.	46
33	Reguljära värden och utprövningsvärden för DTK 98A uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal.	47
34	Medelvärde för ej reviderade och reviderade uppgifter i 96B, 97A, 97B, och 98A. Inom parentes anges antal uppgifter.	48
35	Resultat på högskoleprovet 96A, 96B, 97A, 97B och 98A för hela gruppen provdeltagare. Provets placering under provdagen anges inom parentes.	49
36	Differensen mellan utprövningsmedelvärde och reguljärt medelvärde vid provtillfällen 94A-98A. Inom parentes anges antal uppgifter som prövats ut enligt den nya modellen.	52

SAMMANFATTNING

Högskoleprovet har sedan 1977 fungerat som ett urvalsinstrument för sökande till universitets- och högskolestudier. Utprövning av nya provuppgifter har till och med hösten 1995 skett två gånger per år på elever i gymnasieskolans teoretiska linjer, årskurs tre. Det fanns vissa problem med dessa rutiner, varav ett hänger samman med motivationsaspekten. Eleverna, som besvarade utprövningsuppgifterna, visste att det var frågan om utprövning av uppgifter och kunde därför inte förväntas ha maximal motivation. En annan invändning var att de grupper som genomförde utprövningen var alltför små och heterogena för att resultaten skulle bli tillförlitliga. Ett återkommande problem var stora bortfall eftersom det var frivilligt att delta i utprövningarna. Dessutom medförde förändringen av gymnasieskolan, från linje till program, svårigheter att få tid för utprövningar på gymnasieskolorna. För de flesta internationella urvalsprov som liknar högskoleprovet genomförs utprövning av nya uppgifter i samband med det reguljära provgenomförandet eftersom det är det utan jämförelse bästa sättet att erhålla tillförlitliga utprövningsresultat, vilket i sin tur är en förutsättning för ett bra prov.

Sedan våren 1996 sker, som försöksverksamhet, utprövning av nya provuppgifter till högskoleprovet i samband med det reguljära provgenomförandet.

Syftet med föreliggande rapport är att närmare undersöka hur väl de nya utprövningsresultaten fungerar som underlag för att prognostisera reguljära provresultat. Rapporten redovisar de uppgifter som ingått i något av de reguljära proven hösten 1996, våren 1997, hösten 1997 eller våren 1998 och som prövats ut enligt den nya modellen. Vidare görs en jämförelse också med de uppgifter som ingått i något av dessa prov och prövats ut enligt den gamla modellen. Resultatredovisningen baseras på uppgifternas lösningsproportion, svårighetsgrad för män respektive kvinnor samt uppgifternas diskriminationsförmåga.

Resultaten visar att den nya utprövningsmodellen är klart överlägsen den gamla. Differensen mellan utprövning och reguljärt provtillfälle blir mindre totalt och det blir även mindre förändringar av könsskillnaderna. Dessutom möjliggör den nya modellen användning av den nya testteorin Item Response Theory (IRT). Hittills har högskoleprovsdata enbart analyserats utifrån klassisk testteori.

I samband med utvärderingen har det framkommit en del nya faktorer som är av betydelse för utfallet, vilka medfört att ett antal studier påbörjats för att studera betydelsen av dessa nya faktorer.

INLEDNING

En mycket viktig del vid konstruktionen av högskoleprovet är utprövning av nya provuppgifter. Varje enskild uppgift måste testas minst en gång innan den ingår i ett reguljärt prov. Utprövning av uppgifter har två huvudsyften. Resultaten av utprövningarna används dels som underlag för uppgifts-analyser och dels som underlag för prognos av reguljärprovets svårighetsgrad. Uppgifts-analys används för utvärdering av utprövningsresultaten; vilket bl.a. handlar om att fastställa hur svarsförslagen till en uppgift fungerar, hur svåra eller lätta uppgifterna är, om uppgifterna diskriminerar mellan hög- och lågpresterande. Utprövningarna är också en förutsättning för att proven ska kunna göras parallella mellan provtillfällena¹ och för att skillnader mellan grupper ska kunna kontrolleras.

Till och med hösten 1995 prövades uppgifter ut på elever i gymnasieskolan. Det fanns vissa problem med dessa rutiner, varav ett hänger samman med motivationsaspekten. Eleverna, som besvarade utprövningsuppgifterna, visste att det var frågan om utprövning av uppgifter och kunde därför inte förväntas ha maximal motivation. En annan invändning var att de grupper som genomförde utprövningen var alltför små och heterogena för att resultaten skulle bli tillförlitliga. Ett återkommande problem var stora bortfall eftersom det var frivilligt att delta i utprövningarna. Dessutom medförde förändringen av gymnasieskolan, från linje till program, svårigheter att få tid för utprövningar på gymnasieskolorna. För de flesta internationella urvalsprov som liknar högskoleprovet genomförs utprövning av nya uppgifter i samband med det reguljära provgenomförandet eftersom det är det utan jämförelse bästa sättet att erhålla tillförlitliga utprövningsresultat, vilket i sin tur är en förutsättning för ett bra prov.

Vid sammanträden med dåvarande Verket för högskoleservice (VHS) 1994-06-01 och 1994-10-05 diskuterades förändring av utprövningsrutinerna och under våren 1995 beslöt VHS att utprövning av nya provuppgifter skulle få ske i samband med det reguljära provgenomförandet, med början våren 1996. VHS fattade även beslut om att en utvärdering av de nya utprövningsrutinerna skulle genomföras efter två år. De förändrade utprövningsrutinerna innebär att provdeltagarna måste besvara ett antal uppgifter som inte ingår i deras resultat och som inte heller offentliggörs efter provgenomförandet. En beskrivning av bakgrunden till förändringen av utprövningsrutiner finns i rapporten *Utprövning av provuppgifter. Förändring av högskoleprovets utprövningsrutiner* (Stage, 1995), som även beskriver olika tänkbara modeller för utprövning av uppgifter.

¹ Med parallella prov menas att olika provversioner har samma psykometriska egenskaper (t.ex. att de är lika svåra och att de rangordnar provdeltagare på samma sätt). För att ytterligare säkerställa kravet på parallellitet genomförs, en normering av den reguljära provpoängen (råpoäng), som syftar till att jämställa resultaten från olika provtillfällen. Normeringen innebär att provpoängen, som varierar inom intervallet 0-122, överförs till en skala från 0.0 till 2.0 (s.k. normerad poäng), vilket innebär att det är lika lätt eller svårt att erhålla en viss normerad poäng oavsett sammansättningen på provdeltagargruppen även om provets svårighetsgrad skulle fluktuera.

SYFTE

Det är mycket viktigt att proven är jämförbara mellan provtillfällena vad avser svårighetsnivå, könsskillnader och diskriminering mellan hög- och lågpresterande. För att kunna uppfylla nämnda krav är det av yttersta vikt att ha god kännedom om hur säkra utvärderingsresultaten är för att kunna prognostisera de reguljära provresultaten. Syftet med föreliggande rapport är att närmare undersöka hur väl de nya utvärderingsresultaten fungerar som underlag för att prognostisera reguljära provresultat.

DEN NYA UTPRÖVNINGSMODELLEN

Förändringen av utvärderingsrutinerna kom även att innebära en omstrukturering av provet eftersom provdagen inte kunde förlängas och det bedömdes som angeläget att få en enhetlig längd på varje provpass. Efter ingående diskussioner med det internationella vetenskapliga rådet för högskoleprovet, de svenska ledamöterna av det vetenskapliga rådet och samtliga involverade provkonstruktörer och forskare uppnåddes enighet om att den bästa kompromissen mellan mättekniska och praktiska/administrativa överväganden innefattade en omstrukturering av provet, som samtidigt innebar att delprovssammansättningen förändrades. Detta innebar att ett delprov fick utgå och antalet uppgifter och provtider förändrades för några av delproven (Stage, 1995). Den förändring som kom att ske för respektive delprov var att;

- AO - allmänorientering - utgick
- ORD - ordförståelse - ökades från 30 till 40 uppgifter med provtiden 15 minuter
- ELF - engelsk läsförståelse - minskades från 24 till 20 uppgifter med provtiden 35 minuter
- NOG - kvantitativa resonemang - ökades från 20 till 22 uppgifter med provtiden 50 minuter
- LÄS - svensk läsförståelse - minskades från 24 till 20 uppgifter med provtiden 50 minuter
- DTK - tolkning av diagram, tabeller och kartor - förblev oförändrat, 20 uppgifter med provtiden 50 minuter

Provdagen indelades i fem provblock om vardera 50 minuter varav ett block utgörs av utvärderingsuppgifter. ELF och ORD ingår alltid i samma block, medan de övriga delproven utgör var sitt block. Ordningföljden mellan blocken varierar emellertid mellan olika provtillfällen. I tabell 1 visas ett exempel på hur en provdag kan se ut samt antal uppgifter och provtid för respektive delprov. I tabell 2 visas delprovets placering vid de reguljära provtillfällena våren 1996-våren 1998. I följande redovisning anges provtillfälle på våren som A och på hösten som B.

Tabell 1. Exempel på en provdag.

	Delprov	Antal uppg	Tid (min)	Totaltid (min)
Block 1	DTK	20	50	50
Block 2	LÄS	20	50	50
Block 3	NOG	22	50	50
Block 4	ELF + ORD	20 + 40	35 + 15	50
Block 5	Ngt av ovanst	20/22/20+40	50	50
Totalt		122		250

Tabell 2. Delprovets placering i reguljärproven 96A-98A.

	96A	96B	97A	97B	98A
BLOCK 1	NOG	Utprovning	ELF+ORD	Utprovning	DTK
BLOCK 2	Utprovning	NOG	DTK	NOG	Utprovning
BLOCK 3	ELF+ORD	LÄS	Utprovning	DTK	NOG
BLOCK 4	DTK	DTK	LÄS	LÄS	ELF+ORD
BLOCK 5	LÄS	ELF+ORD	NOG	ELF+ORD	LÄS

De nya utprovningrutinerna togs i bruk våren 1996, vilket innebär att i dagsläget har fem reguljära provomgångar genomförts sedan dess och att det nu finns ett större antal uppgifter som ingått i de reguljära proven och prövats ut enligt den nya modellen.

Den nya modellen innebär, som framgår av tabell 1, att samtliga provdeltagare får göra två versioner av något av blocken. Innehållet i utprovningsblocket varierar mellan högskoleenheterna så att en högskoleenhet har t.ex. ELF- och ORD-provet som utprovningsblock medan en annan har NOG-provet som utprovningsblock. Även mellan de högskoleenheter som har samma delprov som utprovningsblock varierar innehållet eftersom ett stort antal uppgifter prövas ut för varje delprov och fördelas därmed på olika utprovningversioner. Om tre högskoleenheter har t.ex. NOG-provet som utprovningsblock innebär det att de tre högskoleenheterna har var sin version av delprovet. Att uppgifterna prövas ut inom olika högskoleenheter innebär också att man måste ta i beaktande att uppgifterna är utprovade i olika utprovningsgrupper. Om man jämför medelvärdet på reguljärprovet för de olika utprovningsgrupperna finner man att

vissa utprövningsgrupper har ett högre resultat än totalgruppen medan vissa grupper har ett lägre resultat. Utprövningsvärdena härrör alltså från en högskoleenhet medan reguljärvärdena gäller hela populationen provdeltagare.

I och med de förändrade utprövningsrutinerna har arbetet med konstruktion och granskning av uppgifter ökat kraftigt. Detta har medfört betydligt högre krav på de uppgifter som prövas ut, vilket har inneburit att arbetet med utprövningsuppgifter kommit att bli en mycket arbetskrävande fas i provkonstruktionsarbetet. Vid varje provtillfälle prövas ungefär 600 nykonstruerade uppgifter ut fördelade på de fem delproven. Varje nykonstruerad uppgift granskas numera av en lokal granskningsgrupp som förutom ansvarig konstruktör består av såväl ämnes- som språkexperter. Granskningen sker på lite olika sätt beroende på vilket delprov som granskas. Delproven DTK och LÅS har förutom uppgifterna även figurmaterial respektive lästexter som bedöms ur olika aspekter, vilket innebär att dessa granskningar är ganska omfattande. Efter de lokala granskningarna sker revideringar i enlighet med de synpunkter som framkommit och materialet bearbetas på samma sätt som de reguljära provversionerna. Det är viktigt att utprövningsmaterialet är så väl granskat och genomarbetat att provdeltagarna inte kan avgöra vad som är utprövnings- eller reguljära uppgifter. Det är även viktigt ur andra aspekter att utprövningsuppgifterna är väl granskade innan utprovningen. Dels är det betydligt svårare att prognostisera reguljära värden när uppgifterna är reviderade och dels har i och med den nya utprövningsmodellen även möjligheterna att använda en ny testteori (IRT)² tillkommit, vilken förutsätter att utprövningsuppgifterna inte revideras på något sätt efter utprovning.

Varje uppgift som ingår i ett högskoleprov har prövats ut minst en gång på ungefär 2 000 personer. Efter utprovning sker såväl statistiska som logiska analyser, som sedan ligger till grund för en eventuell revidering av uppgifterna. Revidering kan innebära såväl ett nytt svarsförslag som en språklig förändring eller ett förtydligande. Att revidera uppgifter är en del av utvärderings- och konstruktionsarbetet, som visserligen kan komplicera prognosen av uppgifterna, men eftersom ett av huvudsyftena med utprovning är att göra analyser och utifrån dessa revidera det som anses nödvändigt så är revideringar ofrånkomliga. Revideringar kan även bli aktuella när en reguljär provversion ska sättas samman. Dessa är då oftast av annan karaktär och kan till en del beskrivas som mer av "kosmetiska" skäl som att t.ex. ändra ett svarsförslag. Ett exempel på en sådan kosmetisk förändring är uppgift 34 i ORD-provet hösten 1996. Ett av de felaktiga svarsförslagen, *takveranda*, ändrades till *glasveranda*. Det nya svarsförslaget blev .02 svårare. Den "tysta kunskap", som finns hos samtliga konstruktörer, innebär i det här fallet att konstruktören vet att en sådan revidering oftast inte har någon betydelse för utfallet utan har snarare sk. face-validity. Detta är en kunskap som är svår att fästa ner på papper, som är svår att verbalisera men som finns där som en stor tillgång (jfr. Billgren, et al. 1990).

² Item Response Theory (IRT) är modern testteori (se t.ex. Hambleton, 1991, Ramstedt, 1996). Denna analysmodell har använts bl.a. vid analys av konsekvenser av de förändringar av sammansättningen av högskoleprovet som genomfördes våren 1996 (Ramstedt & Stage, 1997).

Det finns dock forskning som visar att små språkliga förändringar ibland kan ha stor betydelse så man kan inte generellt säga att till synes små förändringar i uppgifter alltid ger små förändringar i lösningsproportion.

Det är mycket viktigt att utprövningsuppgifterna är ordentligt granskade för att minimera behovet av revideringar, men som tidigare nämnts är det nödvändigt ibland. Vid sammansättningen av reguljära prov måste man alltid ta hänsyn till i vilken grad uppgifterna är reviderade. Det innebär att man i princip delar in uppgifterna i tre kategorier; helt oförändrade, något förändrade och mycket förändrade. Blir förändringen alltför omfattande kan en omprövning krävas eftersom man inte anser sig kunna prognostisera resultatet efter förändringen. Eftersom utprövningen nu sker i samband med det reguljära provtillfället omprövas uppgifter mycket restriktivt eftersom omprövning innebär att uppgiften exponeras ytterligare en gång. Detta innebär att konstruktören skattar reviderade men ej omprövade uppgifters lösningsproportion för att kunna göra en prognos av reguljära proven. För att kunna göra en bra prognos är det därför av yttersta vikt att ha god kännedom om olika faktorer som kan påverka resultatet.

Vid sammansättningen av de reguljära delproven försöker konstruktörerna dels undvika att ta med helt nyutprovade uppgifter och dels att ta med mer än fyra uppgifter som ingått i samma provversion vid utprövningen. Anledningen är att minska riskerna för att provdeltagare ska känna igen uppgifter. Detta betyder att i det prov som gavs hösten 1996 fanns ytterst få uppgifter som prövats ut enligt den nya modellen. I det reguljära prov som gavs våren och hösten 1997 samt våren 1998 återfanns däremot ett större antal uppgifter.

I tabell 3 redovisas antal uppgifter per delprov som har prövats ut enligt den nya modellen och som ingått i något av proven hösten 1996, våren 1997, hösten 1997 eller våren 1998.

Tabell 3. Antal uppgifter som prövats ut enligt den nya utvärderingsmodellen och ingått i något av proven 96B, 97A, 97B eller 98A. Uppgifterna är fördelade på delprov och utvärderingstillfälle.

Delprov Reguljärt tillfälle	Utvärderingstillfälle				Totalt	
	96A	96B	97A	97B		
ORD	96B	9	-	-	-	9
	97A	20	11	-	-	31
	97B	13	6	-	-	19
	98A	5	11	15	-	31
NOG	96B	-	-	-	-	-
	97A	1	-	-	-	1
	97B	3	-	-	-	3
	98A	2	5	-	-	7
LÄS	96B	4	-	-	-	4
	97A	16	-	-	-	16
	97B	16	4	-	-	20
	98A	4	8	8	-	20
DTK	96B	2	-	-	-	2
	97A	2	-	-	-	2
	97B	3	-	1	-	4
	98A	5	2	-	-	7
ELF	96B	-	-	-	-	-
	97A	14	-	-	-	14
	97B	7	13	-	-	20
	98A	7	12	-	-	19

Endast 15 uppgifter, vilket motsvarar 12 procent av uppgifterna, i reguljärprovet hösten 1996 hade prövats ut enligt den nya utvärderingsmodellen. Vid provtillfällena våren 1997 och hösten 1997 var motsvarande siffra 64 respektive 66 uppgifter, dvs. drygt hälften av uppgifterna. Vid reguljärprovet våren 1998 hade antalet uppgifter som prövats ut enligt den nya modellen ökat till 84, vilket motsvarar 69 procent av uppgifterna. Delproven NOG och DTK har utnyttjat få uppgifter som har prövats ut enligt den nya modellen, medan delproven LÄS och ELF hösten 1997 och våren 1998 enbart innehöll uppgifter utprovade enligt den nya modellen (såväl som på en uppgift för ELF-provet våren 1998).

För att få underlag för utvärderingen av den nya utvärderingsmodellen redovisas i följande avsnitt tämligen detaljerat utfallet av de uppgifter som prövats ut enligt den nya modellen

i varje delprov för sig. En jämförelse görs även med de uppgifter som ingick i något av proven 96B, 97A, 97B eller 98A och som prövats ut enligt den gamla modellen. Med den gamla modellen menas den utprövningsmodell som nyttjades före våren 1996, dvs. utprövning av provuppgifter på elever på gymnasieskolans teoretiska linjer, årskurs tre. I detta avsnitt sker även en detaljerad redovisning av de revideringar som gjorts för de uppgifter som prövats ut enligt den nya modellen. Den som inte är intresserad av alla detaljer hänvisas till den sammanfattning som presenteras efter varje delprov och provtillfälle.

Som avslutning på rapporten redovisas och diskuteras olika faktorer som kan vara av betydelse för förändring i uppgifternas lösningsproportion mellan utprövning och reguljärt provtillfälle.

RESULTATREDOVISNING; reguljärt - utprovning

Följande redovisning omfattar delproven – ORD, NOG, LÄS och DTK - hösten 1996, våren 1997, hösten 1997 och våren 1998. Eftersom utfallet av delprovet ELF redovisats i skriften *Från utprovning till färdigt prov* (Hellekant, 1997) görs ingen redovisning av detta delprov i föreliggande rapport.

Redovisningen är i huvudsak inriktad på de uppgifter som prövats ut enligt den nya modellen. Därför görs en utförlig redovisning av dessa uppgifters statistiska värden från såväl utprovningen som från det reguljära provtillfället. I anslutning till kommentarer av vissa statistiska värden framträder även nya faktorer och problem som uppkommit i och med förändringen av utprovningensrutinerna.

De värden som är viktiga ur prognossynpunkt är uppgifternas svårighetsgrad, som här anges med lösningsproportion (p-värde), svårighetsgrad för män respektive kvinnor (m/k) samt uppgifternas diskriminationsförmåga, som anges med biserial korrelation (rbis). Dessa värden bör, för en god prognos, vara så lika som möjligt vid utprovning och reguljärt prov.

En mängd faktorer, såväl kända som okända, kan ge upphov till skillnader mellan utprovningensvärden och reguljära värden. För att få en systematisk och överblickbar redovisning som samtidigt inkorporerar några av alla de faktorer som kan ge upphov till skillnader mellan utprovningensvärden och reguljära värden, redovisas för varje uppgift även uppgiftens placering vid såväl utprovning som reguljärt provtillfälle samt vilka revideringar som har gjorts av uppgifterna efter utprovningen. I anslutning till uppgiftens nummer har en siffra angetts, vilken motsvarar en viss revidering (se nedan). I de fall där svarsförslagen bytt ordning anges det i texten. Dessutom redovisas skillnaden i medelvärde på motsvarande reguljära delprov för totalgruppen provdeltagare och den aktuella utprovningensgruppen (diff M reg/utpr) samt differensen i lösningsproportion för reguljärgruppen och utprovningensgruppen (diff p reg/utpr).

Revideringar:

- ¹ 1 distraktor
- ² 2 distraktorer
- ³ 3 distraktorer
- ⁴ Samtliga distraktorer
- ⁵ Det rätta svaret
- ⁶ 1 distraktor + rätt svar
- ⁷ 2 distraktorer + rätt svar
- ⁸ Frågestam
- ⁹ Frågestam + 1 distraktor
- ¹⁰ Frågestam + 2 distraktorer
- ¹¹ Frågestam + 1 distraktor + rätt svar
- ¹² Figurtext förenklad

För att kunna jämföra de uppgifter som prövats ut enligt den nya och den gamla modellen redovisas reguljärmedelvärdet och utprövningsmedelvärdet för de uppgifter som prövats ut enligt respektive modell. Dessutom görs en sammanställning av uppgifter som prövats ut enligt den nya modellen uppdelade på reviderade och icke reviderade. Samtliga reguljära värden som redovisas i föreliggande rapport är medelvärden för totalgruppen. Redovisningen sker per delprov och för respektive provtillfälle.

DELPROVET ORD (ordkunskap)

ORD-provet avser att mäta förståelse av ord och begrepp. Provet innehåller 40 uppgifter med fem svarsförslag till varje uppgift. ORD-uppgifterna kan delas in i två huvudgrupper; synonymer och hyponymer. Synonymer kännetecknas av att orden har samma eller nästan samma innebörd. Hyponymer kännetecknas av att det ena ordet betecknar en delmängd av den klass av företeelser som det andra ordet, t.ex. förhållandet ryssja - fiskeredskap.

Orden tillhör någon av ordklasserna substantiv, verb, adjektiv eller adverb. I provet ingår både arvord, dvs. ord som funnits i språket sedan äldsta tid, och lånord. I varje prov ingår också ett eller ett par ålderdomliga ord, dvs. ord som var vanligare förr.

ORD 96B

I det reguljära provet hösten 1996 ingick nio ORD-uppgifter som prövats ut våren 1996 fördelade på fem olika högskoleenheter. Medelvärdet på reguljärprovet våren 1996 var 25.41 för totalgruppen medan medelvärdet på samma prov varierade mellan 24.82 och 25.26 för de fem utprövningsgrupperna, vilket innebär att i detta fall hade samtliga utprövningsgrupper ett lägre resultat på reguljärprovet än totalgruppen. Den största skillnaden mellan totalgruppen och utprövningsgruppen var .59. I genomsnitt räknat borde en sådan uppgift bli .015 lättare vid det reguljära provtillfället förutsatt att ingen betydande revidering har gjorts. Eftersom samtliga utprövningsgrupper hade ett lägre resultat på provet än totalgruppen skulle en uppräknad av utprövningsvärdet kunna göras för varje uppgift med .007, vilket är ett genomsnittsvärde av skillnaden mellan totalgruppens och utprövningsgruppernas medelvärde. Detta genomsnittsvärde blir då ett korrigerat värde för utprövningsgrupperna och är att betrakta som ett schablonpåslag. För att slutligen få en riktig prognos måste varje konstruktör göra en skattning av de revideringar som gjorts. Denna skattning lämnas i detta sammanhang därhän. Redovisningen kommer fortsättningsvis inte att ta hänsyn till sådana skattningar utan endast redovisa utprövningsvärden och reguljära värden. I följande redovisning, som tidigare nämnts, finns uppgifter i form av siffror som visar vilka uppgifter som reviderats och vilken typ av revidering som gjorts (se sid 8).

Tabell 4. Reguljära värden och utprövningsvärden för de ORD-uppgifter i 96B som prövats ut enligt den nya utprövningsmodellen.

Uppgift nr Reg (utpr)	Reguljärt 96B			Utprövning			Diff M reg/utpr	Diff p reg/utpr
	p	m/k	rbis	p	m/k	rbis		
2 ² (22)	.79	76/82	.34	.82	80/84	.39	+.59	-.03
5 (31)	.73	78/69	.64	.75	81/70	.68	+.43	-.02
8 (22)	.69	76/63	.51	.70	76/64	.52	+.20	-.01
13 ² (19)	.69	76/63	.68	.74	81/68	.68	+.15	-.05
21 ⁴ (10)	.57	42/69	.52	.62	48/75	.49	+.20	-.05
26 ² (25)	.48	45/51	.69	.47	43/49	.64	+.59	+.01
34 (25)	.41	39/43	.48	.40	37/43	.52	+.20	+.01
36 (23)	.60	61/59	.57	.54	56/53	.56	+.20	+.06
40 ¹ (39)	.65	73/59	.55	.63	73/55	.56	+.15	+.02
	5.61	5.66/5.58		5.67	5.75/5.61			-.06

Av tabell 3 framgår att fem av de nio uppgifterna som prövats ut enligt den nya modellen har reviderats efter utprövningen. I uppgift 40 var en distraktor reviderad medan samtliga distraktorer var reviderade i uppgift 21. I var och en av uppgifterna 2, 13 och 26 var två distraktorer reviderade.

Medelvärden för de nio uppgifterna var 5.61 vid utprövningen och blev 5.67 (.06 svårare) reguljärt. Skillnaden mellan utprövnings- och reguljärvärden är således liten. Fem uppgifter har blivit svårare och ligger i intervallet .01-.05 och fyra uppgifter har blivit .01-.06 lättare. Uppgift 36 har förändrats mest med avseende på lösningsproportion fastän den inte har reviderats. I uppgift 21 har samtliga distraktorer reviderats, men trots detta har inte uppgiften förändrats mer än .05. Men som tidigare nämnts kan ändringar göras av mer "kosmetiska" skäl och därför förändras lösningsproportionen ibland inte nämnvärt.

Korrelationen för lösningsproportionen mellan utprövning och reguljärt provtillfälle är 0.97 för de nio uppgifterna.

För att kunna jämföra den nya och den gamla utprövningsmodellen redovisas i tabell 4 reguljära värden och utprövningsvärden för ORD 96B uppdelade på uppgifter utprovade enligt den nya och den gamla utprövningsmodellen. De värden som redovisas under den gamla modellen är korrigerade värden³, vilket även gäller vid redovisningen av ORD 97A, 97B och 98A.

³ Denna korrigerings är gjord på basis av sk. ankaruppgifter som ingick i utprövningarna före 1996. Med hjälp av dessa uppgifter, som tidigare ingick i ett reguljärt prov, definierades utprövningsgruppernas nivå. En beskrivning av ankaruppgifternas funktion finns i Henriksson, W. (1992). *Modell för prognos av provutfall utifrån delprovspecifika ankaruppgifter*.

Tabell 5. Regulerade värden och utprövningsvärden för ORD 96B uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.

	ORD 96B						
	Nya modellen (9)			Gamla modellen (31)		Hela provet	
	M	m/k	M	m/k	M	m/k	
Reguljärt	5.61	5.66/5.58		18.21	18.25/18.17	23.82	23.91/23.75
Utprovning	5.67	5.75/5.61		17.15	16.56/17.76	22.82	22.31/23.37
Diff	-.06			+1.06		+1.00	

Skillnaden mellan utprövnings- och reguljärvärden blev endast .06 för de nio uppgifter som prövats ut enligt den nya modellen. För de 31 uppgifter som prövats ut enligt den gamla modellen var motsvarande skillnad 1.06. Det här tyder på att de nya utprövningsvärdena är mera stabila än vad de gamla utprövningsvärdena är.

Skillnaden i utprövningsresultat mellan kvinnor och män var .14, till männens fördel, för de nio uppgifterna som prövats ut enligt den nya modellen. Reguljärt blev skillnaden något mindre (.08). För de 31 uppgifter som prövats ut enligt den gamla modellen var skillnaden i resultat mellan män och kvinnor 1.20 till kvinnornas fördel. Reguljärt blev skillnaden .08 till männens fördel för dessa 31 uppgifter. Även avseende könsdifferensen är således de nya utprövningsvärdena mer stabila.

Sammanfattningsvis kan sägas att utprövningsvärdena är betydligt mer stabila för de uppgifter som har prövats ut enligt den nya modellen. Skillnaden i medelvärde på de nio uppgifter som prövats ut enligt den nya modellen var endast .06, i genomsnitt .007 per uppgift. För de 31 uppgifterna som prövats ut enligt den gamla modellen var motsvarande skillnad 1.06, i genomsnitt .03 per uppgift. Könskillnaden förändrades marginellt mellan utprovning och reguljärt provtillfälle för uppgifter som prövats ut enligt den nya modellen, medan könskillnaden förändrades kraftigt (1.28) mellan utprovning och reguljärt provtillfälle för de uppgifter som prövats ut enligt den gamla modellen.

ORD 97A

I det reguljära provet våren 1997 ingick 31 ORD-uppgifter som prövats ut enligt den nya modellen varav 20 prövats ut våren 1996 och 11 hösten 1996. Uppgifterna som prövats ut våren 1996 var fördelade på samma fem högskoleenheter som uppgifterna som ingick i 96B, dvs. medelvärdet på reguljärprovet våren 1996 var 25.41 för totalgruppen medan medelvärdet på samma prov varierade mellan 24.82 och 25.26 för de fem utprövningsgrupperna.

Uppgifterna som prövats ut hösten 1996 var även de fördelade på fem högskoleenheter. Medelvärdet på reguljärprovet hösten 1996 var 23.82 för totalgruppen och för utprövningsgrupperna varierade medelvärdet mellan 22.59 och 23.70. Den största skillnaden mellan totalgrupp och utprövningsgrupp var 1.23. I genomsnitt räknat borde en sådan uppgift bli .03 lättare reguljärt. För de uppgifter som ingick i 97A hade samtliga 31 uppgifter prövats ut i utprövningsgrupper som hade ett lägre resultat än totalgruppen, det gäller såväl de vår- som höstutprovade uppgifterna.

Eftersom samtliga utprövningsgrupper hade sämre resultat på reguljärprovet än totalgruppen skulle en uppräknning av utprövningsvärdet kunna göras för varje uppgift med .012, ett genomsnittsvärde av skillnaden mellan totalgruppens och utprövningsgruppernas medelvärde.

Tabell 6. Reguljära värden och utprovningvärden för de ORD-uppgifter i 97A som prövats ut enligt den nya utprovningssmodellen.

Uppgift nr Reg (utpr)	Reguljärt 97A			Utprovning			Diff M reg/utpr	Diff reg-utpr
	p	m/k	rbis	p	m/k	rbis		
Vårutprovade:								
1 ² (8)	.70	73/67	.62	.73	76/71	.65	+0.15	-0.03
4 ¹ (20)	.71	66/76	.48	.79	72/84	.47	+0.20	-0.08
5 ⁷ (39)	.71	66/76	.40	.79	76/82	.36	+0.20	-0.08
9 (18)	.70	78/64	.50	.66	74/58	.51	+0.20	+0.04
10 ¹ (36)	.72	62/80	.58	.75	66/82	.56	+0.10	-0.03
11 ¹ (27)	.81	75/86	.42	.79	74/84	.40	+0.15	+0.02
15 ¹ (36)	.58	55/60	.44	.73	72/73	.49	+0.20	-0.15
16 ³ (14)	.71	72/69	.56	.65	64/65	.50	+0.15	+0.06
19 ¹ (5)	.41	44/39	.45	.46	49/44	.50	+0.59	-0.05
23 (16)	.63	64/63	.47	.65	65/64	.43	+0.59	-0.02
24 (38)	.55	53/57	.38	.57	56/58	.53	+0.59	-0.02
25 (12)	.58	58/59	.64	.51	51/51	.62	+0.20	+0.07
27 (24)	.64	64/65	.41	.69	70/68	.44	+0.15	-0.05
28 ⁶ (4)	.42	50/36	.57	.53	60/48	.61	+0.43	-0.11
29 (4)	.42	39/45	.42	.42	39/44	.42	+0.20	-----
35 ¹ (5)	.37	39/35	.53	.31	33/30	.40	+0.15	+0.06
36 ¹ (37)	.62	72/54	.49	.71	77/66	.54	+0.43	-0.09
38 ¹ (6)	.46	41/50	.40	.41	35/45	.40	+0.59	+0.05
39 ³ (28)	.39	40/38	.41	.27	28/27	.36	+0.43	+0.12
40 ² (39)	.31	34/29	.31	.31	33/29	.32	+0.59	-----
Höstutprovade:								
3 ³ (8)	.41	39/42	.45	.63	60/66	.27	+0.63	-0.22
8 ⁴ (34)	.64	66/61	.30	.40	43/38	.22	+1.23	+0.24
12 ⁶ (13)	.75	69/81	.54	.76	70/81	.58	+0.12	-0.01
14 (27)	.53	65/43	.65	.43	54/34	.66	+0.98	+0.10
18 ¹ (19)	.68	75/63	.56	.75	80/71	.57	+0.12	-0.07
26 (4)	.35	32/38	.27	.40	38/43	.26	+0.89	-0.05
30 ⁵ (4)	.42	40/44	.44	.37	33/40	.45	+0.98	+0.05
31 ¹ (6)	.51	60/43	.54	.46	56/38	.52	+0.98	+0.05
32 ¹ (6)	.74	67/79	.51	.69	65/73	.41	+0.63	+0.05
34 (30)	.65	65/66	.45	.64	63/64	.43	+0.63	+0.01
37 ⁷ (34)	.55	58/53	.48	.39	42/37	.30	+0.89	+0.16
	17.67	17.81/17.61		17.65	17.74/17.58			+0.02

Att uppgifter som reviderats på ett eller annat sätt förändras i svårighetsgrad är inte så konstigt. Det är däremot svårt att kunna uppskatta exakt hur mycket en förändring betyder. Om man jämför uppgift 10 och 15 där en distraktor har reviderats i vardera uppgift ser man att för uppgift 10 har revideringen inte påverkat lösningsproportionen nämnvärt (-.03). Däremot har revideringen av en distraktor i uppgift 15 haft stor betydelse (-.15). Detta är ett exempel på skillnader i fråga om revideringars betydelse i förhållande till avsikten med revideringarna.

Uppgifter som inte reviderats på något sätt kan också förändras i lösningsproportion om man jämför värdena mellan utprovning och reguljärt provtillfälle. Uppgift 14 har blivit .10 lättare fastän uppgiften inte reviderats. Få uppgifter som ingår i ett reguljärt prov har utprovats inom samma högskoleenhet. Av erfarenhet vet vi att det skiljer i provresultat mellan högskoleenheter, upp till 10 poäng på totalprovet mellan den högskoleenhet med det bästa och det sämsta resultatet (jfr sid 50). Det kan var en anledning till att det skiljer så pass mycket mellan utprovnings- och reguljärvärde. Medelvärdet på det reguljära ORD-provet var 23.82 för totalgruppen och 22.84 för denna utprovningsgrupp, dvs. utprovningsgruppen var en svag grupp. Uppgiften prövades ut hösten 1996 och gick reguljärt våren 1997. Uppgifter som prövas ut på hösten uppvisar också ett något sämre resultat än uppgifter som prövas ut på våren (Stage et. al, 1996, 1997 och Ramstedt & Stage, 1997). Detta kan innebära att lösningsproportionen blir något högre för en uppgift som är utprovad på hösten och ingår i ett vårprov.

Medelvärdet för de 31 uppgifterna var 17.65 vid utprovningen och 17.67 vid det reguljära provtillfället. Skillnaden mellan utprovningsvärde och reguljärt värde var således mycket liten trots att 22 av uppgifterna hade reviderats efter utprovningen.

Korrelationen för lösningsproportionerna mellan utprovning och reguljärt provtillfälle är .83 för de 31 uppgifterna. Korrelationen för lösningsproportionerna mellan utprovning och reguljärt provtillfälle är .92 för de 20 uppgifterna som var utprovade på våren 1996 och .66 för de 11 uppgifterna som var utprovade på hösten 1996. En anmärkningsvärt stor skillnad i korrelationen mellan vår- och höstutprovade uppgifter, som bör beaktas vid analys av kommande prov.

Tabell 7. Reguljära värden och utprovningsvärden för ORD 97A uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.

	ORD 97A					
	Nya modellen (31)		Gamla modellen (9)		Hela provet	
	M	m/k	M	m/k	M	m/k
Reguljärt	17.67	17.81/17.61	5.49	5.56/5.38	23.16	23.37/22.99
Utprovning	17.65	17.74/17.58	5.44	5.11/5.63	23.09	22.85/23.21
Diff	+.02		+.05		+.07	

Skillnaden mellan utprövningsvärden och reguljära värden blev endast .02 för de 31 uppgifterna som prövats ut enligt den nya modellen. För de nio uppgifterna som prövats ut enligt den gamla modellen blev också skillnaden mycket liten (.05). För hela delprovet blev skillnaden därmed endast .07. Resultatet mellan män och kvinnor förändrades marginellt mellan utprövning och reguljärt provtillfälle för de 31 uppgifterna som prövats ut enligt den nya modellen, från .16 till .20 till männens fördel. För de uppgifter som prövats ut enligt den gamla modellen skedde en förändring så tillvida att kvinnorna hade .52 högre resultat än männen vid utprövningen, men vid det reguljära provtillfället blev skillnaden .18 till männens fördel.

Sammanfattningsvis kan sägas att beträffande medelvärdet för de 31 uppgifter som prövats ut enligt den nya modellen överensstämde utprövningsvärdet mycket väl med det reguljära utfallet. Skillnaden i medelvärde var endast .02. För de nio uppgifter som prövats ut enligt den gamla modellen var skillnaden också mycket liten, .05. Skillnaden i resultat mellan män och kvinnor förändrades marginellt mellan utprövning och reguljärt provtillfälle för de uppgifter som prövats ut enligt den nya modellen medan skillnaden i resultat för män och kvinnor förändrades från .52 till kvinnornas fördel till .18 till männens fördel för de uppgifter som prövats ut enligt den gamla modellen.

ORD 97B

I det reguljära provet hösten 1997 ingick 19 ORD-uppgifter som prövats ut enligt den nya modellen. 13 uppgifter hade prövats ut våren 1996 fördelade på fem högskoleenheter och sex uppgifter hade prövats ut hösten 1996 fördelade på fyra högskoleenheter. De uppgifter som prövats ut våren 1996 hade prövats ut i samma utprövningsgrupper som redovisats för 96B och 97A. Medelvärdet på reguljärprovet våren 1996 var 25.41 för totalgruppen och medelvärdet varierade mellan 24.82 och 25.26 för de olika utprövningsgrupperna.

För de sex uppgifter som prövats ut hösten 1996 gäller samma utprövningsgrupper som finns redovisade för provtillfället 97A. Medelvärdet för totalgruppen var hösten 1996 23.82 och medelvärdet på reguljärprovet varierade mellan 22.59 och 23.70 för utprövningsgrupperna.

Skulle man korrigera värdet på uppgifterna för att utprövningsgrupperna var sämre än totalgruppen borde värdet för de vårutprövade uppgifterna öka i genomsnitt 0.008 per uppgift och de höstutprövade uppgifterna borde öka i genomsnitt med 0.02 per uppgift.

Tabell 8. Reguljära värden och utprövningsvärden för de ORD-uppgifter i 97B som prövats ut enligt den nya utprövningsmodellen.

Uppgift nr Reg (utpr)	Reguljärt 97B			Utprövning			Diff M reg/utpr	Diff p reg/utpr
	p	m/k	rbis	p	m/k	rbis		
Vårutprovade:								
2 (16)	.68	66/69	.36	.69	66/72	.37	+20	-.01
4 ¹ (5)	.89	92/86	.45	.61	63/61	.38	+43	+.28
5 ² (17)	.67	70/64	.48	.71	76/67	.42	+20	-.04
15 ² (31)	.58	63/53	.41	.71	77/67	.44	+15	-.13
16 ¹ (11)	.48	50/46	.41	.60	61/58	.39	+20	-.12
18 (32)	.74	70/77	.62	.77	74/80	.61	+43	-.03
23 ⁶ (30)	.53	50/56	.45	.54	54/53	.45	+20	-.01
25 (4)	.55	56/54	.60	.61	64/59	.57	+59	-.06
31 ² (30)	.62	63/61	.69	.61	64/58	.69	+20	+.01
34 ¹ (35)	.54	48/59	.57	.48	42/53	.48	+20	+.06
37 ² (20)	.56	62/51	.54	.41	44/39	.53	+43	+.15
38 (34)	.46	44/48	.62	.42	40/44	.62	+43	+.04
40 (14)	.74	71/76	.58	.84	83/85	.58	+43	-.10
Höstutprovade:								
20 (30)	.66	49/80	.44	.67	47/84	.42	+89	-.01
27 ¹ (5)	.59	54/63	.74	.52	45/58	.64	+89	+.07
28 ² (25)	.47	46/48	.37	.54	51/56	.22	+12	-.07
29 ² (17)	.57	60/54	.70	.58	61/55	.61	+89	-.01
32 ¹ (36)	.59	60/58	.51	.44	45/43	.47	+98	+.15
39 ¹ (38)	.51	50/52	.63	.43	43/43	.62	+1.23	+.08
	11.43	11.24/11.55		11.18	11.00/11.35			+.25

Av de 19 uppgifter som prövats ut enligt den nya modellen hade 13 uppgifter reviderats. För sex uppgifter hade en distraktor reviderats, för ytterligare sex hade två distraktorer reviderats och för en uppgift hade samtliga svarsförslag inklusive korrekt svar reviderats.

Det ord som har förändrats mest i lösningsproportion är *Millennium* (uppgift 4). Det skiljer .28 mellan utprövnings- och reguljärvärde. Millennium prövades ut våren 1996 och gick reguljärt hösten 1997. En distraktor reviderades och avsikten var att den distraktor som reviderades skulle ”locka” mindre och det blev också resultatet, vilket kan förklara .06, dvs. skillnaden i lösningsproportion mellan den utbytta och den nya distraktorn. Den troliga förklaringen till uppgiften blivit så mycket lättare är att ordet ligger i tiden, vi närmar oss ett millenniumskifte och dessutom fanns det en TV-serie som hette Millennium vid tidpunkten för reguljärprovet.

En annan uppgift som ändrats ganska mycket i lösningsproportion är uppgift 32, som blev .15 lättare. Anledningen är en ny distraktor och att den lockat mindre än den ursprungliga, vilket även var avsikten. Vid revideringar av distraktorerna kan man jämföra lösningsproportionerna före och efter revidering. Om man däremot skall fundera över vad som ligger till grund för en skillnad mellan utprövningsvärde och reguljärt värde i de fall uppgiften inte har reviderats kan man som tidigare nämnts titta på hur ”duktig” utprövningsgruppen var, om uppgiften var vår- eller höstutprövad, men även position i utprövningsversionen jämfört med i reguljärprovet är av intresse att studera. Uppgift 40 hade inte reviderats alls men blev ändå .10 svårare, vilket troligen kan tillskrivas uppgiftens placering i provhäftet. Vid det reguljära provtillfället var uppgiften placerad som sista uppgift medan den i utprövningsversionen var placerad i mitten av provhäftet. ORD-provet består ju som tidigare nämnts av 40 uppgifter och provtiden är 15 minuter, vilket innebär att en ”speedfaktor” kan komma in. Betydelsen av position/placering i ORD-provet är ett område som är föremål för pågående studier.

Korrelationen för lösningsproportionerna mellan utprövning och reguljärt provtillfälle är .61 för de 19 uppgifterna. Uppdelat på vår- och höstutprövade uppgifter är korrelationen .60 respektive .52.

Tabell 9. Reguljära värden och utprövningsvärden för ORD 97B uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.

	ORD 97B		Gamla modellen (21)		Hela provet	
	Nya modellen (19)		M	m/k	M	m/k
Reguljärt	11.43	11.24/11.55	10.70	10.85/10.62	22.13	22.09/22.17
Utprövning	11.18	11.00/11.35	12.22	11.91/12.44	23.40	22.91/23.79
Diff	+.25		-1.52		-1.27	

Medelvärde för de 19 uppgifterna blev reguljärt 11.18 och i utprövningen 11.43, dvs. .25 lättare reguljärt. En liten skillnad med hänsyn till att 13 av uppgifterna hade reviderats och att uppgift 4 blev .28 lättare. Däremot blev skillnaden stor när det gäller de uppgifter som prövats ut enligt den gamla modellen. Dessa uppgifter blev 1.52 svårare reguljärt jämfört med vid utprövningen. Hela provet blev därmed 1.27 svårare reguljärt. Resultatet mellan män och kvinnor förändrades marginellt för de 19 uppgifterna som prövats ut enligt den nya modellen. Kvinnorna hade ett högre resultat vid båda tillfällena. För de uppgifter som prövats ut enligt den gamla modellen förändrades skillnaden så tillvida att kvinnorna hade .53 bättre resultat än männen vid utprövningen, men vid det reguljära tillfället blev skillnaden .23 till männens fördel.

Sammanfattningsvis kan sägas att medelvärdet är mer stabilt när uppgifterna har prövats ut enligt den nya modellen. Skillnaden i medelvärde mellan utprövning och reguljärt prov var .25 för de 19 uppgifter (i genomsnitt .01 per uppgift) som prövats ut enligt den nya

modellen och 1.52 för de 21 uppgifter (i genomsnitt .07 per uppgift) som prövats ut enligt den gamla modellen. Medelvärdet på uppgifterna som prövats ut enligt den gamla modellen blev 1.52 svårare reguljärt än vid utprovningen. Nu bör man komma ihåg att det görs revideringar i uppgifterna och dessa har inte prognostiserats utan det är rena utprovningvärden som ligger till grund för redovisningen. Dessutom redovisas inte heller, i denna sammanställning, i vilken omfattning de uppgifter som prövats ut enligt den gamla modellen var reviderade. Könsskillnaden förändrades marginellt mellan utprovning och reguljärt provtillfälle för de uppgifter som prövats ut enligt den nya modellen medan skillnaden i resultat mellan män och kvinnor för de uppgifter som prövats ut enligt den gamla modellen förändrades från .53 till kvinnornas fördel till .23 till männens fördel.

ORD 98A

I det reguljära ORD-provet våren 1998 ingick 31 uppgifter som prövats ut enligt den nya modellen. Fem uppgifter hade prövats våren 1996 fördelade på tre högskoleenheter. Medelvärdet på reguljärprovet våren 1996 var, som tidigare nämnts, 25.41 för totalgruppen medan medelvärdet varierade mellan 25.20 och 25.49 för de olika utprovningsgrupperna. Den största skillnaden mellan totalgrupp och utprovningsgrupp var .21, i genomsnitt räknat borde en sådan uppgift bli .005 lättare reguljärt.

11 uppgifter hade prövats ut hösten 1996 fördelade på fem högskoleenheter och samma utprovningsgrupper som redovisats för 97A och 97B. Medelvärdet på reguljärprovet hösten 1996 var 23.82 för totalgruppen och medelvärdet varierade mellan 22.59 och 23.70 för de olika utprovningsgrupperna.

15 uppgifter hade prövats ut våren 1997 fördelade på fem högskoleenheter. Medelvärdet på reguljärprovet våren 1997 var 23.16 för totalgruppen och medelvärdet varierade mellan 22.66 och 24.39 för de olika utprovningsgrupperna. Skillnaden mellan totalgrupp och den utprovningsgrupp som hade det lägsta resultatet var .50. I genomsnitt räknat borde en sådan uppgift bli .01 lättare reguljärt. Skillnaden mellan totalgrupp och den utprovningsgrupp som hade det högsta resultatet var 1.23. I genomsnitt räknat borde en sådan uppgift bli .03 svårare reguljärt.

20 av de 31 uppgifterna som prövats ut enligt den nya modellen var utprovade på våren, antingen våren 1996 eller våren 1997. Hälften av dessa uppgifter hade prövats ut i utprovningsgrupper som hade ett sämre resultat än totalgruppen och hälften hade prövats ut i grupper som hade högre resultat än totalgruppen. 11 uppgifter hade prövats ut på hösten och samtliga utprovningsgrupper hade sämre resultat än totalgruppen. Skulle man korrigera värdet för de vårutprovade uppgifterna för att vissa utprovningsgrupper hade ett högre resultat än totalgruppen och för att vissa hade ett lägre resultat borde värdet för dessa vårutprovade uppgifter sänkas med i genomsnitt .004 per uppgift. De höstutprovade uppgifterna borde istället öka med .022 per uppgift eftersom samtliga utprovningsgrupper hade lägre resultat än totalgruppen.

Tabell 10. Regulerade värden och utprövningsvärden för de ORD-uppgifter i 98A som prövats ut enligt den nya utprövningsmodellen.

Uppgift nr Reg (utpr)	Regulerat 98A			Utprövning			Diff M reg/utpr	Diff p reg/utpr
	p	m/k	rbis	p	m/k	rbis		
Vårutprovade:								
1 ⁷ (34)	.71	73/69	.55	.76	78/75	.58	-1.23	-.05
2 (27)	.75	85/68	.60	.80	88/74	.64	-.37	-.05
3 (3)	.65	48/79	.39	.69	54/83	.38	-1.23	-.04
5 (22)	.77	75/78	.50	.76	74/78	.52	-1.23	+.01
6 ² (32)	.73	72/73	.46	.76	76/75	.56	+.21	-.03
10 ¹ (34)	.67	69/64	.68	.62	68/57	.65	+.20	+.05
11 (30)	.82	67/94	.41	.67	53/81	.38	+.29	+.15
12 (4)	.54	55/53	.37	.60	59/61	.36	+.50	-.06
14 (24)	.53	58/48	.45	.58	62/54	.47	+.29	-.05
15 (38)	.70	68/72	.40	.69	69/70	.49	+.21	+.01
16 ¹ (8)	.73	67/77	.33	.67	61/72	.33	-.08	+.06
18 (33)	.57	48/65	.54	.53	44/61	.52	-.37	+.04
19 (3)	.71	69/72	.61	.68	67/69	.59	+.50	+.03
20 ¹ (2)	.49	58/42	.33	.48	56/42	.39	-.25	+.01
23 ¹ (8)	.40	36/43	.39	.50	47/53	.41	-.25	-.10
24 (7)	.57	60/55	.52	.58	61/55	.54	+.50	-.01
27 (9)	.38	35/40	.56	.37	35/38	.62	-1.23	+.01
30 (40)	.74	71/77	.72	.71	68/74	.68	-.37	+.03
31 ¹ (19)	.49	49/49	.41	.46	46/46	.43	+.21	+.03
33 ⁶ (32)	.60	54/65	.58	.48	41/54	.57	+.29	+.12
Höstutprovade:								
7 (13)	.83	84/82	.55	.79	81/77	.56	+.98	+.04
8 ¹ (37)	.64	62/66	.43	.66	65/67	.42	+.63	-.02
17 (32)	.84	72/94	.52	.82	71/90	.47	+1.23	+.02
21 ⁵ (37)	.70	74/67	.51	.63	65/60	.39	+.89	+.07
26 (15)	.57	56/59	.44	.55	57/54	.44	+.98	+.02
29 (21)	.65	68/64	.55	.57	61/54	.55	+1.23	+.08
32 (14)	.50	53/48	.60	.43	47/40	.53	+1.23	+.07
36 ¹ (38)	.39	41/37	.46	.40	43/38	.37	+.98	-.01
37 ¹ (29)	.37	35/38	.39	.39	39/40	.45	+1.23	-.02
38 (6)	.28	28/27	.47	.33	35/32	.42	+.12	-.05
39 (17)	.49	49/49	.57	.46	45/46	.58	+.12	+.03
	18.81	18.39/19.14		18.42	18.16/18.70			+.39

Av de 31 uppgifter som prövats ut enligt den nya modellen hade 12 uppgifter reviderats. För åtta uppgifter hade en distraktor reviderats och för fyra av dessa hade dessutom

placering av rätt svar ändrats. För ytterligare nio uppgifter hade placeringen av rätt svar ändrats. De uppgifter för vilka placeringen av rätt svar hade ändrats är; 1,2,3, 5, 7, 8, 10, 17, 18, 19, 20, 23 och 24.

Medelvärde för de 31 uppgifter som prövats ut enligt den nya modellen var 18.42 vid utprovningen och blev 18.81 reguljärt. Skillnaden mellan utprovning och reguljärt blev .39, dvs. provet blev .39 lättare reguljärt. Skillnaden är inte speciellt stor med tanke på att en del uppgifter reviderats och att för drygt 40 procent av uppgifterna har placering av rätt svar ändrats. Det finns vissa indikationer som tyder på att om man ändrar på placeringen av svarsförslagen kan för vissa typer av uppgifter lösningsproportionen ändras. Anledningen till att man ändrar i placeringen av svarsförslagen är ibland att ett svarsförslag blir så frekvent som rätt svar jämfört med de övriga svarsförslagen eftersom uppgifterna som ska ingå i ett reguljärt prov plockas samman från olika utprovningversioner. Ändring av placeringen av rätt svar görs även av stilistiska skäl, vilket varit anledningen till flertalet av ändringarna i detta ORD-prov. Vid tidpunkten när reguljärprovet 98A sattes samman och färdigställdes hade inte betydelsen av förändring av placering av rätt svar uppmärksammats. Studier har påbörjats för att se vilken betydelse förändring av placering av rätt svar har.

Det ord som förändrats mest i lösningsproportion är uppgift 11, *Sarong*, som förändrats .15 från utprovning till reguljärt provtillfälle, fastän ingen revidering har skett. Ytterligare ett ord som ligger i tiden (jfr sid 16, millennium), såtillvida att det är ett klädesplagg som blivit alltmer populärt under våren 1998. Studerar man utfallet för de övriga uppgifterna i tabell 10 ser man att ungefär tre fjärdedelar av uppgifterna (23 st) har blivit antingen .01-.05 svårare eller lättare. Två uppgifter har blivit .06-.10 svårare medan resterande sex uppgifter har blivit .06-.15 lättare (inkl. uppg. 11). Trots dessa skillnader mellan enskilda uppgifters utprovnings- och reguljärvärden blev skillnaden endast .39.

Korrelationen för lösningsproportionerna mellan utprovning och reguljärt provtillfälle är .93 för de 31 uppgifterna. Korrelationen för lösningsproportionerna mellan utprovning och reguljärt provtillfälle är .88 för de 20 uppgifter som prövats ut våren 1996 och 1997 och .98 för de 11 som prövats ut hösten 1996.

Tabell 11. Reguljära värden och utprovningvärden för ORD 98A uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.

	ORD 98A		Gamla modellen (9)		Hela provet	
	Nya modellen (31)		M	m/k	M	m/k
Reguljärt	18.81	18.39/19.14	4.24	4.16/4.33	23.05	22.55/23.47
Utprovning	18.42	18.16/18.70	3.85	3.67/3.98	22.27	21.83/22.68
Diff	+.39		+.39		+.78	

Medelvärde för de 31 uppgifter som prövats ut enligt den nya modellen blev 18.81 reguljärt och var vid utprövningen 18.42, dvs. provet blev .39 lättare reguljärt. En ganska liten skillnad med tanke på att 23 av dessa uppgifter var reviderade i någon form. Dessutom hade placeringen av rätt svar ändrats från utprövning till reguljärt provtillfälle för fyra av dessa uppgifter och för ytterligare nio uppgifter hade också placeringen ändrats. För en av dessa nio uppgifter hade även en distraktor reviderats och rätt svar hade genomgått en mindre revidering. Detta innebär att för åtta uppgifter hade positionen av rätt svar ändrats, men inga övriga revideringar i uppgiften hade skett. Medelvärde för de nio uppgifterna som prövats ut enligt den gamla modellen var vid utprövningen 3.85 och blev 4.24 reguljärt. Skillnaden mellan utprövningsvärde och reguljärt värde blev lika stor för de nio uppgifterna som prövats ut enligt den gamla modellen som för de 31 uppgifter som prövats ut enligt den nya modellen. Hela provet blev .78 lättare reguljärt än vad utprövningsvärdena visade. Resultatet mellan män och kvinnor förändrades .21, från .54 till .75 till kvinnornas fördel för de 31 uppgifterna som prövats ut enligt den nya modellen. För de nio uppgifter som prövats ut enligt den gamla modellen var resultatet till kvinnornas fördel vid båda tillfällena och resultatet ändrades från .31 till .17.

Sammanfattningsvis kan sägas att utprövningsvärdena för de uppgifter som prövats ut enligt den nya modellen är mer stabila än för de uppgifter som prövats ut enligt den gamla modellen, trots att det görs revideringar. Skillnaden i medelvärde mellan utprövning och reguljärt var .39 för de 31 uppgifter (i genomsnitt .01 per uppgift) som prövats ut enligt den nya modellen. Skillnaden var densamma, dvs. .39 för de nio uppgifter (i genomsnitt .04 per uppgift) som prövats ut enligt den gamla modellen. Resultatet mellan män och kvinnor var till kvinnornas fördel vid såväl utprövningen som reguljärt provtillfälle för de 31 uppgifter som prövats ut enligt den nya modellen. Skillnaden ökade från .54 till .75. För de nio uppgifter som prövats ut enligt den gamla modellen var resultatet också till kvinnornas fördel vid såväl utprövning som reguljärt provtillfälle. Men den hade minskat något, från .31 vid utprövningen till .17 vid reguljärprovet.

En jämförelse av ej reviderade och reviderade uppgifter i ORD 96B-98A

I redovisningen av ORD-proven 96B, 97A, 97B och 98A kan man se att på uppgiftsnivå är variationerna mellan utprövningsvärden och reguljära värden ganska stora, där en del uppgifter har blivit lättare medan andra har blivit svårare. Att svårighetsgraden för reviderade uppgifter förändras är inte oväntat, eftersom revideringar ofta görs av den anledningen att man vill göra en uppgift lättare eller svårare. Att vissa revideringar inte påverkar svårighetsgraden är dock inte lika självklart vilket hänger samman med att de görs av "kosmetiska" skäl. I tabell 12 redovisas en sammanställning av de uppgifter som prövats ut enligt den nya utprövningsmodellen och ingår i något av proven 96B, 97A, 97B eller 98A. Här har en uppdelning på icke reviderade och reviderade uppgifter gjorts.

Tabell 12. Medelvärde för ej reviderade och reviderade uppgifter som prövats ut enligt den nya modellen och ingick i ORD 96B, 97A, 97B och 98A. Inom parentes anges antal uppgifter.

	ORD 96B		ORD 97A		ORD 97B		ORD 98A	
	Ej rev. M (4)	Rev. M (5)	Ej rev. M (9)	Rev. M (22)	Ej rev. M (6)	Rev. M (13)	Ej rev. M (19)	Rev. M (12)
Reguljärt	2.43	3.18	5.05	12.62	3.83	7.60	11.89	6.92
Utprovning	2.39	3.28	4.97	12.68	4.00	7.18	11.61	6.81
Diff	+0.04	-0.10	+0.08	-0.06	-0.17	+0.42	+0.28	+0.11

I ORD 96 B ingick nio uppgifter som prövats ut enligt den nya modellen. Fyra av dessa uppgifter hade inte reviderats och fem hade reviderats. I genomsnitt blev de ej reviderade uppgifterna .01 lättare reguljärt jämfört med utprovningen och de reviderade uppgifterna blev i genomsnitt .02 svårare reguljärt.

I ORD 97A ingick 31 uppgifter som prövats ut enligt den nya modellen varav 22 hade reviderats. I genomsnitt blev de ej reviderade uppgifterna .01 svårare reguljärt jämfört med utprovningen och de reviderade uppgifterna blev endast .06 svårare för sammanlagt 22 uppgifter, vilket blir i genomsnitt 0.003 per uppgift. Om man ser till skillnaden mellan de reguljära värdena och utprovningens värdena är det ett mycket bra resultat. Även om det på uppgiftsnivå kan vara stora skillnader jämnar det ut sig med ökande antal uppgifter.

I ORD 97B ingick 19 uppgifter som prövats ut enligt den nya modellen varav 13 hade reviderats. I genomsnitt blev de ej reviderade uppgifterna .03 svårare reguljärt jämfört med utprovningen och de reviderade blev i genomsnitt .03 lättare reguljärt.

I ORD 98A ingick 31 uppgifter som prövats ut enligt den nya modellen varav 12 hade reviderats. I genomsnitt blev de ej reviderade uppgifterna 0.015 lättare reguljärt jämfört med vid utprovningen och de reviderade uppgifterna blev i genomsnitt .01 lättare reguljärt.

Korrelationen för lösningsproportionen mellan utprovning och reguljärt provtillfälle är .85 för samtliga ORD-uppgifter i 96B, 97A, 97B och 98A som prövats ut enligt den nya modellen.

DELPROVET NOG (kvantitativa resonemang)

Delprovet NOG avser att mäta förmågan att tänka logiskt kring olika typer av matematiska problem. Provet förutsätter kunskaper om vissa grundläggande kvantitativa begrepp och enheter som; aritmetik, reella tal, procent, mätning, enheter, geometri, algebra och funktionslära samt beskrivande statistik och sannolikhetslära. Den uppgift som föreläggs provdeltagaren är att bedöma om tillräcklig information föreligger för att ett givet problem skall kunna lösas entydigt. Uppgiftsformatet är fast, vilket innebär att svarsförslagen är identiska för alla uppgifter i provet. Uppgiften inleds med att ett problem presenteras. Provet består av 22 uppgifter.

För NOG-provets del har hittills endast 11 uppgifter som prövats ut enligt den nya modellen kommit till användning i något av reguljärproven 97A, 97B och 98A. I 96B ingick ingen uppgift som prövats ut enligt den nya modellen.

NOG 97A

I det reguljära provet våren 1997 ingick endast en NOG-uppgift som prövats ut enligt den nya modellen och det var uppgift 11. Den hade prövats ut våren 1996. Medelvärdet på reguljärprovet våren 1996 var 13.83 för totalgruppen och för utprövningsgruppen 13.81, dvs. två mycket jämna grupper.

Tabell 13. Reguljärt värde och utprövningsvärde för den NOG-uppgift i 97A som prövats ut enligt den nya utprövningsmodellen.

Uppgift nr Reg (utpr)	Reguljärt 97A			Utprövning			Diff M	Diff p
	p	m/k	rbis	p	m/k	rbis	reg/utpr	reg/utpr
11 (13)	.22	29/16	.46	.22	29/16	.49	+.02	----

Uppgiften hade inte reviderats efter utprövningen. Ett mycket bra resultat när värdena överensstämmer så väl mellan utprövningsgrupp och reguljärgrupp. Emellertid är överensstämmelsen att betrakta som rimlig när uppgiften inte var reviderad och när utprövningsgrupp och reguljärgrupp kan betraktas som likvärdiga med avseende på prestation.

Även om endast en uppgift har prövats ut enligt den nya modellen görs en jämförelse av den nya och den gamla modellen i tabell 14.

Tabell 14. Regljära värden och utprövningsvärden för NOG 97A uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.

	NOG 97A					
	Nya modellen (1)		Gamla modellen (21)		Hela provet	
	M	m/k	M	m/k	M	m/k
Regljärt	.22	29/16	12.67	14.00/11.56	12.89	14.29/11.72
Utprövning	.22	29/16	11.01	12.12/10.21	11.23	12.41/10.37
Diff	.00		+1.66		+1.66	

Som redan konstaterats blev det ingen skillnad mellan utprövningsvärde och regljärt värde för den enda uppgift som prövats ut enligt den nya modellen. Skillnaden på de 21 uppgifterna som prövats ut enligt den gamla modellen blev däremot 1.66, vilket blev skillnaden för hela NOG-provet 97A.

Skillnaden i utprövningsresultat mellan män och kvinnor var .13 till männens fördel både vid utprövningen och regljärprovet för den enda uppgift som prövats ut enligt den nya modellen. För de 21 uppgifter som prövats ut enligt den gamla modellen var skillnaden i resultat mellan män och kvinnor 1.91 till männens fördel. Regljärt ökade denna skillnad till 2.44.

Sammanfattningsvis kan sägas att medelvärdet på provet förändrades totalt 1.66 och de uppgifterna som prövats ut enligt den gamla modellen stod för denna skillnad. Den enda uppgift som prövats ut enligt den nya modellen hade exakt samma resultat vid regljärprovet som vid utprövningen.

NOG 97B

I det regljära provet hösten 1997 ingick tre NOG-uppgifter som prövats ut enligt den nya modellen. Uppgifterna hade prövats ut våren 1996 fördelade på tre högskoleenheter. Medelvärdet på regljärprovet våren 1996 var, som tidigare nämnts, 13.83 och medelvärdet för de tre utprövningsgrupperna varierade mellan 13.48 och 13.81. Den största skillnaden mellan totalgrupp och utprövningsgrupp var .35. I genomsnitt räknat borde en sådan uppgift bli .016 lättare vid det regljära provtillfället. Eftersom alla tre utprövningsgrupperna hade ett något lägre resultat än totalgruppen borde en uppräknning av utprövningsvärdet göras med .01 för varje uppgift, vilket är ett genomsnittsvärde av skillnaden mellan totalgruppens och utprövningsgruppernas medelvärde.

Tabell 15. Regulerade värden och utprövningsvärden för de NOG-uppgifter i 97B som prövats ut enligt den nya utprövningsmodellen.

Uppgift nr Reg (utpr)	Regulerat 97B			Utprövning			Diff M reg/utpr	Diff p reg/utpr
	p	m/k	rbis	p	m/k	rbis		
4 (4)	.48	54/43	.52	.46	52/41	.54	+ .32	+ .02
10 (5)	.39	48/31	.70	.38	45/32	.66	+ .35	+ .01
13 (16)	.55	61/50	.49	.61	65/57	.47	+ .02	- .06
	1.42	1.63/1.24		1.45	1.62/1.30			- .03

Ingen av de tre uppgifterna som prövats ut enligt den nya modellen hade reviderats. Medelvärde för de tre uppgifterna var 1.42 vid utprövningen och blev 1.45 regulerat. Provet blev .03 svårare regulerat. Skillnaden mellan utprövningsvärden och regulerade värden är således mycket liten.

Tabell 16. Regulerade värden och utprövningsvärden för NOG 97B uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.

	NOG 97B Nya modellen (3)		Gamla modellen (19)		Hela provet	
	M	m/k	M	m/k	M	m/k
Regulerat	1.42	1.63/1.24	11.66	12.69/10.79	13.08	14.32/12.03
Utprövning	1.45	1.62/1.30	9.57	10.61/8.82	11.02	12.23/10.12
Diff	-.03		+2.09		+2.06	

Skillnaden mellan utprövningsvärden och regulerade värden för de tre uppgifterna som prövats ut enligt den nya modellen blev endast .03. För de 19 uppgifter som prövats ut enligt den gamla modellen var motsvarande skillnad 2.09. Totalt blev skillnaden 2.06 mellan utprövning och regulerat provtillfälle, vilket tyder på att de nya utprövningsvärdena är mer stabila än de gamla, även om få uppgifter använts hittills.

Skillnaden i utprövningsresultat mellan män och kvinnor var .32 till mennens fördel för de tre uppgifterna som prövats ut enligt den nya modellen. Regulerat blev skillnaden något större .39. För de 19 uppgifterna som prövats ut enligt den gamla modellen var skillnaden i resultat mellan män och kvinnor 1.79 till mennens fördel vid utprövningen. Regulerat blev skillnaden 1.91.

Sammanfattningsvis kan sägas att medelvärdet på provet förändrades totalt 2.06 och de var de uppgifter som prövats ut enligt den gamla modellen som stod för den största

skillnaden. De tre uppgifter som prövats ut enligt den nya modellen blev .03 (i genomsnitt .01 per uppgift) svårare reguljärt. De övriga 19 uppgifterna som prövats ut enligt den gamla modellen blev 2.09 (i genomsnitt .11 per uppgift) lättare. Könsskillnaden ökade .07 mellan utprovning och reguljärt provtillfälle för de tre uppgifter som prövats ut enligt den nya modellen och för de 19 uppgifter som prövats ut enligt den gamla modellen var ökade könsskillnaden .11. För NOG provets del har skillnaden i resultat mellan män och kvinnor inte varierat så mycket mellan utprovning och reguljärt utfall, vilket vi kan se i tabell 16. Könsskillnaden är oavsett utprovningmodell relativt stor.

NOG 98A

I reguljärprovet våren 1998 ingick sju uppgifter som prövats ut enligt den nya modellen varav två hade prövats ut våren 1996 och fem hösten 1996. Uppgifterna som hade prövats ut våren 1996 var utprovade i varsin högskoleenhet och de uppgifter som prövats ut hösten 1996 var fördelade på fyra högskoleenheter. Medelvärdet på reguljärprovet våren 1996 var 13.83 och medelvärdet för de två utprovningssgrupperna var 13.48 respektive 13.51. Detta innebär att båda utprovningssgrupperna hade ett lägre resultat än totalgruppen. Skillnaden mellan totalgruppen och utprovningssgrupperna var .32 respektive .35. I genomsnitt räknat borde uppgifterna bli .014 respektive .016 lättare reguljärt.

Medelvärdet på reguljärprovet hösten 1996 var 13.51 för totalgruppen och medelvärdet för samma prov varierade mellan 13.87 och 14.37. Den största skillnaden mellan totalgrupp och utprovningssgrupp var .86. I genomsnitt räknat borde en sådan uppgift bli .04 svårare vid det reguljära provtillfället. Eftersom fem av sju uppgifter var utprovade i grupper som hade ett högre resultat än totalgruppen skulle en nedräkning av utprovningssvärdet kunna göras med .011 för varje uppgift, vilket är ett genomsnittsvärde av skillnaden mellan totalgruppen och utprovningssgruppernas medelvärde.

Tabell 17. Reguljära värden och utprövningsvärden för de NOG-uppgifter i 98A som prövats ut enligt den nya utprövningsmodellen.

Uppgift nr Reg (utpr)	Reguljärt 98A			Utprövning			Diff M reg/utpr	Diff p reg/utpr		
	p	m/k	rbis	p	m/k	rbis				
Vårutprovade:										
2 (17)	.58	65/53	.52	.60	69/51	.46	+.35	-.02		
12 (19)	.35	44/27	.49	.35	44/26	.55	+.32	-----		
Höstutprovade:										
1 (12)	.60	67/55	.48	.67	75/62	.41	-.86	-.07		
13 ⁸ (16)	.55	60/51	.49	.44	47/41	.47	+.02	+11		
15 ⁸ (6)	.45	56/36	.59	.54	69/42	.68	-.57	-.09		
18 ⁸ (1)	.29	39/21	.67	.31	39/24	.56	-.36	-.02		
19 (12)	.33	38/29	.42	.41	49/34	.52	-.57	-.08		
			3.15	3.69/2.72				3.32	3.92/2.80	-17

Tre av de sju uppgifterna som prövats ut enligt den nya modellen hade reviderats. För NOG-provets del innebär det att frågestammen har reviderats eftersom svarsförslagen är fasta. Medelvärdet för de sju uppgifterna var 3.32 vid utprövningen och blev 3.15 reguljärt. Skillnaden mellan män och kvinnor blev något mindre vid reguljärprovet än vid utprövningen. Vid utprövningen var skillnaden 1.12 till männens fördel och reguljärt blev den .97. En liten minskning hade skett.

Tabell 18. Reguljära värden och utprövningsvärden för NOG 98A uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.

	NOG 98A					
	Nya modellen (7)		Gamla modellen (15)		Hela provet	
	M	m/k	M	m/k	M	m/k
Reguljärt	3.15	3.69/2.72	9.18	10.21/8.30	12.33	13.90/11.02
Utprövning	3.32	3.92/2.80	7.67	8.60/6.98	10.99	12.52/9.78
Diff	-.17		+1.51		+1.34	

Skillnaden mellan utprövningsvärden och reguljära värden blev .17 för de sju uppgifterna som prövats ut enligt den nya modellen. Reguljärt blev uppgifterna något svårare. För de 15 uppgifter som prövats ut enligt den gamla modellen var motsvarande skillnad 1.51. Totalt blev skillnaden 1.34 mellan utprövning och reguljärt provtillfälle.

Skillnaden i utprövningsresultat mellan män och kvinnor var 1.12 till männens fördel för de sju uppgifterna som prövats ut enligt den nya modellen. Reguljärt blev skillnaden något mindre .97. För de 19 uppgifterna som prövats ut enligt den gamla modellen var skillnaden i resultat mellan män och kvinnor 1.62 till männens fördel vid utprovningen. Reguljärt blev skillnaden 1.91.

Sammanfattningsvis kan sägas att medelvärdet på provet förändrades totalt 1.34 och det var de uppgifter som prövats ut enligt den gamla modellen som stod för den största skillnaden. De sju uppgifter som prövats ut enligt den nya modellen blev .17 (i genomsnitt .02 per uppgift) svårare. De övriga 15 uppgifter som prövats ut enligt den gamla modellen blev 1.51 (i genomsnitt .10 per uppgift) lättare. Könsskillnaden minskade .15 mellan utprovning och reguljärt för de sju uppgifter som prövats ut enligt den nya modellen och för de 15 uppgifter som prövats ut enligt den gamla modellen ökade könsskillnaden med .29. Som tidigare nämnts är könsskillnaden stor oavsett utprovningsmodell.

11 uppgifter som prövats ut enligt den nya modellen har ingått, under försöksverksamheten, i något av NOG-proven. Eftersom endast tre av dessa uppgifter har reviderats mellan utprovning och reguljärt provtillfälle görs ingen redovisning av den typ som gjorts i tabell 12.

Korrelationen för lösningsproportionen mellan utprovning och reguljärt provtillfälle är .91 för samtliga NOG-uppgifter i 97A, 97B och 98A

DELPROVET LÄS (läsförståelse)

Delprovet LÄS avser att mäta svensk läsförståelse i vid mening. Provet består av fem texter och till varje text hör fyra uppgifter, som ska lösas med hjälp av den information som finns i respektive text. Till varje uppgift hör fyra svarsförslag varav ett är rätt. Texterna är varierande vad avser stil och språk. De texter som ingår i LÄS-provet hämtas från såväl vetenskapliga artiklar och uppsatser som böcker, utredningar, recensioner och debattinlägg. De fem texterna ska i så hög grad som möjligt täcka de fem ämnesområdena; teknik och naturvetenskap, administration och ekonomi, vård, kultur och information samt utbildning och yrken.

Till varje text utprövas åtta uppgifter, vilket innebär att samma text ingår i två provversioner. Anledningen till att man måste pröva ut fler än fyra uppgifter är dels att inte fler än två uppgifter bör komma från samma utprovningssversion (för att undvika överexponering) och dels att det blir fler valmöjligheter att sätta samman den slutliga reguljära provversionen.

LÄS 96B

I det reguljära LÄS-provet hösten 1996 ingick fyra uppgifter som prövats ut våren 1996 fördelade på två högskoleenheter. Medelvärdet på reguljärprovet våren 1996 var 12.15 för totalgruppen medan medelvärdet för samma prov var 12.21 respektive 13.18 för de två utprövningsgrupperna, vilket innebär att båda utprövningsgrupperna hade ett högre resultat än totalgruppen. Den största skillnaden mellan totalgrupp och utprövningsgrupp var 1.03, i genomsnitt räknat borde en sådan uppgift bli .05 svårare vid det reguljära provtillfället förutsatt att ingen betydande revidering hade gjorts. Eftersom båda utprövningsgrupperna hade ett högre resultat än totalgruppen skulle en nedräkning av utprövningsvärdet kunna göras för varje uppgift med .027, ett genomsnittsvärde av skillnaden mellan totalgruppens och utprövningsgruppernas medelvärde. Detta blir då ett korrigerat värde för utprövningsgrupperna, ett sk. schablonpåslag. På samma sätt som för ORD- och NOG-proven måste konstruktören göra en skattning av de revideringar som gjorts men även för LÄS-provet lämnas den i detta sammanhang därhän. Redovisningen tar alltså inte hänsyn till sådana skattningar eller schablonpåslag utan redovisar endast utprövnings- och reguljärvärden. De siffror som finns i anslutning till uppgiftens nummer visar vilka uppgifter som reviderats och vilken typ av revidering som gjorts (se sid 8).

Tabell 19. Reguljära värden och utprövningsvärden för de LÄS-uppgifter i 96B som prövats ut enligt den nya utprövningsmodellen.

Uppgift nr Reg (utpr)	Reguljärt 96B			Utprövning			Diff M reg/utpr	Diff p reg/utpr
	p	m/k	rbis	p	m/k	rbis		
17 ¹ (9)	.69	70/68	.54	.90	90/89	.52	-1.03	-.21
18 ² (11)	.57	61/55	.57	.56	58/54	.40	-1.03	+.01
19 ⁴ (12)	.60	59/61	.47	.49	49/49	.46	-.06	+.11
20(10)	.64	64/64	.48	.76	78/75	.46	-.06	-.12
	2.50	2.54/2.48		2.71	2.75/2.67			-.21

Den text och tillhörande uppgifter som prövats ut enligt den nya modellen var placerad som sista text och därmed som uppgifterna 17-20 i LÄS 96B. Vid utprövningstillfället var uppgifterna placerade som nr 9-12, dvs. tillhörde den tredje texten. Medelvärdet för de fyra uppgifterna var 2.71 vid utprövningen och blev 2.50 i reguljärprovet. Tre av uppgifterna hade reviderats efter utprövningen. Den uppgift som inte hade reviderats var uppgift 20, dvs. den sista uppgiften i reguljärprovet, och den blev .12 svårare. När uppgiften prövades ut hade den position 10. En tidigare studie av LÄS-provet har visat att position har betydelse i LÄS-provet (rapport under utarbetande), vilket kan vara en förklaring till att uppgiften blev så pass mycket svårare. Det kan naturligtvis finnas fler tänkbara förklaringar, vilket är något som kommer att diskuteras längre fram.

I tabell 20 görs en jämförelse av de uppgifter i LÄS 96B som prövats ut enligt den nya respektive den gamla modellen.

Tabell 20. Reguljära värden och utprövningsvärden för LÄS 96B uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.

	LÄS 96B					
	Nya modellen (4)		Gamla modellen (16)		Hela provet	
	M	m/k	M	m/k	M	m/k
Reguljärt	2.50	2.54/2.48	10.73	10.86/10.61	13.23	13.40/13.09
Utprövning	2.71	2.75/2.67	8.20	7.49/8.56	10.91	10.24/11.23
Diff	-.21		+2.53		+2.32	

Skillnaden mellan utprövningsvärden och reguljära värden blev .21 för de fyra uppgifter som prövats ut enligt den nya modellen. För de 16 uppgifter som prövats ut enligt den gamla modellen var motsvarande skillnad 2.53. Uppgifter som prövats ut enligt den nya modellen blev .21 svårare reguljärt och uppgifterna utprovade enligt den gamla modellen blev 2.53 lättare reguljärt, dvs. reguljärprovet blev totalt 2.32 lättare än vad utprövningsvärdena visade.

Skillnaden i utprövningsresultat mellan män och kvinnor var .08 till männens fördel, för de fyra uppgifter som prövats ut enligt den nya modellen. Reguljärt blev skillnaden i stort sett densamma (.12). För de 16 uppgifter som prövats ut enligt den gamla modellen var skillnaden i resultat mellan män och kvinnor 1.07 till kvinnornas fördel. Reguljärt blev skillnaden .25 till männens fördel. För dessa 16 uppgifter skedde en ganska stor förändring, från 1.07 till kvinnornas fördel till .25 till männens fördel, vilket innebär att könsskillnaden förändrades totalt 1.32.

Sammanfattningsvis kan sägas att medelvärdet för de få uppgifter som prövats ut enligt den nya modellen är mer stabila än för de uppgifter som prövats ut enligt den gamla modellen. De fyra uppgifter som prövats ut enligt den nya modellen blev .21 (i genomsnitt .05 per uppgift) svårare och de 16 uppgifter som prövats ut enligt den gamla modellen blev 2.53 (i genomsnitt .16 per uppgift) lättare. Könsskillnaden förändrades marginellt mellan utprövning och reguljärt provtillfälle för de uppgifter som prövats ut enligt den nya modellen. Däremot blev förändringen i resultat mellan män och kvinnor för de uppgifter som prövats ut enligt den gamla modellen stor.

LÄS 97A

I reguljärprovet våren 1997 ingick 16 LÄS-uppgifter som prövats ut våren 1996 fördelade på åtta högskoleenheter. Medelvärdet på reguljärprovet våren 1996 var, som tidigare nämnts, 12.15 för totalgruppen och medelvärdet för samma prov varierade mellan 11.56 och 13.18 för de åtta utprövningsgrupperna. Fem av utprövningsgrupperna hade ett högre resultat på reguljärprovet än totalgruppen medan tre hade ett lägre resultat. Den största skillnaden mellan totalgruppen och utprövningsgrupperna var 1.03. I genomsnitt räknat

borde en sådan uppgift bli .05 svårare vid det reguljära provtillfället, förutsatt att ingen annan revidering av betydelse har gjorts. Tre utprövningsgrupper uppvisade ett sämre resultat på reguljärprovet än totalgruppen, vilket innebär att uppgifter som prövats ut i någon av dessa grupper borde bli lättare reguljärt. Skillnaden mellan totalgruppen och den utprövningsgrupp som hade det lägsta resultatet var .59. I genomsnitt räknat borde en sådan uppgift bli .03 lättare reguljärt. Eftersom fem av åtta utprövningsgrupper hade ett högre resultat än totalgruppen skulle en nedräkning av utprövningsvärdet kunna göras med .005 för varje uppgift, vilket är ett genomsnittsvärde av skillnaden mellan totalgruppens och utprövningsgruppernas medelvärde.

Tabell 21. Reguljära värden och utprövningsvärden för de LÄS-uppgifter i 97A som prövats ut enligt den nya utprövningsmodellen.

Uppgift nr Reg (utpr)	Reguljärt 97A			Utprövning			Diff M reg/utpr	Diff p reg/utpr
	p	m/k	rbis	p	m/k	rbis		
5 (5)	.74	74/74	.46	.74	74/74	.40	+.59	---
6 ⁴ (7)	.68	70/66	.53	.31	35/28	.33	-.20	+.37
7 (6)	.70	70/70	.52	.75	74/75	.51	+.59	-.05
8 (8)	.81	83/79	.55	.80	82/79	.52	-.20	+.01
9 (17)	.81	83/80	.58	.63	63/63	.50	-.06	+.18
10 (20)	.68	71/65	.42	.52	51/52	.45	-1.03	+.16
11 (17)	.74	75/73	.29	.61	61/60	.39	-1.03	+.13
12 ¹ (20)	.51	51/50	.49	.44	44/45	.39	-.06	+.07
13 (14)	.64	63/65	.52	.59	57/61	.49	+.42	+.05
14 ² (15)	.49	50/48	.47	.36	37/35	.40	+.42	+.13
15 (14)	.30	30/30	.57	.24	28/22	.46	+.34	+.06
16 ³ (16)	.57	55/59	.47	.28	28/29	.36	+.34	+.29
17 ⁹ (17)	.51	47/54	.51	.36	34/38	.48	-.54	+.15
18 (19)	.75	72/78	.48	.63	60/66	.46	-.54	+.12
19 (19)	.65	62/67	.47	.56	53/58	.46	-.30	+.09
20 (20)	.59	53/64	.51	.56	48/64	.48	-.30	+.03
	10.17	10.09/10.22		8.38	8.29/8.49			+1.79

Den första texten och tillhörande uppgifter hade prövats ut enligt den gamla modellen, de övriga fyra texterna hade prövats ut enligt den nya modellen. Medelvärdet för de 16 uppgifterna i LÄS 97A som prövats ut enligt den nya modellen var 8.38 vid utprövningen och blev 10.17 i reguljärt. Skillnaden i medelvärde mellan de två tillfällena är 1.79, av denna skillnad står de fem reviderade uppgifterna för 1.01. Den uppgift som svarade för den största förändringen i lösningsproportion är uppgift 6, som har förändrats .37, vilket

också var ett förväntat resultat då samtliga distraktorer var reviderade. Uppgift 15 har också förändrats mycket (.29) efter revidering av tre distraktorer.

Det är också av intresse att närmare studera de uppgifter som inte reviderats eftersom de står för .78 av skillnaden mellan utprovning och reguljärt utfall. I tabell 21 kan man se att uppgifterna i den tredje texten står för .47, dvs. de tre uppgifterna som ej reviderats har blivit .47 lättare vid det reguljära provtillfället då de hade position 9-11 jämfört med vid utprovningen då de hade position 17 och 20 (se sid. 29 angående position). Uppgift 18 har blivit .12 lättare reguljärt trots att den vid båda tillfällena tillhörde den sista texten. Här finns anledning att närmare studera dels utprovningens prestationsnivå och dels omkringliggande texter i utprovningen jämfört med i reguljärprovet. Det kan även vara så att texten var utprovad i ett häfte som var svårt dvs. hade ett lågt medelvärde, vilket kan innebära att provdeltagarna inte riktigt hann med den sista texten vid utprovningen (jfr. Kerstiens, 1990). Den erfarenhet som finns är att tidsfaktorn är av betydelse för LÄS-provet (Henriksson & Eriksson, 1992 och Henriksson & Eriksson, 1994). När konstruktören sätter samman såväl ett reguljärt prov som ett utprovningssprov gäller samma regler för att uppnå parallellitet (se sid 1). Men vid sammansättningen av ett utprovningssprov finns inga statistiska värden att stödja sig på och därför är det nästan omöjligt att kunna avgöra hur svårt eller lätt ett prov kommer att bli även om konstruktörerna har mycket goda kunskaper att prognostisera utfall av uppgifter. Det är en av anledningarna till att ett utprovningsshäfte kan bli svårare än ett annat, vilket kan innebära att provdeltagaren inte hinner med den sista texten och tillhörande uppgifter på samma sätt som de övriga. Detta kan då medföra att lösningsproportionen mellan utprovning och reguljärt provtillfälle ändras. Detta är också ett område som är föremål för pågående studier och kommer att redovisas i Enhetens rapportserie.

Som tidigare nämnts undviker man att ta texter/uppgifter från samma utprovningsshäfte för att undvika att provdeltagare känner igen materialet. Det innebär att texterna går tillsammans reguljärt med helt andra texter än vid utprovningen. Dessutom väljs uppgifterna ur två versioner, vilket innebär att de fyra uppgifterna inte heller har prövats ut tillsammans. Det här är ett område som bör mer ingående studeras.

Korrelationen för lösningsproportionerna mellan utprovning och reguljärt provtillfälle är .79 för de 16 uppgifterna som prövats ut enligt den nya modellen.

Tabell 22. Reguljära värden och utprövningsvärden för LÄS 97A uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.

	LÄS 97A					
	Nya modellen (16)		Gamla modellen (4)		Hela provet	
	M	m/k	M	m/k	M	m/k
Reguljärt	10.17	10.09/10.22	2.53	2.83/2.30	12.70	12.92/12.52
Utprövning	8.38	8.29/8.49	2.34	2.54/2.07	10.72	10.83/10.56
Diff	+1.79		+.19		+1.98	

Skillnaden mellan utprövningsvärden och reguljära värden blev 1.79 för de 16 uppgifterna som prövats ut enligt den nya modellen och .19 för de fyra som utprövats enligt den gamla modellen. Uppgifterna utprovade enligt den nya modellen blev 1.79 lättare reguljärt och uppgifterna utprovade enligt den gamla modellen blev också något lättare (.19). Reguljärprovet 97A blev 1.98 lättare än vad utprövningsvärdena visade. Skillnaden i utprövningsresultat mellan män och kvinnor var .20 till kvinnornas fördel för de 16 uppgifterna som prövats ut enligt den nya modellen. Reguljärt blev skillnaden något mindre, .13 även nu till kvinnornas fördel. För de fyra uppgifterna som prövats ut enligt den gamla modellen var skillnaden i resultat mellan män och kvinnor .47 till männens fördel vid utprövningen och den ökade något, till .53 vid reguljärprovet.

Sammanfattningsvis kan sägas att medelvärdet på provet förändrades 1.98 mellan utprövning och reguljärt provtillfälle trots att 80 procent av provet hade prövats ut enligt den nya modellen. De 16 uppgifter som prövats ut enligt den nya modellen blev 1.79 (i genomsnitt .11 per uppgift) lättare och de fyra som prövats ut enligt den gamla modellen blev .19 (i genomsnitt .05 per uppgift) lättare. Resultaten mellan män och kvinnor var däremot mer stabila. För de 16 uppgifterna som prövats ut enligt den nya modellen var medelvärdet för kvinnorna högre vid såväl utprövning som reguljärt provtillfälle och för de fyra uppgifterna som prövats ut enligt den gamla modellen var männens resultat högre vid de båda tillfällena. Totalt blev provet ändå ”manligt” eftersom skillnaden var större för de uppgifter som prövats ut enligt den gamla modellen än den nya.

LÄS 97B

I det reguljära LÄS-provet hösten 1997 hade samtliga uppgifter prövats ut enligt den nya modellen varav 16 utprovades våren 1996 och fyra hösten 1996. Uppgifterna som prövats ut våren 1996 var fördelade på sex högskoleenheter och de uppgifter som prövats ut på hösten var fördelade på två högskoleenheter. Medelvärdet på reguljärprovet våren 1996 var, som tidigare nämnts, 12.15 för totalgruppen och medelvärdet för samma prov varierade mellan 11.56 och 13.18 för de sex utprövningsgrupperna. Endast en av utprövningsgrupperna hade ett lägre resultat än totalgruppen. De övriga fem hade ett högre resultat på reguljärprovet än totalgruppen. Den största skillnaden mellan

totalgruppen och utprövningsgrupp var 1.03. I genomsnitt räknat borde en sådan uppgift bli .05 svårare vid det reguljära provtillfället förutsatt att ingen annan revidering av betydelse gjorts. Skillnaden mellan totalgruppen och den utprövningsgrupp som hade det lägsta resultatet var .59. I genomsnitt räknat borde en sådan uppgift bli .03 lättare reguljärt.

Medelvärdet på reguljärprovet hösten 1996 var 13.23 för totalgruppen och för utprövningsgrupperna var medelvärdet 12.98 respektive 13.07. Skillnaden mellan totalgrupp och utprövningsgrupp var .25 respektive .16. Båda utprövningsgrupperna hade ett lägre resultat än totalgruppen. I genomsnitt räknat borde uppgifter utprovade i gruppen med medelvärdet 12.98 räknas upp med .012 och uppgifter utprovade i den andra gruppen borde räknas upp med .008.

För de uppgifter som ingick i LÄS 97B hade 14 uppgifter prövats ut i utprövningsgrupper som hade ett bättre resultat än totalgruppen och dessa var utprovade på våren 1996. För ytterligare två uppgifter som prövats ut våren 1996 och ingick i 97B hade utprövningsgrupperna ett sämre resultat än totalgruppen. Detsamma gällde för de fyra uppgifterna som prövats ut hösten 1996. Korrigerar man värdet för att vissa utprövningsgrupper hade ett högre resultat och för att vissa hade ett lägre borde utprövningsvärdet minska med .009 per uppgift.

Tabell 23. Reguljära värden och utprövningsvärden för de LÄS-uppgifter i 97B som prövats ut enligt den nya utprövningsmodellen.

Uppgift nr Reg (utpr)	Reguljärt 97B			Utprövning			Diff M reg/utpr	Diff p reg/utpr
	p	m/k	rbis	p	m/k	rbis		
Vårutprovade:								
1 (4)	.69	68/70	.45	.59	58/60	.52	-1.03	+1.10
2 ² (3)	.71	72/71	.36	.63	65/61	.41	-1.03	+0.08
3 (4)	.78	78/78	.56	.73	74/72	.51	-.06	+0.05
4 ¹ (2)	.72	73/71	.53	.57	58/56	.43	-.06	+0.15
5 (9)	.48	48/48	.43	.46	46/47	.43	-.54	+0.02
6 ¹ (12)	.45	50/40	.62	.47	51/44	.60	-.54	-.02
7 ¹ (9)	.48	50/46	.53	.45	46/44	.51	-.30	+0.03
8 ¹ (10)	.58	58/58	.42	.61	60/61	.42	-.30	-.03
9 ⁸ (18)	.48	47/49	.22	.37	36/38	.28	+0.59	+0.11
10 (20)	.77	75/79	.45	.70	67/72	.40	+0.59	+0.07
11 ¹ (17)	.83	83/83	.51	.80	80/80	.48	+0.59	+0.03
12 (20)	.83	84/81	.64	.69	69/68	.61	-.20	+0.14
17 ¹ (2)	.52	56/48	.62	.64	68/60	.61	-.54	+0.12
18 (3)	.50	51/49	.49	.61	61/60	.48	-.54	-.11
19 (4)	.30	30/29	.38	.58	60/56	.41	-.54	-.28
20 ² (4)	.78	78/78	.52	.83	86/80	.48	-.30	-.05
Höstutprovade:								
13 ⁶ (3)	.76	76/76	.47	.55	56/54	.43	+0.16	+0.21
14 ¹¹ (2)	.71	73/69	.50	.72	72/71	.39	+0.16	-.01
15 ⁵ (2)	.52	55/50	.57	.58	62/56	.53	+0.25	-.06
16 ¹ (1)	.63	63/63	.56	.56	57/55	.49	+0.16	+0.07
	12.52	12.68/12.38		12.14	12.32/11.95			+0.38

Samtliga uppgifter i LÄS 97B var utprovade enligt den nya modellen. Utprövningsmedelvärdet var 12.14 och reguljärt blev medelvärdet 12.52. En skillnad på .38 mellan de två tillfällena, dvs. provet blev .38 lättare reguljärt. Ungefär hälften av denna skillnad står de fyra uppgifterna som prövats ut hösten 1996 för. 13 av uppgifterna hade reviderats efter utprövningen och sammantaget står de för hela skillnaden mellan utprövning och reguljärt provtillfälle. Tittar man däremot på enskilda uppgifter som inte har reviderats finner man att t.ex. uppgift 18 och 19 har blivit .11 respektive .28 svårare reguljärt än vad utprövningsvärdena visade. Båda dessa uppgifter tillhörde en första text vid utprövningen. Uppgifterna 9-12 var utprovade som uppgifter till en sista text (uppgift

17-20) och samtliga fyra uppgifterna blev lättare reguljärt, vilket ytterligare styrker att position vid utprovning respektive reguljärt provtillfälle har betydelse i LÄS-provet.

Korrelationen för lösningsproportionerna mellan utprovning och reguljärt provtillfälle var .69.

Sammanfattningsvis kan sägas att medelvärdet på provet förändrades .38, i genomsnitt .019 per uppgift. Skillnaden i resultat mellan män och kvinnor var stabilt manligt. Trots att revideringar i uppgifterna görs, olika utprovningssgrupper har olika prestationsnivåer, position av texter och uppgifter revideras blev skillnaden inte större än .38 mellan utprovning och reguljärt provtillfälle.

LÄS 98A

I det reguljära provet våren 1998 hade samtliga uppgifter i LÄS-provet prövats ut enligt den nya modellen varav fyra prövades ut våren 1996, åtta hösten 1996 och resterande åtta våren 1997. Uppgifterna 1-8 var utprovade hösten 1996, uppgifterna 9-12 och 17-20 var utprovade våren 1997 och uppgifterna 13-16 var utprovade våren 1996. Uppgifterna som prövats ut våren 1996 var fördelade på två högskoleenheter, de uppgifter som prövats ut på hösten 1996 var fördelade på fyra högskoleenheter och de uppgifter som prövats ut våren 1997 var även fördelade på fyra högskoleenheter. Medelvärdet på reguljärprovet våren 1996 var, som tidigare nämnts, 12.15 för totalgruppen och medelvärdet för samma prov var 11.56 respektive 12.35 för de två utprovningssgrupperna. Detta innebär att den ena gruppen hade .59 lägre resultat än totalgruppen och den andra gruppen hade .20 högre resultat än totalgruppen. I genomsnitt räknat borde uppgifter utprovade i dessa grupper bli .03 lättare respektive .01 svårare reguljärt.

Medelvärdet på reguljärprovet hösten 1996 var, som även tidigare nämnts, 13.23 för totalgruppen och medelvärdet varierade mellan 12.91 och 13.07 för utprovningssgrupperna, vilket innebär att alla fyra utprovningssgrupperna hade ett lägre resultat än totalgruppen. Den största skillnaden mellan utprovningssgrupp och totalgrupp var .32. I genomsnitt borde en sådan uppgift bli .016 lättare reguljärt.

Medelvärdet på reguljärprovet våren 1997 var 12.70 för totalgruppen och medelvärdet varierade mellan 12.36 och 13.55 för de fyra utprovningssgrupperna. En av utprovningssgrupperna hade ett högre resultat än totalgruppen och det var den grupp som hade den största skillnaden mellan utprovningssgrupp och totalgrupp. Skillnaden var .85, i genomsnitt räknat borde en sådan uppgift bli .04 svårare vid det reguljära provtillfället. Skillnaden mellan totalgrupp och den utprovningssgrupp som hade det lägsta resultatet var .34. I genomsnitt räknat borde en sådan uppgift bli .017 lättare reguljärt.

För de uppgifter som ingick i LÄS 98A hade tre uppgifter prövats ut i utprovningssgrupper som hade ett högre resultat än totalgruppen och två av dessa var utprovade på våren 1996 och en våren 1997. Resterande 17 uppgifter hade prövats ut i utprovningssgrupper som hade sämre resultat än totalgruppen. Korregerar man värdet för att vissa

utprövningsgrupper hade ett högre resultat och för att vissa hade ett lägre resultat borde utprövningsvärdet öka med .008 per uppgift.

Tabell 24. Regljära värden och utprövningsvärden för de LÄS-uppgifter i 98A som prövats ut enligt den nya utprövningsmodellen.

Uppgift nr Reg (utpr)	Regljärt 98A			Utprövning			Diff M reg/utpr	Diff p reg/utpr
	p	m/k	rbis	p	m/k	rbis		
Vårutprovade:								
9 (8)	.73	72/73	.52	.74	73/75	.46	+08	-.01
10 ¹ (6)	.70	72/68	.65	.61	64/58	.61	+08	+09
11 ⁸ (6)	.27	32/24	.46	.44	48/42	.46	+28	-.17
12 ⁸ (5)	.58	58/59	.50	.46	47/46	.53	+08	+12
13 (13)	.59	62/56	.54	.48	51/46	.49	+59	+11
14 (16)	.61	66/58	.52	.45	48/42	.51	-.20	+16
15 (15)	.75	80/71	.63	.66	70/62	.61	-.20	+09
16 (16)	.64	64/64	.50	.58	57/58	.49	+59	+06
17 ¹ (17)	.58	54/62	.44	.49	45/53	.37	+34	+09
18 (19)	.37	37/37	.46	.34	34/34	.46	+34	+03
19 ¹ (18)	.34	35/33	.43	.33	34/32	.53	-.85	+01
20 (20)	.53	53/53	.42	.54	54/54	.41	+34	-.01
Höstutprovade:								
1 ² (1)	.66	66/67	.45	.67	65/68	.38	+24	-.01
2 (3)	.81	83/79	.63	.75	77/73	.52	+32	+06
3 ¹ (2)	.64	65/64	.49	.54	55/53	.39	+24	+10
4 ⁸ (4)	.61	60/61	.51	.61	60/61	.45	+24	----
5 ¹⁰ (7)	.63	64/62	.46	.42	45/41	.36	+24	+21
6 (6)	.52	57/48	.61	.59	67/53	.53	+24	-.07
7 ¹ (8)	.49	50/48	.30	.58	59/58	.40	+24	-.09
8 ⁸ (6)	.41	41/42	.37	.41	40/42	.30	+16	----
	11.46	11.71/11.29		10.69	10.93/10.51			+77

Samtliga uppgifter i LÄS 98A var utprovade enligt den nya modellen. Utprövningsmedelvärdet var 10.69 och regljärt blev medelvärdet 11.46. En skillnad på .77 mellan de två tillfällena, dvs. provet blev .77 lättare regljärt. Något mindre än hälften av denna skillnad står de 11 reviderade uppgifterna för. För fem av uppgifterna har en distraktor reviderats, för en annan uppgift har två distraktorer reviderats, för fyra uppgifter har frågestammen reviderats och för ytterligare en uppgift har två distaktorer och frågestammen reviderats. Om man jämför de fyra uppgifterna för vilka frågestammen har

reviderats finner man att för uppgifterna 4 och 8 har ingen förändring i lösningsproportion skett mellan utprövning och reguljärt provtillfälle medan för uppgift 11 och 12 har däremot lösningsproportionen ändrats så att uppgift 11 blev .17 svårare och uppgift 12 blev .12 lättare. Här kan man tydligt se att samma typ av revidering kan ge väldigt olika utfall. Det finns naturligtvis grad av revideringar inom varje typ, men även en mindre förändring kan i vissa fall påverka utfallet ganska mycket. Som redan nämnts står de reviderade uppgifterna för nästan hälften av skillnaden i medelvärdet mellan utprövning och reguljärt prov. Studerar man tabell 24 finner man att uppgifterna 13-16 har förändrats sammanlagt .42 fastän inga revideringar gjorts. Dessa uppgifter prövades ut våren 1996 i två olika utprövningsgrupper, vilket innebär att uppgifterna inte gått tillsammans vid utprövningen. Positionen för texten/uppgifterna var densamma vid det reguljära provtillfället som vid utprövningen. En rimlig förklaring till att dessa uppgifter förändrats så pass mycket kan vara att de texter som ingick i LÄS 98A inte var utprövade tillsammans och att sammansättningen av ett provhäfte har en stor betydelse. Detta är ett område som bör studeras utförligt.

Korrelationen för lösningsproportionerna mellan utprövning och reguljärt provtillfälle är .78 för samtliga uppgifter i LÄS 98A. Korrelationen för lösningsproportionerna mellan utprövning och reguljärt provtillfälle är .83 för de 12 uppgifter som prövats ut våren 1996 och våren 1997 och .67 för de 8 uppgifter som prövades ut hösten 1996.

Sammanfattningsvis kan sägas att samtliga uppgifter i LÄS 98A var utprövade enligt den nya modellen. Utprövningsmedelvärdet var 10.69 med en könsskillnad på .42 och reguljärt blev medelvärdet 11.46 med en könsskillnad på .42. Provet blev .77 (i genomsnitt .038 per uppgift) lättare reguljärt än vad utprövningsvärdena visade och skillnaden i resultat mellan män och kvinnor var .42 till männens fördel såväl vid utprövningen som reguljärprovet. Ganska kort tid efter det att den nya utprövningsmodellen tagits i bruk och att uppgifter börjat användas i LÄS-provet fann konstruktören att position var av betydelse, vilket har resulterat i att endast en text och tillhörande uppgifter i detta prov hade annan position än vid utprövningen.

En jämförelse av ej reviderade och reviderade uppgifter i LÄS 96B-98A

I tabell 25 görs en sammanställning av de uppgifter som prövats ut enligt den nya modellen och ingick i något av LÄS-proven 96B, 97A, 97B eller 98A. Materialet uppdelat på ej reviderade och reviderade uppgifter.

Tabell 25. Medelvärde för ej reviderade och reviderade uppgifter i 96B, 97A, 97B, och 98A. Inom parentes anges antal uppgifter.

	LÄS 96B		LÄS 97A		LÄS 97B		LÄS 98A	
	Ej rev.	Rev.	Ej rev.	Rev.	Ej rev.	Rev.	Ej rev.	Rev.
	M (1)	M (3)	M (11)	M (5)	M (7)	M (13)	M (9)	M (11)
Reguljärt	.64	1.86	7.41	2.76	4.35	8.17	5.55	5.91
Utprovning	.76	1.95	6.63	1.75	4.36	7.78	5.13	5.56
Diff	-.12	-.09	+.78	+1.01	-.01	+.39	+.42	+.35

I LÄS 96B ingick endast fyra uppgifter som prövats ut enligt den nya modellen varav tre var reviderade. Här kan man se att den uppgift som ej var reviderad blev .12 svårare reguljärt än vid utprovningen och de tre som var reviderade blev i genomsnitt .03 svårare.

I LÄS 97A ingick 16 uppgifter som prövats ut enligt den nya modellen varav fem hade reviderats. I genomsnitt blev de ej reviderade uppgifterna .07 lättare reguljärt jämfört med vid utprovningen och de reviderade uppgifterna blev i genomsnitt .20 lättare.

I LÄS 97B var samtliga uppgifter utprovade enligt den nya modellen. 13 uppgifter var reviderade och de blev i genomsnitt .03 lättare reguljärt än vid utprovningen. Medelvärdet för de sju ej reviderade uppgifterna förändrades endast .01.

I LÄS 98A var samtliga uppgifter utprovade enligt den nya modellen. Nio uppgifter var ej reviderade och de blev i genomsnitt .05 lättare reguljärt än vid utprovningen och de 11 uppgifter som var reviderade blev i genomsnitt .03 lättare.

Korrelationen för lösningsproportionerna mellan utprovning och reguljärt provtillfälle är 0.69 för samtliga uppgifter som prövats ut enligt den nya utprovningensmodellen och ingår i något av LÄS-proven 96B, 97A, 97B och 98A.

DELPROVET DTK (tolkning av diagram, tabeller och kartor)

Delprovet DTK avser att mäta förmågan att tolka och hämta information ur diagram, kartor, tabeller och andra grafiska framställningar. Provet ska täcka de fem ämnesområdena som gäller för högskoleprovet som helhet. Materialet hämtas från såväl vetenskapliga och populärvetenskapliga publikationer som tidsskrifter och informationsmaterial. Provet består av 10 figuruppsättningar och 20 uppgifter, två till varje figuruppsättning. Till varje figuruppsättning prövas fyra uppgifter ut, detta innebär att samma figuruppsättning prövas ut i två olika provversioner. Anledningen till att man måste pröva ut fler än två uppgifter är de statistiska krav på uppgifterna som inte alltid uppfylls och att när man sätter samman provet måste det finnas ett varierat utbud av uppgifter med avseende på lösningsfrekvens, lösningsstrategier, figurtyp mm.

DTK 96B

I DTK-provet hösten 1996 ingick endast två uppgifter som prövats ut enligt den nya modellen. Uppgifterna hade prövats ut våren 1996 i var sin högskoleenhet. Medelvärdet på reguljärprovet våren 1996 var 11.68 för totalgruppen och medelvärdet för samma prov var för de två utprövningsgrupperna 11.26 respektive 11.35. Detta innebär att båda utprövningsgrupperna hade ett lägre resultat än totalgruppen. Skillnaden mellan totalgruppen och utprövningsgrupperna var .42 respektive .33. I genomsnitt räknat borde uppgifterna bli .02 respektive .016 svårare reguljärt. Korrigerar man värdet för att utprövningsgrupperna var något sämre än totalgruppen borde utprövningsvärdet för de två uppgifterna räknas upp med .036 tillsammans. På samma sätt som för de övriga delproven gör konstruktören en skattning av de revideringar som gjorts, vilket för närvarande lämnas därhän.

Tabell 26. Reguljära värden och utprövningsvärden för de DTK-uppgifter i 96B som prövats ut enligt den nya utprövningsmodellen.

Uppgift nr Reg (utpr)	Reguljärt 96B			Utprövning			Diff M reg/utpr	Diff p reg/utpr
	p	m/k	rbis	p	m/k	rbis		
15 ⁸ (12)	.73	80/68	.57	.68	76/61	.63	+.33	+.05
16 ¹ (12)	.23	26/21	.34	.28	32/24	.41	+.42	-.05
	.96	1.06/.89		.96	1.08/.85			-----

Bägge uppgifterna hade reviderats efter utprövningen. I uppgift 15 hade frågestammen förtydligats, vilket gav det förväntade resultatet att uppgiften blev något lättare. I uppgift 16 var en distraktor ny, vilket medförde att rätt svar blev placerat som svarsförslag C vid reguljärprovet från att ha varit B vid utprövningen. Dessutom hade svarsförslagen A och

E bytt plats, dvs. A blev E och E blev A. Uppgiften blev .05 svårare reguljärt. Vid en tidigare studie framkom det att ändringar i placering av svarsförslagen kan medföra att uppgiftens svarsmönster delvis bibehålls eftersom de som inte har kunskaper att lösa uppgiften väljer de första svarsförslagen (Ögren, 1996). Detta är ytterligare en viktig faktor som man bör skaffa sig mera kunskap om och därför har en studie påbörjats enbart i syfte att studera huruvida förändring i placering av svarsförslagen påverkar lösningsfrekvensen.

Medelvärdet för de båda uppgifterna tillsammans blev .96 reguljärt som vid utprovningen. Bägge uppgifterna hade förändrats i lösningsproportion, dock lika mycket, men den ena uppgiften blev .05 lättare och den andra blev .05 svårare. Skillnaden mellan män och kvinnor hade minskat något, från .23 vid utprovningen till .17 reguljärt.

Tabell 27. Reguljära värden och utprovningvärden för DTK 96B uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal uppgifter.

	DTK 96B					
	Nya modellen (2)		Gamla modellen (18)		Hela provet	
	M	m/k	M	m/k	M	m/k
Reguljärt	.96	1.06/89	10.96	12.16/10.02	11.92	13.22/10.91
Utprovning	.96	1.08/85	9.93	10.37/9.58	10.89	11.45/10.43
Diff	.00		+1.03		+1.03	

Det blev ingen skillnad mellan utprovningvärden och reguljära värden för de två uppgifterna som prövats ut enligt den nya modellen. Skillnaden för de 18 uppgifter som prövats ut enligt den gamla modellen blev däremot 1.03. Reguljärt blev provet alltså 1.03 lättare än vad utprovningvärdena visade.

Skillnaden i utprovningresultat mellan män och kvinnor var .23 till männens fördel, för de två uppgifterna som utprovats enligt den nya modellen. Reguljärt blev motsvarande skillnad något mindre (.17). För de 18 uppgifter som prövats ut enligt den gamla modellen var skillnaden i resultat mellan män och kvinnor .79 till männens fördel. Reguljärt ökade denna skillnad till 2.14 till männens fördel.

Sammanfattningsvis kan sägas att medelvärdet på provet förändrades totalt 1.03 (i genomsnitt .06 per uppgift), vilket kan hänföras till de 18 uppgifter som prövats ut enligt den gamla modellen. Endast två uppgifter hade prövats ut enligt den nya modellen och värdet på dessa två uppgifter tillsammans blev samma reguljärt som vid utprovningen. Ett mycket bra resultat, som visserligen endast gäller två uppgifter, men som är värt att uppmärksamma. Skillnaden i utprovningresultat mellan män och kvinnor var 1.02 för samtliga 20 uppgifter. De två uppgifter som prövats ut enligt den nya modellen stod för .23 av skillnaden. Reguljärt blev skillnaden för samtliga uppgifter 2.31, varav de två uppgifterna som prövats ut enligt den nya modellen stod för endast .17. Detta är ett

vanligt förekommande resultat, att uppgifter som prövats ut enligt den gamla modellen förändras mycket i lösningsproportion mellan utprovning och reguljärt provtillfälle vad gäller skillnaden i resultat mellan män och kvinnor (Stage et al. 1996, 1997).

DTK 97A

Även i DTK-provet våren 1997 ingick endast två uppgifter som prövats ut enligt den nya modellen. Uppgifterna hade, som för 96B, prövats ut våren 1996 och i var sin högskolenhet. Medelvärdet på reguljärprovet våren 1996 var, som tidigare nämnts, 11.68 för totalgruppen och medelvärdet för samma prov var för de två utprövningsgrupperna 11.07 och 11.77. En utprövningsgrupp hade ett lägre resultat än totalgruppen och en hade ett högre resultat. Skillnaden mellan totalgruppen och utprövningsgrupperna var .61 respektive .09. I genomsnitt räknat borde uppgifterna bli .03 lättare respektive .004 svårare. Korrigerar man värdet för utprövningsgrupperna borde utprövningsvärdet för de två uppgifterna räknas upp med .026.

Tabell 28. Reguljära värden och utprövningsvärden för de DTK-uppgifter i 97A som prövats ut enligt den nya utprövningsmodellen.

Uppgift nr Reg (utpr)	Reguljärt 97A			Utprovning			Diff M reg/utpr	Diff p reg/utpr
	p	m/k	rbis	p	m/k	rbis		
5 ¹ (9)	.19	23/16	.49	.24	26/22	.47	-.09	-.05
6 (10)	.66	68/64	.47	.59	60/58	.46	+.61	+.07
	.85	91/80		.83	86/80			+.02

Uppgift 5 hade reviderats i form av en ny distraktor och ny ordning på svarsförslagen. Den revidering som genomfördes blev inte i önskad riktning, dvs. avsikten var att uppgiften skulle bli lättare men den blev istället .05 svårare. Här kan liksom i 96B förändringen i placeringen av svarsförslagen också haft viss betydelse. Uppgift 6 hade inte reviderats alls, trots det blev den .07 lättare reguljärt. En anledning kan vara att utprövningsgruppen hade .61 lägre resultat än totalgruppen och skulle man korrigera för gruppen borde det reguljära värdet öka med .03.

Medelvärdet på de två uppgifterna tillsammans var .83 vid utprovningen och blev .85 reguljärt. Skillnaden mellan utprovning och reguljärt provtillfälle blev endast .02 för de två uppgifterna. Skillnaden mellan män och kvinnor hade ökat något, från .06 vid utprovningen till .11 reguljärt. Vid båda tillfällena var det till männens fördel.

Tabell 29. Reguljära värden och utprövningsvärden för DTK 97A uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Antal uppgifter anges inom parentes.

	DTK 97A					
	Nya modellen (2)		Gamla modellen (18)		Hela provet	
	M	m/k	M	m/k	M	m/k
Reguljärt	.85	91/80	9.45	10.65/8.45	10.30	11.56/9.25
Utprövning	.83	86/80	10.07	10.55/9.68	10.90	11.41/10.48
Diff	+.02		-.62		-.60	

Det blev en mycket liten skillnad mellan utprövningsvärde och reguljärt värde för de två uppgifterna som prövats ut enligt den nya modellen. Skillnaden för de 18 uppgifter som prövats ut enligt den gamla modellen blev .62 svårare reguljärt än vid utprövningen. Totalt blev provet .60 svårare reguljärt, vilket är ett mycket ovanligt resultat och har diskuterats i ett paper angående normeringen för reguljärprovet 97A (Stage & Hamrén, 1997).

Skillnaden i utprövningsresultat mellan män och kvinnor var .06 till männens fördel för de två uppgifter som prövats ut enligt den nya modellen. Reguljärt blev motsvarande skillnad något större (.11). För de 18 uppgifter som prövats ut enligt den gamla modellen var skillnaden i resultat mellan män och kvinnor .87 till männens fördel. Reguljärt ökade denna skillnad till 2.31 till männens fördel.

Sammanfattningsvis kan sägas att medelvärdet på provet förändrades .60, dvs. provet blev .60 svårare reguljärt än vid utprövningen. Endast två uppgifter hade prövats ut enligt den nya modellen och värdet på dessa uppgifter förändrades endast .02 mellan utprövning och reguljärt provtillfälle. I genomsnitt förändrades uppgifterna som utprövats enligt den nya modellen .01 per uppgift och uppgifterna som utprövats enligt den gamla modellen förändrades förändrades i genomsnitt .03 per uppgift. Skillnaden i utprövningsresultat mellan män och kvinnor var .93 för samtliga uppgifterna och de två uppgifterna som prövats ut enligt den nya modellen svarade för .06. Reguljärt blev skillnaden för samtliga uppgifter 2.31, varav de två uppgifter som prövats ut enligt den nya modellen svarade för .11. Både vid utprövningen och reguljärprovet hade männen högre resultat än kvinnorna.

DTK 97B

I det reguljära provet hösten 1997 ingick fyra uppgifter som prövats ut enligt den nya modellen. Uppgift 8, 19 och 20 hade prövats ut våren 1996 fördelade på tre olika högskoleenheter och uppgift 10 hade prövats ut våren 1997. Medelvärdet på reguljärprovet våren 1996 var, som tidigare nämnts, 11.68 för totalgruppen medan medelvärdet på samma prov varierade mellan 11.07 och 11.77 för de tre utprövningsgrupperna. Detta innebär att två utprövningsgrupper hade ett lägre resultat än

totalgruppen och en utprövningsgrupp hade ett högre resultat. Den största skillnaden mellan totalgrupp och utprövningsgrupp var .61, i genomsnitt räknat borde en sådan uppgift bli .03 lättare vid det reguljära provtillfället förutsatt att ingen annan revidering har gjorts. Medelvärde på reguljärprovet våren 1997 var 10.30 för totalgruppen och för utprövningsgruppen var medelvärdet för samma prov 9.79, vilket innebär att den uppgiften borde bli .025 lättare reguljärt. Eftersom tre utprövningsgrupper hade ett lägre resultat än totalgruppen och en utprövningsgrupp hade ett högre resultat skulle en uppräknning av utprövningsvärdet kunna göras med .017 per uppgift, vilket är ett genomsnittsvärde av skillnaden mellan totalgruppens och utprövningsgruppernas medelvärde.

Tabell 30. Reguljära värden och utprövningsvärden för de DTK-uppgifter i 97B som prövats ut enligt den nya utprövningsmodellen.

Uppgift nr Reg (utpr)	Reguljärt 97B			Utprövning			Diff M reg/utpr	Diff p reg/utpr
	p	m/k	rbis	p	m/k	rbis		
8 ² (3)	.71	72/71	.33	.48	55/42	.39	+.33	+.33
10 (18)	.40	42/38	.36	.38	42/36	.34	+.51	+.02
19 ¹² (5)	.38	47/32	.64	.54	60/49	.63	-.09	-.16
20 ¹ (6)	.29	32/27	.43	.25	29/21	.41	+.61	+.04
	1.78	1.93/1.68		1.65	1.86/1.48			+.13

Tre av de fyra uppgifterna som prövats ut enligt den nya utprövningsmodellen hade reviderats efter utprövningen. I uppgift 8 hade två svarsförslag reviderats p.g.a. att båda kunde vara lika rätt. Svaret innehöll siffror och det rätta svaret låg mellan dessa två svarsförslag. Det är anledningen till att uppgiften blev så mycket lättare reguljärt, vilket också var ett förväntat resultat. I uppgift 19 hade figurtexten förenklats. Det borde ha medfört att uppgiften blivit lättare men den blev istället .16 svårare. En trolig orsak är att uppgiften var placerad tidigt, som nr 5, i utprövningsversionen medan den i reguljärprovet var som näst sista uppgift, dvs. uppgift 19. Liksom för LÄS-provet finns en sk. speedfaktor i DTK-provet och även för detta delprov pågår studier på området.

Medelvärdet för de fyra uppgifterna var 1.65 vid utprövningen och blev 1.78 reguljärt. Skillnaden mellan de två tillfällena blev endast .13 fastän en uppgift hade blivit så mycket lättare reguljärt jämfört med vid utprövningen. Skillnaden i resultat mellan män och kvinnor hade minskat från .38 vid utprövningen till .25 vid reguljärprovet. Vid båda tillfällena var det till männens fördel.

Tabell 31. Reguljära värden och utprövningsvärden för DTK 97B uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Antal uppgifter anges inom parentes.

	DTK 97B				Hela provet	
	Nya modellen (4)		Gamla modellen (16)		M	m/k
	M	m/k	M	m/k		
Reguljärt	1.78	1.93/1.68	9.54	10.54/8.67	11.32	12.47/10.35
Utprövning	1.65	1.86/1.48	9.08	9.74/8.54	10.73	11.60/10.02
Diff	+.13		+.46		+.59	

Skillnaden mellan utprövnings- och reguljärvärden blev endast .13 för de fyra uppgifterna som prövats ut enligt den nya modellen. För de 16 uppgifter som prövats ut enligt den gamla modellen var motsvarande skillnad .46. Även här får man säga en liten skillnad, vilket innebär att skillnaden reguljärt också blev ganska liten. Provet blev reguljärt .59 lättare än vad utprövningsmedelvärdet visade.

Skillnaden i utprövningsresultat mellan män och kvinnor var .38 till männens fördel för de uppgifter som prövats ut enligt den nya modellen. Reguljärt minskade motsvarande skillnad till .25. För de 16 uppgifter som prövats ut enligt den gamla modellen var skillnaden i resultat mellan män och kvinnor 1.20 till männens fördel vid utprövningen. Reguljärt blev motsvarande skillnad 1.87 till männens fördel.

Sammanfattningsvis kan man säga att medelvärdet på provet förändrades totalt .59, dvs. provet blev .59 lättare reguljärt än vid utprövningen. Skillnaden i medelvärde mellan utprövning och reguljärt provtillfälle var .13 (i genomsnitt .03 per uppgift) för de fyra uppgifterna som prövats ut enligt den nya modellen och för de 16 uppgifter som prövats ut enligt den gamla modellen var skillnaden .46 (i genomsnitt .03 per uppgift). Resultatet mellan män och kvinnor var till männens fördel vid såväl utprövning som reguljärt provtillfälle oavsett utprövningsmodell. Könsskillnaden blev något större per uppgift mellan utprövning och reguljärt provtillfälle för de uppgifter som prövats ut enligt den gamla modellen.

DTK 98A

I det reguljära provet våren 1998 ingick sju uppgifter i DTK-provet som prövats ut enligt den nya modellen. Fem uppgifter hade prövats ut våren 1996 fördelade på tre olika högskoleenheter och två hade prövats ut hösten 1997. De uppgifter som prövats ut våren 1996 hade prövats ut i samma utprövningsgrupper som redovisats för 97B och även till viss del för 96B och 97A. Medelvärdet på reguljärprovet våren 1996 var alltså 11.68 för totalgruppen medan medelvärdet på samma prov varierade mellan 11.07 och 11.77 för de olika utprövningsgrupperna. För de två uppgifter som prövats ut hösten 1997 gäller att medelvärdet på reguljärprovet hösten 1997 var 11.92 för totalgruppen och för

utprövningsgrupperna var medelvärdet för samma prov 11.75 respektive 12.44, vilket innebär att den ena uppgiften borde bli .008 lättare reguljärt och den andra .03 svårare reguljärt. Eftersom tre utprövningsgrupper hade ett lägre resultat än totalgruppen och fyra utprövningsgrupper hade ett högre resultat skulle en uppräknig av utprövningsvärdet kunna göras med .007 per uppgift, vilket är ett genomsnittsvärde av skillnaden mellan totalgruppens och utprövningsgruppernas medelvärde.

Tabell 32. Reguljära värden och utprövningsvärden för de DTK-uppgifter i 98A som prövats ut enligt den nya utprövningsmodellen.

Uppgift nr Reg (utpr)	Reguljärt 98A			Utprövning			Diff M reg/utpr	Diff p reg/utpr	
	p	m/k	rbis	p	m/k	rbis			
Vårutprovade:									
11 (1)	.66	68/65	.49	.65	66/65	.35	+.61	+.01	
12 (2)	.34	35/44	.44	.28	30/26	.46	-.09	+.06	
13 (15)	.68	75/62	.61	.59	68/51	.55	+.61	+.09	
14 ⁸ (16)	.52	60/45	.53	.40	49/32	.45	-.09	+.12	
16 ⁸ (16)	.56	61/51	.49	.51	57/45	.47	+.33	+.05	
Höstutprovade:									
5 ⁸ (15)	.82	84/81	.50	.72	76/70	.46	+.17	+.10	
6 ⁸ (16)	.83	87/80	.66	.49	54/45	.41	-.52	+.34	
		4.41	4.70/4.28			3.64	4.00/3.34		
								+.77	

Fyra av de sju uppgifterna som prövats ut enligt den nya utprövningsmodellen hade reviderats efter utprövningen. Revideringarna innebar för dessa fyra uppgifter att frågestammen reviderats. För uppgift 6 har en större revidering i frågestammen gjorts medan för uppgifterna 5, 14 och 16 har endast en mindre justering eller ett tillägg gjorts. Uppgift 16 har förändrats mycket i lösningsproportion mellan utprövning och reguljärt provtillfälle. Visserligen hade frågestammen reviderats ganska mycket men denna kan knappast förklara den stora skillnaden. En tänkbar förklaring är att uppgifterna 5 och 6 har prövats ut i olika utprövningsversioner och vid närmare genomgång av uppgifterna finner man att de båda uppgifterna är likartade. Bägge uppgifterna är avläsningsuppgifter med överslagsberäkning och i grunden skall samma information avläsas och jämföras innan provdeltagaren kan lösa respektive uppgift. Dessutom innehåller svarsförslagen samma typ av information. Tidigare studier har visat att det är fördel vid lösandet av en uppgift, om uppgiften innan efterfrågar samma typ av information (Ögren, 1996).

Medelvärdet för de sju uppgifterna var 3.64 vid utprövningen och blev 4.41 reguljärt. Skillnaden i resultat mellan män och kvinnor hade minskat från .66 till .42 till männens fördel.

Tabell 33. Regulerade värden och utprövningsvärden för DTK 98A uppdelade på uppgifter som prövats ut enligt den nya respektive den gamla modellen. Inom parentes anges antal.

	DTK 98A		Gamla modellen (13)		Hela provet	
	Nya modellen (7)		M	m/k	M	m/k
Reguljärt	4.41	4.70/4.28	7.35	8.22/6.52	11.76	12.92/10.80
Utprövning	3.64	4.00/3.34	6.77	7.28/6.40	10.41	11.28/9.74
Diff	+.77		+.58		+1.35	

Skillnaden mellan utprövnings- och reguljärvärden blev .77 för de sju uppgifter som prövats ut enligt den nya modellen. En ganska stor skillnad eftersom inga större revideringar i uppgifterna hade gjorts förutom i uppgift 6. För de 13 uppgifter som prövats ut enligt den gamla modellen var motsvarande skillnad .58. Totalt blev provet 1.35 lättare reguljärt än vad utprövningsvärdena visade.

Skillnaden i utprövningsresultat mellan män och kvinnor var .66 till männens fördel för de uppgifter som prövats ut enligt den nya modellen. Reguljärt minskade motsvarande skillnad till .42. För de 13 uppgifter som prövats ut enligt den gamla modellen var skillnaden i resultat mellan män och kvinnor .88 till männens fördel vid utprövningen. Reguljärt blev motsvarande skillnad 1.70 till männens fördel.

Sammanfattningsvis kan man säga att medelvärdet på provet förändrades totalt 1.35, dvs. provet blev 1.35 lättare reguljärt än vid utprövningen. De sju uppgifterna som prövats ut enligt den nya modellen blev .77 (i genomsnitt .11 per uppgift) lättare reguljärt än vid utprövningen. De övriga 13 uppgifterna som prövats ut enligt den gamla modellen blev .58 (i genomsnitt .04 per uppgift) lättare. Könsskillnaden som var manlig vid såväl utprövning som reguljärt tillfälle för de båda modellerna hade minskat .24 för de uppgifter som hade prövats ut enligt den nya modellen medan den hade ökat .82 för de uppgifter som prövats ut enligt den gamla modellen.

En jämförelse av ej reviderade och reviderade uppgifter i DTK-provet 96B-98A

Totalt har hittills 15 DTK-uppgifter som prövats ut enligt den nya modellen ingått i något av reguljärproven 96B, 97A, 97B eller 98A. I tabell 34 görs en sammanställning av dessa uppgifter, uppdelat på provtillfälle och huruvida de är reviderade eller ej.

Tabell 34. Medelvärde för ej reviderade och reviderade uppgifter i 96B, 97A, 97B, och 98A. Inom parentes anges antal uppgifter.

	DTK 96B		DTK 97A		DTK 97B		DTK 98A	
	Ej rev.	Rev.	Ej rev.	Rev.	Ej rev.	Rev.	Ej rev.	Rev.
	M	M (2)	M (1)	M (1)	M (1)	M (3)	M (3)	M (4)
Reguljärt	---	.96	.66	.19	.40	1.38	1.68	2.73
Utprovning	---	.96	.59	.24	.38	1.27	1.52	2.12
Diff		.00	+.07	-.05	+.02	+.11	+.16	+.61

I DTK 96B ingick endast två uppgifter som prövats ut enligt den nya modellen. Båda dessa uppgifter var reviderade och förändrades lika mycket. Den ena uppgiften blev .05 lättare och den andra blev .05 svårare.

I DTK 97A ingick även endast två uppgifter som prövats ut enligt den nya modellen varav en hade reviderats. Den ej reviderade uppgiften blev .07 lättare reguljärt än vid utprovningen och den reviderade blev .05 svårare reguljärt.

Även i DTK 97B ingick få uppgifter (4 st) som prövats ut enligt den nya modellen. Tre av de fyra uppgifterna var reviderade. Den uppgift som ej var reviderad blev .02 lättare reguljärt än vid utprovningen och de tre som var reviderade blev i genomsnitt .07 lättare reguljärt.

I DTK 98A var sju uppgifter utprovade enligt den nya modellen. Tre uppgifter var ej reviderade och de blev i genomsnitt .05 lättare reguljärt än vid utprovningen och de fyra som var reviderade blev i genomsnitt .15 lättare.

Korrelationen för lösningsproportionerna mellan utprovning och reguljärt provtillfälle är 0.84 för samtliga uppgifter som prövats ut enligt den nya utprovningsmodellen och ingår i något av DTK-proven 96B, 97A, 97B och 98A.

AVSLUTANDE REFLEKTIONER

Det som framkommit, vid genomgång och analys av de uppgifter som hittills ingått i något av reguljärproven 96B, 97A, 97B eller 98A och prövats ut enligt den nya utvärderingsmodellen, är att den nya utvärderingsmodellen är klart överlägsen den gamla. Differensen mellan utvärdering och reguljärt provtillfälle blir mindre totalt och det blir även mindre förändringar av könsdifferensen. Dessutom möjliggör den nya modellen användning av den nya testteorin IRT (se sid 4), vilket måste ses som ett stort framsteg. Hittills har högskoleprovsdata enbart analyserats utifrån klassisk testteori.

Omstruktureringen av provet har medfört att det har uppkommit nya faktorer som kan påverka utfallet t.ex. delprovets placering under provdagen som nu varierar från provtillfälle till provtillfälle, i vilken högskoleenhet som uppgiften prövats ut, placering av uppgifter i delprovet. En prognos av uppgifternas utfall ställer därför stora krav på kunskap om hur dessa nya faktorer påverkar utfallet.

Tabell 2 i inledningen visar hur provdagen varierat sedan den nya utvärderingsmodellen togs i bruk våren 1996. Det har inneburit att t.ex. NOG ingick som block 1 vid provtillfället 96A medan det ingick som block 5 vid provtillfället 97A. Här väcks frågor som; Vilken betydelse har det att provblockens placering varierar mellan provtillfällena? Presterar provdeltagaren bättre på uppgifter som ingår i första provblocket än i det sista? Vilken betydelse har det att en provdeltagare får göra två versioner av ett delprov i rad eller överhuvudtaget får göra två versioner av ett delprov? Vilken betydelse har utvärderingsblockets placering gentemot det reguljära? I 96A ingick NOG i block 1 och för vissa provdeltagare även i block 2, utvärderingsblocket. I block 3 ingick ELF och ORD reguljärt, vilket innebär att vissa provdeltagare hade redan i block 2, utvärderingsblocket, gjort ett ELF- och ett ORD-prov. I tabell 35 redovisas medelvärden för de olika delproven och provtillfällena 96A-98A.

Tabell 35. Resultat på högskoleprovet 96A, 96B, 97A, 97B och 98A för hela gruppen provdeltagare. Provets placering under provdagen anges inom parentes.

	96A	96B	97A	97B	98A
ORD	25.41 (B3)	23.82 (B5)	23.16 (B1)	22.13 (B5)	23.05 (B4)
NOG	13.83 (B1)	13.51 (B2)	12.89 (B5)	13.08 (B2)	12.33 (B3)
LÄS	12.15 (B5)	13.23 (B3)	12.70 (B4)	12.52 (B4)	11.46 (B5)
DTK	11.68 (B4)	11.92 (B4)	10.30 (B2)	11.32 (B3)	11.76 (B1)
ELF	13.14 (B3)	12.88 (B5)	11.54 (B1)	13.58 (B5)	13.63 (B4)
Totalt	76.21	75.34	70.59	72.65	72.23

Delprovet NOG ingick i block 1 vid provtillfället 96A och i block 5 vid provtillfället 97A. Endast 1 uppgift i 97A var utprövad enligt den nya modellen, dvs. sammansättningen och prognos av utfallet på båda proven är gjorda utifrån samma betingelser. Medelvärde på delprovet NOG var 13.83 vid provtillfället 96A och 12.70 vid provtillfället 97A. Huruvida denna skillnad kan hänföras till provets placering är ännu för tidigt att uttala sig om. Däremot är detta en viktig faktor som man framgent bör ha under kontroll. På motsvarande sätt kan vi notera att delprovet ORD 97B blev mycket svårare än vid provtillfällena 96A-97A. För ORD-provet gäller dock att ett flertal uppgifter är utprövade enligt den nya modellen och att ett flertal uppgifter är reviderade (se tab. 8). Detta medför att det är svårt att tillförlitligt förklara utfallet av dessa uppgifter. En språklig förändring, ett förtydligande, en eller flera reviderade distraktorer, förändring i placering av svarsförslagen och rätt svar kan vissa gånger innebära stora förändringar i lösningsproportion.

Uppgifter utprövade enligt den nya modellen innebär, som tidigare nämnts, att uppgifter utprövats inom olika högskoleenheter. Det kan vara stora skillnader i provresultat på det reguljära provet mellan olika högskoleenheter, vilket är något som har betydelse och som man måste ta hänsyn till när man sätter samman och gör prognoser för de reguljära proven. Skillnaden mellan högskoleenheterna har varierat mellan 9.82 och 6.99 poäng för de fem provomgångarna 96A-98A, vilket redovisas i bilaga 1. På hela provet är det i det närmaste 10 poäng som skiljer mellan den högskoleenhet som hade det lägsta och det högsta resultatet vid provtillfället våren 1996 respektive våren 1997 medan höstproven 1996 och 1997 uppvisar en något mindre skillnad. Våren 1998 var skillnaden minst, 6.99 poäng, under den redovisade perioden. ELF-provet står för den största skillnaden i förhållande till antal uppgifter vid samtliga provtillfällen, därefter följer ORD, NOG, LÄS och DTK. Vi kan även notera att det är universitetsstäderna som har de bästa resultaten.

Ytterligare faktorer som man bör beakta när man sätter samman provet är när uppgifterna prövats ut, dvs. antal vår- och höstutprövade uppgifter. I den gamla utprövningsmodellen, med gymnasieelever, togs alltid hänsyn till om uppgiften var höst- eller vårutprövad. Eftersom dessa elever var i utbildning så presterade de ett högre resultat på våren i årskurs tre än på hösten. Eftersom vi sedan tidigare vet att det finns ett starkt samband mellan utbildning och högskoleprovresultat, dvs. ju längre utbildning desto bättre resultat, är det viktigt att ha en sådan faktor under kontroll (jfr. Stage et al., 1996, 1997). Nu innebar det visserligen bara en termins skillnad i gymnasiet mellan höst- och vårutprövade uppgifter, men trots detta visade det sig att höstutprövade uppgifter förändrades mer när de återfanns i ett reguljärt prov på våren än på hösten. Denna skillnad mellan prestationerna höst och vår finns troligen fortfarande kvar. I den reguljära provgruppen är ungefär 45 procent av provdeltagarna 20 år eller yngre och ytterligare ungefär 30 procent är 21-24 år, dvs. en stor del av provandegruppen är under utbildning (Stage et al., 1997).

Att uppgifter prövas ut och placeras i en s.k. uppgifts-bank för att flera år senare komma att användas i ett reguljärt prov är även en faktor som kan påverka resultatet. Tiden mellan utprövning och reguljärt tillfälle kan ha viss betydelse. Det framgår tydligt i ORD-provet hösten 1997 där ordet *Millennium* ingick. Vid det reguljära provtillfället klarade hela 89 procent uppgiften medan endast 61 procent hade svarat rätt vid utprövningen av

uppgiften. Visserligen hade en distraktor förändrats i uppgiften men detta hade sannolikt ringa betydelse eftersom den aktuella distraktorn inte hade dragit någon större andel vid utprovningen och kom ej heller att göra det vid reguljärprovet. Av större betydelse är snarare att millennium är att betrakta som ett trendord. Detta kan även ses som ett exempel på det vanskliga i att dra alltför stora slutsatser utifrån analys av enskilda uppgifter där risken för andra faktors genomslagskraft är stor. Hade ordet millennium varit en utprovningssuppgift vid provtillfället 97B hade det inte kommit med i ett senare reguljärprov eftersom uppgiften blev så lätt. I reguljärprovet 98A kan man finna ytterligare ett ord som ligger i tiden, nämligen *Sarong*, som är ett klädesplagg som blivit alltmer populärt under våren 1998. Vid det reguljära provtillfället klarade 82 procent uppgiften jämfört med 67 procent vid utprovningen. Uppgiften var inte på något sätt reviderad.

Tidsfaktorn är en annan faktor som är av betydelse för vissa av delproven. Delproven LÄS och DTK inbegriper någon form av "speed-faktor" förutom den specifika förmåga som respektive prov mäter (Henriksson & Eriksson, 1994). Efter den nya provsammansättningen då ORD-provet utökades med 10 uppgifter och oförändrad provtid kan även för detta prov en speed-faktor ha kommit att ingå. Ett exempel som bidrar till sådana funderingar är uppgift 40 i ORD 97B som blev .10 svårare reguljärt än vid utprovningstillfället då ordet hade position 14. Uppgiften var inte på något sätt reviderad. För delprovet LÄS visar pågående studier att uppgifters position i utprovning respektive i reguljärt prov har stor betydelse.

I delproven LÄS och DTK undviker man att ta texter/uppgifter – figurer/uppgifter från samma utprovningssversion för att provdeltagarna inte ska känna igen materialet, vilket innebär att utprovningssvärdena i vissa fall kan bli missvisande då texter/figurer och uppgifter inte är utprovade tillsammans. Det kan faktiskt vara så att i en utprovningssversion är andelen ej svar stor p.g.a. att det var en svår utprovningssversion. Detta måste tas hänsyn till, då dessa uppgifter kan få högre svarsfrekvens i ett annat provversion och i en annan position. Vi kan även se i tabell 32 att uppgifter som är likartade förändras i lösningsproportion när de förs samman i en slutlig provversion.

De faktorer som vi hittills har blivit uppmärksam på och som kan ha betydelse är:

- Provblockens varierande innehåll från provtillfälle till provtillfälle
- Utprovningssblockets placering under provdagen
- Utprovningssblockets placering jämfört med den reguljära versionen av samma delprov
- Högskoleenhet, dvs. var uppgiften är utprovad
- Uppgifters position vid utprovningen jämfört med vid reguljärprovet
- Flyttning av svarsförslag
- Flyttning av rätt svar
- Tid mellan utprovning och reguljärt prov, med tanke på t.ex trendord
- Likartade uppgifter från olika provversioner
- Vår- eller höstutprovade uppgifter

- Närliggande texter
- Utprövningsversionens svårighetsgrad (ej svar)
- Revideringar i frågestam och svarsförslag

Syftet med denna studie har varit att undersöka hur väl de nya utprövningsresultaten fungerar som underlag för att prognostisera reguljära provresultat. Vad som hittills framkommit är att för vissa av delproven har den nya modellen inneburit att överensstämmelsen mellan utprövningsvärden och de reguljära värdena är klart bättre än tidigare. För LÄS-provet har förbättringen inte varit lika uppenbar, men allteftersom konstruktörerna fått kunskap om de nya faktorer som framkommit under utvärderingsarbetet och kan påverka utfallet tycks överensstämmelsen bli bättre även för detta delprov. De reguljära proven sätts samman ungefär ett år innan det skall genomföras, vilket medför att det inte alltid är möjligt att ta hänsyn till faktorer som visat sig vara av betydelse förrän vid utplock och sammansättning av kommande reguljära prov. I tabell 36 redovisas differensen mellan utprövningsmedelvärde och reguljärt medelvärde.

Tabell 36. Differensen mellan utprövningsmedelvärde och reguljärt medelvärde vid provtillfällen 94A-98A. Inom parentes anges antal uppgifter som prövats ut enligt den nya modellen.

Regulj. tillf.	Delprov					Totalt
	ORD	NOG	LÄS	DTK	ELF	
94A	2.74	2.69	1.18	1.76	1.63	10.00
94B	4.05	1.63	0.46	0.39	2.09	8.62
95A	3.80	2.20	0.17	1.10	1.63	8.90
95B	3.34	0.79	1.35	1.12	2.02	8.62
96A	3.09	1.76	1.52	1.08	1.34	8.79
96B	1.00 (9)	2.21	2.32 (4)	1.03 (2)	1.08	7.64 (15)
97A	0.07 (31)	1.66 (1)	1.98 (16)	-0.60 (2)	0.24 (14)	3.35 (64)
97B	-1.27 (19)	2.06 (3)	0.38 (20)	0.59 (4)	0.66 (20)	2.42 (66)
98A	0.78 (31)	1.34 (7)	0.77 (20)	1.35 (7)	1.03 (19)	5.27 (84)

Tidigare kunde man göra en förhållandevis bra prognos. Bra, på det viset att den var stabil och det var tack vare den kunskap som konstruktörerna hade, men med bättre statistiskt underlag bör det vara möjligt att på ett mer systematiskt sätt prognostisera utfallet på provet. Den tysta kunskap som konstruktörerna utvecklat kan då kopplas med mer strikta metodiska ansatser.

Prognosarbetet kan förbättras ytterligare när vi fått utökade kunskaper om de nya faktorer som tillkommit. Ett flertal av dessa är nu föremål för pågående studier. Som inledningsvisades i detta avsnitt är dock vår uppfattning att den nya utprovningssmodellen så här långt är klart överlägsen den gamla och vi anser att utprovningarna också fortsättningsvis bör göras i samband med det reguljära provtillfället.

REFERENSER

- Billgren, B., Chekol, I-M., Linder, K., Falck, L. & Gebru, K. (1990). *Tyst kunskap –något att lyssna till i framtidens vård*. FoU rapport 33. SHSTF, Forskning och utveckling.
- Hambleton, R.K., Swaminathan, H. & Rogers, H.J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Hellekant, J. (1997). *Från utprovning till färdigt prov*. Skrift nr 16. Avdelningen för språkpedagogik, Göteborgs universitet.
- Henriksson, W., & Eriksson, S. (1992). *LÄS 92. LÄS-prov med sex texter*. PM nr 58. Avdelningen för pedagogiska mätningar, Pedagogiska institutionen, Umeå universitet.
- Henriksson, W. (1992). *Modell för prognos av provutfall utifrån delprovsspecifika ankaruppgifter*. Pm nr 64. Avdelningen för pedagogiska mätningar, Pedagogiska institutionen, Umeå universitet.
- Henriksson, W., & Eriksson, S. (1994). *Försök med utökad provtid på LÄS-provet*. PM nr 81. Avdelningen för pedagogiska mätningar, Pedagogiska institutionen, Umeå universitet.
- Kerstiens, G. (1990). A slow look at speeded reading comprehension tests. *Review of Research in Developmental Education*, v 7, n 3.
- Ramstedt, K. (1996). *Elektriska flickor och mekaniska pojkar. Om gruppskillnader på prov – en metodutveckling och en studie av skillnader mellan flickor och pojkar på centrala prov i fysik*. Akademisk avhandling. Umeå: Pedagogiska institutionen, Umeå universitet.
- Ramstedt, K. & Stage, C. (1997). *Några jämförelser mellan högskoleproven hösten 1995 och hösten 1996*. Pm nr 128. Enheten för pedagogiska mätningar, Umeå universitet.
- Stage, C. (1995). *Utprovning av provuppgifter. Förändring av högskoleprovets utprovningsrutiner*. Pm nr 106. Enheten för pedagogiska mätningar, Umeå universitet.
- Stage, C., & Jarl, C. (1996). *Högskoleprovet våren 1996. Provdeltagargruppens sammansättning och resultat*. Pm nr 115. Enheten för pedagogiska mätningar, Umeå universitet.
- Stage, C., Ögren, G. & Jarl, C. (1997). *Högskoleprovet hösten 1997. Provdeltagargruppens sammansättning och resultat*. Pm nr 131. Enheten för pedagogiska mätningar, Umeå universitet.

Stage, C. & Hamrén, M. (1997). *Normering av högskoleprovet våren 1997*. Internt material.

Ögren, G. (1996). *Ett försök med olika antal uppgifter per figuruppsättning i DTK-provet*. Pm nr 119. Enheten för pedagogiska mätningar, Umeå universitet.

Ögren, G. (1996). *Ett försök med fyra svarsförslag per uppgift istället för fem i DTK-provet*. Pm nr 119. Enheten för pedagogiska mätningar, Umeå universitet.

Tabell . Skillnad i provresultat mellan olika högskoleenheter

Delprov	M	Min	Max	Diff
96 A	76.21	72.05	81.87	9.82
ORD	25.41	23.78	27.14	3.36
NOG	13.83	13.22	14.83	1.61
LÄS	12.15	11.56	13.18	1.62
DTK	11.68	11.06	12.43	1.37
ELF	13.14	12.17	14.28	2.11
96B	75.34	71.67	79.24	7.57
ORD	23.82	22.42	25.26	2.84
NOG	13.51	12.51	14.37	1.86
LÄS	13.22	12.80	13.89	1.09
DTK	11.92	11.60	12.44	0.84
ELF	12.88	11.76	14.01	2.25
97A	70.59	66.40	75.83	9.43
ORD	23.16	21.88	24.87	2.99
NOG	12.89	12.17	13.69	1.52
LÄS	12.70	12.15	13.55	1.40
DTK	10.30	9.68	11.12	1.44
ELF	11.54	10.42	12.60	2.18
97B	72.65	69.31	77.06	7.75
ORD	22.13	20.61	23.69	3.08
NOG	13.08	12.33	13.94	1.61
LÄS	12.54	12.04	13.25	1.21
DTK	11.32	10.78	11.90	1.12
ELF	13.58	12.59	14.48	1.89
98A	72.23	68.82	75.81	6.99
ORD	23.05	21.86	23.97	2.11
NOG	12.33	11.53	13.11	1.58
LÄS	11.47	10.74	12.18	1.44
DTK	11.76	11.27	12.30	1.03
ELF	13.63	12.50	14.70	2.20