

**PLACERING AV UPPGIFTER OCH
RÄTT SVARSFÖRSLAG I
HÖGSKOLEPROVET**

Positionens betydelse för svårighetsgraden

Ewa Andersson (red)

Since spring 1996 the pretesting of new items for the Swedish Scholastic Aptitude Test (SweSAT) is carried out in connection with the regular test administrations. The new pretesting model brings out possibilities to conduct research on aspects of test construction that has not been possible to investigate earlier. In this report a number of small studies are presented which have been carried out on the issue of the relationship between the position of the items in different subtests and their levels of difficulty. The subtests studied were Vocabulary (WORD), Data Sufficiency (DS), Diagrams, Tables, Maps (DTM) and Reading Comprehension (READ). Furthermore, the relationship between position of the correct answer among the distractors and the item difficulty has been investigated for the subtests WORD and DTM. The results indicate that placement of the item in the subtest as well as placement of the correct answer among the distractors have effect of item difficulty. However the results raises important questions concerning whether the results is generally applicable on different item formats, specific levels of item difficulty or related to different kinds of item solving strategies among different groups of testtakers.

INLEDNING	1
1. HAR UPPGIFTENS POSITION I PROVHÄFTET NÅGON BETYDELSE FÖR DESS SVÅRIGHETSGRAD?	4
1.1 DELPROVET ORD	4
<i>Kerstin Andersson</i>	
1.2 DELPROVET NOG	7
<i>Anders Lexelius</i>	
1.3 DELPROVET DTK.....	13
<i>Birgitta Wallin och Joakim Tegblom</i>	
1.4 DELPROVET LÄS	21
<i>Stig Eriksson och Birgitta Wallin</i>	
2. HAR DET RÄTTA SVARETS POSITION BLAND SVARSFÖRSLAGEN NÅGON BETYDELSE FÖR UPPGIFTENS SVÅRIGHETSGRAD?	31
2.1 DELPROVET ORD	31
<i>Kerstin Andersson</i>	
2.2 DELPROVET DTK.....	35
<i>Birgitta Wallin och Joakim Tegblom</i>	
3. SAMMANFATTANDE KOMMENTAR.....	42
REFERENSER	

Inledning

Ett högskoleprov är giltigt i fem år, vilket innebär att vid ett ansökningstillfälle kan det finnas resultat från tio olika provtillfällen. Detta ställer stora krav på att poäng från olika provtillfällen är likvärdiga. Provkonstruktörerna har som uppgift att sammanställa parallella versioner av varje delprov. Begreppet parallellitet består i detta avseende av två delar. För det första ska varje delprov anpassas till en teoretisk modell för hur provet ska se ut med avseende på ämnesområden, innehållskategorier, uppgiftsformat, informationsmängd och lösningsprocedur. För det andra ska varje delprov anpassas till en empirisk modell där såväl enskilda uppgifter som provet i sin helhet ska ha definierade svårighetsgrader. Av avgörande betydelse är att uppgifterna fungerar på ett sådant sätt att provdeltagare som presterat en hög poäng på delprovet även ska kunna lösa svårare uppgifter. Varje uppgift ska m.a.o. diskriminera på ett tillfredsställande sätt och ha balans vad avser felaktiga svarsförslag. Slutligen ska gruppskillnader i resultat vara under kontroll.

För att de enskilda delproven ska kunna anpassas till den empiriska modellen måste alla kommande provuppgifter genomgå en empirisk prövning. Tidigare genomfördes utprövningarna av kommande provuppgifter med hjälp av elever i årskurs 3 i gymnasieskolan, vilket innebar att antalet personer i varje utprövningsgrupp var begränsat. Eftersom eleverna var medvetna om att resultaten inte hade någon betydelse för dem personligen var motivationen i vissa fall tämligen bristfällig. Sedan våren 1996 sker utprövningarna i samband med de reguljära provtillfällena där provdeltagarna genomför fem block,¹ varav ett inte ingår i det reguljära provet.

Den utvärdering som gjorts av den nya utprövningsmodellen visar på en rad fördelar (Ögren, 1998). Den största fördelen är att betingelserna för utprövningsblock och reguljära block är identiska, eftersom provdeltagarna vid genomförandet inte vet vilket block som innehåller utprövningsmaterial. Detta betyder att prestationsnivå och motivation är densamma som för målgruppen och det betyder också att antalet personer som genomför utprövningsuppgifterna är tillräckligt stort för att ge säkra skattningar av uppgifternas svårighetsgrad. Denna ökade säkerhet hos utprövningsresultaten har lett till att nya faktorer, som inte kunde undersökas vid det tidigare utprövningsförfarandet, nu har uppmärksammats. Föreliggande rapport kan således även betraktas som ett bidrag till utvärderingen av den nya utprövningsmodellen.

¹ Fyra block innehåller ett delprov och ett block innehåller två delprov, ELF och ORD.

Det faktum att provdeltagarna vid de olika högskoleenheterna i genomsnitt erhåller olika poäng på högskoleprovet² ger emellertid anledning att fundera över hur utprövningsdata används i samband med sammansättningen av de reguljära provversionerna. Frånsett att korrigeringar måste göras av utprövningsdata i förhållande till utprövningsgruppernas medelvärde på det reguljära provet, har den nya utprövningsmodellen aktualiserat en rad andra faktorer. Här avses exempelvis aspekter som att de olika delproven genomförs vid olika tidpunkter under dagen vid olika provtillfällen. Vidare är inte uppgifternas position densamma i den reguljära provversionen som i utprövningsversionen. Det hänger samman med att uppgifterna till ett reguljärt prov är hämtade från ett stort antal utprövningsversioner i syfte att minimera antalet uppgifter som en provdeltagare eventuellt kan ha mött vid genomförandet av utprövningsversionerna. I och med införandet av den nya utprövningsmodellen skärptes ytterligare de regler som finns för varje delprov för hur många uppgifter som får hämtas från en och samma utprövningsversion. Därtill kommer andra krav som ställs på den reguljära provversionen och som påverkar uppgifternas ordningsföljd, t.ex. avseende allsidighet och svårighetsgrad.

Förändringen av utprövningarna har även kommit att innebära att en ökad tonvikt har lagts vid de granskningsrutiner som föregår utprövningarna vilket har minskat behovet av omprövningar av uppgifter. Trots att ambitionen är att revidera uppgifter så lite som möjligt efter utprövning ställs denna ambition ibland mot andra krav. Det kan exempelvis bli nödvändigt att efter utprövning revidera en uppgift på ett sådant sätt att det rätta svarets position bland svarsförslagen förändras.

Den nya utprövningsmodellen aktualiserar därmed en rad frågor, dels till följd av praktiska och administrativa skäl, dels till följd av att den nya modellen innebär säkrare skattningar vilket medger möjlighet att undersöka aspekter som tidigare inte var möjliga att upptäcka. Utprövningsvärdena är dock av flera anledningar inte direkt överförbara för att skatta de värden som enskilda uppgifter kommer att uppvisa vid kommande reguljära provtillfällen.

I föreliggande rapport presenteras en rad mindre studier som har genomförts till följd av frågor som har rests sedan den nya utprövningsmodellen togs i

² Noterbart är att skillnaden i genomsnittlig provpoäng mellan högskoleenheterna kan uppgå till 10 poäng (Ögren, 1998), vilket kan jämföras med den genomsnittliga skillnaden i provpoäng mellan kvinnor och män som sedan våren 1996 uppgått till 5-6 poäng till männens fördel. Denna skillnad har tidigare, fr.o.m. vårprovet 1977 t.o.m. vårprovet 1996, uppgått till i genomsnitt 8 poäng till männens fördel (Stage, Ögren & Konradsson, 1999).

bruk våren 1996. Detta innebär att de studier som presenteras har sin upprinnelse i de problemställningar som är förknippade med en praktiskt tillämpad verksamhet. Studierna ingår samtliga som delar i det forsknings- och utvecklingsarbete som kontinuerligt bedrivs i anslutning till framställningen av högskoleprovet³ vid Enheten för pedagogiska mätningar vid Umeå universitet. Mot bakgrund av att de olika studierna baserar sig på olika delprov som till sin karaktär skiljer sig från varandra samt att studierna genomförts vid olika tidpunkter är inte resultaten från de olika studierna direkt jämförbara med varandra.

Rapporten är disponerad enligt följande: I kapitel 1 presenteras fyra olika delstudier som har undersökt vilken betydelse uppgifters position i provhäftet har för uppgifternas svårighetsgrad i delproven ORD, NOG, DTK och LÄS. I kapitel 2 presenteras två delstudier som behandlar betydelsen av det rätta svarets position bland svarsförslagen för delproven ORD och DTK. I kapitel 3 ges en kort sammanfattning och kommentar till de presenterade studierna.

³ För en utförlig beskrivning av högskoleprovets delprov, se t.ex. Andersson (1999).

1. Har uppgiftens position i provhäftet någon betydelse för dess svårighetsgrad?

Syftet med de studier som presenteras i avsnitt 1.1–1.3, vilka omfattar delproven ORD, NOG och DTK, är att undersöka om placeringen av uppgifter i provhäftet har någon betydelse för deras svårighetsgrad. I avsnitt 1.4 presenteras resultaten från en studie som på motsvarande sätt har genomförts av delprovet LÄS. Eftersom uppgifterna i delprovet LÄS är kopplade till en specifik text syftar denna studie till att undersöka betydelsen av texternas placering i provhäftet.

1.1 Delprovet ORD

Som en följd av att AO-provet utgick fr.o.m. våren 1996 utökades ORD-provet med 10 uppgifter, från 30 till 40, men med samma provtid, 15 minuter. Detta innebär att provdeltagarna har fått kortare tid per uppgift. Syftet med denna studie var att undersöka om det har någon betydelse för ORD-uppgifternas relativa svårighetsgrad om de placeras i början eller i slutet av provhäftet.

Metod

Material

Vid provtillfället 98B (höstprovet 1998) genomfördes ett försök för att ta reda på om ORD-uppgifternas position i provhäftet påverkade deras svårighetsgrad. Försöket innebar att exakt samma uppgift fick olika position i två olika provhäften, antingen i början eller slutet av provet. Varje uppgift prövades alltså ut två gånger. Försöket omfattade nio uppgifter. Fem av dem ingick i häfte 2 och 3 och fyra i häfte 4 och 5. I häfte 2 och 3 ingick tre substantiv, ett verb och ett adjektiv. Av sekretesskäl presenteras inte uppgifterna här utan kallas hädanefter S1, S2, S3, V1 respektive A1. I häfte 4 och 5 gjordes motsvarande försök med tre substantiv och ett verb, hädanefter kallade S4, S5, S6 respektive V2.

Försökspersoner

Den grupp som genomförde häfte 2 bestod av 5 956 personer, 2 709 män och 3 247 kvinnor. Den grupp som genomförde häfte 3 bestod av 1 367 personer, 613 män och 754 kvinnor. Gruppen som genomförde häfte 2 presterade något bättre på det reguljära ORD-provet än den totala provdeltagargruppen. Deras medelvärdeskillnad var +.92 jämfört med totalgruppen. Den grupp som genomförde häfte 3 var något svagare än totalgruppen. Deras medelvärdeskillnad var -.26. Med hjälp av dessa medelvärdeskillnader kan man korrigera lösningsproportionerna för uppgifter från olika utprovningshäften.

I det här fallet innebär det att uppgifter från häfte 2 korrigeras med -.02 och att uppgifter från häfte 3 korrigeras med +.01. Den grupp som genomförde häfte 4 bestod av 1 600 personer, 658 män och 942 kvinnor. Gruppen som genomförde häfte 5 bestod av 1 709 personer, 801 män och 908 kvinnor. Båda dessa grupper presterade något sämre på det reguljära ORD-provet än totalgruppen. Medelvärdeskillnaden var i båda fallen -.56, vilket innebär en korrigering med +.01.

Resultat

I tabell 1:1 presenteras de huvudsakliga resultaten. Lösningproportionerna, p-värdena, är korrigerade för gruppernas resultat på det reguljära ORD-provet.

Tabell 1:1 Lösningproportioner (p-värden) totalt och för män/kvinnor samt biseriala korrelationer för nio ORD-uppgifter som prövats ut i början och i slutet av ett provhäfte.

Uppgift	Nr	Häfte	p korr	p korr män	p korr kv	rbis
V1	1	2	.60	.61	.59	.54
	30	3	.57	.59	.56	.51
S1	1	3	.86	.88	.84	.60
	40	2	.85	.88	.83	.56
S2	2	2	.92	.93	.92	.37
	38	3	.94	.95	.93	.37
S3	3	2	.34	.32	.35	.45
	37	3	.31	.30	.31	.37
A1	4	2	.27	.30	.25	.34
	36	3	.24	.25	.23	.29
S4	3	5	.11	.12	.10	.49
	39	4	.12	.13	.12	.35
S5	1	4	.81	.87	.76	.40
	38	5	.79	.85	.74	.47
V2	3	4	.97	.97	.96	.59
	40	5	.96	.97	.95	.48
S6	1	5	.91	.92	.91	.44
	38	4	.89	.91	.88	.42

Tabellen visar att lösningproportionerna ökade för sju och minskade för två ORD-uppgifter då de prövades ut i början av ett provhäfte. Däremot ändrades inte svårighetsnivån för uppgifterna; en lätt uppgift var lätt vare sig den hade position 1 eller 40 och en svår uppgift var svår vare sig den presenterades i början eller slutet av provhäftet. Lösningproportionerna för män ökade

för fem uppgifter och för kvinnor för sju uppgifter då uppgifterna placerades i början av ett häfte. Två uppgifter, S2 och S4, blev svårare för både kvinnor och män i början av ett häfte och två uppgifter, S1 och V2, uppvisade inga skillnader för män vare sig de placerades i början eller slutet av ett häfte. För endast en uppgift, S5, har den biseriala korrelationen ökat då uppgiften placerats i slutet av ett provhäfte. Vad detta egentligen betyder är dock inte så lätt att uttala sig om eftersom uppgifterna inte prövats ut i identiskt lika häften. I tabell 1:2 ges en sammanfattning av resultaten.

Tabell 1:2 Sammanfattande resultat för nio ORD-uppgifter placerade i början respektive i slutet av ett provhäfte (korrigerade värden).

	Början		Slutet		Diff	
	Σ	M	Σ	M	Σ	M
p totalt	5.79	.64	5.67	.63	.12	.01
p män	5.92	.66	5.83	.65	.09	.01
p kvinnor	5.68	.63	5.55	.62	.13	.01
rbis	–	.47	–	.42	–	.05

Vad beträffar uppgiftens position i provhäftet kan det alltså utifrån detta försök konstateras att det verkar ha betydelse om uppgifterna presenteras i början eller i slutet av häftet. Det verkar också som om kvinnor gynnas mer än män av att uppgifterna placeras i början av ett häfte. Vad dessa resultat beror på kan man endast spekulera över. Kanske är det med 40 i stället för 30 uppgifter i ORD-provet något svårare att disponera tiden rätt, kanske ägnar man för lång tid åt de första uppgifterna för att sedan hamna i tidsnöd mot slutet. I detta försök rör det sig om alltför få uppgifter för att man skall kunna dra helt entydiga slutsatser. Försöket bör därför upprepas.

1.2 Delprovet NOG

Det övergripande syftet med denna studie är att undersöka huruvida man erhåller samma resultat på ett NOG-prov, uttryckt i antal rätta svar, om man presenterar uppgifterna i omvänd ordning. Kommer svåra uppgifter, som vanligtvis placeras i slutet av provet, att vara lika svåra då de placeras i början av provet? Detta är några frågor som denna studie förhoppningsvis ska ge svar på.

Metod

I samband med det reguljära högskoleprovet hösten 1997 gjordes ett försök med att presentera NOG-uppgifter i ett av utprövningsproven i omvänd ordning. Till försöket valdes 22 uppgifter som innehållsmässigt speglar de kunskaper och färdigheter som mäts via ett reguljärt NOG-prov. Uppgifterna utgjorde med andra ord en delmängd av de matematiska kunskaper som definieras enligt kurs Matematik A för gymnasieskolan. Försökspersonerna utgjordes av dem som deltog i det reguljära högskoleprovet inom två högskoleregioner i Sverige. I vardera regionen deltog ca 2 400 personer i högskoleprovet. De 22 uppgifterna presenterades i två häften, ett för respektive grupp, med den skillnaden att uppgifternas ordning var omvänd i häfte 2. Den uppgift som placerades som nummer 1 i häfte 1 placerades som nummer 22 i häfte 2, osv.

Det reguljära NOG-provet, som genomfördes av 61 060 personer, användes för att bestämma de båda gruppernas prestationsnivå vad avser NOG-provet.

Tabell 1:3 Medelvärde (M) och standardavvikelse (s) på det reguljära NOG-provet hösten 1997, totalt och uppdelat på kön (m/k), för dem som deltog i försöket.

Häfte	M (m/k)	s (m/k)	N (m/k)
1	13.94 (15.21/12.82)	4.07 (3.90/3.89)	2 407 (1130/1277)
2	13.30 (14.56/12.30)	4.20 (3.98/4.10)	2 422 (1066/1356)

Av tabell 1:3 framgår att de som deltagit i försöket är till antalet nästan lika fördelade på de båda häftena. Man kan vidare konstatera att den grupp som gjorde häfte 1 har erhållit 13.94 rätta svar på det reguljära högskoleprovet mot 13.30 rätta svar för den grupp erhöi häfte 2. De som genomförde häfte 1 har således något högre resultat och lägre spridning än den grupp som genomförde häfte 2. Det är värt att notera att samtliga deltagare i försöket var bättre än den genomsnittlige provdeltagaren. Den genomsnittlige prov-

deltagaren hade 13.08 rätta svar på NOG-provet hösten 1997. Såväl männen som kvinnorna i de båda grupperna presterade bättre än den genomsnittlige mannen respektive den genomsnittliga kvinnan. Männen hade i genomsnitt 14.26 rätta svar medan kvinnorna hade 12.02 på provet hösten 1997.

Resultat

Tabell 1:4 Medelvärde (M) och standardavvikelse (s) på båda försöksproven, totalt och uppdelat på kön (m/k), för dem som gjorde häfte 1 respektive häfte 2.

Häfte	M (m/k)	s (m/k)	N (m/k)
1	12.85 (13.61/12.17)	3.28 (3.27/3.15)	2 402 (1126/1276)
2	12.20 (12.97/11.60)	3.59 (3.63/3.44)	2 419 (1064/1355)

Av tabell 1:4 framgår att det resultat som framkom för de båda grupperna, enligt tabell 1:3, även har sin giltighet för de båda försöksproven. Man kan konstatera att de som gjorde häfte 1 även på försöksprovet har ett högre medelvärde och en lägre spridning än de som gjorde häfte 2. Vi kan med andra ord konstatera att de som gjorde häfte 1 presterat bättre än dem som gjorde häfte 2 inte bara på det reguljära provet utan även på försöksprovet. Den skillnad som fanns initialt mellan grupperna, .64 poäng i medeltal på de reguljära NOG-provet, kvarstår i försöksprovet. Skillnaden i medeltal i försöksprovet mellan de båda grupperna är .65 poäng till fördel för den grupp som gjorde häfte 1. Om man endast betraktar medelvärden, där man korrigerar för prestation på det reguljära NOG-provet, finns det inget som tyder på att presentationsordning i provet har någon betydelse för totalresultatet⁴. Presentationsordningen förefaller inte heller ha någon betydelse när man konstanthåller variabeln kön. Således går resultatet för totalgruppen och en uppdelning på kön i samma riktning, dvs. att presentationsordningen inte påverkar prestationen på provet. En fråga man då kan ställa sig är om presentationsordningen har någon betydelse för de enskilda uppgifterna.

⁴ Korrigerat (estimerat) värde med hjälp av regression på häfte 1 respektive häfte 2 är 12.01 och 12.07 för samtliga 61 060 personer som deltog i högskoleprovet hösten 1999.

Tabell 1:5 De båda gruppernas medelvärde, totalt och uppdelat på kön (m/k), då uppgifterna grupperats i tre grupper på det NOG-prov som utgjorde utprövningsprovet och som ingick i högskoleprovet hösten 1997.

Häfte 1		Häfte 2		Diff
Uppgift	Poäng (m/k)	Uppgift	Poäng (m/k)	Tot (m/k)
1-7	4.65 (4.73/4.58)	22-16	4.38 (4.47/4.34)	.27 (.26/.24)
8-15	5.03 (5.43/4.69)	15-8	4.75 (5.15/4.46)	.28 (.28/.23)
16-22	3.16 (3.46/2.90)	7-1	3.05 (3.37/2.81)	.11 (.09/.09)

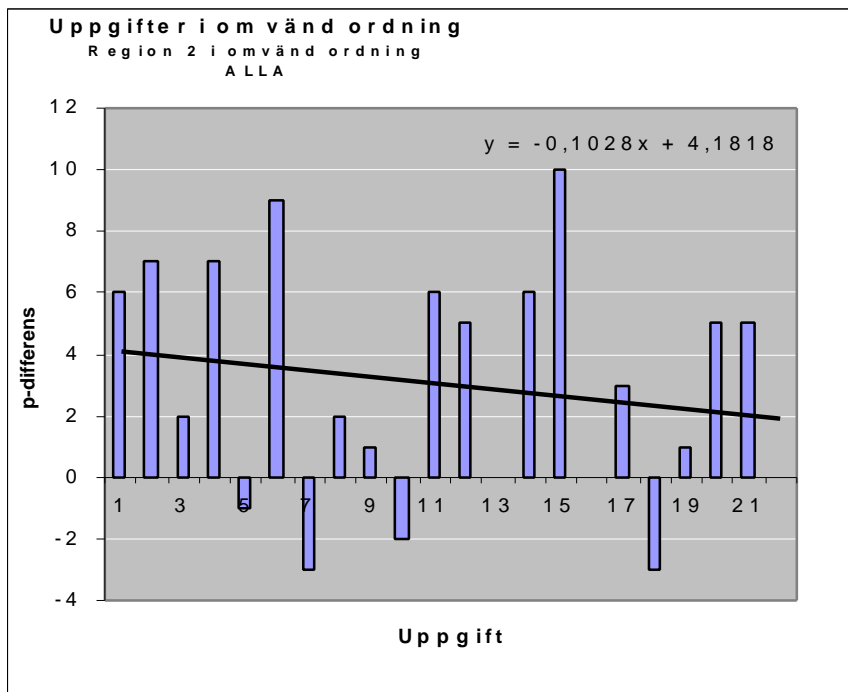
Av tabell 1:5 framgår att även när uppgifterna, utifrån placering, delas in i tre grupper har de som gjort häfte 1 fler rätta svar, i varje grupp, än de som gjort häfte 2. Skillnaden i poäng är dock mindre för de 7 sista uppgifterna i häfte 1, vilket även är de 7 första uppgifterna i häfte 2. Denna skillnad utgör i medeltal endast hälften av skillnaden för de övriga uppgifter. Med andra ord blir skillnaden mellan grupperna mindre ju högre svårighetsgrad uppgifterna har. För svåra uppgifter minskar följaktligen utrymmet för poäng, därav en lägre skillnad. Man kan vidare konstatera att det inte finns något direkt samband mellan gruppresultatet och uppgiftens placering i provet. Svåra uppgifter förblir svåra och lätta förblir lätta oberoende av placering i provet. Även om kvinnorna presterat en lägre poäng än männen så är skillnaden bland männen respektive kvinnorna mellan häfte 1 och häfte 2 tämligen konstant över de tre grupperna.

Tabell 1:6 Lösningsproportion (p-värde) för uppgifterna i häfte 1 och häfte 2 samt rbis. Totalvärden samt uppdelat på kön (m/k).

Häfte 1				Häfte 2			
Nr	p-värde	m/k	rbis	p-värde	m/k	rbis	Nr
1	.62	.64/.60	.42	.56	.59/.54	.49	.22
2	.78	.83/.73	.64	.71	.79/.66	.68	.21
3	.75	.79/.72	.51	.73	.76/.70	.55	.20
4	.80	.78/.81	.42	.73	.74/.72	.47	.19
5	.73	.79/.68	.37	.74	.77/.72	.42	.18
6	.59	.62/.57	.42	.50	.50/.51	.39	.17
7	.38	.28/.47	-.08	.41	.32/.49	.04	.16
8	.27	.36/.19	.53	.25	.35/.17	.56	.15
9	.35	.36/.35	.19	.34	.34/.34	.27	.14
10	.82	.84/.80	.35	.84	.87/.82	.47	.13
11	.64	.71/.59	.57	.58	.64/.53	.56	.12
12	.79	.84/.75	.57	.74	.79/.71	.57	.11
13	.71	.77/.65	.56	.71	.75/.68	.59	.10
14	.92	.95/.89	.79	.86	.89/.83	.75	.9
15	.53	.60/.47	.59	.43	.52/.36	.59	.8
16	.33	.38/.29	.41	.33	.37/.30	.34	.7
17	.71	.73/.68	.44	.68	.72/.66	.51	.6
18	.31	.33/.30	.27	.34	.37/.31	.30	.5
19	.40	.48/.33	.44	.39	.42/.37	.37	.4
20	.66	.72/.61	.61	.61	.68/.55	.59	.3
21	.48	.52/.44	.50	.43	.50/.38	.52	.2
22	.27	.30/.25	.47	.27	.31/.25	.44	.1

Av tabell 1:6 framgår att såväl p-värden som rbis är tämligen konstanta vid en jämförelse mellan de båda häftena.

För varje uppgift har dessutom differensen mellan lösningsproportion för häfte 1 och häfte 2 beräknats. Dessa skillnader framgår av figur 1:1. Uppgift nr 1 i figuren är resultatet av differensen mellan uppgift nr 1 i häfte 1 och uppgift nr 22 i häfte 2. Lösningproportionen på uppgift nr 1 var .62 i häfte 1 medan samma uppgift hade värdet .56 i häfte 2, dvs. differensen är således .06 eller 6 hundradelar. Av figuren framgår även att de prövande som gjorde häfte 1 har presterat bättre än de prövande som gjorde häfte 2 på 15 av 22 uppgifter och sämre på 4 uppgifter. På 3 uppgifter var det ingen skillnad i prestation. På det reguljära NOG-provet var de prövande som gjorde häfte 1 bättre på 20 av de 22 uppgifterna. En uppgift var till fördel för de prövande som gjorde häfte 2 och en uppgift hade lika lösningsproportion i de båda grupperna.



Figur 1:1 Differens i prestation av proportionstal på de 22 uppgifterna för häfte 1 och 2. Till materialet har även en regressionslinje anpassats.

Om man anpassar en rät linje⁵ till differensen i prestation mellan de som gjorde häfte 1 och häfte 2 för varje uppgift finner man att den är som störst i början av provet och att den sedan minskar. Detta kan man dock inte ta som ett tecken på att uppgifternas placering är av betydelse för hur svårighetsgraden förändras med uppgiftens placering. Genom att ändra den första uppgiftens differens från .06 till -.03 erhålls en rät linje utan lutning, dvs. linjens riktningskoefficient = 0. Genom att ändra värdet på en enda uppgift i materialet får man ett resultat som indikerar att uppgiftens placering inte har någon betydelse för hur lätt eller svår uppgiften är. Med andra ord kan en enda uppgift inte påverka resultatet på de andra 21 uppgifterna.

⁵ För männen är regressionslinjen $y = -.1067x + 4.0909$ och för kvinnorna är motsvarande regressionslinje $y = -.1045x + 3.7922$. Regressionslinjen för samtliga framgår av figur 1:1.

En fråga man kan ställa sig är: Varför presterar den "sämre" gruppen en högre provpoäng än den "bättre" gruppen på några av uppgifterna? Den frågan har inte sitt svar i detta arbete utan den får anstå till ett senare tillfälle.

Tabell 1:7 Summa p-värde och genomsnittlig lösningsproportion för 12-uppgifter som placerats i början och i slutet av NOG-provet samt deras differens. Totalt och uppdelat på kön samt medelvärde för rbis.

	Början		Slutet		Diff	
	Σ	M	Σ	M	Σ	M
Häfte 1						
p totalt	4.03	.67	3.91	.65	.12	.02
p män	4.21	.70	4.09	.68	.12	.02
p kvinnor	3.87	.65	3.79	.63	.08	.01
rbis	–	.46	–	.50	–	
Häfte 2						
p totalt	2.66	.44	2.59	.43	.07	.01
p män	2.94	.49	2.84	.47	.10	.02
p kvinnor	2.46	.41	2.37	.40	.09	.02
rbis	–	.46	–	.46	–	

De värden som redovisas i tabell 1:7 är korrigerade med hjälp av resultatet på det reguljära NOG-provet. Således har värdena i det ena häftet korrigeras med -.04 per uppgift medan värdena i det andra häftet korrigerats med -.01 per uppgift. Detta för att initialt jämställa de båda gruppernas prestationsnivå på NOG-provet.

Sammanfattningsvis kan man konstatera att det inte råder någon full klarhet i om uppgifter som presenteras i början av NOG-provet erhåller högre lösningsproportion än uppgifter som placeras i slutet av provet. Visserligen finns det en antydning att uppgifter som placeras i början av provet har en något högre lösningsproportion än uppgifter som placeras i slutet av provet men om det existerar en effekt av placeringsordningen så är den ganska begränsad. Endast fortsatta försök i större skala kan ge svar på frågan om effekten av ordningsföljden för uppgifterna i NOG-provet.

1.3 Delprovet DTK

Det nya utvärderingsförfarandet har ökat säkerheten i prognoserna för det reguljära högskoleprovets utfall. Den förbättrade träffsäkerheten har samtidigt ökat intresset för sådana faktorer som trots allt kan påverka den. Utfallet av de senaste årens provomgångar har till exempel antytt att förändring av en uppgifts position i provhäftet kan påverka uppgiftens svårighetsgrad, dvs. lösningsproportionen (p-värdet). Syftet med den studie som ska beskrivas här var att undersöka om så var fallet vad gäller uppgifter i delprovet DTK.

Metod

Undersökningen genomfördes med hjälp av försöksprov vid två reguljära provtillfällen i följd, våren respektive hösten 1998. Försöksprovet hade samma format som såväl utvärderingsversionerna som de reguljära versionerna av DTK. Det innehöll därmed tio figuruppsättningar och två uppgifter till varje med vardera fem givna svarsförslag, A–E. Provtiden var 50 minuter. Försöksprovet gavs vid båda tillfällena i två varianter, kallade häfte 5 och häfte 6, till var sin grupp provdeltagare. Av försöksprovets 20 uppgifter ägnades tolv – uppgifterna 1–6 och 15–20 – åt att undersöka betydelsen av en uppgifts position.⁶ Figurmaterialiet och uppgiftsinnehållet var identiskt i försöksprovets båda varianter. Däremot var uppgifternas ordningsföljd motsatt i de två häftena, på det sätt som beskrivs i tabell 1:8.

Tabell 1:8 De tolv försöksuppgifternas placering i häftena 5 och 6.

Uppg nr i häfte 5	1	2	3	4	5	6	15	16	17	18	19	20
Uppg nr i häfte 6	19	20	17	18	15	16	5	6	3	4	1	2

⁶ Övriga uppgifter i försöksprovet avsåg att testa betydelsen av det rätta svarets position bland svarsförslagen, se vidare avsnitt 2.2.

I försöksprovet ingick både nykonstruerat och tidigare utprövat, reviderat material som vad gäller figurtyper och ämnesområden kombinerats och varierats enligt samma principer som gäller för utprovningshäften och reguljära DTK-prov. Insamlade data bearbetades på samma sätt som sker efter utprovning av preliminärt material och genomföranden av reguljära prov. Därefter gjordes en traditionell itemanalys.

Försökspersoner

Som tidigare nämnts genomfördes försöksproven vid två provtillfällen och av totalt fyra olika grupper provdeltagare. I tabell 1:9 redovisas antal prövande, medelvärde på det genomförda reguljära DTK-provet (98A respektive 98B) jämfört med medelvärdet för riket samt medelvärde på de tolv försöksuppgifterna för var och en av de fyra grupperna.

Tabell 1:9 Försöksgrupperna och deras medelvärden på det reguljära DTK-provet respektive på de tolv försöksuppgifterna.

Prov-tillfälle	Häfte	N	Män	Kvinnor	M reg prov (riket)	M försöksprov; Uppg 1–6, 15–20
98:04	5	2 045	875	1 170	11.27 (11.76)	4.66
98:04	6	7 483	3 398	4 085	11.87 (11.76)	5.13
98:10	5	1 842	837	1 005	10.49 (10.49)	4.62
98:10	6	1 021	467	554	10.40 (10.49)	4.64

Som framgår av tabell 1:9 hade den grupp som våren 1998 genomförde häfte 6 ett bättre resultat, både på det reguljära DTK-provet och på försöksuppgifterna, än den grupp som vid samma tillfälle genomförde häfte 5. Grupperna som genomförde försöksproven hösten 1998 låg däremot på nästan exakt samma nivå. Medelvärdeskillnaden mellan grupperna som deltog i försöket våren 1998 måste beaktas och vägas in vid granskningen av resultaten.

I tabell 1:10, där de data som samlades in våren 1998 redovisas på uppgiftsnivå, har vart och ett av p-värdena därför korrigerats utifrån de två gruppernas prestation på det reguljära DTK-provet i relation till genomsnittet för riket. För den grupp som genomförde häfte 5 har p-värdena ökat med .02 medan de för den grupp som genomförde häfte 6 har minskats med .01.

Resultat

Tabell 1:10 Erhållna p-värden våren 1998 för de tolv försöksuppgifterna i häfte 5 och 6 samt differenser mellan dessa. Dessutom redovisas skillnaden mellan mäns och kvinnors p-värden samt rbis.

Uppgift nr	Häfte	p	p (korr)	p diff ⁷ (korr)	p diff m/k	rbis
1	5	.57	.59		+ .01	.25
19	6	.53	.52	+ .07	+ .02	.33
2	5	.30	.32		+ .17	.61
20	6	.35	.34	– .02	+ .16	.63
3	5	.76	.78		+ .12	.53
17	6	.72	.71	+ .07	+ .10	.50
4	5	.16	.18		+ .08	.58
18	6	.17	.16	+ .02	+ .08	.59
5	5	.37	.39		+ .11	.40
15	6	.47	.46	– .07	+ .07	.46
6	5	.16	.18		+ .06	.51
16	6	.19	.18	± .00	+ .08	.50
15	5	.51	.53		+ .10	.57
5	6	.60	.59	+ .06	+ .07	.56
16	5	.40	.42		+ .20	.57
6	6	.44	.43	+ .01	+ .18	.55
17	5	.15	.17		+ .07	.37
3	6	.17	.16	– .01	+ .09	.51
18	5	.42	.44		+ .20	.64
4	6	.49	.48	+ .04	+ .24	.65
19	5	.58	.60		+ .17	.56
1	6	.69	.68	+ .08	+ .15	.57
20	5	.28	.30		+ .12	.56
2	6	.31	.30	± .00	+ .16	.62

De största skillnaderna i p-värde beroende på placering i provhäftet uppvisar uppgifterna 1/19, 3/17, 5/15, 15/5 och 19/1.⁸ Av dessa uppgifter har alla

⁷ p diff = (p för uppgift med lägre nummer) – (p för uppgift med högre nummer), dvs. i övre halvan av tabellen (p från häfte 5) – (p från häfte 6) och i nedre halvan av tabellen det omvända. Därmed betyder en positiv differens att lösningsproportionen var större när uppgiften placerats tidigt i provhäftet, medan en negativ differens betyder att lösningsproportionen var större när uppgiften lagts längre bak i häftet.

⁸ Uppgifterna namnges här och i den följande diskussionen enligt principen nummer i häfte 5/nummer i häfte 6.

utom en – 5/15 – ökat i svårighetsgrad när de placerats mot slutet av provhäftet. Uppgifternas förändrade placering tycks på uppgiftsnivå inte ha påverkat skillnaderna mellan mäns och kvinnors p-värden på något uppenbart eller entydigt sätt. Differenserna är genomgående små.

I tabell 1:11 redovisas motsvarande data från försöksomgången hösten 1998. Eftersom de två grupper som vid det tillfället genomförde försöksuppgifterna var praktiskt taget jämbördiga vad avser medelvärde på det reguljära provet och försöksuppgifterna, har någon korrigering av p-värdena inte behövt göras här.

Tabell 1:11 Erhållna p-värden hösten 1998 för de tolv försöksuppgifterna i häfte 5 och 6 samt differenser mellan dessa. Dessutom redovisas skillnaden mellan mäns och kvinnors p-värden samt rbis.

Uppgift nr	Häfte	p	p diff ⁹	p diff m/k	rbis
1	5	.57	+ .04	+ .16	.65
19	6	.53		+ .16	.59
2	5	.24	- .04	+ .05	.29
20	6	.28		+ .04	.32
3	5	.77	+ .09	+ .12	.57
17	6	.68		+ .15	.53
4	5	.17	+ .04	+ .06	.58
18	6	.13		+ .04	.53
5	5	.38	- .04	+ .09	.44
15	6	.42		+ .07	.44
6	5	.16	+ .03	+ .06	.53
16	6	.13		+ .06	.43
15	5	.51	+ .01	+ .06	.54
5	6	.52		+ .06	.52
16	5	.45	+ .01	+ .18	.53
6	6	.46		+ .18	.51
17	5	.16	+ .02	+ .06	.43
3	6	.18		+ .07	.48
18	5	.44	+ .01	+ .21	.66
4	6	.45		+ .23	.60
19	5	.53	+ .09	+ .16	.54
1	6	.62		+ .16	.56
20	5	.24	± .00	+ .10	.62
2	6	.24		+ .14	.58

⁹ p diff = (p för uppgift med lägre nummer) – (p för uppgift med högre nummer), dvs. i övre halvan av tabellen (p från häfte 5) – (p från häfte 6) och i nedre halvan av tabellen det omvända. Därmed betyder en positiv differens att lösningsproportionen var större när uppgiften placerats tidigt i provhäftet, medan en negativ differens betyder att lösningsproportionen var större när uppgiften lagts längre bak i häftet.

Vid försökstillfället hösten 1998 är det framför allt uppgifterna 3/17 och 19/1 som visar tydliga skillnader i p-värde, + .09 vardera, beroende på position. De har båda blivit klart svårare när de placerats mot slutet av provhäftet. Därefter kommer fyra uppgifter vars p-värden förändrats med .04, antingen uppåt eller nedåt. Vad gäller könsskillnaderna har förändringarna i position inte heller här haft några uppenbara effekter på uppgiftsnivå.

I tabell 1:12 jämförs differenserna i erhållna p-värden vid de två försöksomgångarna. Liksom i föregående tabeller har skillnaden beräknats utifrån korrigerade p-värden för de grupper som ingick i försöket våren 1998, medan övriga p-värden redovisats och använts utan justering.

Tabell 1:12 Jämförelse av differenser i erhållna p-värden för de tolv försökssuppgifterna våren och hösten 1998.

Uppg nr	1/19	2/20	3/17	4/18	5/15	6/16
p diff 98:04	+ .07	- .02	+ .07	+ .02	- .07	± .00
p diff 98:10	+ .04	- .04	+ .09	+ .04	- .04	+ .03

Uppg nr	15/5	16/6	17/3	18/4	19/1	20/2
p diff 98:04	+ .06	+ .01	- .01	+ .04	+ .08	± .00
p diff 98:10	+ .01	+ .01	+ .02	+ .01	+ .09	± .00

Som framgår av tabell 1:12 följer differenserna ett likartat mönster i de två försöksomgångarna. I de fall p-värdet förändras nämnvärt går förändringen åt samma håll vid båda tillfällena, och det typiska resultatet är att en uppgifts svårighetsgrad ökar med en placering mot slutet av provhäftet.

Uppgifterna 3/17 och 19/1 visar klara positiva skillnader i p-värde både våren och hösten 1998. Dessa uppgifter har två saker gemensamt: För det första är de de två lättaste bland de tolv försökssuppgifterna (genomsnittligt p-värde i försöksgrupperna är cirka .72 respektive .60). För det andra liknar de varandra till själva formatet. I båda fallen gäller det för provdeltagaren att, med hjälp av information som finns att hämta i figurmaterialet, avgöra vilket av de fem svarsförslagen som uppfyller ett i frågestammen angivet villkor. Sådana uppgifter kan förmodas inbjuda till att systematiskt gå igenom och pröva de fem svarsförslagen ett efter ett. Även om det stämmer att den här typen av uppgifter verkligen är "känsligare" är andra förändringar i position, går det knappast att utifrån det här materialet säga något om orsaken. Tar dessa uppgifter möjligen jämförelsevis lång tid att lösa, och vållar de av den orsaken större svårigheter mot slutet av provtiden än i början?

Uppgiften 5/15 och i viss mån även 2/20 avviker från det gängse mönstret i och med att deras svårighetsgrad vid båda försökstillfällena minskat när de placerats mot slutet av provhäftet. En tänkbar förklaring till den uppenbara skillnad som uppgift 5/15 visar är att det i det material som föregått den när den placerats som nummer 15 kan ha ingått något som indirekt gett vägledning till dess lösning. För fyra av de tolv uppgifterna – 6/16, 16/6, 17/3 och 20/2 – förefaller positionen sakna betydelse för svårighetsgraden.

I tabell 1:13 sammanfattas resultaten av de två genomförda försöken med ändrad uppgiftsposition.

Tabell 1:13 Sammanfattande resultat för 2 x 6 DTK-uppgifter placerade i början respektive i slutet av ett provhäfte (korrigerade värden från provtillfället 98:04).

Provtillfälle 98:04	Början av häfte 5		Slutet av häfte 6		Diff	
	Σ	M	Σ	M	Σ	M
p totalt	2.44	.41	2.37	.40	.07	.01
p män	2.76	.46	2.65	.44	.11	.02
p kvinnor	2.21	.37	2.14	.36	.07	.01
rbis	–	.48	–	.50	–	–.02
Provtillfälle 98:04	Början av häfte 6		Slutet av häfte 5		Diff	
	Σ	M	Σ	M	Σ	M
p totalt	2.64	.44	2.46	.41	.18	.03
p män	3.12	.52	2.96	.49	.16	.03
p kvinnor	2.24	.37	2.09	.35	.15	.02
rbis	–	.58	–	.54	–	.04
Provtillfälle 98:10	Början av häfte 5		Slutet av häfte 6		Diff	
	Σ	M	Σ	M	Σ	M
p totalt	2.29	.38	2.17	.36	.12	.02
p män	2.58	.43	2.45	.41	.13	.02
p kvinnor	2.04	.34	1.93	.32	.11	.02
rbis	–	.51	–	.47	–	.04
Provtillfälle 98:10	Början av häfte 6		Slutet av häfte 5		Diff	
	Σ	M	Σ	M	Σ	M
p totalt	2.47	.41	2.33	.39	.14	.02
p män	2.91	.48	2.75	.46	.16	.02
p kvinnor	2.07	.34	1.98	.33	.09	.01
rbis	–	.54	–	.55	–	–.01

Sammanställningen i tabell 1:13 bekräftar framför allt att svårighetsgraden tenderar att öka när en uppgift flyttas från början till slutet av provhäftet.

Vidare märks en svag men genomgående tendens som inte var synlig på uppgiftsnivå: att svårighetsgraden ökar mer bland män än bland kvinnor, vilket i sin tur gör att könsskillnaden blir något mindre när uppgifter placeras i slutet av provhäftet.

Slutkommentar

Den försiktiga slutsats som kan dras av den studie som beskrivits här är att positionen i provhäftet tycks påverka svårighetsgraden för vissa uppgifter i delprovet DTK, och att graden av denna påverkan förefaller olika bland kvinnor och bland män. I detta försök var det dock bara enstaka uppgifter som visade klara skillnader i p-värde beroende på position. Att dessa uppgifter hade gemensamma drag skulle kunna bilda utgångspunkt för en ny undersökning. Genom att exempelvis använda material där svårighetsgraden redan är känd – dvs. redan utprovade uppgifter – skulle man då vidare kunna utforska tendensen att uppgifter med jämförelsevis högt p-värde förefaller känsligare för förändringar i position än andra.

1.4 Delprovet LÄS

LÄS-provet innehåller för närvarande 20 uppgifter och har en provtid på 50 minuter. Det består av fem olika texter med fyra uppgifter till varje. Texterna följer ämnesmässigt de riktlinjer som gäller för högskoleprovet i dess helhet och avser därmed att täcka fem övergripande områden. De är: 1) kultur (KU), 2) vård (VÅ), 3) naturvetenskap/teknik (NA/TE), 4) administration, ekonomi, samhälle (AES) och 5) utbildning och yrken (UTB/YRK). Detta innebär i praktiken att LÄS innehåller en text från vart och ett av de fem områdena.

I det utprovningssystem som tillämpades fram till 1996 placerades alltid de fem texter som skulle utgöra nästa reguljära LÄS-prov längst fram i sina respektive utprovningssversioner. (Texterna/uppgifterna hade redan valts ut efter en tidigare utprovning och provades därefter på nytt för att få säkrare prognosvärden.) Antagandet var att provdeltagarna vid utprovningen, såväl som vid reguljära tillfället, var mer alerta vid provets början och möjligen också lade ned mer tid på den första texten än den sista. I det nya utprovningssystemet är möjligheterna till omprovning däremot små. Man måste nöja sig med de data man erhåller vid första utprovningstillfället och har inte möjligheten att pröva texterna i någon ny inbördes ordning. Å andra sidan har en mer genomgripande lokal och nationell granskning av materialet – före utprovning – avsevärt minskat behovet att revidera i uppgifter och text, vilket i det avseendet gjort prognosvärdena mera stabila.

Men en jämförelse av utprovnings- och reguljärdatabaser har ändå indikerat förändringar i uppgifters p-värden. När en text som utprovats som förstasida läggs som sista eller näst sista text i en reguljär provversion verkar svårighetsgraden hos de tillhörande uppgifterna förändras, även om inga som helst revideringar gjorts i själva materialet. Det händer alltså något mellan utprovning och reguljärt provtillfälle om placeringen av texter/uppgifter förändras. En hypotes kan vara att tidspressen är hårdare mot slutet av provtiden, dvs att vi har en viss speedeffekt på LÄS-provet.

Metod

För att få ett första begrepp om det föreligger någon ”positionseffekt” – och därigenom indirekt en speedeffekt – på LÄS-provet genomfördes vid de reguljära provtillfällena hösten 1997 och våren 1998 några försök med förändrad text- och uppgiftsposition. Hösten 1997 sammanställdes fem texter (med fyra uppgifter till varje text) till försöksprov 7. I försöksprov 8 placerades texterna i omvänd ordning. (Text nummer 3 blev därmed densamma i bägge proven.)

Försöksproven genomfördes på var sin provort i samband med det reguljära provtillfället hösten 1997. Försöksprov 7 genomfördes av 1 516 personer varav 660 män och 856 kvinnor. Försöksprov 8 genomfördes av 1 882 personer varav 811 män och 1 071 kvinnor.

Våren 1998 upprepades försöket med provhäftena 7 och 8 på två nya provorter. Försöksprov 7 genomfördes då av 2 452 personer varav 1 137 män och 1 315 kvinnor. Försöksprov 8 genomfördes av 1 430 personer varav 645 män och 785 kvinnor.

Våren 1998 sammanställdes dessutom ytterligare fem texter (med tillhörande uppgifter) till nya försöksprov, nummer 9 och 10. Precis som vid tidigare försök var provhäfte 10 en omvänd version av provhäfte 9. Försöksproven genomfördes på var sin provort. Försöksprov 9 genomfördes av 1 042 deltagare varav 501 män och 541 kvinnor. Försöksprov 10 genomfördes av 2 409 deltagare, 1 072 män och 1 337 kvinnor.

Resultat

Tabellerna 1:14, 1:15 och 1:16 visar utfallet av försöken hösten 1997 och våren 1998 i form av p-värden, dvs. summan av p-värdena för de fyra uppgifterna per text. P-värdena är inte korrigerade med avseende på att de två försöksproven (i varje enskilt försök) genomfördes av två olika grupper. Men av alla tre tabellerna framgår att de två deltagande gruppernas prestationer på samma reguljära LÄS-prov i stort sett kan anses likvärdiga. Med positionsförändring avses den skillnad i p-värdessumman som uppstår som ett resultat av textens - och tillhörande uppgifters - förändrade position.

Tabell 1:14 visar resultaten från försöket hösten 1997. Som vi ser i kolumnen "positionsförändringar" kan man utläsa att om en text växlar mellan position 1 och 5 i provhäftet sker en notabel förändring av p-värdet för de tillhörande uppgifterna.

Tabell 1:15 bekräftar samma tendenser när försöket från hösten 1997 upprepades på två nya provgrupper.

Tabell 1:16 visar resultaten från försöket med provhäftena 9 och 10 våren 1998. Resultaten är analoga med de från de tidigare försöken.

Tabell 1:14 Resultat av omvänd ordning i text- och uppgiftsplacering. Försöksprov 7 och 8, hösten 1997.

Text	LÄS 7		LÄS 8		Positions- förändring
	Posi- tion	Tot p-värde 4 uppgifter	Posi- tion	Tot p-värde 4 uppgifter	
Journalism...	1	1.84	5	1.57	-.27 1? 5
Entreprenören...	2	2.30	4	2.01	-.29 2? 4
Dansbaneeländet	3	2.50	3	2.45	-.05 3? 3
Atomteorin	4	2.57	2	2.67	+.10 4? 2
Rehabilitering	5	2.33	1	2.94	+.64 5? 1
Tot försöks- provet		11.02		11.67	
Gruppens resultat på reguljära LÄS		12.04		12.24	

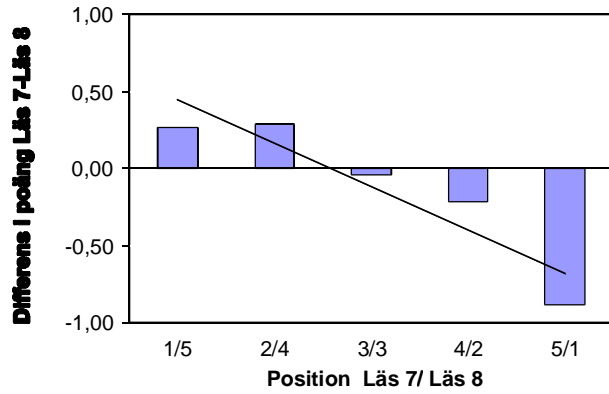
Tabell 1:15 Resultat av omvänd ordning i text- och uppgiftsplacering. Försöksprov 7 och 8, våren 1998.

Text	LÄS 7		LÄS 8		Positions- förändring
	Posi- tion	Tot p-värde 4 uppgifter	Posi- tion	Tot p-värde 4 uppgifter	
Journalism...	1	1.91	5	1.64	-.27 1? 5
Entreprenören...	2	2.30	4	2.16	-.14 2? 4
Dansbaneeländet	3	2.50	3	2.50	.00 3? 3
Atomteorin	4	2.57	2	2.76	+.19 4? 2
Rehabilitering	5	2.33	1	2.99	+.66 5? 1
Tot försöks- provet		11.61		12.05	
Gruppens resultat på reguljära LÄS		11.13		11.12	

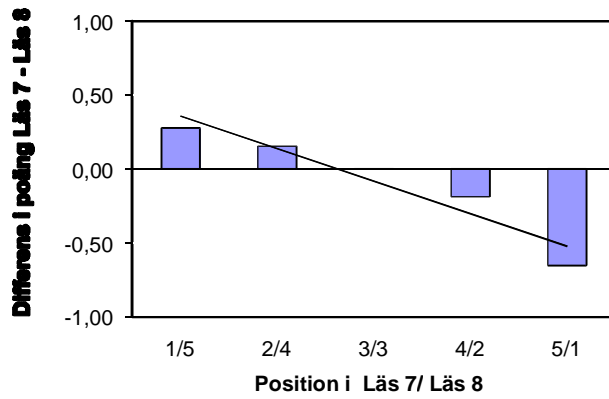
Tabell 1:16 Resultat av omvänd ordning i text- och uppgiftsplacering. Försöksprov 9 och 10, våren 1998.

Text	LÄS 9		LÄS 10		Positions- förändring
	Posi- tion	Tot p-värde 4 uppgifter	Posi- tion	Tot p-värde 4 uppgifter	
Civil olydnad	1	2.05	5	1.88	-.17 1? 5
När jordens aska...	2	2.50	4	2.28	-.22 2? 4
Musicerande...	3	2.37	3	2.28	-.09 3? 3
Att undvika re- former	4	1.88	2	1.93	+.05 4? 2
Flickorna från Maria	5	2.33	1	2.87	+.54 5? 1
Tot försöksprovet		11.13		11.24	
Gruppens resultat på reguljära LÄS		10.97		11.10	

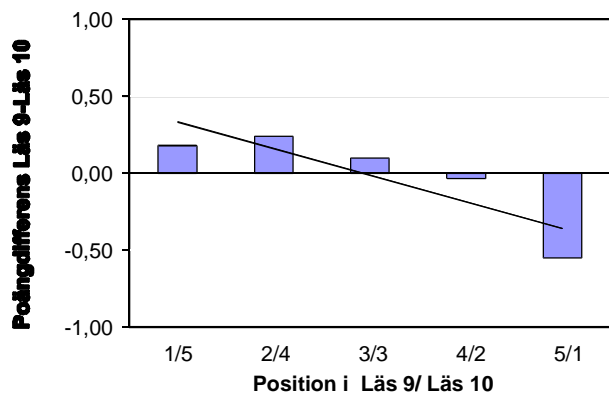
Förändringarna kan också åskådliggöras grafiskt. Figuren 1:2 till 1:4 visar hur p-värdena för varje text (dess fyra uppgifter) förändras beroende på position. Framför allt framstår effekten av förändringen från position 1 till 5 tydligt. De tre försöksomgångarna indikerar samtliga att en texts position i provhäftet påverkar de tillhörande uppgifternas svårighetsgrad och därmed provresultatet.



Figur 1:2 Omvänd ordning på 5 lästexter. Läs 7 – Läs 8, 97B.



Figur 1:3 Omvänd ordning på 5 lästexter. Läs 7 – Läs 8, 98A.



Figur 1:4 Omvänd ordning på 5 lästexter. Läs 9 – Läs 10, 98A.

Nytt försök

För att få ytterligare vetskap om positionseffekten användes texterna och uppgifterna i försöksproven 9 och 10 till att utforma en ny försöksdesign. Denna framgår av tabell 1:17. Alla fem försöksproven, numrerade 11 till och med 15, innehöll samma fem texter med tillhörande uppgifter, men textpositionen varierade på ett konsekvent sätt så att alla texter utprovades i alla fem tänkbara positioner, 1 till 5. Teoretiskt kan man således behandla de fem texterna i en och samma position som ett provhäfte, trots att texterna i praktiken ingått i fem olika häften. Om vi läser tabellen vågrätt får vi alltså fem prov, där gemensamt för de fem ingående texterna är att de utprovats i en och samma position.

Tabell 1:17 Sammansättningen av de fem försöksproven LÄS 11 till och med 15, med varierande text- och uppgiftsplacering (hösten 1998).

Position	LÄS 11	LÄS 12	LÄS 13	LÄS 14	LÄS 15
1	Civil olydnad	Flickorna från Maria	Att undvika reformer	Musicerande som samspel	När jordens aska sprids
2	När jordens aska sprids	Civil olydnad	Flickorna från Maria	Att undvika reformer	Musicerande som samspel
3	Musicerande som samspel	När jordens aska sprids	Civil olydnad	Flickorna från Maria	Att undvika reformer
4	Att undvika reformer	Musicerande som samspel	När jordens aska sprids	Civil olydnad	Flickorna från Maria
5	Flickorna från Maria	Att undvika reformer	Musicerande som samspel	När jordens aska sprids	Civil olydnad

Försöksproven 11 till och med 15 utprovades hösten 1998 i fem olika provgrupper (och därmed på fem olika provorter) i samband med genomförandet av det reguljära högskoleprovet.

Försöksprov 11 genomfördes av 2 628 deltagare, 1 161 män och 1 467 kvinnor. Försöksprov 12 genomfördes av 1 404 deltagare varav 641 män och 763 kvinnor. Försöksprov 13 gavs för 2 719 deltagare varav 1 198 män och 1 521 kvinnor. Försöksprov 14 genomfördes av 2 994 deltagare varav 1 466 män

och 1 528 kvinnor. Försöksprov 15, slutligen, genomfördes av 2 950 deltagare, 1 349 män och 1 601 kvinnor. Tabell 1:18 visar utfallet. De fem "teoretiska" proven varierar i p-värde från 11.74 (position 1) ned till 10.71 (position 5); en prydligt fallande "trappa".

Tabell 1:18 Resultat av omvänd ordning i text- och uppgiftsplacering. Försöksprov 11, 12, 13, 14 och 15. Hösten 1998.

Position	LÄS 11	LÄS 12	LÄS 13	LÄS 14	LÄS 15	Tot.	Tot. män	Tot. kv	Diff m-k
1	2.10	2.78	2.01	2.30	2.55	11.74	11.89	11.62	.27
2	2.48	1.97	2.67	1.97	2.40	11.49	11.72	11.31	.41
3	2.34	2.37	1.95	2.65	2.09	11.40	11.48	11.29	.19
4	1.99	2.28	2.26	1.89	2.67	11.09	11.21	11.04	.17
5	2.53	1.90	2.17	2.07	2.04	10.71	10.77	10.64	.13
Tot. försök	11.44	11.29	11.05	10.88	11.75	Diff pos. 1 → 5			
Tot. reg. prov	10.84	10.39	10.47	10.34	11.23	1.03	1.12	.98	

P-värdena bör dock korrigeras med avseende på skillnaderna mellan de fem olika provgruppernas prestationsnivåer. Alla grupperna genomförde ju även det reguljära LÄS-provet, och deras respektive resultat på detta prov utgör riktnivån för korrigeringen. I tabell 1:19 och 1:20 redovisas alltså samma data som i tabell 1:18, men i form av korrigerade värden. Utfallet är dock i princip detsamma.

Notabelt är också att effekten av positionsförändringen, uttryckt i poängmedelvärde, är relativt könsneutral. Kvinnor kan t.ex. inte anses missgynnade av effekten. I tabell 1:18 framgår att skillnaden för hela försöksprovet mellan position 1 och 5 är 1.03 för alla provtagare, 1.12 för männen och .98 för kvinnorna. Könsskillnaden (för alla fem texterna) i position 1 är .27 jämfört med .13 i position 5. Det verkar alltså snarare som om könsskillnaden minskar något med den positionsförändringen.

Tabell 1:19 Resultat av omvänd ordning i text- och uppgiftsplacering. Försöksprov 11, 12, 13, 14 och 15. Hösten 1998. Korrigerade värden.

Position	LÄS 11	LÄS 12	LÄS 13	LÄS 14	LÄS 15	ptot
1	2.10	2.86	2.09	2.42	2.47	11.94
2	2.48	2.05	2.75	2.09	2.32	11.69
3	2.34	2.45	2.03	2.77	2.01	11.60
4	1.99	2.36	2.34	2.01	2.59	11.29
5	2.53	1.98	2.25	2.19	1.96	10.91
Prov ptot	11.44	11.70	11.46	11.48	11.35	

Om man sedan tar varje text för sig och följer den över alla fem positionerna (tabell 1:20) kan man tydligt se att p-värdet för texten (dvs. summan av de fyra uppgifternas p-värde) hela tiden sjunker ju längre bak i provet texten ligger.

Tabell 1:20 Summa p-värde och genomsnittlig lösningsproportion när varje text (4 uppgifter) placerats i början respektive i slutet av försöksproven samt differens. Totalt och uppdelat på kön samt medelvärde för rbis.

	Början		Slutet		Diff	
Civil olydnad	Σ	M	Σ	M	Σ	M
p totalt	2.10	.52	1.96	.49	.14	.03
p män	2.17	.54	2.03	.51	.14	.03
p kvinnor	2.06	.52	1.90	.48	.16	.04
rbis	–	.49	–	.50	–	–.01
	Början		Slutet		Diff	
När jordens aska	Σ	M	Σ	M	Σ	M
p totalt	2.86	.72	2.53	.63	.33	.09
p män	2.81	.70	2.46	.62	.35	.08
p kvinnor	2.88	.72	2.58	.64	.30	.08
rbis	–	.36	–	.39	–	–.03
	Början		Slutet		Diff	
Musicerande	Σ	M	Σ	M	Σ	M
p totalt	2.09	.52	1.98	.50	.11	.02
p män	2.08	.52	1.95	.49	.13	.03
p kvinnor	2.11	.53	2.00	.50	.11	.03
rbis	–	.35	–	.35	–	.00
	Början		Slutet		Diff	
Att undvika reformer	Σ	M	Σ	M	Σ	M
p totalt	2.42	.60	2.25	.56	.17	.04
p män	2.42	.60	2.20	.55	.22	.05
p kvinnor	2.44	.61	2.29	.57	.15	.04
rbis	–	.26	–	.26	–	.00
	Början		Slutet		Diff	
Flickorna från Maria	Σ	M	Σ	M	Σ	M
p totalt	2.47	.62	2.19	.55	.28	.07
p män	2.61	.65	2.33	.58	.28	.07
p kvinnor	2.34	.58	2.07	.52	.27	.06
rbis	–	.38	–	.38	–	.00

Diskussion

De försök som genomförts och här redovisats visar att text/uppgiftsposition har betydelse för svårighetsgraden, uttryckt i p-värden. Ju längre bak i provhäftet en och samma text med tillhörande uppgifter läggs, ju lägre blir p-värdet. De förändringar i p-värden som noterats är däremot inte stora nog för

att man ska kunna säga att LÄS-provet är ett speedat prov, enligt vedertagna definitioner. Men de är så stora att man – för parallellitetens skull – bör försöka hålla positionsfaktorn under kontroll.

Redan på basis av de resultat som här framkommit kan vissa åtgärder genomföras för att minimera effekten av förändrad textposition. Delprovet LÄS följer liksom övriga delprov, vilket tidigare nämnts, de riktlinjer i ämnestäckning som fastlagts sedan högskoleprovets införande. Det är fem områden som ska finnas representerade i provet: 1) kultur (KU), 2) vård (VÅ), 3) naturvetenskap/teknik (NA/TE), 4) administration, ekonomi, samhälle (AES) och 5) utbildning och yrken (UTB/YRK).

Då LÄS-provet för närvarande innehåller fem texter (med fyra uppgifter till varje) är det mycket enkelt att täcka av dessa områden – en text för varje område. Från och med hösten 1998 utprövas också texterna i den ordning som angivits ovan. I varje LÄS-prov (utprövningsversioner såväl som reguljära provhäften) placeras KU-texter först, VÅ-texter som nummer två och så vidare. Det innebär att texterna vid sammanställning av ett reguljärt prov får samma positioner som de utprövats i. Vill man byta ut en text mot en annan inom samma ämnesområde, så har man med hjälp av ovan beskrivna åtgärd försäkrat sig om att ersättningstexten utprövats i samma position som den gamla, och att utprövningsdata därmed blir mer tillförlitliga för prognosen.

2. Har det rätta svarets position bland svarsförslagen någon betydelse för uppgiftens svårighetsgrad?

I föreliggande kapitel presenteras resultat från två delstudier, som med utgångspunkt i delproven ORD (avsnitt 2.1) och DTK (avsnitt 2.2), har undersökt om det rätta svarets position bland svarsförslagen har någon betydelse för uppgiftens svårighetsgrad.

2.1 Delprovet ORD

Syftet med denna studie var att undersöka om det har någon betydelse för ORD-uppgifternas relativa svårighetsgrad om det rätta svaret presenteras som första (A) eller sista (E) svarsförslag.

Metod

Vid två provtillfällen, våren 1998 (98A) och hösten 1998 (98B), genomfördes försök för att ta reda på om det rätta svarets position i en ORD-uppgift påverkar dess svårighetsgrad. Hypotesen var att uppgifter där det rätta svaret presenterades som svarsförslag A skulle få en högre lösningsproportion, dvs. bli lättare, än uppgifter där det rätta svaret presenterades som svarsförslag E.

Material

Vid det första försöket, 98A, prövades fem uppgifter ut i två provhäften, häfte 4 och 5. Det var exakt samma uppgifter med exakt samma svarsförslag. De hade också exakt samma position, dvs. uppgiftsnummer, i de båda häftena. Skillnaden var att det rätta svaret presenterades som svarsförslag A i det ena häftet och som svarsförslag E¹ i det andra häftet. De fem uppgifterna utgjordes av tre substantiv, ett verb och ett adjektiv. Av sekretesskäl presenteras inte uppgifterna här utan kallas hädanefter S7, S8, S9, V3 respektive A2. Uppgift S7 representerar gruppen hyponymer, dvs. den speciella betydelserelation som finns mellan ord som *tulpan* (hyponym) och *blomma* (hyperonym). Det ena ordet betecknar alltså en delmängd av det andra ordet. De fyra övriga uppgifterna representerar gruppen synonymer, dvs. ord som betyder samma eller nästan samma sak som uppgiftsordet. De här fyra uppgifterna är också exempel på ord där man kan göra en s.k. morfemanalys, dvs. dela upp ordet i mindre delar, där var och en har sin egen betydelse. Detta kan vara till hjälp för att identifiera hela ordets betydelse.

Vid provtillfället 98 B gjordes ett nytt försök med nio uppgifter. Tre av dem ingick i häfte 2 och 3. De övriga ingick i häfte 4 och 5. De tre uppgifterna i häfte 2 och 3 utgjordes av ett substantiv och två verb, hädanefter kallade

¹ I ett fall som D.

S10, V4 och V5. Två av dem, S10 och V5, är främmande ord medan V4 är ett s.k. arvord, dvs. ett ord som sedan urminnes tider funnits i vårt språk. För uppgift V5 kan en morfemanalys göras. I häfte 4 och 5 gjordes ett försök med sex uppgifter, tre substantiv, två verb och ett adjektiv, hädanefter kallade S11, S12, S13, V6, V7 respektive A3. Uppgifterna S11 och V6 tillhör arvorden medan de fyra övriga orden tillhör gruppen främmande ord för vilka en morfemanalys kan göras.

Försökspersoner

Försökspersonerna vid provtillfället 98A utgjordes av två olika utprövningsgrupper. Den ena gruppen, som genomförde häfte 4, bestod av 1 859 personer, 1 040 kvinnor och 819 män. Den andra gruppen, som genomförde häfte 5, bestod av 1 991 personer, 1 084 kvinnor och 907 män. Tabell 2:1 visar att båda dessa grupper presterade något sämre på det reguljära ORD-provet än den totala provdeltagargruppen. Gruppen som genomförde häfte 4 hade en medelvärdeskillnad på $-.71$ jämfört med totalgruppen. Gruppen som genomförde häfte 5 hade en medelvärdeskillnad på -1.08 jämfört med totalgruppen. Med hjälp av dessa medelvärdeskillnader korrigeras lösningsproportionerna för uppgifter från häfte 4 med $+.02$ och för uppgifter från häfte 5 med $+.03$. För en beskrivning av försökspersonerna vid provtillfället 98B, se kapitel 1, sid. 4–5.

Resultat

I tabell 2:1 presenteras de huvudsakliga resultaten av det försök som genomfördes 98A. Lösningproportionerna, p-värdena, är korrigerade för gruppernas resultat på det reguljära ORD-provet.

Tabell 2:1 Lösningproportioner (98A) totalt och för män/kvinnor samt biseriala korrelationer för fem ORD-uppgifter då det rätta svaret presenterades som svarsförslag A eller E.

Uppgift	Nr	Rätt Svar	Häfte	p korr	p korr män	p korr kv	rbis
S7	2	A	4	.93	.95	.91	.34
	2	E	5	.97	1.00	.93	.45
S8	21	A	4	.80	.74	.84	.50
	21	E	5	.71	.63	.77	.43
V3	25	A	4	.60	.67	.54	.41
	25	E	5	.47	.51	.42	.49
S9	28	A	5	.38	.36	.40	.44
	28	D	4	.30	.30	.30	.45
A2	38	A	4	.97	.99	.95	.74
	38	E	5	.95	.97	.94	.64

Tabell 2:1 visar att tre av uppgifterna, S8, V3 och S9, blev klart lättare när det rätta svaret presenterades som svarsförslag A. Könsskillnaderna förändrades endast marginellt för dessa tre uppgifter. De biseriala korrelationerna, rbisvärdena, ökade för en av dessa uppgifter, V3, och minskade för en annan uppgift, S8, då det rätta svaret presenterades som svarsförslag E. De två övriga uppgifterna, S7 och A2, blev så lätta att det inte är meningsfullt att uttala sig om dem.

I tabell 2:2 presenteras de huvudsakliga resultaten av det försök som genomfördes 98B. Lösningproportionerna, p-värdena, är korrigerade för gruppernas resultat på det reguljära ORD-provet.

Tabell 2:2 Lösningproportioner (98B) totalt och för män/kvinnor samt biseriala korrelationer för nio ORD-uppgifter då det rätta svaret presenterades som svarsförslag A eller E.

Uppgift	Nr	Rätt svar	Häfte	p korr	p korr män	p korr kv	rbis
S10	7	A	3	.83	.86	.81	.62
	7	E	2	.78	.80	.76	.69
V4	26	A	2	.94	.95	.94	.48
	26	E	3	.94	.95	.93	.44
V5	29	A	3	.88	.91	.85	.73
	29	E	2	.84	.87	.82	.71
V6	5	A	4	.46	.47	.46	.54
	5	E	5	.27	.28	.26	.48
S11	8	A	4	.75	.85	.69	.59
	8	E	5	.78	.86	.72	.59
A3	12	A	4	.47	.53	.43	.59
	12	E	5	.41	.48	.35	.55
V7	33	A	5	.18	.20	.17	.36
	33	E	4	.22	.23	.21	.33
S13	35	A	5	.96	.98	.95	.50
	35	E	4	.94	.97	.92	.61
S12	17	A	4	.83	.85	.81	.61
	17	E	5	.76	.75	.77	.56

Av tabell 2:2 framgår att fem av uppgifterna, S10, S12, S13, V6 och A3, blev lättare då det rätta svaret presenterades som svarsförslag A. Uppgift V7 uppvisar ett helt annat mönster. Denna uppgift blev något lättare då det rätta svaret presenterades som svarsförslag E. Denna uppgift är emellertid så svår att en del av lösningproportionen kan tillskrivas gissning vilket innebär att den bör utelämnas från detta resonemang. De tre övriga uppgifterna, S13, V4

och V5, blev så lätta att det inte är meningsfullt att uttala sig om dem. Könsskillnaderna är desamma i samtliga fall utom för S12 och V5 där de ändrat riktning från manligt till kvinnligt. De biseriala korrelationerna, rbisvärdena, minskade för sex uppgifter då det rätta svaret presenterades som svarsförslag E.

Tabell 2:3 Sammanfattande resultat för fem ORD-uppgifter (98A) och nio ORD-uppgifter (98B) då det rätta svaret placerats i början respektive i slutet bland svarsförslagen i en uppgift.

	Början		Slutet		Diff	
	Σ	M	Σ	M	Σ	M
98A						
p totalt	3.68	.74	3.40	.68	.28	.06
p män	3.71	.74	3.41	.68	.30	.06
p kvinnor	3.64	.73	3.36	.67	.28	.06
rbis	–	.49	–	.49	–	.00
98B						
p totalt	6.30	.70	5.94	.66	.36	.04
p män	6.60	.73	6.19	.69	.41	.04
p kvinnor	6.11	.68	5.74	.64	.37	.04
rbis	–	.56	–	.55	–	.01

Sammanfattningsvis kan sägas att vad beträffar det rätta svarets position i en ORD-uppgift visar dessa försök att det tycks ha betydelse var det placeras, speciellt om det är en uppgift där en morfemanalys kan göras. Sådana uppgifter blir i de flesta fall lättare då det rätta svaret presenteras som svarsförslag A. Detta är dock ett ganska begränsat försök med få uppgifter och bör därför upprepas för att man skall kunna bekräfta resultatet.

2.2 Delprovet DTK

Som nämndes inledningsvis kan det ibland bli nödvändigt att efter utprovning revidera en provuppgift så att det rätta svarets position bland svarsförslagen ändras. Syftet med den här studien var att undersöka om och hur en sådan förändring påverkar svårighetsgraden, dvs. lösningsproportionen (p-värdet), hos uppgifter i delprovet DTK.

Metod

Undersökningen genomfördes med hjälp av försöksprov vid två reguljära provtillfällen i följd, våren respektive hösten 1998. Försöksprovet hade samma format som såväl utprovningssversionerna som de reguljära versionerna av DTK. Det innehöll därmed tio figuruppsättningar och två uppgifter till varje med vardera fem givna svarsförslag, A–E. Provtiden var 50 minuter. Försöksprovet gavs vid båda tillfällena i två varianter, kallade häfte 5 och häfte 6, till var sin grupp provdeltagare. Av försöksprovets 20 uppgifter ägnades åtta – uppgifterna 7 till och med 14 – åt att undersöka betydelsen av det rätta svarets position.² Innehållet i uppgifterna 7–14 och det tillhörande figurmaterialet var identiskt i försöksprovets båda varianter, med undantag för att uppgifternas svarsförslag lagts i omvänd ordning i häfte 6 jämfört med häfte 5. Ett A-svar i häfte 5 låg därmed som E-svar i häfte 6 och ett D-svar i häfte 5 som B-svar i häfte 6. Svarsförslag C var detsamma i båda häftena.³

I försöksprovet ingick både nykonstruerat och tidigare utprovat, reviderat material som vad gäller figurtyper och ämnesområden kombinerats och varierats enligt samma principer som gäller för utprovningshäften och reguljära DTK-prov. Valet av det material som ingick i försöket gjordes vidare utifrån tanken att det rätta svarets position skulle kunna ha större betydelse i vissa typer av uppgifter än i andra. Exempelvis är så kallade beskrivningsuppgifter en återkommande uppgiftstyp i DTK. Där gäller det för deltagaren att med hjälp av information ur figurmaterialet avgöra vilket av de fem svarsförslagen (till exempel fem län, fem fågelarter eller fem årtal) som bäst svarar mot en beskrivning som ges i uppgiften. En sådan uppgift inbjuder förmodligen till att gå igenom och ”kolla av” svarsförslagen ett efter ett i högre grad än en uppgift där deltagaren ombeds avläsa/beräkna en procentuell förändring eller liknande och sedan ska välja rätt svar i form av ett numeriskt värde. De uppgifter som valdes ut att ingå i försöksprovet varierades därför även utifrån denna aspekt.

² Övriga uppgifter avsåg att testa i vilken mån en DTK-uppgifts svårighetsgrad påverkas av dess position i provhäftet, se vidare kapitel 1.

³ Ingen av de uppgifter som ingick i försöket hade för övrigt C som rätt svar.

Insamlade data bearbetades på samma sätt som sker efter utprovning av preliminärt material och genomföranden av reguljära prov. Därefter gjordes en traditionell itemanalys.

Försökspersoner

Som tidigare nämnts genomfördes försöksprovet vid två provtillfällen och av totalt fyra olika grupper provdeltagare. I tabell 2:4 redovisas antal prövande, medelvärde på det genomförda reguljära DTK-provet (98A respektive 98B) jämfört med medelvärdet för riket samt medelvärde på de åtta försöksuppgifterna för var och en av de fyra grupperna.

Tabell 2:4 Försöksgrupperna och deras medelvärden på det reguljära DTK-provet respektive på de åtta försöksuppgifterna.

Prov-tillfälle	Häfte	N	Män	Kvinnor	M reg prov (riket)	M försöksprov, uppg 7–14
98:04	5	2 045	875	1 170	11.27 (11.76)	4.45
98:04	6	7 483	3 398	4 085	11.87 (11.76)	4.75
98:10	5	1 842	837	1 005	10.49 (10.49)	4.37
98:10	6	1 021	467	554	10.40 (10.49)	4.37

Som framgår av tabell 2:4 hade den grupp som våren 1998 genomförde häfte 6 ett bättre resultat, både på det reguljära DTK-provet och på försöksuppgifterna, än den grupp som vid samma tillfälle genomförde häfte 5. Grupperna som genomförde försöksproven hösten 1998 låg däremot på nästan exakt samma nivå. Medelvärdeskillnaden mellan grupperna som deltog i försöket våren 1998 måste beaktas och vägas in vid granskningen av resultaten. I tabell 2:5, där de data som samlades in våren 1998 redovisas på uppgiftsnivå, har vart och ett av p-värdena därför korrigerats utifrån de två gruppernas prestation på det reguljära DTK-provet i relation till genomsnittet för riket. För den grupp som genomförde häfte 5 har p-värdena ökat med .02 medan de för den grupp som genomförde häfte 6 har minskats med .01.

Resultat

Tabell 2:5 Erhållna p-värden våren 1998 för de åtta försöksuppgifterna i häfte 5 och 6 samt differenser mellan dessa. Dessutom redovisas skillnaden mellan mäns och kvinnors p-värden samt rbis.

Uppg nr	Häfte	Rätt svar	p	p (korr)	p diff ⁴ (korr)	p diff m/k	rbis
7	5	B	.57	.59	+ .05	– .01	.31
	6	D	.56	.55		+ .03	.36
8	5	E	.38	.40	+ .05	+ .18	.59
	6	A	.46	.45		+ .17	.56
9	5	A	.74	.76	– .04	+ .22	.62
	6	E	.81	.80		+ .15	.60
10	5	D	.15	.17	+ .05	+ .11	.66
	6	B	.23	.22		+ .15	.67
11	5	D	.74	.76	+ .03	+ .11	.43
	6	B	.80	.79		+ .07	.48
12	5	B	.59	.61	± .00	+ .24	.58
	6	D	.62	.61		+ .17	.60
13	5	A	.82	.84	+ .06	+ .05	.48
	6	E	.79	.78		+ .04	.54
14	5	E	.46	.48	– .01	+ .21	.63
	6	A	.48	.47		+ .20	.62

De största skillnaderna i p-värde uppvisar uppgifterna 7, 8, 10 och 13. Dessa skillnader går alla åt samma håll och visar att uppgiftens svårighetsgrad ökat när rätt svar haft positionen D eller E. Könsskillnaden vad avser p-värdet tycks variera i omfattning beroende på det rätta svarets position, men något entydigt mönster går inte att urskilja. Här finns exempel på att könsskillnaden ökat när rätt svar varit D eller E, men också på motsatsen.

I tabell 2:6 redovisas de motsvarande data som samlades in hösten 1998. Eftersom de två grupper som vid det tillfället genomförde försöksuppgifterna var praktiskt taget jämbördiga vad avser medelvärde på det reguljära provet och försöksuppgifterna, har någon korrigering av p-värdena inte behövt göras här.

⁴ p diff = (p för uppgift med rätt svar A/B) – (p för uppgift med rätt svar D/E). Därmed betyder en positiv differens att lösningsproportionen var större när rätt svar var A eller B, medan en negativ differens betyder att lösningsproportionen var större med D eller E som rätt svar.

Tabell 2:6 Erhållna p-värden hösten 1998 för de åtta försöksuppgifterna i häfte 5 och 6 samt differenser mellan dessa. Dessutom redovisas skillnaden mellan mäns och kvinnors p-värden samt rbis.

Uppg nr	Häfte	Rätt svar	p	p diff ⁵	p diff m/k	rbis
7	5	B	.58	± .00	+ .02	.33
	6	D	.58		+ .06	.33
8	5	E	.40	+ .02	+ .14	.57
	6	A	.42		+ .11	.50
9	5	A	.77	- .01	+ .20	.63
	6	E	.78		+ .16	.60
10	5	D	.11	+ .06	+ .09	.65
	6	B	.17		+ .12	.62
11	5	D	.77	- .01	+ .04	.40
	6	B	.76		+ .11	.48
12	5	B	.57	+ .03	+ .15	.54
	6	D	.54		+ .17	.59
13	5	A	.78	+ .07	+ .02	.50
	6	E	.71		+ .08	.56
14	5	E	.39	+ .02	+ .19	.62
	6	A	.41		+ .15	.57

Här är det bara två av uppgifterna, 10 och 13, som uppvisar tydliga skillnader i p-värde beroende på det rätta svarets position. Även här har svårighetsgraden ökat med D eller E som rätt svar. För övriga uppgifter är skillnaderna i stort sett obetydliga. Könsskillnaderna varierar något i omfattning men gör det också här på ett till synes planlöst sätt, även i förhållande till resultaten från våren 1998.

I tabell 2:7 jämförs differenserna i erhållna p-värden vid de två tillfällena när försöket genomfördes. Liksom i tidigare tabeller har skillnaden beräknats utifrån korrigerade p-värden för de grupper som ingick i försöket våren 1998, medan övriga p-värden redovisats och använts utan justering.

⁵ p diff = (p för uppgift med rätt svar A/B) – (p för uppgift med rätt svar D/E). Därmed betyder en positiv differens att lösningsproportionen var större när rätt svar var A eller B, medan en negativ differens betyder att lösningsproportionen var större med D eller E som rätt svar.

Tabell 2:7 Jämförelse av differenser i erhållna p-värden för de åtta försöksuppgifterna våren och hösten 1998.

Uppg nr	7	8	9	10	11	12	13	14
p diff 98:04	+ .05	+ .05	- .04	+ .05	+ .03	± .00	+ .06	- .01
p diff 98:10	± .00	+ .02	- .01	+ .06	- .01	+ .03	+ .07	+ .02

Som tabell 2:7 visar sammanfaller resultaten från de två försöksomgångarna ganska väl. Uppgifterna 11, 12 och 14 uppvisar små differenser i p-värde vid båda tillfällena. Dessa uppgifter har också det gemensamt att de varken till innehåll eller utformning kan förväntas inbjuda till någon systematisk genomgång och prövning av de fem svarsförslagen. I uppgift 11 exempelvis ställs frågan "Hur stor var skillnaden mellan X och Y när den var som störst?". Här gäller det alltså för deltagaren att avläsa data i figurmaterialet, använda dessa för att göra en beräkning och därefter välja rätt svar i form av ett numeriskt värde.

Uppgifterna 10 och 13 uppvisar vid båda försökstillfällena däremot klara skillnader i p-värde beroende på det rätta svarets position. Också mellan dessa uppgifter finns likheter i utformning och innehåll. I uppgift 10 gäller det att via en jämförelsevis svår division identifiera den kategori läkemedel, bland de totalt 14 som ingår i figurmaterialet, som bäst uppfyller ett angivet villkor. I uppgift 13 ska deltagaren på liknande sätt identifiera ett visst län (av 24 möjliga). Eftersom figurmaterialet innehåller så många fler valmöjligheter än de fem som anges i svarsförslagen, förefaller det rimligt att provdeltagare här oftare väljer att utgå från svarsförslagen och pröva dem ett efter ett. Intressant att notera här är dessutom att uppgift 10 är den klart svåraste av de åtta försöksuppgifterna medan uppgift 13 tillhör de allra lättaste. För andra uppgifter ter sig resultaten från de två försökstillfällena mera vaga och motstridiga. Vad detta kan bero på är svårt att säga något om.

I tabell 2:8 sammanfattas resultaten av de två genomförda försöken med förändring av det rätta svarets position.

Tabell 2:8 Sammanfattande resultat för 2 x 4 DTK-uppgifter där det rätta svaret placerats i början respektive i slutet bland svarsförslagen (korrigerade värden från provtillfället 98:04).

Provtillfälle 98:04	Början (A/B), häfte 5		Slutet (D/E), häfte 6		Diff	
	S	M	S	M	S	M
p totalt	2.80	.70	2.74	.68	.06	.02
p män	3.10	.78	2.96	.74	.14	.04
p kvinnor	2.60	.65	2.57	.64	.03	.01
rbis	–	.50	–	.52	–	–.02
Provtillfälle 98:04	Början (A/B), häfte 6		Slutet (D/E), häfte 5		Diff	
	S	M	S	M	S	M
p totalt	1.93	.48	1.81	.45	.12	.03
p män	2.25	.56	2.16	.54	.09	.02
p kvinnor	1.66	.42	1.55	.39	.11	.03
rbis	–	.58	–	.58	–	.00
Provtillfälle 98:10	Början (A/B), häfte 5		Slutet (C/D), häfte 6		Diff	
	S	M	S	M	S	M
p totalt	2.70	.67	2.61	.65	.09	.02
p män	2.91	.73	2.86	.72	.05	.01
p kvinnor	2.52	.63	2.39	.60	.13	.03
rbis	–	.50	–	.52	–	–.02
Provtillfälle 98:10	Början (A/B), häfte 6		Slutet (C/D), häfte 5		Diff	
	S	M	S	M	S	M
p totalt	1.76	.44	1.67	.42	.09	.02
p män	2.02	.50	1.94	.48	.08	.02
p kvinnor	1.53	.38	1.48	.37	.05	.01
rbis	–	.54	–	.56	–	–.02

Ovanstående sammanställning (tabell 2:8) av data från försöken ger ökat stöd för slutsatsen att svårighetsgraden tenderar att öka när det rätta svaret placeras i slutet bland svarsförslagen. Det går däremot inte att urskilja att manliga respektive kvinnliga deltagare konsekvent skulle reagera olika på en förändring av det rätta svarets position.

Slutkommentar

I en tidigare studie har Wester-Wedman (1992) visat att provdeltagarnas strategier för att lösa DTK-uppgifter dels ofta avviker från vad provkonstruktören ”tänkt sig”, dels varierar starkt både mellan uppgifter och mellan individer. Inte minst förhåller sig provdeltagare olika till de givna svarsförslagen – vissa löser uppgiften först för att därefter se efter om det bland dem finns något svar som passar, medan andra tar svarsförslagen till utgångs-

punkt för att lösa uppgiften. Den kunskapen manar till försiktighet vad gäller generella påståenden om hur uppgifter och svarsförslag fungerar och hantearas i en konkret provsituation. Undersökningen som beskrivits här kan ändå sägas ha visat att det rätta svarets position har betydelse för svårighetsgraden i vissa typer av DTK-uppgifter. Denna slutsats skulle kunna prövas vidare, förslagsvis då med hjälp av ett större antal uppgifter och än större medvetenhet och variation vad gäller deras utformning och innehåll. Tills vidare kan resultatet av den här studien, i all sin enkelhet, förhoppningsvis tillvaratas i konstruktionsarbetet som ännu en erfarenhet.

3. Sammanfattande kommentar

Har uppgiftens position i provhäftet någon betydelse för uppgiftens svårighetsgrad? Detta är den övergripande fråga som har ställts i kapitel 1 och som baseras på ett antagande om att provdeltagarna löser uppgifterna i den ordningsföljd som de presenteras i provhäftet. Svaret på frågan tycks inte vara entydigt utan varierar mellan delproven. Resultaten visar att uppgiftens position i provhäftet förefaller vara av betydelse främst för delprovet LÄS men även för delproven ORD och DTK samt i viss mån NOG även om resultaten inte är entydiga. Vad detta beror på är, utifrån de genomförda studierna, svårt att uttala sig om p.g.a. att antalet uppgifter är alltför litet för att någon mer ingående analys ska kunna göras på uppgiftsnivå.

Beträffande NOG-provets avvikelse från de övriga delproven kan denna delvis förklaras av att erfarenheten har visat att provdeltagarna har förhållandevis gott om tid för att lösa uppgifterna. Uppgiftens position kan m.a.o. antas vara av underordnad betydelse för uppgiftens svårighetsgrad. En specifik aspekt för delproven NOG och ORD är emellertid att uppgifterna saknar koppling till ett figur- eller textmaterial. Provdeltagarna kan därför i högre utsträckning, än vid genomförandet av övriga delprov, välja att lösa uppgifterna i den ordningsföljd de föredrar. Till skillnad från NOG-provet är ORD-provet ett kunskapsprov vilket är en faktor som kan påverka tillämpningen av eventuella lösningsstrategier. Huruvida provdeltagarna använder sig av strategier för att lösa uppgifter i högskoleprovet samt hur dessa eventuellt ser ut är emellertid ett område där kunskapen ännu är tämligen begränsad. Tidigare studier av provdeltagarnas lösningsstrategier har t.ex. visat att olika provdeltagare använder sig av olika strategier för att lösa samma uppgift (jfr Wester-Wedman, 1992).

En annan aspekt som i sammanhanget bör beaktas är uppgifternas utformning. Av den studie som genomförts av delprovet DTK framgår bl.a. att det förefaller som om vissa uppgiftstyper - vilka kan betraktas som mer tidskrävande än andra - tycks vara känsligare för förändringar i position. Vidare kan det anses rimligt att uppgifter på en viss svårighetsnivå kan vara mer känsliga för förändringar i position. Av den studie som genomförts av delprovet DTK framgår exempelvis att lättare uppgifter förefaller mer känsliga för förändringar i position än uppgifter med en lägre lösningsproportion. Även i denna delstudie är antalet uppgifter samt deras representation med avseende på andra former av uppgiftskaraktäristika (t.ex. uppgiftstyp) alltför begränsat för att några generella slutsatser ska kunna dras från de analyser som har genomförts. Delprovet LÄS är i nuläget det enda delprov där position fått genomslag i hur sammansättningen av utprövningshäften och reguljära prov sker. Konkret har detta kommit att innebära att de fem ämnesområden som

de olika texterna representerar återkommer i samma ordning i såväl utprovningensversioner som reguljära provversioner.

Beträffande könsskillnaden förefaller de olika studierna ge olika resultat. Medan kvinnorna för delprovet ORD tycks gynnas av att uppgifterna placeras i början av häftet förefaller könsskillnaden i delproven DTK och LÄS minska om uppgifterna placeras i slutet av häftet. Däremot tycks uppgiftens position i NOG-provet sakna betydelse för könsskillnaden. Underlaget är dock i nuläget alltför begränsat för att några mer generella slutsatser ska kunna dras.

Av de studier som har undersökt betydelsen av det rätta svarets placering, och som presenteras i kapitel 2, framgår för delproven ORD och DTK att placeringen tycks ha betydelse. För delprovet ORD var detta speciellt framträdande för den uppgiftstyp där en morfemanalys kunde göras. Däremot är det inte möjligt att dra några slutsatser av de genomförda studierna med avseende på vilka effekter placeringen av det rätta svaret har för könsskillnaden. Genomgående gäller att resultaten bör tolkas med stor försiktighet eftersom underlaget ännu är alltför osäkert. Det kan även vara på sin plats att påpeka att de studier som presenterats i kapitel 1 och 2 har varit inriktade mot att undersöka eventuella förändringar av uppgifternas svårighetsgrad, vilket *inte* är samma sak som att undersöka förändrade provprestationer hos provdeltagarna.

Avslutningsvis och i sammanfattning kan sägas att de resultat som presenterats i föreliggande rapport ger stöd för att placering av uppgifter och rätt svarsförslag i högskoleprovet förefaller vara av betydelse för uppgifternas svårighetsgrad. Ännu återstår dock att undersöka om detta är något som gäller generellt eller om det är kopplat till specifika delprov, olika lösningsstrategier, skilda uppgiftsformat, vissa typer av uppgifter eller uppgifter på en viss svårighetsnivå. För att besvara denna typ av frågor fordras vidare studier som specifikt fokuserar denna typ av aspekter för olika grupper av provtagare, t.ex. kvinnor och män. Att undersöka olika provdeltagares lösningsstrategier för olika delprov och uppgiftstyper i förhållande till placering av uppgifter och rätt svarsförslag kan vara ett sätt att gå vidare för att erhålla mer fördjupad kunskap.

Referenser

Andersson, K. (Red.). (1999). *Högskoleprovet. Konstruktion, resultat och erfarenheter* (Pm, Nr 153). Umeå: Umeå universitet, Enheten för pedagogiska mätningar.

Stage, C., Ögren, G., & Konradsson, L. (1999). *Högskoleprovet hösten 1999. Provdeltagargruppens sammansättning och resultat* (Pm, Nr 159). Umeå: Umeå universitet, Enheten för pedagogiska mätningar.

Wester-Wedman, A. (1992). *Hur tänker provtagarna – egentligen? En studie av lösningsprocessen vid genomförandet av DTK-prov* (Pm från H-gruppen Nr 51). Umeå: Umeå universitet, Pedagogiska institutionen.

Ögren, G. (1998). *Utprovning av uppgifter till högskoleprovet. Utvärdering av försöksverksamhet med en ny utprovningsmodell*. (Pm, Nr 140). Umeå: Umeå universitet, Enheten för pedagogiska mätningar.