PER GUNNEMYR

DEN FJÄRDE KVADRANTENS DILEMMA

Kunskapsbedömning i en föränderlig historiekultur

DEN FJÄRDE KVADRANTENS DILEMMA

Skrifter med historiska perspektiv, No 27

© Copyright Per Gunnemyr 2021
ISBN 978-91-7877-234-6 (tryckt)
ISBN 978-91-7877-235-3 (pdf)
ISSN 1651-4513
Holmbergs, Malmö 2021

PER GUNNEMYR

DEN FJÄRDE KVADRANTENS DILEMMA

Kunskapsbedömning i en föränderlig historiekultur

Malmö universitet, 2021
Fakulteten för lärande och samhälle

Malmö Studies in history (doctoral dissertations)

- Helena Tolvhed, *Nationen på spel. Kropp, kön, svenskhet i populärpressens representation av olympiska spel 1948-1972*, 2008.
- Stefan Nyzell, *"Striden ägde rum i Malmö". Möllevångskravallerna 1926. En studie av politiskt våld i mellankrigstidens Sverige*, 2010.
- Vanja Lozic, *I historiekansons skugga. Historieämne och identifikationsformering i 2000-talets Sverige*, 2010.
- Carolina Jonsson Malm, *Att plantera ett barn*, 2011.
- Martin Kjellgren, *Taming Prophets, Astrology, and the World of God in the Early Moderns Sweden*, 2011.
- Matilda Svensson, *När något blir annorlunda. Skötsamhet och funktionsförmåga i berättelser om poliosjukdom*, 2012.

Malmö Studies in history and history teachings and uses of history (doctoral dissertations)

- Anders Lindh, *"Unity Pervades All Activity As water Every Wave". Principal Teachings and Philosophy of Maharishi Yogi*, 2014.
- Joakim Glaser, *Fotboll från Mielke till Merkel. Kontinuitet, brott och förändring i supporterkultur i östra Tyskland*, 2015.
- Susan Lindholm, *Remembering Chile. An entangled History of Hip-hop in-between Sweden and Chile*, 2016.
- David Rosenlund, *History Education as Content, Methods or Orientation?* 2016.
- Fredrik Alvéén, *Tänka rätt och tycka lämpligt. Historieämnet i skärningspunkten mellan att fostra kulturbärare och förbereda kulturbyggare*, 2017.
- Robin Ekelund, *Retrospektiva modernister. Om historiens betydelse för nutida mods*, 2017.
- Joel Rudnert, *Bland stenxor och tv-spel. Om barn, historisk tid och när unga blir delaktiga i historiekulturen*, 2019.
- Jessica Jarhall, *Historia från kursplan till klassrum. Perspektiv på lärares historieundervisning från Lpo 94 till Lgr 11*, 2020.
- Magnus Bergman, *Att höra det förflutna till. Adlig minneskultur och kollektiv identitet efter ständssamhällets upplösning, 1869-1976*, 2021.

Publikationen finns även elektroniskt,
se mau.diva-portal.org

Till Lucy

INNEHÅLL

FÖRORD.....	9
1. INLEDNING.....	10
1.1 Två minnesbilder och ett perspektivskifte.....	10
1.2 Kunskapsbedömning i en föränderlig historiekultur.....	13
1.3 Den fjärde kvadrantens dilemma	14
1.4 Relevans.....	15
1.5 Kontext.....	16
2. FORSKNINGSPROJEKTET	23
2.1 Studiens syfte och forskningsfrågor.....	23
2.2 Studiens bärande hypotes	24
2.3 Studiens disposition	25
3. TIDIGARE FORSKNING	26
3.1 Ett forskningsfält tar form	26
3.2 Olika perspektiv inom forskningsfältet	31
3.3 Tidigare forskning i relation till föreliggande studie	40
4. TEORIER OCH BEGREPP	42
4.1 Historiekultur – det primära portalbegreppet.....	42
4.2 Historiekultur som analysverktyg.....	45
4.3 Historiekulturella traditioner	53
4.4 Den historiekulturella förhandlingen	58
4.5 Bedömningskultur – det sekundära portalbegreppet...	64
4.6 Bedömningskulturella traditioner	65
4.7 Ett bedömningsteoretiskt interludium	68
4.8 Den bedömningskulturella förhandlingen	73
4.9 Den fjärde kvadrantens dilemma	82
5. METOD	92
5.1 Datainsamling	92
5.2 Dataanalys	98
6. KONSTRUKTIONEN	104
6.1. HLP	104
6.2 KP	124

6.3 Resultatdiskussion.....	142
7. GENOMFÖRANDET.....	146
7.1 HLP	146
7.2 KP.....	151
7.3 Resultatdiskussion.....	159
8. ANVÄNDNINGEN.....	162
8.1 HLP	162
8.2 KP.....	169
8.3 Resultatdiskussion.....	176
9. PÅVERKAN.....	179
9.1 Påverkansområden	179
9.2 Tidigare forskning	180
9.3 HLP	186
9.4 KP.....	192
9.5 Resultatdiskussion.....	206
10. SAMMANFATTNING OCH SLUTDISKUSSION.....	209
10.1 De empiriska resultaten	209
10.2 Det teoretiska ramverket	210
10.3 Empiri möter teori	212
10.4 Slutdiskussion.....	220
SUMMARY.....	226
REFERENSER	232
BILAGOR.....	245

FÖRORD

Det här projektet påbörjades 2015 och många personer har på olika sätt varit involverade under arbetsprocessen. Jag vill dock rikta ett särskilt tack till några av dem.

- Min doktorandkollega Jessica Jarhall. Tack för att du har gjort doktorandtiden så mycket roligare. Våra samtal har varit en ständig glädjekälla.
- Min diskutant på slutseminariet Henrik Åström Elmersjö. Tack Henrik för viktiga kommentarer som förde arbetet framåt.
- Mina arbetskollegor. Tack för att ni har gjort bunkern till en fantastisk arbetsplats.
- Mina föräldrar. Tack för att ni finns och ett särskilt tack för all hjälp med korrekturläsningen.
- Mina handledare Ingmarie Danielsson och Anders Jönsson. Tack för alla kloka synpunkter och glada tillrop.
- Min huvudhandledare Per Eliasson. Tack för allt! Utan dig hade jag inte klarat detta.
- Min livskamrat. Tack Lucy för att du har stöttat mig under dessa år. Nu börjar det roliga!

Hörby, 11 november 2021

Per Gunnemyr

1. INLEDNING

1.1 Två minnesbilder och ett perspektivskifte

I oktober 2006 befann jag mig i Riga med anledning av ”De svenska historiedagarna”. Skolledningen på den gymnasieskola där jag var verksam som lärare i historia och svenska hade välvilligt låtit mig och några av mina kollegor inom ämneslaget få möjlighet att resa dit som ett led i vår fortbildning. En av föreläsningarna handlade om ”Skåneprovet” eller ”Historielärarnas prov” (HLP), som det sedermera kom att kallas. Initiativtagaren till provet, Lars Andersson Hult, berättade under föreläsningen att HLP var ett frivilligt kursprov för gymnasieskolans A-kurs i historia som hade konstruerats och använts av historielärare i Skåneregionen sedan 2004.¹ HLP:s tillkomst var enligt Andersson Hult en reaktion på då aktuell betygsstatistik där det framkommit att de nationella genomsnittsbetygen för A-kursen hade ökat markant inom loppet av bara några få år. Detta samtidigt som det fanns få indikatorer på att de höga betygen verkligen avspeglade en motsvarande ökning av elevernas faktiska kunskaper i historia.

Väl hemkomna till vardagen tog jag och mina kollegor inom ämneslaget ett beslut att låta även våra elever skriva ett prov som en avslutning på A-kursen. När vi konstruerade kursprovet utgick vi i hög grad från HLP, men vi gjorde vissa korrigeringar för att uppgifterna i provet skulle ligga i linje med vår lokala tolkning av kursplanen. Snart kunde vi märka att såväl beslutet att införa ett avslutande kursprov som processen med att utarbeta provet gav upphov till intensiva diskussioner inom ämneslaget. Det var den här reaktionen som utgjorde den initiala utgångspunkten för mitt övergripande intresse för bedömning i allmänhet och för externa prov i synnerhet.² Jag kunde se att ett externt prov i historia påverkade mig själv och mina kollegor, men jag ville veta mer om denna påverkan. På vilka olika sätt påverkades vi? Varför påverkades vi? I förlängningen

¹ Den dåvarande A-kursen i historia motsvaras idag av kursen Historia 1b.

² För en mer precis definition av begreppet externa prov, se avsnitt 4.7.1.

handlade mitt intresse om vilka konsekvenser användningen av externa prov i historia får för ämnesutvecklingen på skolan.

I min licentiatuppsats från 2011 var det just denna påverkansdimension som utgjorde det centrala forskningsområdet. Uppsatsen var en intervju-baserad mikrostudie där jag undersökte hur och varför gymnasielärare i historia uppfattar att deras didaktiska verksamhet påverkas av externa prov i historia. Studien hade en komparativ ansats som tog sig uttryck i att jag dels undersökte hur svenska gymnasielärare uppfattade att de påverkades av HLP, dels hur finländska gymnasielärare uppfattade att de påverkades av det finländska ämnesrealprovet i historia.³ Genom det komparativa anslaget fick jag möjlighet att undersöka hur och varför *olika slags* externa prov påverkar historielärares didaktiska verksamhet på skilda sätt. I analysen framkom att externa prov i historia, beroende på deras konstruktion och tillämpning, har olika former av påverkanspotential. De finländska lärarna uppfattade att ämnesrealprovet i hög grad påverkade deras undervisnings- och bedömningspraktik, medan de svenska lärarna i första hand uppfattade att HLP påverkade deras kollegiala interaktion.⁴

När jag på hösten 2015 påbörjade arbetet med min doktorsavhandling var jag inte längre verksam som gymnasielärare i Gävle utan jag var sedan drygt ett år tillbaka anställd som provkonstruktör på Malmö högskola, numera Malmö universitet. Högskolan hade av Skolverket fått i uppdrag att konstruera kursprov för två av gymnasieskolans historiekurser, Historia 1a1 och historia 1b.⁵ Under åren 2015-2018 konstruerade jag och mina kollegor tre kursprov för Historia 1a1 och fyra kursprov för Historia 1b.

³ Det finländska ämnesrealprovet i historia är en del av den studentexamen som har funnits i Finland sedan 1852. Den realämnesreform som genomfördes i Finland 2006 innebar att det tidigare realprovet ersattes med tolv olika ämnesspecifika realprov. De tolv ämnesrealproven är: evangelisk-luthersk religion, ortodox religion, livsåskådningskunskap, filosofi, psykologi, samhällslära, fysik, kemi, biologi, hälsokunskap, geografi och historia. Ungefär tio procent av de finländska studenterna väljer att skriva ämnesrealprovet i historia.

⁴ Gunnemyr (2011).

⁵ Skolverkets hade i sin tur fått uppdraget att utveckla provet i enlighet med följande regeringsdirektiv: "Inom ramen för den s.k. provbanken för gymnasieskolan tillhandahåller Skolverket olika kursprov som håller samma kvalitet som de obligatoriska nationella proven. De är inte obligatoriska att använda i gymnasieskolan men kan utnyttjas flexibelt under året för att stödja lärares betygssättning. Regeringen anser att sådana kursprov kan bidra till rättvis och likvärdig betygssättning och bedömer därför att ytterligare kursprov bör utvecklas och tillhandahållas i provbanken fr.o.m. 2015, företrädesvis i kurserna Engelska 7 och i Historia 1. Regeringen föreslår därför att det avsätts 5 miljoner kronor för 2014 för att utveckla proven samt beräknar 1 miljon kronor årligen därefter." Prop. 2013/14:1 Utgiftsområde 16 (2013/14), s.63.

Det är kursprovet för den sistnämnda kursen som utgör analysobjekt i föreliggande projekt och den kommer fortsättningsvis att benämnas KP. När jag var verksam som provkonstruktör för KP minns jag att mina tankar ofta gick till de historielärare runt om i landet som valde att ladda ner de prov som jag och mina kollegor i arbetsgruppen hade konstruerat. Jag visste att proven i viss utsträckning laddades ner, men jag ville veta mer. Hur uppfattar historielärare KP:s konstruktion? Hur genomför historielärare KP? Hur använder historielärare KP?

En bärande tanke i det nya forskningsprojektet var att avhandlingen skulle bygga vidare på den tidigare licentiatuppsatsen, något som var lättare sagt än gjort, delvis på grund av att mitt eget perspektiv hade förändrats. Under arbetet med licentiatuppsatsen var jag verksam som historielärare och när jag skulle påbörja arbetet med avhandlingen var jag verksam som provkonstruktör. Detta perspektivskifte innebar forskaretska implikationer eftersom jag under avhandlingsarbetet beforskar det prov som jag själv varit med om att konstruera. Den här problematiken kommer att behandlas i metodavsnittet. Perspektivskiftet innebar även att jag kom att ställa delvis nya frågor i avhandlingen. I licentiatuppsatsen handlade den centrala forskningsfrågan om historielärares uppfattningar om de externa provens påverkanseffekt. I avhandlingen kompletteras denna forskningsfråga med tre andra forskningsfrågor där historielärares uppfattningar om provens konstruktion, genomförande och användning uppmärksammas. De fyra forskningsfrågorna kommer att preciseras närmare i kapitel 2.

En annan viktig skillnad mellan licentiatuppsatsen och avhandlingen är att det komparativa anslaget har förändrats. I licentiatuppsatsen var komparationen geografisk medan den i avhandlingen är temporal. I studien utgör 2011 års skolpolitiska reformer, manifesterad i läroplanen för gymnasieskolan (Gyl1), brytpunkten för den temporala komparationen. HLP konstruerades och genomfördes under åren 2003-2013 och HLP-lärarna intervjuades 2009. KP konstruerades och genomfördes under åren 2015-2018 och KP-lärarna intervjuades 2016. Den temporala komparationen är central för avhandlingen eftersom den gör det möjligt att undersöka förändring över tid.

1.2 Kunskapsbedömning i en föränderlig historiekultur

Forskningsprojektets övergripande syfte är att undersöka historiekulturell förändring över tid med avseende på skolämnet historia, och då särskilt beakta vilken betydelse kunskapsbedömning i allmänhet och externa prov i synnerhet kan ha för denna förändring. Syftesformuleringen tydliggör att begreppet historiekultur kommer att utgöra avhandlingens portalbegrepp. Det är ett omdiskuterat begrepp, men inom ramen för detta forskningsprojekt avses det kommunikativa sammanhang inom vilket olika sätt att uppfatta det förflutna lyfts fram och ges mening. Eftersom studien har en uttalat praktknära ansats är det i första hand den historiekultur som kommer till uttryck i skolan som utgör fokus för studien. Skolans historiekultur manifesteras i skolämnet historia, närmare bestämt i historieämnets vad- och varför-dimension, det vill säga epistemologiska uppfattningar om vilka kunskaper som inryms i skolämnet historia och varför dessa kunskaper är meningsfulla. För en mer utförlig redogörelse av begreppet historiekultur, se avsnitt 4.1.

Förändringar inom skolans historiekultur kan belysas utifrån flera olika perspektiv, men i syftesformuleringen framkommer att fokus för denna studie ligger på vilken betydelse skolans former för kunskapsbedömning har för dessa förändringar.⁶ Begreppet bedömningskultur kommer därför att utgöra studiens sekundära portalbegrepp. Även bedömningskultur är ett omdiskuterat begrepp, men inom ramen för detta projekt avses det kommunikativa sammanhang inom vilket olika sätt att uppfatta kunskapsbedömning lyfts fram och tillskrivs mening. Skolans bedömningskultur kommer till uttryck i bedömningens hur- och varför-dimension, det vill säga metodologiska uppfattningar om hur bedömningen av skolämnets kunskaper ska genomföras och varför denna bedömning är meningsfull. För en mer utförlig redogörelse av begreppet bedömningskultur, se avsnitt 4.5.

⁶ Det finns förhållandevis gott om studier som har intresserat sig för skolrelaterad historiekulturell förändring över tid, men där perspektivet och analysobjektet har varit ett annat. Exempel på sådana studier är Zander (2001) med fokus på den offentliga debatten om historia; Danielsson Malmros (2012) med fokus på läroböcker i historia; Rosenlund (2016) med fokus på styrdokument i historia och Åström Elmersjö (2017) med fokus på läromedelsgranskningar i historia. Grundläggande för min egen och dessa studier är dock, som Zander poängterar i sin avhandling, att det finns en relation mellan historiekulturen som den kommer till uttryck i skolan och samhällsutvecklingen i stort. Zander (2001), s. 460.

1.3 Den fjärde kvadrantens dilemma

Det är viktigt att betona att forskningsprojektet har en tydlig teoriutvecklande ambition. Genom att i mitt teoretiska ramverk fläta samman historiedidaktisk och bedömningsteoretisk teoribildning på ett systematiskt och i vissa avseenden tidigare oprövat sätt har jag utarbetat en analytisk verktygslåda som manifesteras i såväl ett antal konceptuella modeller som i en stringent begreppsapparat. I linje med den teoriutvecklande ambitionen avser jag att pröva den övergripande hypotesen att relationen mellan skolans historiekulturella vad-dimensionen och bedömningskulturella hur-dimensionen inrymmer ett latent spänningsförhållande, fortsättningsvis benämnt *den fjärde kvadrantens dilemma* (se figur 1). Den fjärde kvadrantens dilemma kommer att diskuteras mer utförligt i teorikapitlet men i korthet bygger hypotesen på det grundläggande antagandet att det alltid finns ett samband mellan *vad* som bedöms och *hur* någonting bedöms, i det här sammanhanget mellan den historiekulturella vad-dimensionen och den bedömningskulturella hur-dimensionen.

Figur 1: Den fjärde kvadrantens dilemma

I fyrfältsmatrisen utgör den historiekulturella vad-dimensionen den vertikala axeln och den bedömningskulturella hur-dimensionen den horisontella axeln. Den historiekulturella vad-dimensionen är mångfacetterad, men rent analytiskt är det möjligt att urskilja två breda epistemologiska traditioner: Memory History-traditionen (MH-traditionen) och Historical Thinking-traditionen (HT-traditionen). Den förra är förankrad i en mer

objektivistisk och den senare i en mer konstruktivistisk syn på historisk kunskap.

Även den bedömningskulturella hur-dimensionen är komplex, men inom det bedömningsteoretiska fältet är det vanligt särskilja mellan en mer psykometriskt förankrad mättradition och en mer pedagogiskt förankrad tolkningstradition. Lite förenklat kan man säga att den förra i högre utsträckning metodologiskt begränsar bedömarens tolkningsutrymme medan den senare i högre utsträckning bejakar bedömarens tolkningsutrymme.

Hypotesen bygger som sagt på det grundläggande antagandet att det alltid finns ett samband mellan bedömningens vad- och hur-dimension. Ytterligare ett antagande är att det finns särskilt spänningsförhållande mellan den historiekulturella HT-traditionen och den bedömningskulturella mättraditionen. Det innebär med andra ord att det är problematiskt att förena HT-traditionell epistemologi med mättraditionell metodologi.⁷

1.4 Relevans

Genom att låta den fjärde kvadrantens dilemma utgöra avhandlingens bärande hypotes tydliggörs forskningsprojektets historiedidaktiska relevans. Hypotesen ger för handen att det är problematiskt att göra kunskaper i historia mätbara. Är det över huvud taget möjligt att bedöma historisk kunskap på ett mätsäkert sätt utan att det påverkar vårt sätt att uppfatta vad som menas med historisk kunskap? Den här frågeställningen kan beaktas från såväl ett kunskapsfilosofiskt perspektiv som ett konsekvensrelaterat perspektiv.

Utifrån ett kunskapsfilosofiskt perspektiv kan man ställa sig frågan om det över huvud taget är möjligt att bedöma kunskaper i historia. Går det till exempel att bedöma elevernas historiemedvetande?⁸ Det här är en grundläggande kunskapsfilosofisk fråga som såväl skolpolitiker som historielärare måste förhålla sig till. Jag ska inte föra några fördjupade

⁷ Alla ämnesdiscipliner inrymmer epistemologiska spänningar av olika slag, men den fjärde kvadrantens dilemma aktualiseras kanske framför allt i ämnesdiscipliner som saknar ett absolut sanningsanspråk. Dilemmat har uppmärksamrats inom det historiedidaktiska fältet, men det har inte formulerats explicit på detta sätt, se t.ex. Anderson Hult (2021) s. 57–60. För en utförligare redogörelse av tidigare studier med koppling till den fjärde kvadrantens dilemma, se avsnitt 3.3.4.

⁸ Den här frågan har diskuterats av historiedidaktiker inom fältet. Den finländska historiedidaktikern Arja Virta menar att det varken är möjligt eller önskvärt att mäta elevernas historiemedvetande. Virta (1997), s. 181–182. Per Eliasson delar delvis Virtas uppfattning. Han menar dock att elevens historiemedvetande kvalificeras av förågor, vilka i sin tur är möjliga att bedöma. Eliasson (2009), s. 323.

kunskapsfilosofiska resonemang i det här sammanhanget, men en viktig lärdom som vuxit sig stark under de år jag har forskat inom ämnesområdet är att bedömning av kunskaper i historia inte är någon exakt vetenskap.

Utifrån ett mer konsekvensrelaterat perspektiv kan man ställa sig frågan vilka konsekvenser ett ökat fokus på mätsäker kunskapsbedömning får för de aktörer vars kunskaper bedöms? Den amerikanske bedömnings-teoretikern Georg F. Madaus anspelar på Heisenbergs osäkerhetsprincip när han menar att all form av bedömning inom skolan påverkar det som bedöms: "Any measurement of the status of an educational institution, no matter how well designed and well intentioned inevitably changes its status."⁹ Här framträder skillnaden mellan naturvetenskap och humanvetenskap som betydelsefull. En sten som får sin vikt uppmätt påverkas inte av mätningen. Det gör däremot en elev som får sina kunskaper i historia bedömda eller en historielärare som får sin undervisningsförmåga utvärderad.¹⁰ Det är viktigt att betona att denna påverkan inte behöver uppfattas som något negativt. I sammanhanget bör man även ställa sig frågan vilka konsekvenser som följer av att *inte* bedöma elevers kunskaper eller att *inte* utvärdera lärares undervisningsförmåga.

Relationen mellan historiekulturella uppfattningar om vad som menas med kunskaper i historia och bedömningskulturella uppfattningar om hur och varför dessa kunskaper ska bedömas är som framgått komplex och väcker många frågor. Vad händer med skolämnet historia om fokus för bedömningen glider över på mätning och om målen för undervisningen anpassas efter det mätbara? Vad händer med de mål som inte är mätbara? Vad händer till exempel med ambitionen att utveckla elevens historiemedvetande?

Innan jag tar mig an uppgiften att ytterligare precisera avhandlingens syfte och frågeställningar vill jag avsluta det här inledande kapitlet med att placera in forskningsprojektet i en bredare kontext

1.5 Kontext

1.5.1 Skolämnet historia

Skolämnet historia hade länge svårt att hitta sin plats i den skola som växte fram i Sverige under andra hälften av 1900-talet. Ämnet kom allt

⁹ Madaus (1988), s. 89.

¹⁰ Gipps (1994), s. 63–65.

mer att uppfattas som en främmande fågel i det progressiva landskap som hade börjat ta form och det var istället det nya ämnet samhällskunskap som skulle fostra det framväxande välfärdssamhällets medborgare.

Historieämnets legitimitet blev på allvar ifrågasatt redan 1946 i samband med skolkommissionens utredning om den framtida skolan. Skolämnet historia uppfattades som bakåtsträvande och kom att betraktas med kritiska blickar eftersom förändring snarare än kontinuitet var ledstjärnan i det svenska modernitetsprojektet. Det historieämne som tog form i spåren av skolkommissionen präglas dock inte av några större epistemologiska förändringar. Det var fortfarande ett ämne stadigt förankrat i MH-traditionen som levde vidare, men med den skillnaden att det fostrande nationella narrativ som tidigare stått i centrum för ämnet hade ersatts med andra narrativ.¹¹ Englund beskriver historieämnet som ett traditionellt skolämne i ett annars progressivt landskap.¹²

Den följande utvecklingen fram till 1994 års reformer kan liknas vid en långsam förblödning av historieämnet. Trots skolpolitiska försäkringar om vikten av historiska kunskaper fick ämnet allt färre undervisningstimmar. 1994 års reformer innebar en symbolisk dödsstöt i och med att ämnet inte blev ett kärnämne i den nya gymnasieskolan. Historieämnet hade blivit ett litet och anonymt karaktärsämne med låg status och få undervisningstimmar.¹³ Det karaktärsämne som trots allt skrevs fram i 1994 års kursplaner hade dock, lite i skymundan, förändrats i grunden. Det var inte längre ett MH-traditionellt ämne, med fokus på historiska narrativ, som skrevs fram, utan ett ämne med en tydlig HT-traditionell förankring. Det kanske tydligaste tecknet på denna förändring var att det HT-traditionella begreppet historiemedvetande hade letat sig in i de nya kursplanerna.¹⁴

Historieämnets återinträde i den offentliga debatten i Sverige under 1990-talet är ett omskrivet fenomen.¹⁵ Det har varit vanligt att koppla

¹¹ Östling (2008) menar att det nationella narrativet fick en ny dimension. Se även Zander (2001).

¹² Englund (1986). Till skillnad från Englund vill Samuelsson (2021) tona ner spänningsförhållandet mellan traditionalisterna och de progressiva. Han tycker sig se att progressivismen i första hand får ett metodologiskt genomslag, men att historieämnet förblir förankrat i en traditionell epistemologi.

¹³ Se t.ex. Schüllerqvist (2005); Larsson (2001) och Zander (2001). Larsson tolkar händelseutvecklingen under reformåren 1991–1994 som en maktförskjutning där fackpedagoger i en ohelig allians med kameralt intresserade kommunalpolitiker tog kontroll över ämnesutvecklingen. Larsson (2009), s. 414–417.

¹⁴ Se t.ex. Rosenlund (2016), s. 77–79. Åström Elmersjö (2017) lyfter dock fram att framtidsperspektivet omnämns redan i supplementet till 1981 års kursplan.

¹⁵ Se t.ex. Karlsson (2009), s. 34–37. Den skolpolitiska debatten om historieämnet som växte fram under slutet av 1990-talet kan jämföras med motsvarande debatt under slutet av 1940-talet. Även då var det historiska

återuppståndelsen till en bredare kontext med fokus på den vilsenhet som tog vid efter det kalla krigets slut i kombination med en ökad osäkerhet i vilken riktning det svenska välfärdssamhället var på väg.¹⁶ Det ökade historieintresset återspeglas även inom skolan. Skolämnet historia började nu lämna skuggtillvaron bakom sig och kliva in i det mediala och politiska strålkastarljuset. Det skolpolitiska intresset för historieämnet fick ytterligare bränsle i samband med att de nedslående resultaten av en undersökning om svenska elevers kunskaper om förintelsen som publicerades 1997. En diskussion tog nu sin början huruvida historieämnet åter skulle välkomnas in i kretsen av kärnämnen. När detta skedde i samband med reformerna 2011 var det ämne som i många avseenden byggde vidare på det HT-traditionella vägval som 1994 års kursplaner hade stakat ut.¹⁷

1.5.2 Externa prov i skolan

Det skolpolitiska intresset för extern kunskapsbedömning har varierat över tid. Ett genomgående mönster har dock varit att intresset blossat upp när likvärdighetsfrågor stått högt upp på den skolpolitiska agendan. Under 1960-talet, i samband med reformeringen av läroverken, fick likvärdighetsfrågan hög prioritet. Den fria intagningen till universiteten hade upphört och frågan var hur de sökande skulle ges rättvisa möjligheter att bli antagna till högre utbildning. I en mättraditionell tankevärld framstod externa standardiserade centralprov utformade på vetenskapliga grunder som en framkomlig väg för att lösa problemet. Dessa prov, i kombination med ett relativt betygssystem, skulle dessutom säkerställa en konstant genomströmning av elever via gymnasiet till högre utbildning. Pär Widén menar att reformerna, som i praktiken innebar att studentexamen avskaffades och att en ny läroplan för gymnasiet infördes, speglar en övergång från ett bedömningsparadigm till ett annat. Det tidigare tolkningstraditionella paradigmat, med autonoma läroverkslärare som satte absoluta be-

perspektivet i fokus, men tolkningen av det förflutna gav andra svar. Ett återkommande mönster är att historieämnets relevans tenderar att eskalera i samband med borderline events och i efterdyningarna av berlinmurens fall och 11 september-attackerna återfick ämnet något av sin forna relevans.

¹⁶ Anderson Hult sätter in historieämnets återuppståndelse i övergången från modernitet till senmodernitet. Han argumenterar för att det moderna samhället hittade lösningen på samhällsproblemen genom ett mer framtidsinriktat perspektiv medan det senmoderna samhället finner lösningarna i ett mer tillbakablickande perspektiv. Andersson Hult (2021), s. 49–53.

¹⁷ I vilken utsträckning 2011 års kursplaner skiljer sig från 1994 års kursplaner råder det dock delade meningar om inom det historiedidaktiska fältet, se avsnitt 6.2.1.

tyg utan någon tydlig extern kontroll, ersattes med ett mättraditionellt paradigmbaserat på ett relativt betygssystem och centralt utformade externa prov. I linje med de nya intentionerna ersattes det tidigare censorssystemet med ett system där Skolöverstyrelsens gymnasieinspektörer utgjorde huvudaktörer.¹⁸

De centrala prov som infördes i samband med reformeringen av läroverken hade haft en begränsad legitimitet redan från början, men kritiken mot proven tilltog i styrka under 1970-talet för att kulminera runt 1980. Motståndet mot de centrala proven var allmänt utbrett, men den skarpaste kritiken kom från den progressiva rörelse som hade växt sig stark under samma period. Pedagoger med en progressiv syn på skolan uppfattade de mättraditionella centrala proven som ett hinder för att förändra gymnasieskolan i önskvärd riktning. Likvärdighetsfrågan hamnade allt mer i skuggan av andra frågor. Högst upp på den progressiva dagordningen fanns visionen om en decentraliserad skola genomsyrad av en konstruktivistisk kunskapssyn. I kampen för en ökad decentralisering fick de progressiva krafterna stöd från ett oväntat håll. De nyliberala idéer som genomsyrate 1980-talets idépolitiska debatt kom att få stort genomslag inom den offentliga sektorn. Den progressiva visionen om en decentraliserad skola med eleven i centrum fick dock i dessa kretsar andra, nyliberala, förtecken i form av slagord som ”ökad valfrihet” och ”ökad effektivitet”.¹⁹ I samband med det krympande ekonomiska utrymmet under 1980-talet avtog det tidigare politiska intresset för att med traditionell social ingenjörskonst detaljstyra skolans verksamhet och den progressiva visionen om en decentraliserad skola blev verklighet under reformåren i början av 1990-talet.²⁰

De externa prov som kom att ersätta de centrala proven i samband med reformerna under åren 1991–1994 speglade initialt den progressiva agendan.²¹ De mättraditionella ambitioner som genomsyrate 1970- och 1980-

¹⁸ Widén (2010), s. 72–75. Widén menar att orsakerna till paradigmskiftet, vid sidan av likvärdighetsaspekten, även kan kopplas till ekonomiska och pedagogiska motiv. Den nya gymnasieskolan var en betydligt större organisation och censorssystemet, med bland annat muntliga förhör, var ett dyrt sätt att säkerställa likvärdigheten. Den pedagogiska tankegången var att avslutande examinationer var negativa för inläringen. De centrala proven skulle därför osynliggöras och integreras så att större fokus hamnade på undervisningen.

¹⁹ Det är i denna brytpunkt som grunden läggs för den väg av resultatbaserad styrning som under benämningen New public management (NPM) har präglat svensk offentlig sektor sedan början av 1990-talet.

²⁰ Jfr Lundahl (2006) och Widén (2010).

²¹ Stenlås (2009), s. 9–14.

talets centrala prov tonas ner och lärarna återfår kontrollen över bedömningen. En viktig aspekt i sammanhanget är att kunskapsbedömningens betydelse förblev lågprioriterad i de läroplaner som presenterades 1994. Av det bedömningsteoretiska paradigmskifte med tolkningstraditionella (formativa) förtecken som enligt Caroline Gipps ägde rum i den internationella pedagogiska debatten under 1990-talet märks föga i dessa styrdokument.²² Christian Lundahl menar att frånvaron av en tydlig bedömningsteoretisk riktlinje i de nya läroplanerna kan förklaras av att de progressiva pedagogerna hade en tydlig uppfattning *mot* det tidigare mättraditionella systemet, men att de inte hade något nytt tydligt och genomarbetat tolkningstraditionellt alternativ. Kunskapsbedömningen blev därför inte någon integrerad del av läroplanen.²³ I teorin hade staten delegerat kunskapskontrollen till Skolverket, men Skolverket hade å sin sida tydliga direktiv att ”göra halt vid kommungränsen”. Resultatet blev en skola med mycket stor bedömningsautonomi och ett stort tolkningsutrymme för de enskilda lokala aktörerna. Det här kan tolkas som att aktörerna bakom reformen hade en stark tilltro till professionen, men det kan också tolkas som att de skapade en reform med orealistiska förväntningar på de lokala aktörernas bedömningskompetens.

Det visade sig att just kontrollfunktionen eller kvalitetssäkringen kom att bli den svaga punkten i 1994 års reformer. Likvärdighetsfrågan hamnade efter hand allt högre upp på den skolpolitiska dagordningen, vilket innebar att synen på de nationella provens funktion gradvis kom att förändras. Lundahl beskriver processen som en smygreformerings.²⁴ Jan Morawski beskriver samma utveckling som en process där den tidigare progressiva uppfattningen att centralt utformade externa prov begränsar lärarnas professionella frihet, kom att ersättas med en utbredd mättraditionell uppfattning att en extern kontrollregim behövdes för att upprätta en kvalitativt god skola.²⁵ I Morawskis beskrivning av processen utgör Leif Davidssons betänkande från 2007, *Tydliga mål och kunskapskrav i grundskolan*, den skriftliga manifestationen för denna omsvängning. Betoningen av de externa provens kontrollfunktion lyfts fram på ett mycket tydligt sätt i Davidssons betänkande.

²² Gipps (1994), s.1.

²³ Lundahl (2006), s. 356–358.

²⁴ Lundahl (2006), s. 392–394.

²⁵ Morawski (2010), s.218.

Proven måste ha en sådan tyngd att de fungerar inte bara vägledande i betygssättningen utan även styrande. Det förutsätter att kursplaner utformas så att de i högre grad styr såväl undervisning och provkonstruktioner. De nationella proven utprövas med kvalificerade metoder och bedöms därför vara mer träffsäkra i sin uppgift att pröva elevens kunskaper inom respektive ämne än de flesta lärargjorda prov.²⁶

I samband med 2011 års skolpolitiska reformer kom antalet nationella prov i grundskolan och på gymnasiet att öka, vilket indikerar att Davidssons betänkande hade fått stor betydelse för den skolpolitiska färdriktningen. De nationella proven hade blivit ett viktigt mättraditionellt verktyg för att säkerställa likvärdigheten i den nya gymnasieskolan.

1.5.3 Externa prov i skolämnet historia

I inledningen till Hjalmar Bergmans roman *Markurells i Wadköping* sitter lektorn i historia, Ivan Barföth, och blickar ut över staden. Det är fortfarande tidig morgon och det kommer att dröja åtskilliga timmar innan årets kull av studenter ska avlägga sin studentexamen. Studenternas kunskaper kommer att prövas i ett flertal ämnen, däribland historia. De kommer att få visa sina kunskaper dels genom skriftliga prov, dels genom muntliga förhör där externa censorer spelar en avgörande roll. Den här formen av extern kunskapsbedömning i historia förändrades över tid, men levde ändå vidare i någon form ända fram till 1968 när studentexamen och det därtill hörande censorssystemet slutgiltigt gick i graven.²⁷

År 2003, efter närmare fyrtio års frånvaro, återuppstod en extern kunskapsbedömning i historia på regional nivå när det första Skåneprovet, sedermera omdöpt till Historielärarnas prov genomfördes. Konstruktionen av HLP var ett projekt som tog form i slutet av den period inom svensk skola som inleddes med reformerna 1994 och som avslutades med reformerna 2011.²⁸ På sätt och vis kan HLP betraktas som en triumf för

²⁶ SOU (2007:28), s. 273–274.

²⁷ Samuelsson (2010) menar att de skolpolitiska reformerna under 1960-talet innebar att pedagoger med mättraditionella (psykometriska) förtecken tog kontroll över det nationella bedömningssystemet. Hans tes är att övergången från ett censorssystem till ett system med centrala prov och gymnasieinspektörer innebar en övergång från en akademisk till en byråkratisk kontroll över historieämnet. Det var först i och med denna övergång som läroplanen blev ett effektivt verktyg i händerna på staten och de byråkratiska tjänstemännen – förkroppsligade i gymnasieinspektörerna.

²⁸ För utförligare historik om HLP se webbplatsen <http://historieprov.se/sida13.html>.

de decentraliseringstankar som präglade perioden. De skånska historielärarna visade att den svenska yrkesprofessionen hade vilja och förmåga att axla det tolkningsansvar som tilldelades dem i samband med decentraliseringen. Det går dock att tolka HLP:s tillkomst i mörkare färger. I frånvaron av central styrning och med en påtaglig oförmåga hos de kommunala aktörerna att kvalitetssäkra skolans verksamhet, agerade de skånska lärarna med ryggen mot väggen. Likvärdigheten vad gäller bedömning betvivlades och HLP kan uppfattas som ett isolerat rop på hjälp från en övergiven profession.

Det frivilliga kursprovet i historia som konstruerades under perioden 2015–2018 kan uppfattas som ett skolpolitiskt svar på detta skånska nöd-rop. KP:s tillkomst innebar att samtliga historielärare på gymnasiet fick tillgång till ett externt prov på nationell nivå.²⁹ KP:s historia blev dock kortvarig och eftersom även konstruktionen av HLP upphörde 2013 innebär det att det idag saknas externa prov i historia i den svenska gymnasieskolan.³⁰

²⁹ För utförligare historik om KP, se Ammert & Eliasson (2019).

³⁰ På grundskolan finns det dock sedan 2013 ett nationellt prov i historia för åk 9.

2. FORSKNINGSPROJEKTET

2.1 Studiens syfte och forskningsfrågor

Det övergripande syftet med forskningsprojektet är som tidigare framgått att undersöka historiekulturell förändring över tid med avseende på skollämnets historia, och då särskilt beakta vilken betydelse kunskapsbedömning i allmänhet och externa prov i synnerhet kan ha för denna förändring. Den temporala brytpunkten för denna förändringsprocess är 2011 års skolpolitiska reformer manifesterade i Gy11.

Det övergripande syftet operationaliseras i en fyrdelad studie som omfattar nedanstående forskningsfrågor.³¹

- Hur uppfattar historielärare på gymnasiet de frivilliga externa provens **konstruktion**?
- Hur beskriver historielärare på gymnasiet att de **genomför** de frivilliga externa proven?
- Hur beskriver historielärare på gymnasiet att de **använder** de frivilliga externa proven?
- Hur uppfattar historielärare att deras bedömningspraktik **påverkas** av de frivilliga externa proven?

Att på det här sättet använda sig av historielärares uppfattningar om kunskapsbedömning i allmänhet och frivilliga externa prov i synnerhet för att undersöka historiekulturell förändring över tid bygger på det grundläggande antagandet att det finns ett samband mellan *vad* som bedöms och *hur* någonting bedöms. Genom att jag i studien intervjuar HLP-lärare verksamma före Gy11 och KP-lärare verksamma efter Gy 11 genereras ett empiriskt underlag som möjliggör en analys av historiekulturell förändring över tid.

³¹ De tre första forskningsfrågorna är teoretiskt förankrade i den konceptuella modell som presenteras på s. 72 i avhandlingens teorikapitel. Den fjärde forskningsfrågan utgjorde som tidigare nämnts den centrala frågeställningen i min licentiatuppsats.

Studien inrymmer flera betydelsebärande avgränsningar. En sådan avgränsning är att studien endast omfattar historielärare verksamma på *gymnasiet*. En annan viktig avgränsning är att studien endast omfattar *frivilliga* externa prov. Den här avgränsningen inrymmer förvisso en pragmatisk dimension eftersom de två externa prov i historia som finns tillgängliga för gymnasiet är frivilliga. Av större betydelse är att provens frivillighet indikerar att de svarar mot någon form av behov eller problem. De historielärare som ingår i studien har av någon anledning *valt* att använda sig av ett frivilligt externt prov i historia. Den här mer relevansbaserade dimensionen av provens frivillighet kan relateras till Carlgrens resonemang nedan, se avsnitt 2.2.

Vad gäller den fjärde forskningsfrågan är det viktigt att tydliggöra att studien endast uppmärksammar den påverkan som kan kopplas till historielärares *bedömningspraktik* och inte till deras didaktiska verksamhet i stort. Jag kommer således inte att beröra utsagor där historielärare ger uttryck för uppfattningar att det externa provet har påverkat deras undervisningspraktik. Inte heller kommer jag att beröra utsagor som handlar om huruvida deras administrativa verksamhet eller deras sociala relationer till elever och kolleger har påverkats av det frivilliga externa provet.³² Det bör även betonas att gränsdragningen mellan undervisningspraktik och bedömningspraktik i många fall är flytande. Att jag avgränsar den fjärde forskningsfrågan till att endast omfatta påverkan på bedömningspraktiken är i första hand avhängigt studiens bedömningskulturella förankring.

2.2 Studiens bärande hypotes

Studien har en uttalat praktikhänsatta ansats. I linje med Ingrid Carlgren menar jag att det finns ett behov av forskning som kan utgöra en brygga mellan grundforskning och lärares professionella yrkesutövning. Hon beskriver att mycket av den forskning som bedrivs inom utbildningsområdet utgörs av forskning *om och på lärare* och efterfrågar i högre utsträckning forskning *för* lärare. Utgångspunkten för en sådan forskningsansats bör

³² Läraryrket är idag ett yrke som inrymmer många olika slags verksamheter. Vid sidan av den didaktiska verksamheten behöver läraren hantera en rad andra uppgifter av administrativ och social karaktär. I den tidigare licentiatuppsatsen var ambitionen att undersöka hur historielärare uppfattar att externa prov påverkar deras didaktiska verksamhet i sin helhet, men jag har i föreliggande avhandling avgränsat undersökningen till att endast omfatta hur de uppfattar att proven påverkar deras bedömningspraktik. För en elaborerad definition av begreppet bedömningspraktik, se t.ex. Samuelsson (2015), s. 55.

tas i ett uppfattat problem i verksamheten eller i en fråga som har anknytning till ett uppfattat problem.³³

En viktig utgångspunkt i föreliggande forskningsprojekt har varit att undersökningen ska kunna förankras i ett sådant praktisknära/verksamhetsrelaterat problem eller dilemma. I projektet prövas därför hypotesen att relationen mellan den historiekulturella vad-dimensionen och den bedömningskulturella hur-dimensionen inrymmer ett latent spänningsförhållande, benämnt *den fjärde kvadrantens dilemma*. På ett generellt plan handlar dilemmat om hur historielärare kan hantera relationen mellan skolans historiekultur och bedömningskultur. Mer precist uttryckt handlar problemet om hur historielärare kan hantera spänningsförhållandet mellan den historiekulturella HT-traditionen och den bedömningskulturella mättraditionen.

2.3 Studiens disposition

Den fortsatta avhandlingstexten är indelad i följande kapitel. I kapitel tre redogör jag för tidigare forskning i relation till forskningsprojektet. I kapitel fyra redogör jag för de teoretiska och begreppsliga utgångspunkter som ligger till grund för forskningsprojektet. I kapitel fem beskriver jag forskningsprojektets metodologiska överväganden. I kapitlen sex till nio redovisar jag forskningsprojektets empiriska resultat. I det tionde och avslutande kapitlet sammanfattar jag forskningsprojektets empiriska resultat och därefter förs en avslutande diskussion där resultaten relateras till projektets övergripande syfte och bärande hypotes.

³³ Carlgren (2011). Carlgrens resonemang ligger i linje med Portellis reflektion gällande relationen mellan den intervjuande forskaren och den som intervjuas: "I was not *studying* them but *learning* from them". Portelli (1991), s. xi.

3. TIDIGARE FORSKNING

Forskningsfältet som ligger till grund för avhandlingen är till sin natur tvärvetenskapligt och kan i bred bemärkelse benämnas som *historiedidaktisk forskning med inriktning mot bedömning*. Inom forskningsfältet är det möjligt att urskilja en mer teoriutvecklande forskning, där fokus ligger på vad som menas med kunskaper i historia och hur dessa kunskaper kan kvalitetsbestämmas, och en mer praktikinära forskning som i högre utsträckning fokuserar på hur bedömning av kunskaper i historia kommuniceras och tillämpas.³⁴ Den teoriutvecklande forskningen presenteras i nästa kapitel och kommer därför inte att behandlas nedan. Det bör även förtydligas att tidigare forskning som har en direkt koppling till avhandlingens resultat behandlas i anslutning till respektive empiriskt kapitel.

Kapitlet är indelat i tre delar. I den första delen redovisas forskningsfältets bredare konturer. Detta görs genom att ett antal internationella och nationella forskningsöversikter redovisas och kommenteras. Syftet med denna del är dels att ge en överblick, dels att tydliggöra den specifika problematik som har uppmärksamats på internationell respektive nationell nivå. I den andra delen redovisas studier med särskild relevans för forskningsprojektet. Till stöd för denna redovisning används ett kategoriseringsverktyg som möjliggör en tredelad kategorisering av dessa studier. I den tredje och avslutande delen förtydligas hur forskningsprojektet kan relateras till tidigare forskning.

3.1 Ett forskningsfält tar form

3.1.1 En internationell överblick

Forskning med inriktning mot bedömning är och har varit en eftersatt dimension inom det historiedidaktiska fältet. Denna frånvaro av bedömningsrelaterad forskning har varit ett återkommande tema i många internationella forskningsöversikter. I en forskningsöversikt från 2006 noteras

³⁴ Gränsdragningen är i många fall flytande, men för att göra kapitlet mer överskådligt har jag försökt att särskilja dem åt.

att det råder brist på forskning om ”progression of thought processes required to move from novis to expert”.³⁵ I översikten konstateras förvisso att historieämnet har kommit längre än övriga samhällsorienterande ämnen, men samtidigt betonas att historieämnet ligger långt efter andra ämnesdiscipliner, till exempel matematikämnet.

I en senare forskningsöversikt från 2015 skriver Monte-Sano och Reisman om behovet av bedömningsrelaterad forskning inom fältet:

The urgency for robust, valid assessments of historical thinking and understanding should be self-evident. First, in their absence, multiple-choice assessments focused only on factual recall will remain the norm. But, more importantly in the context of this review, until the field has precise and valid measures that capture agreed-upon student outcomes, the research on student historical learning will lag.³⁶

Enligt artikelförfattarna har bedömningsforskningen inom det historiedidaktiska fältet legat i träda sedan britternas forskning under 1970-talet.³⁷ Den brittiska bedömningsforskningen från 1970-talet och framåt var till stor del förankrad i två skolrelaterade forskningsprojekt: Chata, School History Project (SHP) samt Cambridge A-levet History project. Den här forskningen handlade dock i första hand om stadie- eller åldersrelaterad progression och de konceptuella modeller som lanserades var tänkta att användas formativt, inte summativt.³⁸

Bilden av att den historiedidaktiska bedömningsforskningen har varit eftersatt är således utbredd inom fältet, men kanske är en förändring på väg.³⁹ I antologin *New Directions in Assessing Historical Thinking* från 2015 presenteras den senaste forskningen av ledande europeiska och nordamerikanska forskare inom fältet.⁴⁰ I förordet till antologin argumenterar antologins redaktörer för att fältet har breddats på senare år, men att

³⁵ VanSledright & Limon (2006), s. 557.

³⁶ Monte-Sano & Reisman (2015), s. 290.

³⁷ Monte-Sano & Reisman (2015), s. 290.

³⁸ Lee & Ashby (2000), s 213. Jfr Shemilt (2009), s. 183–187. För motsvarande forskningsansats i Sverige, se Persson (2018). Jfr Karlsson (2014), s. 83.

³⁹ Se t.ex. Seixas et al som konstaterar att det förvisso har lanserats många forskningsbaserade modeller och begrepp med koppling till HT-traditionen, men att bedömningsdimensionen i många fall har saknats. Seixas et al (2015), s. 102 ff. Även Radinsky et al uppfattar att bedömningsforskningen är eftersatt inom det historiedidaktiska fältet. De betonar att tydlighet gällande progression krävs för att bedömningsverktygen i sin helhet ska vara legitima. Radinsky et al (2015), s. 140–141. Se även Grant (2018), s. 434.

⁴⁰ Ercikan & Seixas (red.) (2015).

forskarnas skilda epistemologiska utgångspunkter har skapat ett heterogent forskningsfält.⁴¹ De särskiljer den anglosaxiska mera praktikhäna forskningen, från den kontinentala forskningen med en mer teoretisk/historiefilosofisk förankring. Den senare mera teoretiskt inriktade forskningen kommer som sagt att behandlas i det efterföljande teorikapitlet.

Inom ramen för tidigare praktikhäna forskning finns ett tydligt spår som uppmärksammar hur historielärares undervisningspraktik påverkas av kunskapsbedömning i allmänhet och externa prov synnerhet.⁴² Resultaten i dessa studier har visat att historielärare påverkas av externa prov, men ett återkommande tema är att forskarna har haft svårt att urskilja några tydliga mönster på vilka olika sätt de påverkas. Den amerikanska bedömningsteoretikern Sandra Cimbricz tydliggör dock i en forskningsöversikt att tidigare forskningsresultat i stort sett samstämmigt visar att det framför allt är den epistemologiska dimensionen av historielärares undervisningspraktik som påverkas, det vill säga *vad* historielärarna undervisar om. Hon menar att väldigt få undersökningar har visat på någon påverkan inom den metodologiska dimensionen, det vill säga *hur* lärare undervisar.⁴³

Det är betydligt svårare att skaffa sig en överblick av tidigare forskning gällande externa provs påverkan på historielärares bedömningspraktik. Svårigheten beror framför allt på att tidigare forskning sällan särskiljer bedömningspraktiken från undervisningspraktiken. I många empiriska studier är bedömningspraktiken endast en av flera aspekter i en större undersökning med fokus på undervisningspraktiken.

S. G. Grant tycker sig dock kunna urskilja några mönster utifrån ett antal amerikanska studier, varav det mest grundläggande är att historielärares bedömningspraktik påverkas i begränsad omfattning av externa prov. Orsaken till denna begränsade påverkan är enligt Grant att historielärare redan tidigare har använt sig av de uppgiftsformat som återfinns i de externa proven, det vill säga flervalsuppgifter, kortsvarsuppgifter och essäuppgifter. Ett annat mönster som Grant identifierar är att historielärare visserligen uppfattar att det finns problem kopplade till central extern

⁴¹ Seixas & Ercikan (2015), s. 1-12. Se även Radinsky et al (2015), s. 141; Monte-Sano & Reisman (2015), s. 290; VanSledright & Limon (2006), s. 566.

⁴² Forskning kring hur historielärare genomför och använder externa prov är däremot svårare att hitta.

⁴³ Cimbricz (2002), s. 1-21.

bedömning, till exempel hur provresultaten används, vilka slags provuppgifter som används och vilket budskap om historieämnets karaktär som proven förmedlar, men trots dessa problem är det endast ett fåtal historielärare som är uttryckligt negativt inställda till själva idén om centrala externa prov. Grant lyfter fram flera orsaker till denna inställning, men han betonar framför allt att skriftliga prov tycks vara en allmänt accepterad bedömningsform. Det finns en konsensus bland historielärare att skriftliga prov är meningsfulla och att de representerar ett rimligt sätt att pröva elevers kunskaper i historia. Samtidigt tycker sig Grant kunna se att det hos många historielärare även finns en uppfattning att skriftliga prov endast prövar en viss slags kunskaper och färdigheter.⁴⁴

På ett mera generellt plan tycker sig Grant se flera stora metodologiska problem med den praktikinrä forskning som bedrivits inom området. Han menar att den i många fall inte bygger på något empiriskt underlag och i den mån forskningen bygger på ett empiriskt underlag så är det i huvudsak intervjubaserade studier. Den forskning som bygger på klassrumsobservationer är ytterst begränsad.⁴⁵ Grant sammanfattar forskningsläget gällande externa provs påverkan på historielärares undervisningspraktik på följande sätt: ”So despite a huge potential literature, little relates directly to impact on teachers practices.”⁴⁶

3.1.2 En nationell överblick

I Sverige har frånvaron av bedömningsrelaterad historiedidaktisk forskning inte uppmärksamats i samma utsträckning som på internationell nivå. Fokus har istället riktats mot balansen mellan den mer praktikinrä forskningen och den mer samhällsinriktade forskningen. År 1997 utgavs antologin, *Historiedidaktik*, med bidrag från ledande svenska, danska och finländska historiedidaktiker.⁴⁷ Klas-Göran Karlsson, en av redaktörerna för antologin, skriver i ett senare sammanhang att artiklarna i antologin beskriver den förändring som det historiedidaktiska forskningsfältet genomgick under perioden 1982–1997. Han tycker sig se en förändring från

⁴⁴ Grant (2006), s. 309.

⁴⁵ Grant (2006), s. 36–38.

⁴⁶ Grant (2003), s. 130. Grants beskrivning av det amerikanska forskningsfältet går att applicera även i en svensk kontext. Det finns en inom fältet tydlig metodologisk slagsida mot intervjubaserade studier, se t.ex. Nygren (2011); Persson (2017) och Jarhall (2020). Även föreliggande studie förstärker denna bild. Ett undantag är Fribergs studie som bygger på klassrumsobservationer. Friberg (2017).

⁴⁷ Karlegård & Karlsson (red.) (1997).

en mera praktiktäna forskning med fokus på skol- och utbildningsfrågor till en forskning med ”ett tilltagande intresse för såväl historiedidaktisk teoribildning och begreppsanalys som frågor om det förflutnas betydelse för individ och samhälle”. Vidare drar Karlsson en skiljelinje mellan den forskning som i första hand intresserar sig för ”kunskapens form och inlärningspsykologiska överväganden, utvärdering och historieuppfattning” och den som intresserar sig för ”relationen mellan historiedidaktik och historievetskap, historiemedvetandets teori och innebörd, sambanden mellan historia och identitetsbildning samt den historiska berättelsens särdrag”.⁴⁸

När man bläddrar igenom *Historiedidaktik* är det tydligt att de historiedidaktiska gränsdragningar som Karlsson beskriver motsvaras av en geografisk skiljelinje. De två finländska bidragen, skrivna av Sirkka Ahonen och Arja Virta, representerar en mer praktiktäna historiedidaktisk inriktning med betoning på kunskapsprocessens betydelse. Både Ahonen och Virta har dessutom en tydlig bedömningsteoretisk profil. De svenska bidragen, skrivna av Klas-Göran Karlsson, Christer Karlegård och Ulf Zander kan däremot sägas representera en inriktning med ett bredare samhällsligt perspektiv där identifikationsprocessen, med historiemedvetande som ett centralt begrepp, står i centrum.⁴⁹

Bengt Schüllerqvist beskriver i en forskningsöversikt perioden 1987 – 1996 som en stagnationsperiod för svensk historiedidaktisk forskning. Under den efterföljande perioden fram till 2005 ökar sedan intresset för historiedidaktiska frågeställningar, men forskningen har framför allt koncentrerat sig på historiedidaktikens samhällsliga sida.⁵⁰ Genomgående har det varit ett svalt intresse hos de ledande historiedidaktikerna för en mer skolrelaterad forskning. Även Karlsson menar att utvecklingen bort från skolan förstärktes i Sverige under början av 2000-talet och i ett arbete från 2008 efterlyser han ”ett förnyat intresse för skolans historieundervisning, dess praktik, funktioner och förutsättningar”.⁵¹

Karlssons förhoppningar från 2008 har delvis infriats under 2010-talet. Den skolrelaterade forskningen har fått ett uppsving och det gäller särskilt den bedömningsrelaterade forskningen. Kunskapsbedömning i allmänhet

⁴⁸ Karlsson (2008), s. 15.

⁴⁹ Karlegård & Karlsson (red.) (1997).

⁵⁰ Schüllerqvist (2005), s. 14–26.

⁵¹ Karlsson (2008), s. 206.

och extern kunskapsbedömning i synnerhet hade fått ett allt större utrymme såväl generellt som inom det historiedidaktiska forskningsfältet. Den historiedidaktiska forskningen med en mer eller mindre uttalad bedömningsteoretisk inriktning har ökat i omfattning. I Sverige har det på senare år vuxit fram bedömningsinriktade forskningsmiljöer i spåren av de nationella proven. Dessa miljöer kan även kopplas till de nationella forskarskolorna med historiedidaktisk inriktning i Karlstad och Malmö/Lund. Ett tydligt exempel på att fältet har breddats är antologin *Att bedöma i historieämnet* som utkom 2020 och som inrymmer bidrag från många av de ledande forskarna inom fältet.⁵²

3.2 Olika perspektiv inom forskningsfältet

3.2.1 Ett kategoriseringsverktyg

I en forskningsöversikt från 2020 kartlägger Mikael Berg och Anders Persson de senaste tio årens svenska forskning om lärares bedömningspraktik i historie- och samhällskunskapsundervisning. Till stöd för kartläggningen har de utarbetat ett kategoriseringsverktyg som gör det möjligt att särskilja fyra olika perspektiv inom forskningsfältet: 1) det värderande perspektivet 2) det förklarande perspektivet 3) det policy-kritiska perspektivet och 4) det lärar-erfarande perspektivet.⁵³

Med det värderande perspektivet avses forskning där forskaren, utifrån ”en egen färdig idé om hur läroplanen på en formell nivå bör tolkas”, påvisar eventuella diskrepanser mellan läroplanen och historielärares bedömningspraktik.⁵⁴ Med det förklarande perspektivet avses forskning som har en uttalad ambition att förklara variationen mellan historielärares olika sätt att hantera sitt bedömningspraktiska uppdrag. Med det policy-kritiska perspektivet avses forskning som undersöker vilka negativa effekter mättraditionellt förankrade policy-influenser har på historielärares bedömningspraktik. Med det lärar-erfarande perspektivet avses forskning som undersöker historielärares egna erfarenheter av sitt bedömningspraktiska uppdrag.

⁵² Ludvigsson & Andersson Hult (red.) (2021).

⁵³ Berg & Persson (2020). Deras kategoriseringsverktyg bygger i sin tur på delar av det teoretiska ramverket i Perssons avhandling. Persson (2017).

⁵⁴ Berg & Persson (2020), s. 34.

Med stöd av kategoriseringsverktyget argumenterar de för att det svenska forskningsfältet, till skillnad från den internationella, kan beskrivas som ett perspektivfattigt forskningsfält där forskning med ett värderande och/eller förklarande perspektiv dominerar.

Berg och Perssons kategoriseringsverktyg är i många avseenden användbart för att på ett systematiskt sätt kartlägga forskningen inom fältet. Ett problem med analysverktyget är dock att kategoriseringen av perspektiven kan uppfattas som hierarkisk. Det gäller i första hand deras benämning och beskrivning av den första kategorin ”det värderande” perspektivet. Att i vetenskapliga sammanhang på detta sätt få sin forskning förankrad i ett ”värderande” perspektiv är föga eftersträvänsvärt. Ett sätt att undvika denna hierarkisering är att benämna och beskriva detta perspektiv som ett **”alignment-baserat”** perspektiv, det vill säga forskning som uppmärksammar i vilken utsträckning policy-influenser och bedömningspraktik ligger i linje med varandra.⁵⁵

Ett annat problem är benämningen av den andra kategorin, det policykritiska perspektivet. Berg och Persson begränsar detta perspektiv till att endast omfatta forskning som fokuserar på de negativa effekter som mättraditionellt förankrade policy-influenser har på historielärares bedömningspraktik. Genom att istället benämna detta perspektiv som **”effekt-baserat”** inryms forskning som uppmärksammar såväl positiva som negativa effekter.

Ett tredje problem med kategoriseringsverktyget är deras benämning av den fjärde kategorin, ”det lärar-erfarande” perspektivet. Eftersom de flesta studier inom fältet är intervjubaserade och därmed gör anspråk på att beskriva lärarnas egna erfarenheter av sin bedömningspraktik kan benämningen uppfattas som vag. När Berg och Persson förtydligar vad som utmärker det här perspektivet framkommer dock att det snarare handlar om forskning som uppmärksammar hur lärare hanterar olika dilemma/utmaningar/spänningsförhållanden i sin bedömningspraktik. En alternativ benämning skulle därför kunna vara ett **”problem-baserat”** perspektiv.

⁵⁵ Begreppet alignment behandlas mer utförligt i avsnitt 8.3.2.

Ett fjärde, mera övergripande, problem med kategoriseringsverktyget är det perspektiv som Berg och Persson benämner det ”förklarande perspektivet”. Som jag ser det utgör ambitionen att ”förklara” inget perspektiv i sig utan kan snarare beskrivas som ett grundläggande fundament för all vetenskaplig forskning.

I linje med ovanstående resonemang har jag utarbetat ett alternativt tredelat kategoriseringsverktyg enligt nedan.

Alignment-baserat perspektiv
Forskning som uppmärksammar i vilken utsträckning policy-influenser och lärares bedömningspraktik är länkade till varandra.
Effekt-baserat perspektiv
Forskning som uppmärksammar policy-influensernas påverkan på lärares bedömningspraktik.
Problem-baserat perspektiv
Forskning som uppmärksammar hur lärare hanterar olika dilemman/utmaningar/spänningsförhållanden i sin bedömningspraktik.

Med stöd av kategoriseringsverktyget kan jag på ett systematiskt sätt positionera mitt eget forskningsprojekt inom fältet samt redovisa tidigare studier med särskild relevans för detta projekt.⁵⁶

3.2.2 Det alignment-baserade perspektivet

Föreliggande studie i relation till perspektivet

Studier förankrade i det alignment-baserade perspektivet uppmärksammar som sagt förskjutningar och/eller diskrepanser (glapp) mellan policy-influenser och historielärares bedömningspraktik.⁵⁷ Föreliggande studie har inget uttalat alignment-baserat perspektiv, men de tre första forskningsfrågorna och då särskilt den fråga som handlar om historielärares uppfattningar om hur frivilliga externa prov i historia är konstruerade tangerar detta perspektiv. I studien kategoriseras HLP och KP epistemologiskt utifrån MH-traditionen och HT-traditionen. Därefter relateras *kategoriseringen av proven* till historielärarnas *uppfattningar om proven*.

⁵⁶ Notera att de olika perspektiven inte förekommer i renodlad form i en specifik studie. Perspektiven ska uppfattas om idealtyper för att underlätta en kartläggning av forskningsfältet.

⁵⁷ Radinsky et al (2015) resonerar kring alignment i ett historiedidaktiskt sammanhang genom att definiera alignment som sambandet mellan C-I-A (Curriculum, Instruction, Assessment). Se även Rosenlund (2016) för ett elaborerat resonemang kring alignment i ett historiedidaktiskt sammanhang.

Eventuella diskrepanser analyseras sedan i ljuset av den fjärde kvadrantens dilemma.

Relevanta studier förankrade i perspektivet

En studie med ett tydligt alignment-baserat perspektiv är David Rosenlunds avhandling från 2016.⁵⁸ I avhandlingen analyserar han 1994/2000 års styrdokument och positionerar dessa i en kontinental HT-tradition. Därefter analyserar han i vilken utsträckning historielärares skriftliga bedömningsunderlag ligger i linje med gällande styrdokument. I studien framkommer att historielärarna i hög grad konstruerar bedömningsuppgifter som endast mäter innehållsrelaterade minneskunskaper, det vill säga uppgifter med en tydlig MH-traditionell förankring ("the transfer of content"). Den variationsförklarande dimensionen är underordnad i Rosenlunds studie, men han för ett resonemang om att det centrala innehållet var så vagt formulerat i 1994/2000 års kursplaner att lärarna i avsaknad av annan vägledning blev hänvisade till att fortsätta undervisa enligt gamla invanda spår.

Även Johan Samuelssons studie från 2015 har ett tydligt alignment-baserat perspektiv.⁵⁹ På samma sätt som Rosenlund beskriver han 1994/2000 års läroplaner som influerade av HT-traditionen. I studien undersöker han sedan hur åtta historielärares bedömningspraktik förhåller sig till de HT-traditionella styrdokumenten genom att analysera deras skriftliga bedömningsunderlag i kombination med lärarintervjuer. Resultatet ligger delvis i linje med Rosenlund. Tre av lärarna avviker från de HT-traditionella intentionerna i styrdokumentet och Samuelsson beskriver deras bedömningspraktik som "deklarativ allmändidaktisk praxis" där fokus ligger på att eleverna ska "skriva av rätt" och "minnas".⁶⁰ Denna deklarativa kunskapssyn med fokus på "reproduktion av det förflutna" motsvarar i hög grad den MH-traditionella ämnesuppfattningen. Till skillnad från Rosenlund beskriver Samuelsson en större spännvidd mellan lärarnas bedömningspraktik. Två av lärarna tillämpar en "processuell historiedidaktisk praxis", vilket motsvarar den HT-traditionella ämnesuppfattningen, och de övriga tre positioneras i ett mellanläge. Den variationsför-

⁵⁸ Rosenlund (2016).

⁵⁹ Samuelsson (2015).

⁶⁰ Samuelsson (2015), s. 62.

klarande dimensionen hos Samuelsson är på samma sätt som hos Rosenlund underordnad, men det finns en hänvisning till lärares utbildning – där en kortare utbildning tenderar att skapa en deklarativ allmändidaktisk praxis. I slutdiskussionen uppmärksammar Samuelsson även att den deklarativa praxisen tenderar att stärkas när de mättraditionella inslagen av bedömningskulturen stärks.

Noteras kan att Rosenlunds och Samuelssons resultat ligger i linje med den rapport som Skolinspektionen publicerade 2015. I rapporten framkommer att historielärares undervisning på grundskolan i flera avseenden inte ligger i linje med styrdokumentet.⁶¹

3.2.3 Det effekt-baserade perspektivet

Föreliggande studie i relation till perspektivet

Studier med ett effekt-baserat perspektiv uppmärksammar i bred bemärkelse vilken effekt policy-influenser har på historielärares bedömningspraktik.⁶² Berg och Persson benämner som sagt perspektivet som ”det policy-kritiska perspektivet” och de snävar in perspektivet genom att begränsa det till att omfatta forskning som förhåller sig kritiskt till mättraditionellt förankrade policy-influenser. I praktiken innebär det att de lyfter fram forskning som på olika sätt fokuserar på negativa effekter av dokumenterbar kunskap och bedömning. Mättraditionen ställs mot en bildningstradition.⁶³ I föreliggande kategorisering uppfattas perspektivet bredare: att kritiskt granska policy-influensers påverkan kan och kanske bör innebära att uppmärksamma såväl negativa som positiva aspekter. Perspektivet är framträdande i mitt eget forskningsprojekt. Den fjärde forskningsfrågan fokuserar explicit på historielärares uppfattningar om hur deras bedömningspraktik påverkas av externa prov i historia.

Relevanta studier förankrade i perspektivet

Internationellt har det effekt-baserade perspektivet fått ett stort genomslag, framför allt inom den anglosaxiska forskningen. Den amerikanske historiedidaktikern S.G. Grant har i en studie med en uttalat policykritisk ansats undersökt hur ett centralt externt prov i historia, New York Regents

⁶¹ Skolinspektionen (2015).

⁶² Grant identifierar tre olika externa policy-influenser med potential att påverka lärares didaktiska verksamhet: politiska styrdokument, externa prov och läroböcker. Grant (2003), s. 174.

⁶³ Berg & Persson (2020), s.23–24.

Exams (NYRE), påverkar historielärares undervisningspraktik.⁶⁴ Han har intervjuat två lärare, som han i undersökning kallar George Blair och Linda Strait, och sedan observerat deras undervisningspraktik under ett moment där de behandlar den amerikanska medborgarrättsrörelsen. Grant målar upp en bild av två historielärare med diametralt skilda uppfattningar om historieämnet. Han använder en annan terminologi, men Blair beskrivs som MH-traditionell med ett uttalat fokus på etablerade nationella narrativ medan Strait beskrivs som HT-traditionell med större fokus på källhantering och olika tolkningar av historiska narrativ. Grants slutsatser gällande vilken påverkan NYRE har på deras historieundervisning, är att båda två påverkas av provet, men att de påverkas på helt olika sätt. Blair uppfattar att han påverkas i begränsad omfattning eftersom han uppfattar att hans undervisning ligger i linje med det som externa provet efterfrågar. Trots att Blair själv inte uppfattar att hans undervisning påverkas av provet menar Grant att han påverkas indirekt eftersom hans undervisningspraktik blir bekräftad av provet. Strait uppfattar, till skillnad från Blair, provet som problematiskt då hon upplever att hon måste prioritera mellan undervisningsmoment som hon tycker är viktiga och undervisningsmoment som hon uppfattar att det externa provet efterfrågar. Ibland lyckas hon balansera dessa prioriteringar, men inte alltid. Grant menar att i hennes fall har det externa provet en direkt påverkan på hennes undervisningspraktik.⁶⁵

Grants distinktion mellan direkt och indirekt påverkan är intressant, men inte helt oproblematiske. Historielärare som uppfattar att de inte påverkas av ett externt prov kan förvisso vara lärare vars undervisning ligger i linje med det externa provets intentioner, men det behöver inte vara så. Den amerikanske bedömningsteoretikern Don Zancanella har i en studie visat att lärare i litteraturkunskap som uppfattar att de inte påverkas av externa prov antingen har en så pass hög kollegial status (curricular power) att de kan välja att ignorera provet helt och hållet, eller att deras undervisningspraktik ligger så pass nära det externa provets intentioner (curricular alignment) att de uppfattar att de inte behöver justera sin undervisning.⁶⁶

⁶⁴ Grant (2003).

⁶⁵ Grant (2003), s. 124.

⁶⁶ Zancanella (1992), s. 292–294.

Det effekt-baserade perspektivet har inte fått ett lika stort genomslag på en nationell nivå.⁶⁷ Ett viktigt undantag utgör Jessica Jarhalls avhandling från 2020.⁶⁸ I studien intervjuas historielärare i grundskolan före och efter Lgr 11. Jarhall använder begreppet omformningsfaktor för att benämna de faktorer som lärare själva uppfattar som avgörande för hur de utformar sin undervisning. Ett av studiens resultat är att historielärare verksamma efter Lgr 11 uppfattar att de nationella prov i historia som introduceras 2013 utgör en av de viktigaste omformningsfaktorerna. Lärarna uppfattar att det nationella provet påverkar såväl deras val av centralt innehåll som det sätt som de konstruerar uppgifter och bedömer elevernas kunskaper. I studien framkommer att de uppfattar att de nationella proven har påverkat dem att i högre utsträckning betona källhantering och historiebruk i sin undervisning, det vill säga en förskjutning mot en mer HT-traditionell undervisning.⁶⁹

Vid sidan av Jarhalls studie har Skolverket initierat ett antal studier som på en ämnesövergripande nivå har uppmärksammat de nationella provens påverkan på lärares praktik. Särskilt intressant för min forskning är en rapport från 2005, baserad på bland annat lärarintervjuer, som uppmärksammar de nationella provens betydelse för en likvärdig bedömning. Ett av studiens huvudresultat är att de nationella proven förbereds, genomförs och bedöms på ett mycket varierande sätt i Sveriges gymnasieskolor.⁷⁰ I en annan intervjubaserad rapport från 2003, som undersöker de nationella provens betydelse för grundskolan, framkommer att en majoritet av lärarna uppfattar att proven *inte* påverkar deras ämnesinnehåll eller arbets sätt i någon större utsträckning. De uppfattar inte heller att de anpassar sin undervisning särskilt till proven. I samma undersökning uppmärksammas även en skillnad mellan erfarna och oerfarna lärare där de senare uppfattar att de påverkas av provet i högre utsträckning än de förra.⁷¹ I en studie från 2016 initierad av Skolverket framkommer dock en tydlig förändring. En av slutsatserna är att lärarnas undervisning anpassas till, och

⁶⁷ En studie med ett tydligt effektbaserat perspektiv är dock Gunnemyr (2011).

⁶⁸ Jarhall (2020).

⁶⁹ Jarhalls studie inrymmer även ett alignment-baserat perspektiv. Hon visar på en förändring som innebär att lärarprofessionen under perioden 1994-2011 förhåller sig friare till styrdokumentet jämfört med vad de gör efter 2011. Lärarna ger sig själva ett större tolkningsutrymme och förlitar sig på en inofficiell kanon. Jarhall förklarar skillnaden genom att bland annat lyfta fram Skolverkets roll vid implementeringen av 1994 respektive 2011 års läroplaner.

⁷⁰ Skolverket (2005), s. 114.

⁷¹ Skolverket (2003), s. 79.

påverkas av, de nationella proven i mycket högre grad jämfört med vad som framkom i 2003 års studie.⁷²

Att policy-influenser i form av styrdokument och externa prov påverkar historielärares undervisningspraktik är som framkommit väl belagt i tidigare forskning. En viktig fråga är om det finns någon given hierarki mellan dessa båda policy-influenser. Om det finns en epistemologisk diskrepans mellan styrdokument och externa prov så visar amerikansk forskning att provets påverkanseffekt är starkare. Enligt Grant är det legio att amerikanska historielärare uppfattar en sådan bristande överensstämmelse mellan de externa provens budskap och de budskap som förmedlas av de styrdokument som lärarna har att förhålla sig till.⁷³ Grants sammanfattande slutsats är att de externa proven har en högre påverkanseffekt än styrdokumentet, men han betonar samtidigt att policy-influenser generellt sett har en lägre påverkanseffekt än vad många hävdar. De påverkar i första hand *vad* lärare undervisar om, inte *hur*, och de påverkar dessutom olika lärare på högst olika sätt.⁷⁴

Även Löfström et al uppmärksammar relationen mellan policy-influenserna. I en artikel som handlar om samhällskunskapsämnet i Finland ställer de sig frågan vilka aktörer som egentligen sätter kriterierna för ett skolämne. I artikeln driver de tesen att de finländska ämnesplanerna i samhällskunskap är otydligt formulerade och att de inte inrymmer de grundläggande begrepp som konstituerar ”social studies literacy”. Deras slutsats är att de vaga skrivningarna i styrdokumentet har inneburit att de aktörer som konstruerar de finländska ämnesrealproven har fått en nyckelroll.

As the social studies core curriculum has been rather vague on competences also the designers of the social studies matriculation examination have faced a challenging situation, where they, in fact, are the major interpreters of the objectives and aims of social studies teaching and learning. Preparing the exam questions they effectively make authoritative statements about what the core concepts and competence in the subject are. [...] the expert group has an exceptionally central role in translating the diffuse objectives of social studies teaching in

⁷² Arensmeier et al (2016).

⁷³ Grant (2006), s. 313.

⁷⁴ Grant (2003), s. 180–181.

the *Core Curriculum* into a set of more operational concepts.⁷⁵

Artikelförfattarna uppfattar det som problematiskt att det är provkonstruktörerna och inte den finländska utbildningsstyrelsen som är huvudaktörer när skolämnets konstrukt ska konkretiseras och operationaliseras.⁷⁶

3.2.4 Det problem-baserade perspektivet

Föreliggande studie i relation till perspektivet

Studier med ett problem-orienterande perspektiv uppmärksammar hur lärare hanterar olika former spänningsförhållanden eller dilemman relaterade till bedömningspraktiken. Till skillnad från de två övriga perspektiven utgör inte policy-influenser någon bärande utgångspunkt för detta perspektiv. Berg och Persson benämner därför kategorin som det lärar-erfarande perspektivet. Föreliggande studie har ett tydligt problem-baserat perspektiv. Den fjärde kvadrantens dilemma utgör avhandlingens bärande hypotes och samtliga forskningsfrågor analyseras i ljuset av detta dilemma.

Relevanta studier förankrade i perspektivet

Inom fältet finns ett flertal studier som kan beskrivas som problem-baserade. Ett tydligt exempel är Fredrik Alvéns som i sin avhandling från 2017 uppmärksammar spänningen mellan kursplanen i historia och värdegrundsuppdraget, såsom det kommer till uttryck i styrdokument (den formella läroplanen) och bland elever (den upplevda läroplanen). I studien undersöker Alvéns hur 15 historielärare uppfattar och hanterar sitt uppdrag i relation till detta spänningsförhållande mellan lärarens fostrans- och kunskapsuppdrag.⁷⁷

En studie som ligger närmare min egen avhandling och som indirekt uppmärksammar den fjärde kvadrantens dilemma är Anders Perssons avhandling från 2017.⁷⁸ I studien undersöker Persson hur tjugosex svenska grundskolelärare tolkat, hanterat och upplevt 2011 års skolpolitiska refor-

⁷⁵ Löfström, Virta & Van den Berg (2010), s. 9.

⁷⁶ Löfström, Virta & Van den Berg (2010), s. 11–12.

⁷⁷ Alvéns (2017).

⁷⁸ Persson (2017).

mer. Detta gör han i ljuset av vad han beskriver som en internationell policyförskjutning mot mättraditionella ideal, ”policy as numbers”.⁷⁹ Persson kopplar samman policyförskjutningen med lärarnas deprofessionalisering och han förankrar spänningsförhållandet i Stephen J. Balls begrepp ”performativitetskultur”.⁸⁰ Persson argumenterar i sin studie för att detta spänningsförhållande har inneburit ett paradigmskifte i läraruppdraget och han ser en direkt koppling mellan policyförskjutningen och skolans kunskapsuppdrag. Persson uppmärksammar således spänningsförhållandet mellan mättraditionella verktyg och HT-traditionella konstrukt. Notera dock att Persson argumenterar på en hög abstraktionsnivå och att de förändrade förutsättningarna runt bedömning endast utgör en av flera andra undersökta aspekter.

Även Henrik Friberg kontextualiserar sin studie från 2017 på ett sätt som ligger nära den fjärde kvadrantens dilemma.⁸¹ Till skillnad från de andra studierna där fokus ligger på insamlade prov och lärarintervjuer, så bygger Fribergs undersökning i huvudsak på klassrumsobservationer. I studien beskriver han på samma sätt som Persson det spänningsfält som har aktualiserats av de senast genomförda skolreformerna, reformer som har inneburit större fokus på bedömning (grund- och gymnasireformerna med nya ämnesplaner, ny betygsskala, betyg från årskurs sex och fler nationella prov). Fribergs resultat ligger i linje med de ovan refererade alignment-baserade studierna, det vill säga att lärarnas bedömningspraktik tenderar att vara MH-traditionellt förankrad. Han konstaterar exempelvis att det i första hand tycks vara innehållskunskaper som utgör den primära basen för lärarnas bedömningar i klassrummet: ”Innehåll först, förmågor sedan”.⁸²

3.3 Tidigare forskning i relation till föreliggande studie

Av ovanstående genomgång framgår att den bedömningsinriktade forskningen länge har varit ett eftersatt område inom det historiedidaktiska fältet, såväl nationellt som internationellt. En förändring är dock möjlig att spåra och fältet har på senare år breddats. Det är dock ett förhållandevis nyetablerat fält som brottas med olika dilemman. På en internationell nivå

⁷⁹ Persson (2017), s. 4.

⁸⁰ Persson (2017), s. 5–6.

⁸¹ Friberg (2017).

⁸² Friberg (2017), s. 117.

har fokus i första hand riktats mot fältets epistemologiska och/eller terminologiska heterogenitet och utmaningen har i första hand varit att hitta en gemensam grund att utgå från. På en nationell nivå har fokus riktats mot spänningsförhållandet mellan den mer praktikinriktade forskningen och den mer samhällsinriktade forskningen. Berg och Persson tycker sig dessutom se en obalans eller ensidighet inom forskningsfältet. Denna ensidighet kommer till uttryck metodologiskt i en övervikt av intervju-baserade studier och epistemologiskt i en övervikt av alignment-baserade studier.

Min egen studie är förvisso en i raden av intervjubaserade studier, men epistemologiskt kan den uppfattas som ett sätt att berika detta perspektivfattiga fält. Studien har ett tydligt effekt-baserat perspektiv och eftersom den fokuserar på gymnasielärare kan den utgöra ett viktigt komplement till Jarhalls studie baserad på historielärare i grundskolan. Min egen studie inrymmer dessutom ett tydlig problem-baserat perspektiv, manifesterat i hypotesen om den fjärde kvadrantens dilemma. Spänningsförhållandet mellan en mättraditionell kunskapsbedömning och en HT-traditionell ämnesuppfattning har uppmärksamrats tidigare, men i föreliggande studie är analysen empiriskt förankrad i historielärares uppfattningar om externa prov i historia. Det empiriska underlaget är dessutom insamlat vid två olika tillfällen, 2009 och 2016, vilket möjliggör en analys av förändring över tid.

4. TEORIER OCH BEGREPP

I teorikapitlet kommer avhandlingens centrala begrepp att definieras och belysas. Begreppsdefinitioner och begreppsanvändning är i många avseenden grundläggande för all form av vetenskaplig verksamhet. De centrala begreppen utgör de fundamentala byggstenarna i ett teoribygge, byggstenar som skapar ramar och som ger forskningsprojektet en färdriktning. Uppdraget att definiera de centrala begreppen är kanske särskilt viktigt inom vetenskapliga fält som är relativt nyplöjda, som det historiedidaktiska.

Det övergripande syftet för avhandlingen är, som tidigare framgått, att undersöka historiekulturell förändring över tid och då särskilt beakta vilken betydelse kunskapsbedömning i allmänhet och externa historieprov i synnerhet kan ha för denna förändring. I teorikapitlet redovisas ett tredelat teoretiskt ramverk. I den första delen behandlas begreppet historiekultur som kan sägas utgöra ramverkets *primära* portalbegrepp. I den andra delen behandlas begreppet bedömningskultur som utgör det *sekundära* portalbegreppet. I den tredje och avslutande delen av det teoretiska ramverket behandlas den fjärde kvadrantens dilemma, det vill säga det hypotetiska spänningsförhållandet mellan historiekulturen och bedömningskulturen.

4.1 Historiekultur – det primära portalbegreppet

4.1.1 Begreppshistorik

Historiekultur är ett omdiskuterat begrepp som har fått ett stort genomslag inom det historiedidaktiska fältet, såväl internationellt som nationellt. Inom fältet framstår den tyske historiefilosofen Jörn Rüsen som den viktigaste introduktören av begreppet.⁸³ 1994 skriver han att begreppet har börjat användas under de senaste åren och att det utgör en egen begreppslig kategori som inte kan infogas i ”kulturlivets ingrodda underavdelningar”.⁸⁴ I sin egen explicita definition relaterar Rüsen begreppet till det

⁸³ För en mer utförlig begreppshistorik, se Grever & Adriaansen (2017) s. 74–77.

⁸⁴ Rüsen (2004), s. 151.

då väletablerade begreppet historiemedvetande. Han skriver att historiekultur kan definieras ”som en artikulering med praktiska konsekvenser av historiemedvetande i ett samhälle”.⁸⁵ Frånvaron av referenser till tidigare forskning tolkar jag som att Rüsen inte har någon etablerad explicit begreppsdefinition att utgå ifrån.⁸⁶

Historiekultur har sedan Rüsen introducerade begreppet kommit att bli ett centralt begrepp inom den historiedidaktiska forskningen, särskilt i Norden. Det råder det dock ingen konsensus inom fältet hur begreppet ska definieras.⁸⁷ Utan att luta mig mot någon systematisk litteraturstudie tycker jag mig dock kunna identifiera två skilda huvudspår inom den nordiska historiedidaktiken. Det första spåret kopplar jag till Klas-Göran Karlsson.⁸⁸ Det finns en rad forskare, många av dem hemmahörande i en sydsvensk forskningsmiljö med bas i Lund och Malmö, som lutar sig mot Karlsson när de i sina teoribyggen använder sig av historiekulturbegreppet.⁸⁹ Det andra spåret kopplar jag till Peter Aronsson.⁹⁰ Aronssons definition har fått stort genomslag utanför den sydsvenska forskningsmiljön hos såväl svenska som norska och danska historiedidaktiker.⁹¹

4.1.2 Begreppskategorisering

I Karlssons definition framkommer att historiekultur *är* ett specifikt kommunikativt sammanhang.

En historiekultur [...] är det kommunikativa sammanhang inom vilket

⁸⁵ Rüsen (2004), s. 152. Rüsen särskiljer tre olika historiekulturella dimensioner: en politisk, en estetisk och en kognitiv.

⁸⁶ I andra begreppshistoriska genomgångar spåras begreppets ursprung till Gadamer, se t.ex. Grever & Adriaansen (2017), s. 74–77.

⁸⁷ Kvande & Naastad (2015), s. 55. Historiekultur utgör ofta en del av en begreppstrojka där historiemedvetande och historiebruk utgör de två andra delarna. En del av den bristande konsensus som råder inom fältet kring dessa tre begrepp kan kopplas till forskares olika sätt att relatera dem till varandra. Definitionen av ett av begreppen är i många fall avhängigt hur begreppet relateras till de båda andra begreppen. Noteras kan att Thorp (2013) argumenterar för att det inom fältet finns en form av konsensus gällande definitionen av historiemedvetandebegreppet.

⁸⁸ Karlsson har under lång tid intresserat sig för historiekulturbegreppet och att hans explicita begreppsdefinitioner har varierat en aning över tid. Genomgående har han dock varit djupt influerad av Rüsen.

⁸⁹ Se t.ex. Zander (2001); Ammert (2008); Danielsson Malmros (2012) och Sjöland (2016).

⁹⁰ Aronsson (2004). Noteras kan att när Aronsson redogör för sin explicita definition finns det inga referenser till tidigare definitioner, varken till Rüsen eller någon annan. Min tolkning av denna frånvaro av referenser är att Aronssons definition är mer fristående än Karlssons. Ytterligare en notering är att Aronssons bok bygger på reviderade texter från perioden 1997–2004. Det är tydligt att det uppstår betydelseglidningar när Aronsson använder historiekulturbegreppet i de olika avsnitten av boken.

⁹¹ Se t.ex. Nordgren (2016), Kvande & Naastad (2015) och Kayser Nielsen (2010).

historiemedvetandet och minnet verkar och det förflutna tillskrivs mening.⁹²

Karlsson är medveten om att hans definition är bred och därmed svår att använda i ett analytiskt sammanhang. Han särskiljer därför mellan två analytiska underkategorier: historiekulturens text och kontext. Med den förra underkategorin avser han ”historieprodukterna som sådana” och han exemplifierar med ”läroböcker, vetenskapliga monografier, filmer och foton, debatter, monument och ritualer”. Med den senare underkategorin, den historiekulturella kontexten, avser han de olika slags villkor som utgör ramar när historieprodukterna i fråga blir ”’kultur’ och ’lärdomar’”.⁹³

I Aronssons explicita definition *är* historiekultur det som Karlsson benämner som historiekulturens text eller ”historieprodukterna som sådana”.

Historiekultur är de källor, artefakter, ritualer, sedvänjor och påståenden med referenser till det förflutna som erbjuder påtagliga möjligheter att binda samman relationen mellan dåtid, nutid och framtid. I undantagsfall utgör de direkta och uttryckliga tolkningarna av detta samband.⁹⁴

Genom att jämföra Aronssons och Karlssons konkreta exemplifieringar framgår tydligt skillnaden dem emellan. Det som för Aronsson utgör själva historiekulturen är för Karlsson endast en analytisk underkategori. Om man vill kategorisera de två begreppsdefinitionerna skulle man således kunna tala om en snäv respektive bred definition av historiekulturbegreppet.

4.1.3 Begreppsdefinition

I mitt eget sätt att närma mig begreppet historiekultur har jag valt att utgå från Karlsson och hans bredare definition. Jag har även låtit mig inspireras av Henrik Åström Elmersjö som definierar historiekultur genom att särskilja två olika tankeled. Han menar att historiekultur dels är olika sätt att konceptualisera historia beroende på kulturellt sammanhang, dels

⁹² Karlsson (2010), s. 76.

⁹³ Karlsson (2010), s. 76–88.

⁹⁴ Aronsson (2004), s. 17.

kommunikationen av denna konceptualisering via olika kanaler.⁹⁵ Med utgångspunkt från Karlsson och Åström Elmersjö har jag valt att definiera historiekultur som *det kommunikativa sammanhang inom vilket olika sätt att uppfatta det förflutna lyfts fram och ges mening*.⁹⁶

I det här sammanhanget kan det vara på sin plats att resonera kring relationen mellan begreppet historiekultur och skolämnet historia. I linje med Åström Elmersjö uppfattar jag historieämnet som en del av historiekulturen. Historia, historiekultur och historieämnet är sammanvävda. Jag kommer därför fortsättningsvis benämna den historiekultur som kommer till uttryck i skolan som *skolans historiekultur*.⁹⁷ I förlängningen av ovanstående resonemang kan man säga att skolans historiekultur manifesteras i skolämnet historia, närmare bestämt i historieämnets vad- och varför-dimension, det vill säga i de epistemologiska uppfattningarna om vilka kunskaper som inryms i historieämnet och varför dessa kunskaper är meningsfulla.

4.1.4 Sammanfattning

Genom mitt val av historiekultur som portalbegrepp har jag positionerat mig inom det historiedidaktiska fältet. Genom mitt breda sätt att definiera historiekulturbegreppet har jag dessutom positionerat mig själv i Karlssons sydsvenska forskningsmiljö. Värt att notera är dock att för de flesta nordiska historiedidaktiker med ett intresse för historiekultur är det inte Karlsson som utgör utgångsreferensen, utan det är Aronsson och hans snäva definition som lyfts fram och modifieras.⁹⁸

4.2 Historiekultur som analysverktyg

4.2.1 Analytiska underkategorier

Nackdelen med breda definitioner är att de behöver brytas ner i analytiska underkategorier för att kunna operationaliseras i en empirisk studie.⁹⁹

⁹⁵ Åström Elmersjö (2017), s. 18.

⁹⁶ Till skillnad från Karlsson har jag valt att inte precisera definitionen med stöd av begreppen minne och historiemedvetande. Anledningen är att jag inte vill laborera alltför mycket med dessa svårhanterliga begrepp i mitt teoribygge.

⁹⁷ Åström Elmersjö (2017), s. 19. Aronsson menar skolundervisningen är en av historiekulturens former. Han varnar dock för att använda begreppet historiekultur i ett sammanhang där även analysobjektet är historia. Aronsson (2004), s. 43.

⁹⁸ Se not 90.

⁹⁹ Ytterligare en nackdel med breda begreppsdefinitioner är att de tenderar att blir så allomfattande att deras förklaringsvärde blir begränsat.

Karlssons distinktion mellan text och kontext uppfattar jag som ett klargörande steg i en sådan nedbrytning, men till skillnad från Karlsson har jag valt att benämna dessa båda underkategorier som "produkt" respektive "arena".¹⁰⁰ En annan skillnad jämfört med Karlsson är att jag har valt att låta "process" utgöra en tredje analytisk underkategori. Schematiskt kan ovanstående resonemang illustreras i nedanstående modell (se figur 2).

Figur 2. Analytiska underkategorier

För att kunna tillämpa dessa underkategorier som analytiska verktyg behöver de dock preciseras ytterligare, vilket kommer att göras i de följande tre avsnitten.

4.2.2 Historiekulturella produkter

Som framgått tidigare menar Karlsson att historiekulturen kan analyseras textuellt och att i centrum för en sådan analys står en historieprodukt.¹⁰¹ Jag uppfattar att benämningen historieprodukt är en synonym till vad Karlsson i ett annat avsnitt av sitt teoribygge benämner historisk artefakt. I detta avsnitt klargör han att "en historisk artefakt är ett objekt eller icke-materiellt historiskt fenomen med koppling till ett förflutet som vi upplever att vi har en relation till".¹⁰² En viktig distinktion i detta sammanhang är skillnaden mellan en historisk produkt och en historiekulturell produkt.

¹⁰⁰ Med text avser Karlsson historieprodukterna i sig, t.ex. en lärobok. Vad gäller kategorin kontext menar han att den kan förstås på tre olika sätt: processuellt, strukturellt och funktionellt. Karlsson betonar dock att de två sistnämnda perspektiven är näraliggande varandra och att man ofta talar om ett strukturfunktionalistiskt betraktelsesätt. Karlsson (2010), s. 86 ff.

¹⁰¹ Karlsson (2010), s. 86 ff.

¹⁰² Karlsson (2010), s. 77.

Jag menar att epiteten "historisk" och "historiekulturell" är betydelsebärande. Genom att använda det senare epitetet markeras att produkten, med Karlssons formulering, inte bara har en koppling till ett förflutet, utan att den har en koppling *till ett förflutet som vi upplever att vi har en relation till*. Skillnaden ligger i graden av relevans. Om man använder sig av en skafferimetafor skulle det innebära att de historiska produkterna är alla de produkter som finns tillgängliga i det historiska skafferiet, medan de historiekulturella produkterna är de produkter som i ett givet sammanhang plockas fram och brukas. Det är således först när den historiska produkten aktualiseras och aktiveras som den kan benämnas som en historiekulturell produkt.¹⁰³

Inom det historiedidaktiska fältet finns exempel på flera historiekulturellt förankrade studier där någon form av historiekulturell produkt utgjort analysobjekt. Ett exempel är Niklas Ammert som i sin avhandling har analyserat läroböcker i historia och i sin analys satt in läroböckerna i en historiekulturell kontext. En lärobok i historia, menar Ammert, utgör en förmedlande historieprodukt mellan läroboksförfattaren och eleven.¹⁰⁴ Ett annat exempel är Åström Elmersjö som för ett liknande resonemang i sin studie av förhandsgranskningar av svenska läromedel. I hans studie utgör läromedelsgranskarnas sakkunnighetsutlåtanden den historiekulturella produkten.¹⁰⁵ I föreliggande studie kommer ett externt prov i historia att på motsvarande sätt uppfattas som en historiekulturell produkt.

4.2.3 Historiekulturella arenor

Begreppet arena

I en bok från 2014 lanserar Karlsson en delvis omarbetad definition av begreppet historiekultur. I boken definieras historiekultur som "den kulturella arena där ett samhälle kommunicerar och värderar vilken historia som ska ägnas uppmärksamhet [...]".¹⁰⁶ Det är en definition där begreppet arena lyfts fram på ett tydligare sätt än i den tidigare definitionen. En viktig förutsättning för att jag ska kunna använda historiekulturell arena

¹⁰³ En annan viktig distinktion är den mellan en historiekulturell produkt och en historiekulturell referens. Se t.ex. Nordgren (2016), s. 481.

¹⁰⁴ Ammert (2008), s. 25. Jfr Malmros (2012).

¹⁰⁵ Åström Elmersjö (2017), s. 39.

¹⁰⁶ Karlsson (2014), s. 65. Det jag uppfattar som problematiskt med denna definition är singularisformen av arenabegreppet. Karlsson är makrohistoriskt förankrad och har en preferens att utgå från nationen som autonom analytisk enhet. Han tar stöd mot Andersons begrepp *imagined communities* för att spåra stora meningsbärande berättelser som formar historiekulturen. Karlsson (2010) s. 20.

som en analytisk underkategori i mitt teoribygge är att jag kan skilja mellan historiekulturella arenor på olika nivåer. För att möjliggöra en sådan nivåindelning har jag tagit stöd i teoribildning hämtad från det läroplans-teoretiska fältet.

Arenor på olika nivåer

Bo Lindensjö och Ulf P. Lundgren har i sin forskning undersökt förhållandet mellan de politiska och administrativa aktörernas intentioner, som de kommer till uttryck i styrdokument, och den faktiska styrning som dessa intentioner sedan får ute i skolorna. I sin analys använder de sig av en styrteoretisk modell som är uppbyggd kring tre arenor: en formuleringsarena, en transformerings- och medieringsarena samt en realiseringsarena. Lite förenklat kan man säga att modellen tydliggör att på formuleringsarenan fattas besluten och på realiseringsarenan omsätts dessa beslut i handling. Medierings- och transformeringsarenan utgör ett mellanled i relation till de båda andra arenorna. De styrdokument som tar form på formuleringsarenan förmedlas, tolkas och transformeras av mediala aktörer innan de når fram till realiseringsarenan.¹⁰⁷

Genom att tillämpa delar av Lindensjö och Lundgrens teoribygge vill jag tydliggöra min teoretiska utgångspunkt att ett externt prov i historia är en historiekulturell produkt som hanteras på två olika arenor: en historiekulturell formuleringsarena och en historiekulturell realiseringsarena.¹⁰⁸

Verksamheten på arenorna

Som framgått ovan menar Lindensjö och Lundgren att verksamheten på formuleringsarenan handlar om att fatta beslut och att verksamheten på realiseringsarenan handlar om att genomföra dessa beslut. Problemet är dock att besluten som fattas på formuleringsarenan måste konkretiseras i form av skriftliga produkter innan de kan genomföras på realiseringsarenan.¹⁰⁹ I linje med Lindensjö och Lundgren uppfattar jag att verksam-

¹⁰⁷ Lindensjö & Lundgren (2000), s. 171–179.

¹⁰⁸ Inom det historiedidaktiska fältet används ofta Goodlads (1979) mer elaborerade läroplansteoretiska modell/teoribygge, se t.ex. Alvén (2017) och Nygren (2011).

¹⁰⁹ Lindensjö & Lundgren (2000), s. 147. Lindensjö och Lundgren tar sig an problematiken genom att beskriva två olika stadier av besluts genomförande: dels "guideline-writing" då ämbetsmän tolkar de fattade besluten och transformerar dem till skriftliga produkter i form av styrdokument, dels "field-implementation" då lärare

heten på formuleringsarenan inrymmer två dimensioner: en beslutsfattande politisk dimension och en operationell administrativ dimension. I samband med ett externt prov i historia innebär det att provbeställaren fattar beslut om att provet ska konstrueras; att provkonstruktörerna konkretiserar beslutet i ett prov och att historielärarna sedan implementerar beslutet genom att genomföra detta prov.

Aktörerna på arenorna

Lindensjö och Lundgren har i sin forskning uppmärksammat relationen mellan de politiska och administrativa aktörerna, och de menar som framgått ovan att de senares roll är att konkretisera politikernas intentioner i form av skriftliga styrdokument, en konkretisering som de menar innebär en tolkningsprocess. De administrativa ämbetsmännens friutrymme i den här tolkningsprocessen är en empirisk fråga, men Lindensjö och Lundgren menar att deras handlingsutrymme generellt sett är stort.

Demokratiska system präglas av avstånd mellan politik, val och partikoalitioner å ena sidan och den ”verkliga” styrningsprocessen, som väsentligen sker genom byråkrati, å den andra sidan. [...] [Det] ökande antalet statliga uppgifter och de uppskruvade ambitionerna med dem [har] bidragit till att ”befria” förvaltningens ämbetsmän från mycket av det traditionella beroendet av deras politiska ledning.¹¹⁰

Problematiken kring ämbetsmännens, eller i min modell provkonstruktörernas, friutrymme leder vidare till komplicerade resonemang där begrepp som ramfaktorer och påverkansfaktorer har stor betydelse. I föreliggande forskningsprojekt ligger fokus i huvudsak på den mottagande parten i kommunikationsprocessen, det vill säga historieläraren, men jag kommer även att beröra problematiken kring provkonstruktörernas friutrymme.¹¹¹

omsätter styrdokumentens intentioner i praktisk handling. Jfr med Lipskys (2010) teori om gräsrotsbyråkrater, det vill säga offentliga tjänstemän som har direktkontakt med medborgarna.

¹¹⁰ Lindensjö & Lundgren (2000), s. 149.

¹¹¹ Jfr Andersson Hults resonemang om provkonstruktörens autonomi. Andersson Hult (2021), s. 51.

4.2.4 Historiekulturella processer

Historiekulturens dubbla processer

Karlsson menar att en historiekultur kan uppfattas som en dubbel process: en förmedlande och en värderande process. Om det är den förmedlande processen som står i centrum för analysen riktas fokus i första hand mot den kommunikativa kedjan mellan avsändare, förmedlare och mottagare. Om det är den värderande processen som står i centrum är det i första hand makten över vilka budskap som förmedlas i denna kommunikativa kedja som utgör analytiskt fokus.¹¹²

Jag delar Karlssons uppfattning att historiekulturen inrymmer två olika processuella dimensioner. Till skillnad från Karlsson och i linje med Åström Elmersjö väljer jag dock att använda mig av begreppet förhandling för att beskriva den värderande dimensionen.¹¹³ Historiekulturens dubbla processer i form av förmedling och förhandling utgör en viktig byggsten i mitt teoretiska ramverk och kommer därför att preciseras ytterligare.

Historiekultur som förmedling

Ett sätt att närma sig den förmedlande dimensionen av den historiekulturella processen framkommer i Peter Aronssons historiedidaktiska teori-bygge. Han konstaterar att vissa historiekulturella produkter har en alldeles bestämd upphovsperson. Det kan till exempel vara en skulptör av ett historiskt monument eller en avhandlingsförfattare i historia. Han konstaterar vidare att både skulptören och avhandlingsförfattaren är brukare av historia i ett avsändarled, men han betonar samtidigt att produkten även brukas i mottagarledet. Aronssons poäng är att mottagarna i sitt bruk av den historiekulturella produkten visserligen använder och tolkar produkten på sitt eget sätt, men att tolkningen inte är helt godtycklig; det finns alltid någon form av koppling till upphovspersonens intention. Denna koppling innebär att det uppstår en budskapsförmedlande länk mellan avsändare och mottagare.¹¹⁴

Jag delar Aronssons uppfattning att mottagarens tolkning av den historiekulturella produkten inte är godtycklig och menar vidare att hans resonemang kan överföras till extern provkonstruktion i historia. Ett externt

¹¹² Karlsson (2014), s. 65–66.

¹¹³ För ett utförligt resonemang om begreppet förhandling, se Åström Elmersjö (2017), s. 23–25.

¹¹⁴ Aronsson (2004), s. 98–101.

prov i historia kan uppfattas som en historiekulturell produkt i en budskapsförmedlande historiekulturell process där provkonstruktörerna är avsändare och historieläraren är mottagaren av dessa budskap.¹¹⁵ Eva Forsberg och Christian Lundahl använder en annan terminologi, men deras beskrivning av kunskapsbedömningar ”som ett signalsystem, det vill säga som ett system som skickar signaler om vad som ytterst är att betrakta som centrala kunskaper och valida realiseringar av dessa”, ligger nära mitt sätt att uppfatta externa prov i historia.¹¹⁶ I ett svenskt historiedidaktiskt sammanhang har framför allt Rosenlund uppmärksammat kunskapsbedömningens budskapsförmedlande funktion. Han uppmärksammar dock i första hand historielärens interna bedömning och den betydelse lärarens konstruktion av bedömningsverktyg får för elevens uppfattningar om historieämnet.¹¹⁷

Historiekultur som förhandling

Den historiekulturella förhandlingen om skolämnet historia kretsar kring två centrala frågeställningar: 1) Varför ämnet historia i skolan? 2) Vad ska uppfattas som historisk kunskap? På formuleringsarenan manifesteras utfallet av den här historiekulturella förhandlingen i skolans styrdokument för historieämnet och i förlängningen i de externa prov som baseras på dessa styrdokument. På realiseringsarenan manifesteras utfallet i historielärares undervisnings- och bedömningspraktik.

Om man på nytt använder sig av skafferimetaphoren handlar frågeställningarna om *vad* det är som plockas fram ur det historiekulturella skafferiets och *varför* just dessa produkter plockas fram. Frågeställningarna är intimt sammanlänkade. Vad som plockas fram är avhängigt i vilket syfte skafferidörren öppnas. Om syftet med historieämnet är att fostra kritiska medborgare är det kanske inte i första hand ett ensidigt urval av nationella narrativ som plockas fram utan snarare kontrasterande narrativ som möjliggör såväl perspektivbyten och som ett kritiskt tänkande.¹¹⁸

¹¹⁵ Det finns en omfattande forskning som har uppmärksammat att externa prov har en kommunikativ funktion. Se t.ex. Gipps (1994); Grant (2006); Madaus (1988); Vogler & Virtue (2007).

¹¹⁶ Forsberg & Lundahl (2006), s. 24. De uppmärksammar att kunskapsbedömningar, som ett medium, kan fylla både en styrande och en kommunikativ funktion. Forsberg & Lundahl (2006), s. 7–29.

¹¹⁷ Rosenlund (2016), s. 54 ff. Se även Jarhall (2020).

¹¹⁸ Jfr Alvén (2017).

4.2.5 Sammanfattning - en konceptuell modell

För att sammanfatta och tydliggöra mina resonemang kring begreppet historiekultur har jag konstruerat en enkel konceptuell modell.¹¹⁹ Modellen ska uppfattas som ett heuristiskt redskap för att tydliggöra den grundläggande tanken att det med nödvändighet uppstår en kommunikativ relation mellan provets externa aktörer och den undervisande historieläraren i samband med att ett externt prov konstrueras, genomförs och används (se figur 3).

Figur 3. Externa prov som del i en historiekulturell kommunikationsprocess

I modellen tydliggörs att den kommunikativa relationen mellan å ena sidan provbeställare och provkonstruktörer verksamma på en historiekulturell formuleringsarena och å andra sidan historielärare verksamma på en historiekulturell realiseringsarena omfattar dubbla processer. Det är dels en budskapsförmedlande process och dels en budskapsförhandlande process. Modellen tydliggör med andra ord den för avhandlingen grundläggande tankeoperationen att ett externt prov i historia förmedlar olika slags historiekulturella budskap till historielärarna samtidigt som det föranleder en förhandling om dessa budskap.

¹¹⁹ Min syn på konceptuella modeller är synnerligen pragmatisk. En konceptuell modell har inget egenvärde utan den ska vara användbar i ett forskningssammanhang – en bra konceptuell modell är en användbar modell. Ovanstående konceptuella modell uppfattar jag som ett heuristiskt redskap som hjälper mig att ställa frågor till det empiriska materialet och den ger mig dessutom de "glasögon" jag behöver för att kunna tolka empirin och sätta in mina resultat i ett större sammanhang. Konceptuella modeller innebär med andra ord alltid en förenkling, men de underlättar framskrivandet av analytiska poänger.

Externa prov i historia har således alltid en dubbel funktion: dels som ett bedömningsverktyg för att generera data om elevers kunnsighet i historia, dels som en historiekulturell produkt som ligger till grund för såväl förmedling av som förhandling om vad kunnsighet i historia innebär. En extern provkonstruktör kan genom sitt sätt att utforma provet förmedla budskap till historielärarna om vad som ytterst är att betrakta som centrala kunskaper inom historieämnet samt vad som utmärker progression gällande historiska kunskaper. Dessa budskap kan uppfattas som mer eller mindre normativa, det vill säga ligga till grund för en mer eller mindre öppen förhandling. I avhandlingen kommer den övergripande förhandlingen om skolämnet historia att uppfattas som en förhandling mellan två historiekulturella traditioner: MH-traditionen och HT-traditionen. Dessa kommer att beskrivas utförligt i nästa avsnitt.

4.3 Historiekulturella traditioner

4.3.1 Memory-history vs. Historical thinking

Vad-dimensionen gällande skolans historieämne handlar i grunden om vad som menas med historisk kunskap och i förlängningen om vad det innebär att vara kunnig i historia. Historisk kunskap kan teoretiseras utifrån två olika aspekter, en epistemologisk och en funktionell aspekt.

Epistemologiskt är det möjligt att göra en åtskillnad mellan en mer objektivistisk respektive en mer konstruktivistisk syn på historisk kunskap. Den objektivistiska synen på historisk kunskap innebär att en redogörelse för ett historiskt fenomen kan betraktas som antingen sann eller falsk. Den konstruktivistiska synen på historiska kunskaper innebär däremot att olika tolkningar av samma historiska fenomen, kan bedömas vara möjliga, samtidigt som en viss tolkning kan vara mer rimlig. Den grundläggande epistemologiska skillnaden är således graden av tolkningsutrymme i relation det historiska fenomenet. I grunden handlar den epistemologiska aspekten om sanningsanspråk.¹²⁰

¹²⁰ Jfr med Alvesson & Sköldbärgs resonemang om det trilaterala sanningsbegreppet. Objektivismen utgår från ett representativt sanningsbegrepp där det finns en direkt korrespondens mellan sanning och verklighet medan konstruktivismen utgår från ett signifikativt sanningsbegrepp som man når genom tolkning. Alvesson & Sköldbärg 2008, s. 35–39. På ett mera övergripande plan har objektivismen starka kopplingar till den evidensbaserade (kognitiva) forskningstraditionen. Den underliggande premisen för denna forskningstradition är farhågorna för kunskapsrelativism och i förlängningen ett samhälle präglad av faktaresistens. Konstruktivismen har starka kopplingar till den progressiva rörelsen och socialkonstruktivismen, ofta med ett uttalat makt-perspektiv. Se även Rosenlund som utifrån ett historiedidaktiskt perspektiv nyanserar skillnaden objektivismen och konstruktivismen. Rosenlund (2016) s. 43–44. Se även Virta (1997), s. 163–173.

Den funktionella aspekten av historisk kunskap handlar om hur kunskapen manifesteras. Jag kommer under analysarbetet att särskilja mellan två olika manifestationer av historisk kunskap: det historiska referatet och det historiska resonemanget. Det historiska referatet korrelerar i högre grad med den objektivistiska synen på historisk kunskap medan det historiska resonemanget korrelerar i högre grad med den konstruktivistiska synen på historisk kunskap.¹²¹

Den kanadensiske historiedidaktikern Stéphane Lévesque konkretiserar ovanstående resonemang när han särskiljer två olika traditioner att hantera den historiekulturella vad-dimensionen i skolämnet historia: den mer objektivistiskt orienterade "memory-history"-traditionen (MH-traditionen) och den mer konstruktivistiskt orienterade "historical thinking"-traditionen (HT-traditionen).¹²²

MH-traditionen

Lite förenklat kan man säga att MH-traditionen har sina historievetenskapliga rötter hos de tyska historiker, med Ranke i spetsen, som under 1800-talet tog avstånd från den då rådande traditionen att skriva historia med moraliska förtecken. Historikerns uppgift skulle istället vara att med stöd av vetenskaplig källkritisk metod ta reda på "vad som egentligen hände". Den grundläggande tankeoperationen var att det var möjligt att rekonstruera det förflutna och att det är historikerns uppgift att konstruera historiska narrativ som gör anspråk på att vara sanna.¹²³

I en MH-traditionell tolkning av skolämnet historia uppfattas historisk kunskap därför i grunden som minneskunskaper av olika slag.¹²⁴ Det kan handla om att minnas rena faktakunskaper som årtal och personer, men även mer krävande minneskunskaper som att kunna återberätta ett händelseförlopp, det vill säga att med stöd av minnet kunna rekonstruera ett

¹²¹ Den funktionella aspekten av historisk kunskap kan relateras till "de fyra F:n": fakta, förståelse, färdighet och förtrogenhet. De fyra F:n kan, som kan beskrivas som fyra olika kunskapsformer, har utgjort det kunskapsteoretiska fundamentet i den svenska skolan sedan 1994 års läroplansreformer.

¹²² Lévesque (2008), s. 3-17. Inom det historiedidaktiska fältet finns det ett flertal olika benämningar på MH-traditionen. Se t.ex. Duquette som använder benämningen "Traditional history". Duquette (2015), s. 52. Shemilt som använder benämningen "Historical knowledge". Shemilt (2018), s. 451. Rosenlund som använder benämningen "transfer-of-content tradition". Rosenlund (2016) s. 41.

¹²³ Se t.ex. Reisman (2019).) Se även Uggla (2002), s. 315-374.

¹²⁴ Samuelsson använder begreppet deklarativ kunskap när han beskriver vad som kännetecknar minnesbaserad kunskap i historia: "Historisk kunskap är något som ska överföras från skolan till en föreställt passiv elev." Samuelsson (2015), s. 64. Shemilt utgår från begreppet deklarativ kunskap när han resonerar om olika former av minneskunskaper inom historieämnet. Han menar att det är meningsfullt att särskilja mellan minneskunskaper i form av "recognition" och "recall". Shemilt (2018), s.451. Se även Virta (1997), s. 163-167.

givet historiskt narrativ, till exempel händelseförloppet i samband den franska revolutionen.¹²⁵

Det handlar således inte bara om att minnas, det innefattar även ett görande som manifesteras i det historiska referatet. Att referera komplexa historiska narrativ ska inte uppfattas och reduceras till en enkel kognitiv tankeoperation. Det goda historiska referatet kräver till exempel genomtänkta överväganden gällande historisk signifikans.¹²⁶

HT-traditionen

Återigen lite förenklat kan man säga att HT-traditionen har sina historievetenskapliga rötter hos de historiker som från olika håll under 1900-talet problematiserade det objektivistiska sanningsanspråket. De ifrågasatte om det var möjligt att ta reda på ”vad som egentligen hände” eftersom historikern inte kan frigöra sig från sina egna tolkande glasögon. Därmed inte sagt att det saknas ett sanningsanspråk inom HT-traditionen, snarare handlar det om att sanningsanspråket ligger på en annan nivå. De fakta som underbygger ett historiskt narrativ kan och bör provas mot det objektivistiska sanningsanspråket. Skillnaden jämfört med det MH-traditionella sanningsanspråket är med andra ord att det inte är narrativet i sig som är bärare av ett sanningsanspråk utan de fakta som underbygger narrativet.¹²⁷

I en HT-traditionell tolkning av skolämnet historia innebär historisk kunskap således något mer än bara minneskunskaper. Lévesques kollega Peter Seixas gör en jämförelse med en teaterföreställning när han särskiljer dessa traditioner. Han menar att en historielärare med ambitionen att förmedla minneskunskaper kan nöja sig med att låta eleverna ta del av föreställningen på scenen. Om läraren däremot har en ambition att utveckla kunskaper med koppling till HT-traditionen så ska eleven inte bara passivt ta del av föreställningen utan även bli medveten om de processer som äger rum bakom scenen och som möjliggör föreställningen.¹²⁸ I grunden handlar HT-traditionen om att eleven ska utveckla kunskaper

¹²⁵ Rosenlund uppmärksammar att det även är möjligt att referera (minnas) *någon annans* historiska resenemang. Rosenlund (2011).

¹²⁶ MH-traditionen manifesterad i det historiska referatet kan kopplas till historia som ett (allmän)bildningsämne. Det kan även kopplas till historia som ett led i en identitetsföstran. Det historiska referatet kan uppfattas som ett narrativ vilket innebär att grundläggande krav på intrig och sensmoral eftersträvas. Se Reisman (2019) s. 536.

¹²⁷ Se t.ex. Reisman et al (2019). Se även Uggle (2002), s. 315–374.

¹²⁸ Seixas (2008), s. viii.

som möjliggör en förståelse för att ett händelseförlopp (narrativ) endast är en konstruktion och inte en direkt avspeglning av en verklighet. Det handlar också om att eleven ska utveckla kunskaper/förmågor som möjliggör egna analytiska resonemang och inte bara återgivande av andras analytiska slutsatser. HT-traditionen kräver således ett *görande* på ett annat sätt än vad som krävs inom MH-traditionen.¹²⁹

En trepartsförhandling?

Det historiedidaktiska fältet har på senare tid berikats med en rad elaborerade kognitiva modeller förankrade i HT-traditionen som har det gemensamt att de på olika sätt problematiserar den MH-traditionella synen på historisk kunskap.¹³⁰ Inom fältet har sedan stort utrymme ägnats åt att särskilja dessa HT-traditionella modeller sinsemellan, ofta genom att särskilja två huvudspår: ett anglosaxiskt spår med fokus på historisk metod och ett kontinentalt spår med fokus på historiemedvetande (temporal orientering).¹³¹ I förlängningen innebär den här interna HT-förhandlingen att det är möjligt att urskilja två olika sätt att beskriva den epistemologiska förhandlingen inom fältet, antingen som en tvåpartsförhandling eller som en trepartsförhandling.

I en trepartsförhandling utgör det anglosaxiska respektive kontinentala spåret två separata parter och MH-traditionen utgör den tredje parten. Ett tydligt exempel på detta sätt att beskriva förhandlingen framkommer i Rosenlunds avhandling *History Education as Content, Methods or Orientation?* I avhandlingen tydliggör Rosenlund skillnaden mellan det anglosaxiska respektive kontinentala spåret inom HT-traditionen eller, med Rosenlunds terminologi, ”the empirical approach” respektive ”the orienting approach”. I en historieundervisning med en ”empirical approach” ligger fokus på historisk metod och på empiriska studier av vad som har hänt i det förflutna. I en historieundervisning med en ”orienting ap-

¹²⁹ HT-traditionen manifesterad i det historiska resonemanget kan kopplas till historia som ett temporalt orienterande och identitetsskapande ämne.

¹³⁰ För exempel på kognitiva modeller med HT-traditionell förankring, se t.ex. Lévesque & Clark (2018); van Bostel & van Drie (2018); Clark & Grever (2018); VanSledright (2014). Ett exempel på en svensk kognitiv modell med HT-traditionell förankring är Karlsson (2014). Kölbl & Konrad för ett resonemang om för- och nackdelar med att utarbeta den här typen av kognitiva modeller. De menar att i förlängningen kan alltför rigida ”guidelines” begränsa historielärarnas kreativitet och därmed försvåra lärarens didaktiska uppdrag. Kölbl & Konrad (2015), s. 26.

¹³¹ Se t.ex. Rosenlund (2016).

proach” ligger fokus i högre utsträckning på ämnets existentiella dimension och på hur historiska kunskaper om det förflutna kan användas för att skapa temporala samband mellan dåtid, nutid och framtid.¹³² Dessa båda parter relateras sedan till den tredje, MH-traditionella parten, där historieundervisningens fokus ligger på ”the transfer of content”.¹³³

Att beskriva den historiedidaktiska förhandlingen inom fältet som en tvåpartsförhandling innebär i praktiken att skillnaderna mellan den anglosaxiska respektive kontinentala traditionen tonas ner. Susan Brookhart argumenterar för att skillnaden mellan dessa båda traditioner i praktiken är begränsad och att de två traditionerna har mycket gemensamt. Den huvudsakliga skillnaden är enligt Brookhart att den kontinentala tolkningen av HT-traditionen i högre utsträckning inkluderar elevens egen livsvärld i den praktiska undervisningen och därmed tydligare betonar historieämnets identitetsskapande dimension.¹³⁴

I linje med Brookhart uppfattar även jag den historiekulturella förhandlingen om skolämnet historia som en tvåpartsförhandling. Jag är medveten om att det finns betydande relevansrelaterade och epistemologiska skillnader mellan det anglosaxiska och det kontinentala spåret inom HT-traditionen. Om man återigen plockar fram teatermetaforen är det tydligt att den kontinentala HT-traditionalisten i högre grad än den anglosaxiska är intresserad av vilken eller vilka föreställningar som ska visas under besöket samt i vilket syfte dessa föreställningar visas. Denna skillnad överskuggas dock av den grundläggande epistemologiska skillnaden mellan den mer objektivistiskt förankrade MH-traditionen och de båda HT-traditionerna som båda är förankrade i konstruktivismen. Skillnaden innebär att MH-traditionalistens elever återvänder till klassrummet efter föreställningens slut medan HT-traditionalisternas elever stannar kvar och undersöker vad som finns bakom scenen. Den grundläggande epistemologiska förhandlingen sker således mellan företrädare för den MH-traditionella objektivismen respektive den HT-traditionella konstruktivismen.

¹³² Andersson Hult argumenterar för att det inom den kontinentala HT-traditionen i sin tur kan urskiljas två huvudlinjer med fokus på identitet respektive emancipation. Andersson Hult (2021), s.48–49.

¹³³ För en mer elaborerad redogörelse av trepartsförhandlingen, se t.ex. Shemilt (2018), s. 451–453. Ett alternativt sätt att beskriva trepartsförhandlingen är att betrakta den som två separata tvåpartsförhandlingar. En generell tvåpartsförhandling mellan den objektivistiska MH-traditionen och den konstruktivistiska HT-traditionen och en intern konstruktivistisk tvåpartsförhandling mellan den kontinentala och den anglosaxiska tolkningen av HT-traditionen.

¹³⁴ Brookhart (2015), s.183 ff. Hon hänvisar i sin tur till Cuban (2009) när hon särskiljer ”the heritage tradition” från ”the historical tradition”. Cuban menar att de flesta historielärare tillämpar någon form av pragmatism och att de håller fast vid heritage-traditionen.

4.4 Den historiekulturella förhandlingen

4.4.1 Konstrukt-begreppet

För att möjliggöra en mer precis beskrivning av förhandlingen mellan MH-traditionen och HT-traditionen behöver jag introducera begreppet konstrukt i det teoretiska ramverket. I ett skolrelaterat historiekulturellt sammanhang är konstrukt-begreppet ett användbart, men omdiskuterat begrepp, som underlättar hanteringen av historieämnets epistemologi, det vill säga ämnets vad-dimension.

Begreppet är förankrat i den psykometriska teoribildningen och det är först på senare år som det har börjat användas mera frekvent i en ämnesdidaktisk kontext. En ofta använd definition av ett konstrukt är att det omfattar ”begreppet eller karaktäristiken som ett prov avser att mäta”.¹³⁵ I ett skolrelaterat sammanhang manifesteras konstruktet i kurs- och ämnesplanerna.

För att precisera ett skolämnets konstrukt krävs någon form av kognitiv konceptuell modell. Att konstruera kognitiva konceptuella modeller som preciserar ett konstrukt är dock problematiskt. En del av problematiken ligger i att ett konstrukt kan inrymma flera dimensioner. Jag särskiljer tre olika dimensioner: en domänlogik, en emfaslogik och en progressionslogik.¹³⁶ I skolämnet historia kan konstruktets tre dimensioner preciseras enligt nedan:

- Domänlogik – vilka substantiella kunskaper och procedurala förmågor omfattar skolämnet historia?
- Emfaslogik – vilka substantiella kunskaper och procedurala förmågor är centrala i skolämnet historia?
- Progressionslogik – vad särskiljer/utmärker olika kvalitetsnivåer i skolämnet historia?

Vad som avses med dessa tre dimensioner kommer att förtydligas i de nästföljande avsnitten, men jag vill understryka att det inom fältet inte råder någon konsensus kring huruvida kognitiva konceptuella modeller som syftar till att precisera ett givet konstrukt utöver domänlogiken även bör omfatta någon form av emfaslogik och/eller progressionslogik.¹³⁷

¹³⁵ American Educational Research Association (2014), s. 11.

¹³⁶ Modell of Cognition (MoC) är frekvent begrep inom den historiedidaktiska bedömningsforskningen. Se t.ex. Seixas & Ercikan (2015), s.1–3.

¹³⁷ Jfr t.ex. Pellegrino et al (2001) som särskiljer två olika dimensioner: en epistemologisk dimension och en progressionsdimension.

4.4.2 Förhandlingen om domänlogiken

Domänlogiken utgör konstruktets epistemologiska fundament. Inom fältet särskiljs ofta mellan substantiella (innehållsbaserade) kunskaper och procedurala (färdighetsbaserade) förmågor.¹³⁸ I en svensk kontext manifesteras de substantiella kunskaperna i kursplanernas centrala innehåll och de procedurala förmågorna i kursplanernas målformulering. Den högljudda debatten 2019 om antikens vara eller inte vara i samband med det nya förslaget till kursplan för historieämnet i grundskolan var således en debatt om domänlogikens substantiella dimension.¹³⁹ När förmågan att resonera om hur historia används skrevs in i samma kursplan 2011 var det en genomgripande förändring av domänlogikens procedurala dimension.

I en MH-traditionellt förankrad domänlogik ligger fokus på *rekonstruktionen* av historiska narrativ, det vill säga på den substantiella dimensionen av domänlogiken. Rekonstruktionen kan manifesteras i ett sammanhängande referat av ett narrativ eller i rekapitulationen av specifika detaljer i ett givet narrativ. Om man tar exemplet antiken så kan det handla om att eleven ska kunna rekonstruera centrala händelseförlopp, till exempel de grekiska-persiska krigen. De historiska begrepp som lyfts fram i den MH-traditionella domänlogiken är på motsvarande sätt förankrade i den substantiella dimensionen. Det behövs i första hand innehållsbegrepp eller (first-order concepts) för att rekonstruera ett givet narrativ. För att rekonstruera händelseförloppet under perserkrigen behövs till exempel innehållsbegreppet stadsstat. Förmågan att med stöd av minnet kunna rekonstruera ett givet narrativ ska inte förringas. Det är en grundläggande förmåga inom historieämnet som på många sätt kan kopplas till begreppen bildning eller allmänbildning.¹⁴⁰

Även i en HT-traditionellt förankrad domänlogik ligger fokus på historiska narrativ, men till skillnad från den MH-traditionella domänlogiken handlar det inte bara om att kunna rekonstruera ett givet narrativ utan även om att kunna *dekonstruera* (resonera om) givna narrativ samt att kunna *konstruera* egna narrativ (dra egna slutsatser). Det innebär att eleven inte bara ska kunna rekonstruera händelseförloppet under de grekiska-persiska krigen. Eleven ska även kunna resonera om hur krigen har skildrats

¹³⁸ Se t.ex. Lévesque (2008), s. 29–30. Hur dessa båda dimensioner av domänlogiken ska benämnas råder det dock delade meningar om. I en anglosaxisk miljö har benämningarna "content" och "skills" länge varit dominerande. Se Seixas & Ercikan (2015), s. 2.

¹³⁹ Se t.ex. Strannegård (2019).

¹⁴⁰ I 1994 års läroplan (Lpf 94) är strävansmålet ett uttryck för ett bildningsideal. Se Carlgren et al (2009).

i olika källor med koppling till krigen samt med stöd av dessa källor kunna dra underbyggda slutsatser, till exempel slutsatser om vad som orsakade krigen. Den här skillnaden innebär att domänlogikens fokus förskjuts från den substantiella dimensionen till den procedurala dimensionen. Det innebär också att de historiska begrepp som lyfts fram är såväl innehållsbegrepp (first-order concepts) som metahistoriska begrepp (second-order concepts).¹⁴¹ De metakognitiva begreppen behövs framför allt för att kunna dekonstruera (resonera om) ett givet narrativ. För att kunna resonera om hur de grekiska-persiska krigen har skildrats i olika källor behöver eleven kunna hantera grundläggande källkritiska begrepp, till exempel skillnaden mellan första- och andrahandskällor.¹⁴²

4.4.3 Förhandlingen om emfaslogiken

Emfaslogiken kan beskrivas som konstruktets vågskål. Om konstruktets samtliga domänlogiska delkomponenter, det vill säga allt innehåll och samtliga förmågor, väger lika tungt behövs det ingen emfaslogik. Om däremot någon eller några av delkomponenter intar en särställning krävs någon form av emfaslogik.¹⁴³ Emfaslogiken handlar således om huruvida konstruktets domänlogik inrymmer någon form av epistemologisk hierarki (delkomponenternas relativa vikt). Notera att behovet av en emfaslogik minskar om konstruktet inrymmer få domänlogiska delkomponenter. Ett konstrukt vars domänlogik endast omfattar ett narrativ och en förmåga behöver inte någon emfatisk logik. I de nuvarande kursplanerna i historia i Gy11 saknas en explicit emfaslogik. I skrivningarna för det centrala innehållet finns dock implicit en emfaslogik som manifesteras i skillnaden mellan ”ska-satser” och ”bör-satser”.¹⁴⁴

Emfaslogiken inom MH-traditionen omfattar i första hand den substantiella dimensionen av domänlogiken. MH-traditionens fokus på rekonstruktionen av historiska narrativ innebär att kravet på kronologisk bredd

¹⁴¹ Under senare tid har även benämningen third-order concepts börjat användas av forskare inom fältet, se t.ex. Jarhall (2020).

¹⁴² Jfr med Reismans resonemang om ”historical reading”. Hon särskiljer mellan bakgrundskunskaper (rekonstruktion), källkritik i form av trovärdighet/relevans (dekonstruktion) och källhantering i form av dra slutsatser utifrån flera källor (konstruktion). Reisman (2015), s. 29 ff. Waldis et al betonar att dekonstruktionen är en analys och att konstruktionen är syntes. Waldis et al (2015), s. 117.

¹⁴³ Rosenlund argumenterar för att den emfaslogiken i många fall är en normativ fråga. Rosenlund (2016), s. 39 ff.

¹⁴⁴ Notera att emfaslogiken är ämnesspecifik och att den inte ska förväxlas med den bedömningsteoretiska (psykometriska) vikttningslogiken. Den förra handlar om delkomponentens relativa tyngd inom ett givet konstrukt och den senare handlar om prestationsomdömetts relativa tyngd vid en given sammanvägning.

tenderar att öka, vilket i sin tur innebär att det krävs en emfaslogik som preciserar vilket eller vilka narrativ som är minnesvärda (meningsfulla). Behovet av en emfaslogik gällande domänlogikens procedurala dimension är däremot mer begränsat eftersom MH-konstruktets domänlogik är förhållandevis snäv och i första hand omfattar förmågan att rekonstruera historiska narrativ.

Emfaslogiken inom HT-traditionen omfattar i första hand den procedurala dimensionen av domänlogiken. HT-traditionens procedurala dimension är bredare och inrymmer fler delkomponenter än MH-traditionen. Emfaslogiken blir därmed mer komplex och svårare att hantera. Det bör till exempel framgå vilken tyngd rekonstruktionen har i relation till dekonstruktionen och konstruktionen.¹⁴⁵

4.4.4 Förhandlingen om progressionslogiken

Progressionslogiken kan beskrivas som den epistemologiska underbyggnad som gör det möjligt att särskilja olika kvalitetsnivåer i en elevprestation.¹⁴⁶ Om ambitionen endast är att särskilja mellan två kvalitetsnivåer (vet/kan och vet inte/kan inte) behövs ingen progressionslogik eftersom kriterierna för dessa nivåer kan baseras direkt på konstruktets domänlogik, antingen vet/kan du det som efterfrågas i konstruktet eller så vet/kan du det inte. Konstruktets domänlogik utgör med andra ord ett tillräckligt teoretiskt underlag.¹⁴⁷ Om ambitionen däremot är att särskilja mellan flera nivåer krävs någon form av progressionslogik som gör det möjligt att särskilja mellan olika nivåer/kvaliteter av kunskap. Utöver det grundläggande minimikravet tillför således progressionslogiken kompletterande

¹⁴⁵ Det är viktigt att särskilja emfaslogiska förskjutningar *inom* ett givet konstrukt från grundläggande epistemologiska förskjutningar *mellan* konstrukt förankrade i olika historiekulturella traditioner. En emfaslogisk förskjutning inom ett konstrukt kan vara en förskjutning från äldre till moderna narrativ. En grundläggande epistemologisk förskjutning kan vara en förskjutning från rekonstruktion av narrativ till (de)konstruktion av narrativ, det vill säga en förskjutning från en MH-tradition till en HT-tradition. I det senare fallet handlar det om att en eller flera av de centrala procedurala delkomponenterna inom ett givet konstrukt tillkommer eller bortfaller. Den senaste revideringen av grundskolans styrdokument (Lgr22), där uppdraget var att förskjuta fokus från "analys" till "fakta" i de yngre årskurserna, skulle kunna uppfattas som sådan mera grundläggande epistemologisk förskjutning.

¹⁴⁶ VanSledright använder begreppet "weighting rationale" när han betonar vikten av en tydlig progressionslogik för att kunna särskilja olika kvalitetsnivåer i en elevprestation. VanSledright (2015), s. 85.

¹⁴⁷ Huruvida progressionslogiken utgör en separat del (dimension) av konstruktet handlar i grunden om huruvida det är möjligt och/eller önskvärt att skriva fram en domänspecifik progressionslogik eller inte. Om progressionslogiken endast omfattar det grundläggande minimikravet som sedan kompletteras med kvantitativa optimumnormer är det inte meningsfullt att tala om en domänspecifik progressionslogik. Om optimumnormerna däremot omfattar additiva och kvalitativa indikatorer är det rimligt att låta progressionslogiken utgöra en separat del av konstruktet.

optimumnormer.¹⁴⁸ Ett sätt att tydliggöra skillnaden mellan minimikraven och de progressionslogiska optimumnormerna är att göra en jämförelse med en vanlig lampa och en dimmer-lampa. En vanlig lampa har endast två lägen, den är antingen av eller på (minimikravet). En dimmer-lampa har den grundläggande av- och på-funktionen, men den har även en dimmer-funktion som gör det möjligt att reglera ljusstyrkan (optimumnormerna).¹⁴⁹ I en svensk kontext utgör/manifesteras konstruktets progressionslogiska optimumnormer i kursplanernas kunskapskrav.¹⁵⁰

MH-traditionen bygger på en mer objektivistisk kunskapssyn och en mer uttalad strävan efter ett absolut sanningsanspråk (sant/falskt). Det innebär i sin tur att det sanna narrativet ("det som egentligen hände") utgör utgångspunkten för den MH-traditionella progressionslogiken. Minimikravet i ett skolrelaterat sammanhang handlar om huruvida elevens referat är en korrekt rekonstruktion av det narrativ som förmedlats i undervisningen. Vad gäller MH-traditionens progressionslogiska optimumnormer görs antagandet att MH-traditionens mer objektivistiska kunskapssyn korrelerar med användningen av kvantitativa progressionsindikatorer. Historisk kunskap manifesterad i referatets återberättande form lämpar sig bättre att kvantifiera än de HT-traditionella resonemangen. Ett referat som inrymmer flera orsaker till franska revolutionen håller högre kvalitet än det referat som endast inrymmer en orsak.

HT-traditionen bygger på en mer konstruktivistisk kunskapssyn, vilket innebär att det objektivistiska sanningsanspråket tonas ner.¹⁵¹ Det progressionslogiska minimikravet inom HT-traditionen utgår således inte i första hand från "det sanna narrativet" utan från att eleven i sina resonemang visar en förståelse för narrativets konstruktivistiska natur. Vad gäller progressionslogikens optimumnormer görs antagandet att HT-traditionens konstruktivistiska kunskapssyn i högre utsträckning korrelerar med användningen av additiva och kvalitativa progressionsindikatorer.

¹⁴⁸ Dessa optimumnormer kan inrymma fler eller färre kvalitetsnivåer. Rosenlund noterar att det nuvarande svenska betygssystemet med sex olika kvalitetsnivåer är förhållandevis omfattande jämfört med många andra länder. Han noterar även att den ämnesspecifika progressionslogik som delvis utmärker de svenska kunskapskraven är avvikande. Rosenlund (2019) s. 84-85.

¹⁴⁹ Jönsson diskuterar skillnaden mellan stegvisa progressionsmodeller och "glidande" progressionsmodeller. Jönsson (2017), s. 20-21.

¹⁵⁰ Progressionslogikens optimumnormer är ett "svart hål" inom det historiedidaktiska bedömningsfältet (se tidigare forskning). Shemilt har dock presenterat progressionslogiska flerstegsmodeller som laborerar med optimumnormer. Shemilt (2009), s.183-187.

¹⁵¹ Att sanningsanspråket tonas ner innebär inte att det bortfaller. En underbyggnad i form av sakhistoriskt korrekta exempel är en viktig förutsättning för det goda resonemanget.

Med en additiv progressionsindikator avses huruvida ett resonemang som inrymmer en eller flera *olika* tankeoperationer. Ett resonemang som inrymmer såväl kortsiktiga som långsiktiga perspektiv är av högre kvalitet än ett resonemang som endast inrymmer ett av perspektiven. En kvalitativ indikator som ofta förekommer i HT-traditionella progressionsmodeller är begreppsanvändningen. Ett resonemang baserat på en historiedidaktiskt förankrad begreppsanvändning, såväl innehållsbegrepp som metahistoriska begrepp, är av högre kvalitet än ett resonemang baserat på en mera vardaglig begreppsanvändning. Betoningen på additiva och kvalitativa indikatorer innebär inte att kvantitativa progressionsindikatorer saknar betydelse inom HT-traditionen. Antalet historiska exempel som underbygger ett resonemang har betydelse för resonemangets kvalitet. Antalet källor som hanteras för att konstruera ett narrativ (dra en slutsats) har betydelse för narrativets/slutsatsens kvalitet.¹⁵²

4.4.5 Sammanfattning

Som framkommit av de teoretiska överväganden som redovisats ovan särskiljer jag mellan den MH-traditionella objektivismen och den HT-traditionella konstruktivismen. Den grundläggande skillnaden mellan dessa båda traditioner handlar om hur de historiska narrativen ska uppfattas och hanteras. Omsätter man denna grundläggande skillnad till historielärarens bedömningspraktik innebär det att den HT-traditionella historieläraren i första hand behöver konstruera uppgifter som ger eleven möjlighet att med stöd av minnet rekonstruera det historiska narrativet. Den HT-traditionella historieläraren behöver utöver denna typ av uppgifter dessutom konstruera uppgifter som ger eleven möjlighet att med stöd av metahistoriska begrepp dekonstruera det historiska narrativet samt uppgifter som ger eleven möjlighet att med stöd av källor konstruera egna narrativ, alternativt dra slutsatser om det historiska narrativet. Det är viktigt att no-

¹⁵² Se Rosenlund (2019) för ett resonemang kring progressionslogikens epistemologiska förankring. Han identifierar en tredelad progressionslogik i 2011 års kunskapskrav: den begreppsbaseade, den räkneordbaseade och den värdeordbaseade. Han argumenterar sedan för att den begreppsbaseade progressionslogiken kan kopplas till HT-traditionen, medan den räkneordbaseade traditionen kan kopplas till MH-traditionen. Vad gäller den värdeordbaseade progressionslogiken är Rosenlunds slutsats att den inte vilar på någon teoretisk grund och därför inte kan sägas utgöra någon logik över huvud taget. I antologin *New Directions in Assessing Historical Thinking* finns flera ledande företrädare för HT-traditionen representerade. Progressionslogiken i deras exempel på värderingsmatriser inrymmer såväl kvantitativa som kvalitativa och additiva progressionsindikatorer. Ercikan & Seixas (red) (2015).

tera att den HT-traditionella läraren inte kan bortse från den första rekonstruerande uppgiften. Om man på nytt använder sig av Seixas teatermetafor är det inte möjligt att låta eleverna hoppa över teaterföreställningen och gå direkt bakom scenen. Om eleverna inte har sett föreställningen finns det inget narrativ att dekonstruera.

4.5 Bedömningskultur – ett sekundärt portalbegrepp

4.5.1 Begreppsdefinition

Ett externt prov i historia är inte vilken historiekulturell produkt som helst utan det är en historiekulturell produkt *som används för att bedöma historiska kunskaper*. Det innebär att det externa provet, vid sidan av den historiekulturella vad-dimensionen, även inrymmer en bedömningskulturell hur- och varför-dimension. Om historiekultur utgör forskningsprojektets primära portalbegrepp kan därför bedömningskultur sägas utgöra projektets sekundära portalbegrepp.

Som framkommit tidigare definierar jag begreppet historiekultur som det kommunikativa sammanhang inom vilket olika sätt att uppfatta det förflutna lyfts fram och ges mening. Begreppet bedömningskultur uppfattar jag på motsvarande sätt som ett kommunikativt sammanhang, men det är inte uppfattningar om det förflutna, utan uppfattningar om kunskapsbedömning som utgör fokus i detta kommunikativa sammanhang. Explicit formulerat definierar jag således begreppet bedömningskultur som *det kommunikativa sammanhang inom vilket olika sätt att uppfatta kunskapsbedömning lyfts fram och tillskrivs mening*.

4.5.2 Bedömningskultur som analysverktyg

Bedömningskulturen kan analytiskt brytas ner i de kategorier som redovisats gällande historiekulturen, det vill säga i bedömningskulturella produkter, arenor och processer. Inom ramen för detta specifika forskningsprojekt utgörs de bedömningskulturella produkterna av de bedömningsverktyg och bedömningsdirektiv som tillämpas under en extern bedömningsprocess. De bedömningskulturella arenorna utgörs av de arenor där ett samhälle kommunicerar och värderar vilka former av extern kunskapsbedömning som ska ägnas uppmärksamhet. Den bedömningskulturella processen är på samma sätt som den historiekulturella processen tvådelad och utgörs dels av en förmedlande process med fokus på den kommuni-

kativa kedjan mellan avsändare, förmedlare och mottagare, dels av en värderande process med fokus på förhandlingen om vilka former av extern kunskapsbedömning som är värda att uppmärksammas. I avhandlingen kommer denna förhandling om kunskapsbedömningen att uppfattas som en förhandling mellan två bedömningskulturella traditioner: mättraditionen och tolkningstraditionen. Dessa kommer att beskrivas utförligt i nästa avsnitt.

4.6 Bedömningskulturella traditioner

Som tidigare framgått uppfattar jag den värderande dimensionen av den historiekulturella processen som en förhandling. I ett skolrelaterat sammanhang är det en förhandling som handlar om historieämnets epistemologi (vad-frågan).¹⁵³ Motsvarande förhandling kan spåras även i ett skolrelaterat bedömningskulturellt sammanhang, men här handlar förhandlingen om kunskapsbedömningens metodologi (hur-frågan och varför-frågan). 1) Hur bedöms elevernas kunskaper? 2) Varför bedöms elevernas kunskaper?¹⁵⁴

På motsvarande sätt som att förhandlingen om de externa provens epistemologiska dimension kan uppfattas som en tvåpartsförhandling kan förhandlingen om provens metodologiska dimension uppfattas som en förhandling mellan två parter. Inom det historiedidaktiska fältet använder sig till exempel VanSledright av begreppen *testing-culture* och *assessment-culture* för att särskilja dessa båda parter åt.¹⁵⁵ För att få en mer elaborerad bild av denna tvåpartsförhandling är det dock nödvändigt att vända sig till det bedömningsteoretiska fältet.

4.6.1 Mättradition vs. Tolkningstradition

Även inom det bedömningsteoretiska fältet är det vanligt att särskilja mellan olika sätt eller ”traditioner” för att hantera kunskapsbedömningens

¹⁵³ Två läroplaner, Lpf 94 med revidering 2000 och Gyll, bildar ram för undersökningen och därmed även den grundläggande beskrivning av skolans uppdrag som där anges. Denna syftesbeskrivning är i stort oförändrad mellan de olika läroplanerna och återspeglas även på olika sätt i gällande kurs- och ämnesplaner för historieämnet. Eftersom undersökningens fokus är historiekulturell förändring i relation till bedömningskultur har den koncentrerats till vad-frågan och den komplexa historiefilosofiska varför-frågan undvikits i sammanhanget. För tidigare forskning där varför-frågan behandlas se exempelvis Rosenlund (2016).

¹⁵⁴ Det är viktigt att särskilja den bedömningskulturella syftesformulering som skrivs fram i provets syftesreglerande direktiv från den historiekulturella syftesformulering som skrivs fram i styrdokumentet för historieämnet.

¹⁵⁵ VanSledright (2014), s. 113–119.

metodologiska dimension. Anders Jönsson särskiljer mellan en provtradition och en bedömningstradition.¹⁵⁶

Jönsson menar att den grundläggande skillnaden mellan dessa båda traditioner handlar om bedömningens objekt. Det övergripande syftet inom provtraditionen är att dra slutsatser om *elevens kunskaper* medan det övergripande syftet inom bedömningstraditionen är dra slutsatser om kvaliteten i *en elevprestation*. I det förra fallet handlar det om att *indirekt mäta* kvaliteten på något som inte är synligt (elevens kunskap) och i det senare fallet att *direkt tolka* kvaliteten på något konkret (en elevprestation).

Jag delar Jönssons uppfattning att det är fruktbart att med utgångspunkt från bedömningens objekt särskilja dessa båda traditioner, men jag vill tydligare än Jönsson särskilja traditionernas metodologiska förankring. Jag menar att provtraditionen är förankrad i den kvantitativa metodologin medan bedömningstraditionen är förankrad i den kvalitativa metodologin. En annan skillnad mellan mig och Jönsson handlar om terminologin. I mitt teoribygge är ”prov” en benämning på ett specifikt bedömningsverktyg medan ”bedömning” är att betrakta som ett övergripande paraplybegrepp. Jag har därför valt att ersätta Jönssons benämningar med ”mättradition” och ”tolkningstradition”.

I linje med min betoning av den metodologiska skillnaden mellan traditionerna driver jag tesen att den grundläggande hur-dimensionella skillnaden mellan de båda traditionerna handlar om bedömarens/lärarens tolkningsutrymme under bedömningsprocessen. Lite förenklat kan man säga att idealet inom mättraditionen är att reglera/begränsa lärarens tolkningsutrymme medan idealet inom tolkningstraditionen är att bejaka lärarens tolkningsutrymme.¹⁵⁷

¹⁵⁶ Inom det bedömningsteoretiska fältet används ibland begreppet bedömningsparadigm för att komma åt samma dikotomi, se t.ex. Gipps (1994). Inom det läroplansteoretiska fältet har Lundahl & Tveit (2014) diskuterat skillnaden mellan en (amerikansk) psykometrisk bedömningstradition och en (kontinental) professionsförankrad bedömningskultur.

¹⁵⁷ Även Jönsson menar att bedömningsprocessens hur-dimension hanteras på olika sätt inom mättraditionen och tolkningstraditionen. Jönsson (2017), s. 8. Notera att tolkningsutrymmet inrymmer två dimensioner. Ett historiekulturellt epistemologiskt tolkningsutrymme som regleras av bedömningsverktygens konstrukt och ett bedömningskulturellt metodologiskt tolkningsutrymme som regleras i användningsdirektiven samt indirekt i bedömningsverktygens format. Alla former av direktiv och bedömningsverktyg innebär således en begränsning av lärarens tolkningsutrymme, men slutna verktyg begränsar detta friutrymme i högre utsträckning än öppna verktyg.

I detta sammanhang kan det vara klargörande att relatera traditionerna till begreppsparet reliabilitet och validitet. Reliabilitetsbegreppet handlar om huruvida de resultat som genereras i samband med en bedömning är tillförlitliga eller inte. Validitetsbegreppet är på många sätt det mest centrala begreppet inom det bedömningsteoretiska fältet. Detta till trots är det ett begrepp som inte låter sig definieras på ett enkelt sätt. Validitet i positivistisk bemärkelse handlar om huruvida bedömningsverktygen verkligen mäter de kunskaper (det konstrukt) som de är avsedda att mäta. Inom den moderna validitetsteorin handlar begreppet snarare om huruvida de slutsatser som dras och de beslut som fattas i samband med en bedömning är giltiga eller inte. Validitet handlar således om att dra slutsatser utifrån en bedömning eller ett prov, snarare om bedömningen/provet per se. Validitetsbegreppet har blivit bredare och inkluderar även bedömningens konsekvenser.¹⁵⁸

Lite förenklat kan man argumentera för att det mättraditionella idealet att begränsa lärarens tolkningsutrymme korrelerar med uppfattningen att kravet på reliabilitet är överordnat kravet på validitet. Om inte den data som genereras är mätsäker spelar det ingen roll huruvida bedömningen är valid eller inte. På motsvarande sätt kan man argumentera för att det tolkningstraditionella idealet att bejaka lärarens tolkningsutrymme korrelerar med uppfattningen att kravet på validitet är överordnat kravet på reliabilitet. Om de slutsatser som dras och de beslut som tas på basis av dessa slutsatser inte är giltiga är det inte rimligt att genomföra bedömningen över huvud taget oavsett om den data som genereras är mätsäker eller inte. Reliabilitetskravet tonas ner och hanteras på ett annat sätt. Det är återigen viktigt att betona att det hela tiden handlar om fokus och balans. Den mättraditionella bedömaren kan inte bortse från validitetsaspekten och den tolkningstraditionella bedömaren kan inte bortse från reliabilitetsaspekten.

Sammanfattningsvis särskiljer jag således mellan en mättradition med stort fokus på reliabilitet och där idealet är att begränsa lärarens tolkningsutrymme, och en tolkningstradition med stort fokus på validitet och där idealet är att bejaka lärarens tolkningsutrymme. Innan jag går vidare och redovisar en mer precis beskrivning av hur tolkningsutrymmet hanteras

¹⁵⁸ För en mer elaborerad redogörelse av det den moderna validitetsteorin, se Jönsson (2017), s.55 ff.

inom dessa båda traditioner behöver jag dock infoga ett tvådelat bedömningsteoretiskt interludium. I den första delen behandlas begreppet externa prov och i den andra delen behandlas bedömningsprocessen i samband med ett skriftligt externt prov.

4.7 Ett bedömningsteoretiskt interludium

4.7.1 En definition av begreppet externa prov

Inom den anglosaxiska forskarvärlden används begreppet *external tests* flitigt, i många fall kopplat till den debatt som följde i spåren av den bedömningskultur som växte fram i Storbritannien och USA under 1980-talet. Det var en bedömningskultur som präglades av en statlig ambition att använda sig av massproducerade storskaliga prov för att mäta såväl elevernas kunskaper som utbildningssystemens kvalitet.¹⁵⁹ Även i Sverige har begreppet *externa prov* fått allt större utrymme inom det bedömningsteoretiska forskningsfältet, i takt med att intresset för extern kunskapsbedömning har ökat.¹⁶⁰

Trots en omfattande forskning inom ämnesområdet har jag dock haft svårt att finna precisa definitioner i litteraturen. Lundahl särskiljer till exempel intern bedömning från extern bedömning på följande allmänt hållna sätt:

Interna bedömningar produceras och används inom skolorna och för deras elever och föräldrar. Externa bedömningar produceras av skolmyndigheter, forskare och andra, antingen genom egna studier eller genom att utnyttja de interna bedömningarna.¹⁶¹

Den amerikanske bedömningsteoretikern George F. Madaus gör en något mer elaborerad distinktion mellan intern och extern bedömning:

An internal testing program is one which is carried out within a school at the initiative and under the control of the school superintendent, principal, or teacher. [...] External testing programs are those con-

¹⁵⁹ Caroline Gipps 1994, s. 14–15.

¹⁶⁰ Exempel på svenska forskare som har intresserat sig för skolrelaterad extern bedömning: Se t ex Forsberg & Lundahl (2006); Forsberg & Lundahl (2009); Lundahl (2006); Lundahl (2009) och Jönsson (2017). För en översikt av forskningen inom fältet pedagogisk bedömning, se Eva Forsberg 2010.

¹⁶¹ Forsberg & Lundahl (2009), s. 62.

trolled and/or mandated by an external authority, such as a state department of education, the state legislature or a private agency such as the College Entrance Examination Board (CEEB).¹⁶²

Madaus menar således att den grundläggande skillnaden mellan ett internt och externt prov handlar om vilken eller vilka aktörer som initierar och kontrollerar bedömningen. Mitt sätt att uppfatta externa prov ligger i linje med Madaus, men hans definition kan som jag ser det problematiseras i två avseenden: dels genom att diskutera vad som menas med kontroll, dels genom att diskutera betydelsen av på vilken nivå inom utbildnings-systemet som den kontrollerande aktören befinner sig.

Vad innebär det att som undervisande lärare ha kontroll över en bedömningsprocess? En grundläggande aspekt är naturligtvis huruvida den undervisande läraren själv kan besluta om en bedömning ska äga rum eller inte, men vid sidan av denna vetorätt är frågan inte helt lätt att besvara eftersom en bedömningsprocess inrymmer flera olika delprocesser (se figur 4 nedan). I föreliggande forskningsprojekt preciseras därför kontrollaspekten genom tydliggörandet att ett externt prov är *ett prov där en eller flera av bedömningens delprocesser, helt eller delvis, kontrolleras av en extern aktör*.

Den andra aspekten i Madaus definition som jag vill problematisera är som sagt vilken betydelse det har på vilken nivå som de kontrollerande aktörerna befinner sig. Madaus tycks mena att det endast är centrala prov, kontrollerade av aktörer på en nationell nivå, som ska definieras som externa prov – en snäv avgränsning som utesluter alla de kollegiala prov som utformas och kontrolleras gemensamt av lärare på lokal eller regional nivå. I min definition av externa prov är det primära inte på vilken nivå de externa aktörerna befinner sig utan det faktum att den undervisande läraren inte på egen hand kontrollerar alla delar av bedömningsprocessen. Nivåaspekten är således av underordnad betydelse i förhållande till kontrollaspekten och jag vill därför vidga begreppet externa prov till att gälla alla prov, oavsett nivå, som den undervisande läraren inte helt på egen hand kontrollerar.

Sammanfattningsvis, med hänsyn tagen till kontrollaspekten och nivåaspekten, vill jag därför definiera begreppet externt prov som *ett prov där*

¹⁶² Madaus (1988), s. 86.

den undervisande läraren saknar vetorätt och/eller där externa aktörer på lokal, regional eller central nivå, helt eller delvis, kontrollerar en eller flera av bedömningsprocessens delprocesser.

4.7.2 Bedömningsprocessen

Den svenske bedömningsteoretikern Peter Nyström har i en modell schematiskt redogjort för de produkter och de delprocesser som han menar ingår i en bedömningsprocess.¹⁶³ Jag har utgått från Nyström när jag i en egen konceptuell modell tydliggör hur jag uppfattar bedömningsprocessen i samband med ett skriftligt externt prov (se figur 4).¹⁶⁴

Figur 4: Den externa bedömningsprocessens olika delar

I modellen framgår bedömningsprocessens komplexitet. Det framgår också att den terminologi som krävs för att hantera denna komplexitet är omfattande. Jag kommer därför, med stöd av modellen, att förtydliga såväl bedömningsprocessens olika delar som den bedömningsteoretiska terminologi som jag kommer att använda mig av framgent.¹⁶⁵

¹⁶³ Nyström (2004), s. 4.

¹⁶⁴ Modellen bygger på Nyström (2004) och Gunnemyr (2011).

¹⁶⁵ Notera att den presenterade modellen endast är ett sätt att beskriva bedömningsprocessen. Inom det bedömningsteoretiska fältet finns det många modeller där bedömningsprocessen beskrivs på ett mer eller mindre elaborerat sätt. En sådan modell har utarbetats av Crooks, Kane och Cohen (1996). Denna mycket elaborerade modell har fått stort genomslag inom det bedömningsteoretiska fältet och används bland annat av Skolverket. I modellen särskiljs åtta olika delprocesser. En modell som tillämpas av många historiedidaktiker med en bedömningsteoretisk inriktning är den modell (the assessment triangle) som utarbetats av Pellegrino et al (2001). Även Allals (2013) modell om bedömningsprocessens olika nivåer är relevant i detta sammanhang, se vidare not 166.

Delprocesser

I modellen framkommer att bedömningsprocessen kan indelas i tre övergripande delprocesser: konstruktion, genomförande och användning.

Med **konstruktionen** avses den del av bedömningsprocessen då bedömningens verktyg och direktiv konstrueras. Konstruktionsprocessen är fundamental för bedömningen. Utan verktyg av något slag, till exempel en uppgift med tillhörande matris, samt någon form av direktiv som reglerar hur och i vilket syfte dessa verktyg ska användas är det inte möjligt att genomföra en kunskapsbedömning.

Med **genomförandet** avses den del av bedömningsprocessen då bedömningens verktyg och direktiv tillämpas. Genomförande kan indelas i tre underordnade delprocesser: observation, värdering och sammanvägning.

- Med *observationen* avses den del av bedömningsprocessen då eleven ges möjligheter att visa kunskaper och då läraren ges möjlighet att observera elevens kunskaper. Observationen kan genomföras på många olika sätt. En historielärare kan till exempel ge eleven i uppgift att skriva uppsats om franska revolutionen eller be en elevgrupp att iscensätta en dramatisering av stormningen av Bastiljen. Båda sätten ger eleven möjlighet att visa kunskaper i historia och läraren möjlighet att observera dessa kunskaper.¹⁶⁶
- Med *värderingen* avses den del av bedömningen då läraren värderar kvaliteten i en elevprestation och sedan formulerar någon form av prestationsomdöme. En historielärare som har gett sina elever i uppgift att skriva en uppsats om franska revolutionen ska värdera huruvida elevens prestation motsvarar de kunskapskrav (kriterier) som är kopplade till uppgiften. Även värderingen kan genomföras på olika sätt, men ofta används någon form av värderingsmatris. Med stöd av en värderingsmatris kan historieläraren särskilja olika nivåer eller kvaliteter av historisk kunskap i elevens prestation.

¹⁶⁶ En viktig problematik gällande observationen är att det epistemologiska och metodologiska friutrymmet inrymmer såväl ett elevperspektiv som ett lärarperspektiv. Är det elevens eller lärarens friutrymme som regleras av direktiven och verktygen? I det fortsatta resonemanget betonas lärarperspektivet. Kanske kan man särskilja elevens direkta friutrymme från lärarens indirekta friutrymme?

- Med *sammanvägningen* avses den del av bedömningen då läraren väger samman enskilda prestationsomdömen till ett samlat prestationsomdöme. En historielärare som i ett skriftligt prov har gett en elev i uppgift att besvara ett antal frågor om franska revolutionen måste först värdera kvaliteten i varje enskilt elevsvar och därefter väga samman dessa värderingar till ett samlat prestationsomdöme.¹⁶⁷

Med **användningen** avses den del av bedömningsprocessen då läraren använder prestationsomdömena i något syfte. Det kan vara i ett summativt syfte, till exempel som ett stöd för att sätta ett slutbetyg på en kurs. Det kan också vara i ett formativt syfte, till exempel som återkoppling på hur pass väl undervisningen i ett specifikt moment har fungerat.

Bedömningsverktyg och bedömningsdirektiv

I modellen framkommer även att det under konstruktionsprocessen konstrueras två olika slags produkter: bedömningsverktyg och bedömningsdirektiv. Med bedömningsverktyg avses de verktyg som tillämpas under genomförandets tre delprocesser (se ovan). Med bedömningsdirektiv avses dels de processreglerande direktiv som reglerar genomförandet, dels de syftesreglerande direktiv som reglerar användningen.

Med de processreglerande direktiven avses de direktiv som reglerar genomförandets tre delprocesser. Observationsdirektiven kan till exempel reglera om eleven ska skriva uppsatsen i klassrummet eller inte. Värderingsdirektiven kan till exempel reglera om uppsatsen behöver samvärderas eller inte. Även sammanvägningsdirektiven kan reglera huruvida läraren behöver involvera någon kollega när uppsatsens prestationsomdöme ska vägas samman med andra enskilda prestationsomdömen till ett samlat prestationsomdöme.

Med syftesreglerande direktiv avses de direktiv som reglerar i vilket syfte bedömningens resultat och verktyg ska användas. I dessa direktiv

¹⁶⁷ Allal (2013) beskriver sammanvägningen som det tredje ledet i en tredelad tolkningskedja. I det första ledet tolkas styrdokumentet (konstruktet) i sin helhet för att omsättas i konstruktionen av bedömningsuppgifter. I det andra ledet tolkas i första hand kunskapskraven i styrdokumentet (progressionslogiken) för att omsättas i värderingsverktyg som sedan kan tillämpas vid värderingen av elevprestationerna. I det tredje ledet tolkas och tillämpas styrdokumentens direktiv för betygssättningen. Det tredje ledet kan beskrivas som en sammanvägning där belägg från enstaka elevprestationer vägs samman till en kunskapsprofil, ett omdöme eller ett slutbetyg.

kan det till exempel regleras om bedömningen har ett eller flera syften. I de syftesreglerande direktiven för de nuvarande nationella proven framkommer att proven har ett ensidigt summativt syfte, resultatet ska användas (särskilt beaktas) vid betygssättningen. Tidigare hade de nationella proven flera olika syften, både summativa och formativa.¹⁶⁸

Dataformat

Slutligen framkommer i modellen att det under bedömningsprocessen genereras tre olika former av data. Med prestation avses den data som genereras under observationsprocessen och som hanteras under värderingsprocessen. Det kan vara allt ifrån ett kryss i en flervalsuppgift till en skriftlig uppsats. Med prestationsomdöme avses den data som genereras under värderingsprocessen och som hanteras under sammanvägningsprocessen. Det kan vara alltifrån ett betygsomdöme till ett omfattande skriftligt omdöme. Med samlat prestationsomdöme avses data som generas under sammanvägningsprocessen och som hanteras under användningsprocessen. Även här är det möjligt att tänka sig ett enskilt betygsomdöme eller ett omfattande skriftligt omdöme.

Avslutningsvis vill jag understryka att detta bedömningsteoretiska interludium är en del i den teoriutvecklande ambition som genomsyrar forskningsprojektet. Den konceptuella modell gällande den externa bedömningsprocessens olika delar som presenteras och den terminologiska verktygslåda som följer av modellen skapar förutsättningar för en analytisk skärpa när teori möter empiri i studiens empiriska kapitel. Med detta sagt är det hög tid att återkoppla till den bedömningskulturella förhandlingen om tolkningsutrymmet.

4.8 Den bedömningskulturella förhandlingen

Som framgått tidigare omfattar den bedömningskulturella förhandlingen såväl hur-frågan som varför-frågan. Hur-frågan inrymmer i sin tur två olika aspekter: en formatrelaterad aspekt och en processrelaterad aspekt.

¹⁶⁸ Det kan finnas syftesreglerande direktiv som reglerar huruvida prestationsomdömena i första hand ska användas i ett summativt eller formativt syfte. Det kan även finnas direktiv som reglerar att det externa provets verktyg inte bara ska användas för att generera data utan att de även ska fylla en normativ funktion, det vill säga fungera som förtydligande exempel på hur bedömningsverktyg kan konstrueras. De syftesreglerande direktiven kan också reglera huruvida prestationsomdömen ska användas på enskild eller aggregerad nivå samt huruvida omdömena ska offentliggöras eller inte.

Den formatrelaterade aspekten manifesteras i bedömningsverktygens format och den processrelaterade aspekten manifesteras i de processreglerande bedömningsdirektiven. Varför-frågan manifesteras i de syftesreglerande direktiven. Förhandlingen omfattar således tre delar och dessa kommer att behandlas i tur och ordning.

4.8.1 Förhandlingen om formatet

Förhandlingen om bedömningsverktygens format handlar i grunden om graden av tolkningsutrymme. De mättraditionella formaten begränsar tolkningsutrymmet medan de tolkningstraditionella formaten bejaktar tolkningsutrymmet. Jag särskiljer därför mellan de mättraditionella **slutna verktygsformaten** och de tolkningstraditionella **öppna verktygsformaten**.¹⁶⁹ Som framgått tidigare används tre bedömningsverktyg under en bedömningsprocess: observationsverktyg, värderingsverktyg och sammanvägningsverktyg.

Observationsverktygets format

För att genomföra en observation behövs det observationsverktyg. Ett observationsverktyg kan vara en enskild uppgift i ett prov, men det kan även vara ett prov i sin helhet.¹⁷⁰ Vid konstruktionen av en enskild uppgift i historia, finns det en historiekulturell vad-dimension som handlar om vilka historiska kunskaper som ska bedömas, det vill säga till vilken domänlogisk delkomponent inom ett konstrukt som uppgiften ska relateras till. Det finns även en bedömningskulturell hur-dimension som handlar om uppgiftens format, det vill säga *hur* uppgiften ska konstrueras för att eleven i sitt svar ska ges möjlighet att visa de historiska kunskaper som uppgiften avser att bedöma.

Det som utmärker de mättraditionella slutna observationsverktygen är att de genererar kvantitativ data i någon form. Typexemplet är en flervalsuppgift där eleven ska markera rätt svarsalternativ. Det innebär i sin tur att elevens kunskaper observeras/manifesteras indirekt.

De öppna tolkningstraditionella observationsverktygen genererar däremot någon form av kvalitativ data, till exempel en text. Till skillnad från

¹⁶⁹ En vanlig uppfattning inom det bedömningsteoretiska fältet är att det slutna verktygsformatet korrelerar med den mättraditionella betoningen på reliabilitet och att de öppna verktygsformatet korrelerar med den tolkningstraditionella betoningen på validitet. Se t.ex. Rudnert & Bruér (2021), s. 161–164.

¹⁷⁰ Ett annat sätt att särskilja den enskilda uppgiften från provet i sin helhet är att den senare prövar ett konstrukt medan den förra prövar en delkomponent inom ett konstrukt.

de mättraditionella observationsverktygen observeras/manifesteras elevens kunskap direkt i en prestation. Provkonstruktören har två olika tolkningstraditionella format att välja mellan: kortsvarsuppgiften och essäuppgiften.¹⁷¹ I den förra besvarar eleven uppgiften med några få meningar och i den senare med en längre löpande text.

Värderingsverktygets format

För att genomföra värderingen av en elevprestation används värderingsverktyg. Vid konstruktionen av ett värderingsverktyg finns det en historiekulturell vad-dimension som handlar om vad som utmärker kvalitet i en elevprestation, det vill säga vilken progressionslogik som tillämpas. Det finns även en bedömningskulturell hur-dimension som handlar om verktygets format, det vill säga hur verktyget ska utformas för att möjliggöra en värdering av elevprestationen.

De mättraditionella värderingsverktygen utmärks av att de genererar kvantitativ data i någon form. Värderingen sker i två steg. Om värderingsverktyget är konstruerat för att hantera slutna svarsformat det vill säga kvantitativ data, innebär det första steget att ett dikotomt facit (rätt/fel) används för att generera ett statistiskt underlag i någon form. Om värderingsverktyget däremot är konstruerat för att hantera öppna svarsformat, det vill säga kvalitativ data, innebär det första steget att någon form av värderingsmatris används (se nedan) för att generera motsvarande statistiska underlag. I båda fallen relateras sedan det statistiska underlaget till någon form av skala med gränser (cut scores) för att särskilja olika kvalitetsnivåer.¹⁷² Notera att i det första fallet är värderingsprocessen automatiserad och lärarens tolkningsutrymme reducerad till noll.

De tolkningstraditionella värderingsverktygen utmärks av att de hanterar och genererar kvalitativ data i någon form. Kompetensen att som lärare värdera kunskapskvaliteter i öppna svarsformat har traditionellt sett kopplats till svävande begrepp som ”magkänsla” eller ”konnässörförmåga”. En ytterlighet inom tolkningstraditionen är att läraren helt bortser från värderingsverktygen och istället förlitar sig på sin egen magkänsla.

¹⁷¹ Notera att resonemanget gäller skriftliga prov. Kunskaper kan även visas genom praktisk handling.

¹⁷² Datainsamlingen i samband med slutna flervalsuppgifter kan beskrivas som kvantitativ. Jönsson skriver om cut scores och särskiljer mellan dikotoma, kontinuerliga och flersteg. Jönsson (2013), s. 94–97. Det är även möjligt att använda uppgifter med en poängberäkning av typen ”partial credits”, det vill säga att poäng ges på basis av svarets kvalitet. Den här typen av poängberäkning tillämpas t.ex. i Finland i samband med värderingen av uppgifterna i ämnesrealproven. Gunnemyr (2011).

Läraren ersätter bedömningsverktyget. På senare tid har dock allt fler konkreta värderingsverktyg letat sig in i klassrummen.¹⁷³ I föreliggande projekt riktas fokus huvudsakligen mot värderingsmatrisen.¹⁷⁴ Värderingsmatriser används för att tydliggöra vilka kriterier som en elevprestation ska värderas efter. En vanlig kritik mot värderingsmatriser handlar om att de låser tanken och hämmar kreativiteten hos såväl elev som lärare.¹⁷⁵ I grunden handlar denna kritik om att matrisen begränsar både elevens och lärarens tolkningsutrymme. I vilken utsträckning tolkningsutrymmet begränsas kan dock variera. I matriser med en hög abstraktionsgrad är kriterierna allmänt hållna vilket ger elev och lärare ett stort tolkningsutrymme. Genom att förtydliga kriterierna (ökad precisering/ fler detaljer) sjunker abstraktionsgraden vilket begränsar tolkningsutrymmet. Det är således en glidande skala där abstraktionsgraden och tolkningsutrymmet är tätt sammanlänkade.¹⁷⁶

Inom ramen för detta forskningsprojekt är det viktigt att notera att en värderingsmatris kan användas för generera såväl kvantitativ som kvalitativ data, det vill säga både öppna och slutna prestationsomdömen. Värderingsmatriser som används för att hantera ett öppet svarsformat i form av en text (kvalitativ data) och som genererar ett slutet prestationsomdöme i form av ett betygsomdöme (kvantitativ data) befinner sig således i gränslandet mellan mättraditionen och tolkningstraditionen. Som framgår ovan har jag dock valt att kategorisera dessa datatransformerande värderingsmatriser som hemmahörande inom den mättraditionella traditionen.

Sammanvägningsverktygets format

För att genomföra sammanvägningen av enskilda prestationsomdömen till ett samlat prestationsomdöme används sammanvägningsverktyg. Den

¹⁷³ Lindberg 2009, s. 113–117.

¹⁷⁴ Det finns i princip tre olika typer av värderingsverktyg som kan användas för att hantera öppna svarsformat: värderingsmatriser, prestationsexempel och (metakognitiva) värderingskommentarer. Dessa värderingsverktyg kan användas separat eller tillsammans med varandra. Prestationsexempel är konkreta exempel på elevprestationer som motsvarar olika kravnivåer. Genom att kvalitetsbestämma typprestationer (autentiska eller konstruerade) skapas en mall som elevprestationen kan jämföras med. I praktiken används prestationsexempel ofta för att komplettera och konkretisera en värderingsmatris. Även värderingskommentaren kan ses som ett komplement till matrisen. Genom ett metakognitivt resonemang kring värderingen tydliggörs grunderna för värderingen.

¹⁷⁵ Lindberg (2009), s. 113–117.

¹⁷⁶ Inom det bedömningssteoretiska fältet är det vanligt att särskilja mellan holistiska och analytiska värderingsmatriser. Den analytiska matrisen svarar mot en mer deduktiv ansats medan den holistiska svarar mot en mer induktiv ansats.

historiekulturella vad-dimensionen handlar om vilken tyngd de olika delkomponenterna i ett konstrukt har, det vill säga vilken emfaslogik som tillämpas. Den bedömningskulturella hur-dimension handlar om verktygets format, det vill säga hur verktyget ska utformas för att möjliggöra en sammanvägning av enskilda prestationsomdömen till ett samlat prestationsomdöme.

De mättraditionella sammanvägningsverktygen hanterar kvantitativ data i form av en sammanställning av poäng på ett antal enskilda uppgifter och genererar ett kvantitativt prestationsomdöme i någon form, till exempel en totalpoäng eller ett slutbetyg. Om det sammanfattande omdömet består av en totalpoäng måste denna relateras till någon slags skala med gränser (cut scores) för olika provbetyg för att generera ett meningsfullt prestationsomdöme. Denna skala baseras ofta på normalfördelningskurvan. En strikt tillämpning av en normalfördelningskurva utgör en ytterlighet där lärarens tolkningsutrymme reduceras till noll.

De tolkningstraditionella sammanvägningsverktygen hanterar kvalitativ data i form av en sammanställning av enskilda skriftliga omdömen. Dessa är meningsfulla i sig själva men de kan även ligga till grund för ett sammanfattande skriftligt (slut)omdöme. Det öppna formatet skapar ett tolkningsutrymme för läraren i samband med sammanvägningen.

På samma sätt som gäller för de tolkningstraditionella värderingsverktygen finns det dock sammanvägningsverktyg som befinner sig i gränslandet mellan de båda traditionerna. Dessa datatransformerande verktyg, som hanterar öppna prestationsomdömen och genererar slutna samlade omdömen i form av ett betygsomdöme, kategoriseras som mättraditionella.

4.8.2 Förhandlingen om processen

För att reglera genomförande av bedömningen krävs tre olika processreglerande direktiv: observationsdirektiv, värderingsdirektiv och sammanvägningsdirektiv. Den grundläggande skillnaden mellan direktiven handlar återigen om graden av handlings/tolkningsutrymme (processuellt tolkningsutrymme). De mättraditionella formaten begränsar handlings/tolkningsutrymmet och de tolkningstraditionella formaten bejakar handlings/tolkningsutrymmet. Jag särskiljer därför mellan de mättraditionella **standardiserade direktiven** och de tolkningstraditionella **situerade direktiven**.

Observationsdirektiven

Med observationsprocessen avses den del av bedömningsprocessen då eleven ges möjligheter att visa kunskaper och där läraren således ges möjlighet att observera elevens kunskaper. Observationsprocessen hanteras på olika sätt inom mättraditionen och tolkningstraditionen.

Inom mättraditionen är idealet att begränsa lärarens handlingsutrymme. Ett begränsat handlingsutrymme i samband med observationen skapas genom en hög grad av standardisering. Vid en standardiserad observationsprocess särskiljs bedömningen från undervisningen och det ställs höga krav på att skapa lika (neutrala) förutsättningar för de lärare som ska genomföra observationen. Vad gäller provsituationen handlar det om att tid, plats och skrivtid ska vara lika för de elever som ska skriva provet. Standardisering innebär också att läraren inte kan välja ut vilka elever som ska skriva provet. Ytterligare en aspekt av standardiseringen är att de provuppgifter som används ska vara lika, det är således inte förenligt med en standardiserad observationsprocess att ändra på någon enskild uppgift eller att välja ut enskilda uppgifter från ett prov med flera uppgifter. Syftet med att standardisera observationsprocessen är att de elevprestationer som genereras ska kunna användas i ett jämförande/summativt syfte.

Inom tolkningstraditionen är idealet däremot att bejaka lärarens handlingsutrymme. Ett öppet handlingsutrymme skapas genom att kraven på standardisering tonas ned och att läraren ges möjlighet att situera observationen. Vid en situerad observationsprocess integreras observationen med undervisningen i sin helhet. Observationen anpassas till undervisningsgruppens förutsättningar. Syftet med att tona ner kraven på standardisering och därmed ge läraren möjlighet att situera observationen är att de elevprestationer som genereras ska kunna användas i ett formativt syfte.¹⁷⁷

Värderingsdirektiven

Med värdering avses den del av bedömningen då kvaliteten i en elevprestation ska bestämmas. Inom det bedömningsteoretiska fältet särskiljer man ofta mellan tre olika former av värderingsprocesser: intern värdering,

¹⁷⁷ Se Jönsson (2017).

extern värdering och kollegial samvärdering. En intern värdering innebär att det är den undervisande läraren som på egen hand värderar elevprestationen medan en extern värdering innebär att den undervisande läraren helt eller delvis har lämnat över kontrollen över värderingsprocessen till externa aktörer. Kollegial samvärdering innebär att det sker ett kollegialt utbyte av elevprestationer så att den enskilde läraren inte ensam värderar sina egna elevers prestationer.¹⁷⁸ Gränsdragningen mellan dessa tre former är många gånger flytande och komplex, men som ett analytiskt grepp är kategoriseringen motiverad. De tre formerna av värdering korrelerar mer eller mindre väl med de bedömningskulturella traditionerna.

De mättraditionella värderingsdirektiven syftar till att begränsa lärarens handlingsutrymme. Den interna värderingen är i grunden tolknings-traditionellt förankrad, men den kan förskjutas i mättraditionell riktning genom att läraren standardiserar sin egen värderingsprocess, till exempel genom att reglera om elevprestationerna i ett skriftligt prov ska värderas som en helhet eller uppgift för uppgift.¹⁷⁹ Den kollegiala samvärderingen är förankrad i mättraditionen om syftet och formerna för värderingen är att kontrollera den undervisande lärarens initiala värdering.¹⁸⁰ Den externa värderingen är tydligt förankrad inom mättraditionen. Vid en helt igenom extern värdering reduceras lärarens handlingsutrymme till noll.¹⁸¹ Notera att den externa värderingsprocessen kan uppfattas som en avlastning för professionen, men det kan även uppfattas som ett hot. Den undervisande läraren avprofessionaliseras.¹⁸²

De tolkningstraditionella värderingsdirektiven syftar till att bejaka lärarens handlingsutrymme. Den interna värderingen ligger i linje med denna ambition. Till skillnad från mättraditionen uppfattas den undervisande lärarens kontroll över värderingen som något eftersträvänsvärt. Grundtanken är att det är den undervisande läraren som är mest lämpad

¹⁷⁸ Vad gäller kollegial samvärdering finns det enligt Thornberg inget etablerat vetenskapligt begrepp inom det bedömningsteoretiska fältet. Thornberg förordar begreppet social moderation. Med social moderation avses att bedömare försöker nå konsensus i hur bedömning och betygssättning ska genomföras. Thornberg (2014), s. 18. Jfr Jönsson (2013), s. 190 ff.

¹⁷⁹ I det senare fallet ökar värderingens intrabedömarreliabilitet.

¹⁸⁰ Vid en kollegial samvärdering ökar bedömningens interbedömarreliabilitet.

¹⁸¹ Det som är intressant i denna studie är att jag bryter ner den externa bedömningsprocessen i olika delprocesser och att jag således kan se hur lärarens tolknings- och handlingsutrymme kan variera mellan de olika delprocesserna.

¹⁸² Eftersom bedömningsverktyget ersätter läraren är det bedömningskompetensen hos de aktörer som konstruerar mätverktygen som avgör graden av mätsäkerhet.

att hantera värderingen av elevprestationen. Genom att låta den undervisande läraren på egen hand värdera elevprestationerna skapas ett stort handlingsutrymme. Den interna värderingsprocessen ställer således höga krav på lärarens bedömningskompetens.¹⁸³ Den kollegiala samvärderingen är förankrad i tolkningstraditionen om syftet med samvärderingen är att stärka den undervisande lärarens bedömningskompetens.¹⁸⁴ Den externa värderingen har en svag förankring i tolkningstraditionen.

Sammanvägningsdirektiven

Med sammanvägning avses den del av bedömningsprocessen då läraren väger samman enskilda prestationsomdömen till ett samlat prestationsomdöme. Processuellt har sammanvägningen många likheter med värderingen, vilket innebär att det även vid en sammanvägning är möjligt att särskilja mellan interna, externa och kollegiala processer. Det innebär även att det är möjligt att särskilja mellan en mer standardiserad mättraditionell sammanvägning och en mer situerad tolkningstraditionell sammanvägning.¹⁸⁵

Det är viktigt att notera att sammanvägningen är en nödvändig delprocess i en renodlad mättraditionell bedömningsprocess. Enskilda prestationsomdömen i form av poängsummor måste vägas samman och relateras till en skala med gränser (cut scores) för olika provbetyg för att vara meningsfulla och funktionella. Motsvarande resonemang gäller inte för de enskilda prestationsomdömen som genereras i en renodlad tolkningstraditionell bedömningsprocess. Dessa skriftliga omdömen är meningsfulla och funktionella i sig själva, framför allt för formativa syften.

4.8.3 Förhandlingen om syftet

Varför ska man bedöma elevers kunskaper? Varför-frågan i ett skolrelaterat bedömningskulturellt sammanhang handlar om bedömningens syfte. Den bedömningskulturella varför-frågan regleras i provens syftesreglerande direktiv.

¹⁸³ Ett annat sätt att beskriva skillnaden är att vid en automatiserad värderingsprocess eftersträvas verktygsbaserad interbedömarreliabilitet medan vid en professionaliserad värderingsprocess eftersträvas en kompetensbaserad interbedömarreliabilitet.

¹⁸⁴ Jönsson särskiljer mellan jämkning och samvärdering. Den förra är en värderingsprocess som är tydligt förankrad i mättraditionens logik, medan den senare är mer komplex. Jönsson (2013).

¹⁸⁵ Resonemang kring kollegial värdering av elevprestationer är allmänt förekommande inom det bedömnings-teoretiska fältet, men motsvarande resonemang gällande kollegial sammanvägning av betygsunderlag i samband med slutbetygsättning relaterad är inte lika vanligt.

Inom fältet för pedagogisk bedömning används ofta epiteten *summativ* och *formativ* för att benämna olika sätt att hantera bedömningens syfte.¹⁸⁶ Skillnaden mellan en summativ och formativ bedömning handlar i grunden om huruvida bedömningens syfte är att mäta kunskaper vid ett givet tillfälle, bedömning *av* lärande, eller om bedömningens syfte är att fungera som ett pedagogiskt instrument, bedömning *för* lärande. Bedömning för lärande innebär att bedömningen genomförs i syfte att stödja elevens lärande, medan bedömning *av* lärande i första hand syftar till att kontrollera vad eleven kan.¹⁸⁷

Begreppsparet summativ och formativ bedömning kan relateras till mättraditionen och tolkningstraditionen. Ambitionen inom båda traditionerna är att bedömningen ska generera användbar data, men vad som uppfattas som användbara data skiljer sig åt, vilket i sin tur innebär att de datagenerande processerna skiljer sig åt. I mättraditionella sammanhang är bedömningens syfte i första hand summativt. I tolkningstraditionella sammanhang är bedömningens syfte i första hand formativt. Om syftet med bedömningen är summativt följer per automatik en förskjutning mot att kraven på reliabilitet och ett begränsat tolkningsutrymme, något som manifesteras i slutna format och standardiserade processer. Om syftet med bedömningen däremot är formativt tonas kraven på reliabilitet ned och det sker en förskjutning mot validitet och ett ökat tolkningsutrymme, något som manifesteras i öppna format och situerade processer.

4.8.4 Sammanfattning

Det är viktigt att understryka att ingen av de båda bedömningsteoretiska traditionerna i praktiken kan sägas förekomma i en renodlad form, utan att det hela tiden handlar om hybridformer där balansen mellan kraven på reliabilitet och validitet varierar. Traditionerna ska således inte uppfattas som renodlat dikotoma traditioner utan snarare som ”falska” dikotomier. Bättre är kanske att särskilja mellan traditionernas fokus och förgivettagande. Mättraditionen betonar reliabiliteten och förutsätter validiteten medan tolkningstraditionen betonar validiteten och förutsätter reliabiliteten.

¹⁸⁶ Se t.ex. Taras (2007). Shemilt (2018) särskiljer tre olika syften med bedömning: summativ, formativ och diagnostisk.

¹⁸⁷ Begreppsparet summativ/formativ bedömning uppfattas av många som synonyma med begreppsparet *bedömning av lärande* respektive *bedömning för lärande*. Jönsson argumenterar dock för att det finns betydelsebärande skillnader mellan formativ bedömning och bedömning för lärande. Jönsson (2013), s. 14–17.

I praktiken handlar det om att läraren i sin bedömningspraktik ska hantera balansen mellan de ideal som skrivs fram inom respektive tradition. Det handlar som sagt inte om ”antingen eller” utan om ”både och”. Utan någon form av slutna verktyg och standardiserade processer finns det en risk att bedömningen inte genererar ett tillförlitligt/reliabelt dataunderlag till stöd för en summativ användning. Å andra sidan finns det en risk att en renodlad användning av slutna format och standardiserade processer inte genererar ett dataunderlag som motsvarar de krav på validitet (i olika former) som eftersträvas vid en formativ användning.

4.9 Den fjärde kvadrantens dilemma

4.9.1 Ett teoretiskt fundament

I avhandlingens inledande kapitel presenterades en fyrfältsmatris vars grundläggande teoretiska utgångspunkt var att alla former av kunskapsbedömning i historia omfattar en historiekulturell vad-dimensionen och en bedömningskulturell hur-dimension.¹⁸⁸ I de två inledande delarna av teorikapitlet har dessa båda dimensioner behandlats separat. I det här tredje och avslutande avsnittet är tanken att ställa dem i relation till varandra.

Oavsett om man är historielärare eller provkonstruktör måste man hantera dessa båda grundläggande dimensioner i relation till varandra. Relationen mellan skolans historiekultur och skolans bedömningskultur är komplex och i många fall problematisk. Nedanstående figur beskriver i korthet det teoretiska fundament som ligger till grund för det fortsatta resonemanget.

¹⁸⁸ Schwab laborerar med en tredje dimension som han benämner den filosofiska dimensionen. Denna dimension inrymmer metakognitiva överväganden som kan kopplas till den didaktiska varför-frågan. Schwab (1978), s. 229–273. Se även Shemilt (2018), s.450. Det kan i detta sammanhang vara viktigt att understryka att skolans historiekultur är ämnesspecifik medan skolans bedömningskultur i grunden är ämnesgenerell.

Figur 5: Ett teoretiskt fundament

Verktyg och direktiv	Dimensioner	
	<i>Historiekultur</i> Vad-dimensionen (epistemologi)	<i>Bedömningskultur</i> Hur-dimensionen (metodologi)
Bedömningsverktyg	Konstrukt	Format
Bedömningsdirektiv		Process

Av figuren framkommer att den historiekulturella vad-dimensionen manifesteras i det konstrukt som ligger till grund för bedömningsverktygen. Vidare framkommer att den bedömningskulturella hur-dimensionen inrymmer två olika aspekter: en formatrelaterad aspekt och en processrelaterad aspekt. Den formatrelaterade aspekten manifesteras i bedömningsverktygens format. Den processrelaterade aspekten manifesteras i de processreglerande bedömningsdirektiven.

4.9.2 Teoretiska utgångspunkter

I det inledande kapitlet framkommer också att projektet prövar hypotesen att relationen mellan skolans historiekulturella vad-dimension och skolans bedömningskulturella hur-dimension inrymmer ett latent spänningsförhållande.

En första utgångspunkt för hypotesen är antagandet att det finns ett samband mellan dessa båda dimensioner, det vill säga mellan *vad* som bedöms och *hur* någonting bedöms. I det här sammanhanget handlar det om att det finns ett samband mellan den historiekulturella vad-dimensionen och den bedömningskulturella hur-dimensionen. Förändringar inom vad-dimensionen påverkar hur-dimensionen och förändringar inom hur-dimensionen påverkar vad-dimensionen.

En andra utgångspunkt för hypotesen är antagandet att skolans historiekulturella vad-dimension inrymmer både HT-traditionella och MH-traditionella inslag och att skolans bedömningskulturella hur-dimension inrymmer både mättraditionella och tolkningstraditionella inslag. Inom båda dimensionerna handlar det således om hybridformer. Även om det

hela tiden handlar om hybridformer är det dock möjligt att undersöka fokusförskjutningar inom respektive dimension.

En tredje utgångspunkt för hypotesen är antagandet att det finns ett särskilt spänningsförhållande mellan den historiekulturella HT-traditionen och den bedömningskulturella mättraditionen. Spänningsförhållandet kommer till uttryck när HT-traditionella konstrukt förankrade i en konstruktivistisk kunskapssyn möter mättraditionella format och processreglerande direktiv som är utformade för att begränsa bedömarens tolkningsutrymme. Spänningsförhållandet kan visualiseras med stöd av en modell i form av en fyrfältsmatris (se figur 6). Av figuren framgår att relationen mellan bedömningens vad-dimension och hur-dimension blir särskilt problematisk inom en av kvadranterna. Hypotesen och det specifika spänningsförhållandet kommer därför fortsättningsvis att benämnas som *den fjärde kvadrantens dilemma*.

Figur 6: Den fjärde kvadrantens dilemma

De två pilar som är infogade i kvadranten visualiserar en fjärde utgångspunkt för hypotesen, närmare bestämt antagandet att den förhandling som äger rum till följd av spänningsförhållandet kan ta sig två olika uttryck. Om de mättraditionella kraven väger tyngst under förhandlingen sker en vertikal förskjutning i uppåtgående riktning inom fyrfältsmatrisen. En sådan förskjutning i MH-traditionell riktning skulle innebära en ökad betoning på den objektivistiska dimensionen av historisk kunskap. Om däremot de HT-traditionella kraven väger tyngst sker en motsvarande horisontell förskjutning i vänstergående riktning inom fyrfältsmatrisen. En sådan förskjutning i tolkningstraditionell riktning skulle innebära ett ökat

tolkningsutrymme i samband med bedömningen. I det analytiska uppdraget blir det således viktigt att klargöra vilken av dimensionerna som väger tyngst under förhandlingen, det vill säga vilken av dimensionerna som äger tolkningsföreträde.

4.9.3 Praktisk tillämpning

Fyrfältsmatrisen och den fjärde kvadrantens dilemma kan konkretiseras genom att låta fyra idealtypiska historielärare få i uppdrag att konstruera en bedömningsuppgift som ger eleven möjligheter att visa kunskaper om den franska revolutionens konsekvenser, se figur 7 nedan. Den MH-traditionella historieläraren med förankring i mättraditionen kan hantera uppdraget genom att konstruera en flervalsuppgift där det korrekta svarsalternativet utgörs av ett eller flera exempel på konsekvenser av den franska revolutionen. Den MH-traditionella historieläraren med förankring i tolkningstraditionen byter format och konstruerar en essäuppgift där det i uppgiftsformuleringen framgår att eleven i sitt svar ska redogöra för konsekvenserna av franska revolutionen, det vill säga en rekonstruktion baserad på ett eller flera exempel på konsekvenser. Den HT-traditionella historieläraren med förankring i tolkningstraditionen konstruerar också en essäuppgift, men justerar uppgiftsformuleringen så att det framkommer att eleven i sitt svar på något sätt ska föra ett resonemang kring olika tolkningar av revolutionens konsekvenser. Om man till sist låter den HT-traditionella historieläraren med förankring i mättraditionen ta sig an uppdraget ställs hen inför den fjärde kvadrantens dilemma, vilket i det här fallet innebär att konstruera en flervalsuppgift där eleven ska hantera olika tolkningar av revolutionens konsekvenser.

Figur 7: Den fjärde kvadrantens dilemma i praktiken

4.9.4 Hanteringen av den fjärde kvadrantens dilemma

Huruvida mättraditionella uppgiftsformat i form av flervalsuppgifter kan användas för att pröva annat än minneskunskaper är omdiskuterat. Om man vänder sig till det bedömningsteoretiska fältet så finns det enligt Scully en utbredd uppfattning inom detta fält att det inte är möjligt att använda slutna format för att pröva "higher order thinking".

This reflects a general perception that MC items are incapable of assessing cognitive process beyond recall or recognition of knowledge, given that the correct answer is provided amongst the response options.¹⁸⁹

Hennes egen uppfattning är dock att det är möjligt, men att det är såväl utmanande som tidskrävande.¹⁹⁰ Scully preciserar sedan sin uppfattning genom att argumentera för att det är möjligt att konstruera flervalsuppgifter som prövar vissa former av "higher order thinking".¹⁹¹

Även inom det historiedidaktiska fältet finns det en utbredd uppfattning att flervalsuppgifter passar bäst för att pröva MH-traditionella konstrukt

¹⁸⁹ Scully (2017), s. 4. Jönsson et al (2017) betonar att slutna uppgiftsformat ökar risken för konstruktirrelevans, det vill säga att uppgiften prövar andra förmågor än de som är tänkta att prövas som exempelvis läsförmåga eller ordförståelse.

¹⁹⁰ Scully (2017), s. 4.

¹⁹¹ Även Wikström (2013) menar att det går att mäta högre kognitiva nivåer genom uppgifter med slutna format, men att det finns andra format som är bättre.

med fokus på faktakunskaper (rekonstruktion).¹⁹² På samma sätt som inom det bedömningsteoretiska fältet finns det dock röster som menar att det är möjligt att konstruera flervalssuppgifter som prövar konstruktion och dekonstruktion, det vill säga hantera den fjärde kvadrantens dilemma.¹⁹³ Det finns dessutom exempel på historiedidaktiker som tagit sig an uppdraget att hantera den fjärde kvadrantens dilemma i praktiken.

På den internationella arenan är Bruce VanSledrights flervalssuppgifter med viktade svarsalternativ det kanske mest omskrivna exemplet. Hans grundtanke är att i ett tolkningsämne som historia är det både möjligt och önskvärt att de olika svarsalternativen i en flervalssuppgift svarar mot mer eller mindre relevanta tolkningsalternativ. En flervalssuppgift som prövar elevens kunskaper om den franska revolutionens konsekvenser bör därför innehålla ett felaktigt svarsalternativ (en distraktor) samt två eller flera svarsalternativ som är rimliga, men som bedöms ha olika relevans. Hans tankemodell inrymmer även en möjlighet för eleven att i efterhand argumentera för det valda svarsalternativet.¹⁹⁴

På den svenska nationella arenan kan nämnas det nationella provet i historia för åk 9 där det finns exempel på traditionella flervalssuppgifter som gör anspråk på att pröva HT-traditionella konstrukt. I anslutning till uppgiften argumenterar provkonstruktörerna för uppgifternas validitet genom att hävda att de prövar elevernas ”förutsättningar att resonera”. Denna validitetsargumentation ligger i linje med Scullys resonemang om ”pseudoassessment” av komplexa kunskaper med stöd av flervalssuppgifter. Hon menar att flervalssuppgifter har potential att pröva komplexa kunskaper, men det förutsätter att provkonstruktören beaktar ett antal grundläggande strategier vid konstruktionen, till exempel att låta samtliga svarsalternativ vara rimliga, men att ett alternativ är ”mer korrekt” än de övriga.¹⁹⁵

Huruvida den tankeoperation som föranleder eleven att sätta ett kryss i ett av svarsalternativen verkligen bygger på kunskaper i historia har dock ifrågasatts av forskare inom fältet. Reich (2009) har i en historiedidaktisk studie visat att det finns en bristande korrelation (ett glapp) mellan de

¹⁹² Rudnert & Bruér (2021), s. 161-162.

¹⁹³ Se Rudnert & Bruér (2021), s. 161-162. En tydlig fördel med flervalssuppgifterna är också att de går snabbare att besvara vilket möjliggör en bredare täckning av den substantiella dimensionen av domänlogiken, något som ligger i linje med MH-traditionen.

¹⁹⁴ VanSledright är tydlig med att dessa uppgifter har ett formativt syfte på lärnivå. VanSledright (2014).

¹⁹⁵ Scully (2017), s.5-9.

tankeoperationer som elever tillämpar för komma fram till rätt svar och det konstrukt som uppgifterna gör anspråk på att pröva. Hans slutsats är att flervalsuppgifterna i högre grad prövar elevernas provsvarsförmåga (*test wiseness*). Reichs resultat kan dock ställas mot en annan historiedidaktisk studie som genomfördes av Smith et al. Deras resultat indikerar att det förvisso finns inslag av testwisness i elevernas tankeoperationer i samband med flervalsuppgifter, men att eleverna trots allt ändå använder relevant historiskt tänkande när de besvarar uppgifterna.¹⁹⁶

4.9.5 Korrelation mellan format och konstrukt

Vad gäller korrelationen mellan uppgiftens format och konstrukt är det som sagt en vanlig uppfattning inom det historiedidaktiska fältet att det slutna uppgiftsformatet i högre utsträckning korrelerar med MH-traditionella konstrukt medan det öppna uppgiftsformatet i högre utsträckning korrelerar med HT-traditionella konstrukt. Korrelationen mellan uppgiftsformat och konstrukt är dock komplex och väcker många viktiga frågor, inte minst inom historieämnet. Det finns forskare inom fältet som har ifrågasatt huruvida öppna essäuppgifter verkligen är ett effektivt verktyg för att mäta elevers historiska kunskaper. Alexander Maxwell menar att historielärare i högre utsträckning bör använda sig av öppna kortsvarsuppgifter.¹⁹⁷

Även de amerikanska historiedidaktikerna Keith Barton och Linda Levstik, har problematiserat användningen av den öppna essäuppgiften inom historieämnet. De fokuserar på de dokumentbaserade källkritiska essäuppgifterna (DBQ) som har fått en allt större tyngd i många av de externa proven i USA. Förespråkare för DBQ menar att frågorna skapar förutsättningar för en mer autentisk bedömning, men Barton och Levstik är starkt skeptiska till att detta i praktiken verkligen blir fallet. I sin kritik av uppgiftsformatet utgår de från att en meningsfull och autentisk dokumentbaserad undersökning består av tre delar. Vid sidan av den källkritiska prövningen av materialet måste uppgiften även innehålla ett meningsfullt problem eller en meningsfull fråga och den källkritiska prövningen måste resultera i att eleven drar egna slutsatser. De menar att de

¹⁹⁶ Smith et al. (2019) s.137. Reisman et al. noterar att ingen av dessa lösningar beaktar "content knowledge" i sina uppgifter. Reisman et al. (2019), s. 537.

¹⁹⁷ Maxwell (2010), s. 233–241.

dokumentbaserade uppgifterna isolerar den källkritiska färdigheten. Eleverna ställs inför frågor som inte är deras egna och som de många gånger inte uppfattar som meningsfulla. Om frågan inte är meningsfull så saknar svaret betydelse. Det som återstår är ingen meningsfull undersökning utan endast en artificiell analys av dokument.

Many such exercises fail to engage students even in the most minimal evaluation of the usefulness of sources they are examining, for their purpose is not to develop any kind of reasoned judgment, but to help students explain how given primary sources illustrate facts and concepts they have learned elsewhere.¹⁹⁸

För att konkretisera resonemanget gällande essäuppgifter kan det avslutningsvis vara intressant att ta del av finländsk forskning gällande essäuppgifterna i det finländska realprovet i historia. Arja Virta särskiljer en äldre form av essäuppgift från en förnyad form av essäuppgift. Den förra varianten bestod i princip endast av en essärubrik medan den senare varianten ofta var dokumentbaserad och betydligt mer komplex till sin konstruktion.

De [...] fick en ny formulering så att de innehöll en specificerad synvinkel som antydde tankeprocessen eller ett förväntat fokus. Essäuppgiften kunde som utgångspunkt ha en paradox eller ett påstående som eleven skulle kommentera eller bedöma”.¹⁹⁹

Virta skriver vidare att utvecklingen av den ”förnyade” essäuppgiften skedde parallellt med den under 1980-talet framväxande diskussionen om historieämnets egenart och historiskt tänkande. Även Sirkka Ahonen, som var huvudansvarig för historiefrågorna i finländska realprovet, under åren 1994–2000, menar att ett nytänkande i fråga om historiefrågornas utformning i Finland tog sin början under 1990-talet. Då skedde en övergång från de traditionella rubrikuppgifterna till uppgifter som utgick från dokument- och statistikuppgifter. Provet skulle mäta elevens förmåga att dra

¹⁹⁸ Barton & Levstik (2004), s. 199.

¹⁹⁹ Virta (1997), s. 174.

slutsatser och tillämpa sina kunskaper – det skulle inte längre vara tillräckligt med att minnas och upprepa.²⁰⁰

Tom Gullberg menar att den successiva omläggningen av uppgiftsformatet bekräftade övergången från en stoffcentrerad ämnessyn mot en mer funktionell ämnessyn i historia.²⁰¹ Gullbergs slutsats att historieämnets domänlogiska förskjutning i en mer HT-traditionell riktning kan relateras till en formatlogisk förskjutning i tolkningstraditionell riktning ligger i linje med den grundläggande premissen för den fjärde kvadrantens dilemma, det vill säga att det finns ett samband mellan skolans historiekultur och skolans bedömningskultur.

4.9.6 Återkoppling till forskningsprojektet

Avslutningsvis kan det vara på sin plats att återkoppla till forskningsprojektet. Det övergripande syftet med projektet är som sagt att undersöka historiekulturell förändring över tid med avseende på skolämnet historia och då särskilt beakta vilken betydelse kunskapsbedömning i allmänhet och externa prov i synnerhet kan ha för denna förändring. Inom ramen för projektet finns en tydlig teoriutvecklande dimension, vilket bland annat tar sig uttryck i att pröva hypotesen om den fjärde kvadrantens dilemma, det vill säga att relationen mellan skolans historiekulturella vad-dimensionen och bedömningskulturella hur-dimensionen inrymmer ett latent spänningsförhållande.

Eftersom de empiriska resultat som studien genererar genomgående kommer att analyseras i ljuset av den fjärde kvadrantens dilemma är det ett viktigt klargörande att dilemmat i samband med externa prov operationaliseras på olika sätt beroende på vilken arena som utgör fokus för analysen. På den historiekulturella formuleringsarenan operationaliseras dilemmat *vid konstruktionen av provet* medan det på den historiekulturella realiseringsarenan operationaliseras *vid implementeringen av provet*. Som provkonstruktör på formuleringsarenan handlar det om att konstruera bedömningsverktyg som hanterar ett HT-traditionellt konstrukt samtidigt som de genererar mätsäkra resultat. Som historielärare på realiseringsarenan handlar det om att vid implementeringen ta ställning till i vilken utsträckning du förmår vara lojal med dessa bedömningsverktyg.

²⁰⁰ Ahonen (2004), s. 6.

²⁰¹ Gullberg (2006), s. 36–37.

Både provkonstruktörer och historielärare behöver således utarbeta strategier för att hantera den fjärde kvadrantens dilemma. Provkonstruktören kan ställas inför dilemmat om provbeställaren kräver slutna verktygsformat samtidigt som styrdokumentet utgår från ett HT-traditionellt konstrukt. Historielärare kan på motsvarande sätt behöva hantera dilemmat om ett mättraditionellt externt prov introduceras. Historielärare kan dessutom behöva hantera dilemmat om kraven på att deras interna bedömning ska generera mätsäkra resultat ökar.

5. METOD

Föreliggande studie kan metodologiskt beskrivas som en intervjubaserad mikrostudie. Att genomföra en studie med ett begränsat antal informanter innebär att mikroperspektivets fokus på det enskilda och unika betonas i högre utsträckning än makroperspektivets fokus på det allmänna och generella. Det är dock ett viktigt klargörande att mikro- och makrostudier förutsätter och kompletterar varandra. Varje mikrostudie behöver sättas in i en större kontext för att bli begriplig och meningsfull. Relationen mellan mikro- och makrostudier kan med fördel kopplas till begreppet generaliserbarhet, ett begrepp som kommer att behandlas i slutet av detta kapitel (se avsnitt 5.2.6).²⁰²

Metodkapitlet består av två delar. I den första delen beskrivs de metodologiska vägvalen i samband med datainsamlingen och i den andra delen motsvarande vägval för den efterföljande dataanalysen. Gränsdragningen mellan datainsamling och databehandling är dock inte ristad i sten, särskilt inte om man uppfattar datainsamlingen som en *reflexiv praktik* i linje med Helen Thomsson.²⁰³ I en reflexiv praktik är datainsamlingen en process vars förutsättningar förändras allteftersom det insamlade materialet analyseras. Betydelsen av denna pendelrörelse mellan insamling och analys är viktig. Det som eftersträvas i föreliggande projekt är en transparens kring förutsättningarna för denna pendelrörelse.

5.1 Datainsamling

5.1.1 Den kvalitativa forskningsintervjun

I valet av metod för datainsamling har jag tagit hänsyn till en flera olika faktorer, alltifrån teoretiska överväganden till praktiska omständigheter gällande tid och resurser, men grundläggande för mina ställningstaganden har i slutändan ändå alltid varit det syfte och de forskningsfrågor som mitt forskningsprojekt utgår ifrån.

²⁰² För ett elaborerat resonemang kring mikro och makro i en historiekulturellt förankrad studie, se Rudnert (2019), s. 82–85.

²⁰³ Thomsson (2002).

Jag har valt att använda mig av kvalitativa forskningsintervjuer som metod för datainsamlingen. Genom personliga och semistrukturerade intervjuer med fyra HLP-lärare och fyra KP-lärare har jag fått tillgång till ett empiriskt underlag som har gett mig möjlighet att undersöka hur historielärare uppfattar och hanterar frivilliga externa prov i historia. Den kvalitativa intervjun är en metod vars styrka ligger i att den ger mig som forskare möjlighet att få fram detaljerad och djuplodande information kring komplexa frågeställningar. Det är en flexibel metod som gör att jag hela tiden har möjlighet att ställa följdfrågor och följa upp sidospår.²⁰⁴

5.1.2 Den kvalitativa forskningsintervjun i praktiken

I likhet med många andra forskare som använder sig av kvalitativa forskningsintervjuer uppfattar jag att mötet mellan den intervjuande forskaren och den intervjuade informanten som situerat.²⁰⁵ Det innebär att det empiriska underlag som intervjun genererar måste ses mot bakgrund dels av forskarens och informantens *förförståelse*, dels av *interaktionen* mellan dessa båda aktörer. I mina fortsatta metodologiska resonemang gällande datainsamlingen har jag valt att särskilja dessa båda dimensioner.

Förförståelsens betydelse

Betydelsen av forskarens förförståelse kan belysas på åtminstone två olika sätt. Ett sätt är att resonera om betydelsen av *förkunskaper* inom det specifika ämnesområde som studien behandlar. Genom att kunskapsmässigt behärska det aktuella ämnesområdet framstår forskaren som trovärdig i intervjusituationen. Om forskaren saknar ämneskunskaper finns det en risk att informanten uppfattar intervjusituationen som oseriös, vilket kan påverka det empiriska underlag som genereras under intervjun.²⁰⁶ Kraven på forskarens förkunskaper kan och har problematiserats, bland annat genom att ställa sig frågan om alltför omfattande förkunskaper hos forskaren kan innebära att öppenheten för alternativa tolkningar begränsas. Min egen uppfattning är dock att utan omfattande förkunskaper inom

²⁰⁴ Denscombe (2000), s. 161–162.

²⁰⁵ Se t.ex. Glaser (2016).

²⁰⁶ Ritchie nyanserar denna kunskapsmässiga aspekt genom att särskilja mellan intervjuar-kunskap och ämneskunskap. Ritchie (2015), s. 44.

ämneseområdet är det svårt för forskaren att inför intervjun formulera relevanta intervjufrågor och ännu svårare att under intervjun ställa följdfrågor och följa upp sidospår.²⁰⁷

Ett andra sätt att belysa förståelsens betydelse är att resonera om forskarens *position* i relation till det specifika ämneseområde som studien behandlar. Ett sätt att beskriva denna position är att ta stöd av begrepps-paret insider och outsider.²⁰⁸ En forskare med ett insiderperspektiv är själv en del av det undersökta ämneseområdet medan en forskare med ett outsiderperspektiv undersöker ett ämneseområde som hen inte är en del av.²⁰⁹ Värdet av insiderperspektivet kan kopplas till det föregående resonemanget kring vikten av förkunskaper, men det inrymmer som jag ser det en tydlig forskaretsk problematik. Det är i detta sammanhang viktigt att hålla isär begreppen forskaretsk och forskningsetik. Forskaretsk handlar om forskarens relation till själva forskningsuppgiften eller forskningsuppdraget medan forskningsetik behandlar frågor om hur man i forskningen tar hänsyn till och skyddar olika deltagare och försökspersoner som berörs av forskningen.²¹⁰

Den forskaretska problematiken är relevant i föreliggande forskningsprojekt eftersom jag ingick i den arbetsgrupp som av Skolverket hade fått i uppdrag att konstruera KP. Det innebär att jag brottas med åtminstone två forskaretska dilemman. Det första är att jag som provkonstruktör beforskar mitt eget prov och det andra är att KP-lärarna vid en intervjusituation kan påverkas av att intervjuaren och provkonstruktören är en och samma person.

Det förstnämnda dilemmat handlar i grunden om huruvida det är lämpligt att bedriva forskning om ett fenomen/objekt som man som forskare hyser personliga sympatier eller antipatier för. Jag delar dock Lifmarks resonemang kring vad han har formulerat som ”principen om känslomässig neutralitet”. Han menar att det är omöjligt att kräva en sådan känslomässig neutralitet hos forskaren. Det är inte rimligt att på förhand

²⁰⁷ Jfr Ritchie (2015).

²⁰⁸ Se t.ex. Thompson (1980), s. 20 och Ritchie (2015), s. 118.

²⁰⁹ Se t.ex. Glaser (2016).

²¹⁰ I en rapport från Vetenskapsrådet gör rapportförfattarna en distinktion mellan *forskaretsk* som handlar om forskarens relation till själva forskningsuppgiften eller forskningsuppdraget och *forskningsetik* som behandlar frågor om hur man i forskningen tar hänsyn till och skyddar olika deltagare och försökspersoner som berörs av forskningen. Gustafsson et al (2005), s. 19.

diskvalificera en forskare som vill bedriva forskning inom ett visst ämnesområde endast på grundval av en känslomässig bias. Det är forskarens metod och inte forskarens person som ska granskas.²¹¹ I föreliggande projekt kan dilemmat hanteras genom att jag i mitt analysarbete eftersträvar tydlighet och transparens.

Det andra dilemmat, det som handlar om att KP-lärarnas utsagor kan påverkas av mina dubbla roller som forskare och provkonstruktör, är och förblir ett svårlöst dilemma. Interaktionen mellan forskaren och informanten är dock ett problem som gäller generellt vid intervjubaserade studier.

Interaktionens betydelse

Att interaktionen mellan intervjuaren och den intervjuade påverkar den data som genereras under en intervju är en välkänd problematik i samband med intervjubaserade studier. Ofta problematiseras interaktionen genom att relatera den till maktförhållandet mellan forskaren och informanten.²¹² Jag kommer som sagt att behöva hantera maktdimensionen i samband med interaktionen eftersom KP-lärarna vid intervjusituationen kan påverkas av att jag som forskare/intervjuare också är provkonstruktör och därmed indirekt kan uppfattas som Skolverkets förlängda arm.

Ett grundläggande antagande gällande interaktionen är att maktförhållandet påverkar såväl själva intervjun som resultatet av den. Resultatet kan kopplas till begreppet intervjuareffekt.²¹³ En viktig dimension av intervjuareffekten är relationen mellan retorik och praktik, det vill säga att det kan finnas skillnader mellan vad den intervjuade säger sig göra och vad hen faktiskt gör.²¹⁴ Inom det historiedidaktiska fältet är Ronald Evans forskning belysande för detta fenomen. Han har med stöd av två studier som baserar sig på både intervjuer och uppföljande klassrumsobservationer kunnat observera stora skillnader mellan vad historielärare säger sig göra och vad historielärare faktiskt gör i klassrummet.²¹⁵ Evans studier visar att det är en fördel att kunna styrka sina resultat genom triangulering,

²¹¹ Lifmark (2004). Lifmark understryker att forskarens frågor är viktiga att beakta: "Forskarens personliga och yrkesmässiga bakgrund kan prägla perspektivval, stoffurval och tolkningar och det är värt att diskutera i vilken grad denna bakgrund bör synliggöras och problematiseras". Lifmark (2004), s. 195.

²¹² Se t.ex. Portelli (1991), s.35 och s. 42; Thompson (1980), s. 120–127.

²¹³ Gunnemyr (2011), s. 19–20.

²¹⁴ Denscombe (2000), s. 163.

²¹⁵ Evans (1989) och (1990).

det vill säga genom att tillämpa flera metoder inom ett och samma projekt. Att likt Evans följa upp intervjuresultaten med klassrumsobservationer hade varit en styrka, men i föreliggande projekt används endast intervjuer som metod för datainsamling. En konsekvens av denna metodologiska begränsning är att jag tydligt måste särskilja mellan historielärarnas retorik och praktik. De resultat och slutsatser som framkommer i projektet gör därför endast anspråk på att vara gällande för historielärares *uppfattningar om och beskrivningar av* hur de genomför, använder och påverkas av frivilliga externa prov i historia.²¹⁶

5.1.3 Urvalet av historielärare

I studien ingår åtta historielärare – fyra HLP-lärare och fyra KP-lärare. Vilka specifika lärare som jag har valt att intervjua har i hög grad styrts av tillgängligheten. Antalet HLP-lärare var begränsat eftersom HLP endast användes av ett mindre antal skolor i Skåne-regionen. Även antalet KP-lärare var begränsat, eftersom KP var ett frivilligt externt prov.²¹⁷

I en kvalitativ undersökning är det oftast en omöjlighet att eftersträva ett representativt urval i en mera strikt bemärkelse, men jag har haft en ambition att urvalet skulle vara så varierat som möjligt. Jan Trost redogör för ett sätt att uppnå ett strategiskt urval i en kvalitativ undersökning. Tanken är att man väljer ut ett antal variabler hos informanterna och att man sedan anger lämpliga variabelvärden för dessa. Eftersom proven är frivilliga är den grundläggande variabeln att läraren har använt sig av något av proven och att denna användning har varit återkommande. Jag hade dessutom ambitionen att låta kön och erfarenhet av undervisning utgöra variabler, men särskilt den senare variabeln var i praktiken svår att leva upp till.²¹⁸ Fyra av lärarna är kvinnor och fyra är män. Två av lärarna har en mer begränsad undervisningserfarenhet: fem respektive åtta år. De övriga sex har undervisat mellan 24 och 35 år. De intervjuade lärarna är anställda på åtta olika gymnasieskolor belägna i södra Sverige.

²¹⁶ Resonemang kring intervjueffekten är central inom alla former av professionsforskning, men den är även viktig att beakta när mediavana människor intervjuas. Det finns i båda fallen en retorisk dimma i form av ett färdigt (och inövat) narrativ som forskaren måste tränga igenom. Jfr med Quinlans (2011) resonemang om professionsforskning.

²¹⁷ Hur många historielärare som på något sätt använt sig av KP är oklart. Se Ammert & Eliasson (2019), s. 73–74.

²¹⁸ Trost (2005), s. 117–122.

Avslutningsvis vill jag understryka att lärarnas utsagor inte kommer att analyseras som individuella fallstudier. I min undersökning är gruppens uppfattningar viktigare än den enskilda individens uppfattningar. Jag har därför valt att inte ge någon närmare presentation av de åtta lärare som ingår i undersökningen.

5.1.4 Genomförandet av intervjuerna

Lärarna kontaktades via mail och efter ett uppföljande telefonsamtal avtalades tid och plats för intervjun. Sex av de åtta intervjuerna har ägt rum på informantens arbetsplats. En av HLP-lärarna föredrog att låta sig intervjuas på ett café och en av KP-lärarna föredrog att låta sig intervjuas i hemmet.

Vid intervjutillfället uppgav jag den beräknade intervjutiden till en timme. Fem av intervjuerna varade mellan 35 och 60 min. De tre övriga varade mellan 90 och 100 minuter. Orsakerna bakom denna spännvidd kan diskuteras. En grundläggande aspekt är naturligtvis den semistrukturerade intervjun i sig. Två av de långa intervjuerna ägde rum i hemmiljö, vilket kan ha bidragit till att dessa två informanter inte upplevde någon arbetsrelaterad stress och därför svarade utförligare på mina frågor. En annan orsak kan vara min grundläggande inställning att informanten skulle uppleva intervjusituationen mer som ett samtal än som en intervju-situation. Denna inställning innebar att jag medvetet lät bli att avbryta informanten om denne avvek från intervjuguidens struktur genom hoppa mellan frågorna eller genom att komma in på något sidospår.

Intervjuerna var semistrukturerade till sin karaktär och utgick från en intervjuguide (se bilaga 1 och 2). Efter intervjuerna antecknade jag mina omedelbara tankar. De inspelade intervjuerna transkriberades sedan ordagrant, men jag har inte systematiskt markerat pauseringar eller särskilda betoningar.

5.1.5 Forskningsetiska överväganden

Som forskare inom de humanistiska och samhällsvetenskapliga forskningsfälten måste jag förhålla mig till fyra huvudregler: informationskravet, samtyckeskravet, konfidentialitetskravet och nyttjandekravet. De lärare som jag har intervjuat har samtliga informerats om studiens syfte och de har samtyckt till medverkan. Allt intervjumaterial behandlas konfidentiellt och förvaras hos författaren. Lärarna har tilldelats fingerade namn

och de personliga uppgifterna är begränsade till kön och undervisnings-
erfarenhet. Det ska således inte vara möjligt att identifiera den enskilda
informanten.²¹⁹

5.2 Dataanalys

5.2.1 Analysarbetets inledning

Värdet av att systematiskt läsa och koda det empiriska materialet är odiskutabelt. Jag inledde dock analysarbetet med att spåra utsagor som kan beskrivas som nyckelutsagor. Emerson förespråkar användandet av nyckelincidenter (key incidentents) i en begynnelsefas av ett analysarbete. Han definierar en nyckelincident som ett specifikt stycke data som är empiriskt rikt, det vill säga att det rymmer gott om analytiska möjligheter. Emerson skriver vidare att det handlar om att jag som forskare ska få en känsla av att ”här har vi något – vad kan det betyda?” Han skriver även att det i forskarens möte med en nyckelincident infinner sig ”en intuitiv och teoretiskt känslig övertygelse om att något spännande har ägt rum”. Det är således jag som forskare som tillskriver en specifik utsaga status som nyckelutsaga. Tanken är sedan att jag i det fortsatta analysarbetet ska kunna belägga utsagans betydelse genom att hitta andra utsagor som stärker den initiala känslan.²²⁰

5.2.2 Den abduktiva pendelrörelsen

Nästa steg i analysarbetet kan beskrivas som en slags sortering.²²¹ När man inom kvalitativ forskning påbörjar sorteringen av insamlad data är det vanligt att man använder någon form av kodning.²²² I samband med kodningen ställs forskare inför ett vägval: den deduktiva eller induktiva vägen. En kodning med deduktiv ansats laborerar med teoribaserade koder medan en kodning med induktiv ansats laborerar med databaserade

²¹⁹ Se Gustafsson et al (2005), s. 1–17

²²⁰ Emerson (2004), s. 460 ff.

²²¹ Rennstam & Wästerfors (2015). Rennstam & Wästerfors delar in i analysarbetet i tre olika delar: sortering, reducering och argumentering.

²²² Här kan noteras att det råder en viss begreppsförvirring kring vad som menas med kodning, se t.ex. Gibbs (2007), s. 39.

koder. I praktiken tenderar den analytiska ansatsen ofta landa i en abduktiv pendelrörelse, men någonstans måste man börja.²²³ I mitt forskningsprojekt finns en ambition att utgångsläget för pendelrörelsen ska vara den deduktiva positionen, vilket innebär att jag utgår från teoribaserade koder.

En tydlig och i mina ögon intressant skillnad mellan den induktiva och deduktiva ansatsen handlar om *när* under arbetsprocessen som de empiriska resultaten tar form. Med en induktiv ansats genereras empiriska resultat relativt tidigt i den analytiska processen medan den teoretiska påbyggnaden kommer senare i processen. Med en deduktiv ansats är ordningen den omvända. Det inledande teoribygget är tidskrävande och de empiriska resultaten faller ut relativt sent i processen. Den deduktiva ansatsen ställer därför höga krav på forskarens tålamod och tillit till det egna teoribygget.

5.2.3 Kodningen av lärarnas utsagor

Den teoribaserade kodningen av lärarnas utsagor utgår från de fyra forskningsfrågor som ligger till grund för studien. De tre första forskningsfrågorna är baserade på bedömningsprocessens grundläggande delprocesser: konstruktion, genomförande och användning. Dessa grundläggande kategorier bryts sedan ner i underordnade kategorier, baserade på det teoretiska ramverket, och kodas enligt nedanstående tabell.

²²³ Se t.ex. Alvesson & Skoldberg (2008), s. 42.

1. Konstruktionen	
1:1	Uppfattningar om observationsverktygens domänlogik
1:2	Uppfattningar om observationsverktygens formatlogik
1:3	Uppfattningar om värderingsverktygens progressionslogik
1:4	Uppfattningar om värderingsverktygens formatlogik
1:5	Uppfattningar om sammanvägningsverktygens emfaslogik
1:6	Uppfattningar om sammanvägningsverktygens formatlogik
2. Genomförandet	
2:1	Beskrivningar av hur observationen genomfördes
2:2	Beskrivningar av hur värderingen genomfördes
2:3	Beskrivningar av hur sammanvägningen genomfördes
3. Användningen	
3.1	Beskrivningar av hur provet används

Kodningen av det empiriska underlaget till den fjärde forskningsfrågan, den som handlar om hur historielärare uppfattar att externa prov påverkar deras bedömningspraktik, modifierades delvis under analysarbetets gång. Många av lärarna uppfattade att det externa provet hade påverkat den kollegiala interaktionen på olika sätt. Dessa utsagor kom därför, i linje med den abduktiva pendelrörelsen, att utgöra en egen underkategori med en specifik kodning (se nedan).

4. Påverkan	
4.1	Uppfattningar om påverkan på domänlogiken
4.2	Uppfattningar om påverkan på progressionslogiken
4.3	Uppfattningar om påverkan på emfaslogiken
4.4	Uppfattningar om påverkan på observationsverktygens formatlogik
4.5	Uppfattningar om påverkan på värderingsverktygens formatlogik
4.6	Uppfattningar om påverkan på sammanvägningsverktygens formatlogik
4.7	Uppfattningar om påverkan på den kollegiala interaktionen

Analysarbetet tog således sin början i en huvudsakligen deduktivt förankrad kodning av det empiriska underlaget. I det fortsatta analysarbetet positionerades sedan lärarnas utsagor i relation till de bedömnings- och

historiekulturella traditioner som jag har beskrivit i det teoretiska ramverket, det vill säga mät- och tolkningstraditionen respektive MH- och HT-traditionen.

5.2.4 Transparens

Styrande för mitt analysarbete har varit i vilken grad jag lever upp till det uppriktighetskriterium (*sincerity*) som enligt Tracy kan användas för att värdera kvalitativ forskning. I sin redogörelse för hur detta kriterium kan uppfyllas använder hon sig av begreppen självreflexivitet och transparens. Som jag tolkar Tracy handlar kravet på självreflexivitet i första hand om forskarens uppriktighet gällande sin egen roll i forskningsprocessen medan kravet på transparens mer handlar om en uppriktighet gällande forskningsprocessen i sig (själva hantverket).²²⁴

Transparens gällande den sorterande fasen av analysarbetet är ett besvärligt strävansmål eftersom det är svårt att fokusera den abduktiva pendelrörelsens utgångsposition. Lika omöjligt som det är formulera ett empiribaserat utgångsläge frikopplat från teoretisk förförståelse, i linje med Grounded Theory, är det att formulera ett teoribaserat utgångsläge frikopplat från empirin. Det som däremot är möjligt är att tydligt redogöra för de teoribaserade verktyg som man tar med sig in i analysarbetet, det vill säga de kategorier/koder som inte är induktivt ”grundade” i empirin.

Det kan vara klargörande att särskilja mellan transparens i samband med datainsamling och transparens i samband med databehandling. Min uppfattning är att många forskare har lättare att leva upp till kravet på transparens när det gäller datainsamlingen. Anledningen till denna skillnad kan diskuteras, men kanske är det så enkelt som att den senare tankeoperationen är kognitivt mer utmanande än den förra. Eller så finns det en mer pragmatisk dimension som har att göra med att resonemangen gällande datainsamlingen ofta tillkommer relativt tidigt under forskningsprocessen medan resonemangen kring dataanalysen tillkommer senare, under forskningsprocessens slutskede, då tidsbrist är en obehaglig realitet.

²²⁴ Tracy formulerar åtta kriterier som hon menar utmärker god kvalitativ forskning. Att på detta sätt lyfta ut ett av Tracys kriterier är inte helt lätt, eftersom de i många fall glider in i varandra. Transparens är t.ex. en viktig del av kriteriet ”rich rigor”. Det här kriteriet handlar om ”the care and practice of data collection and analysis procedures” och Tracey förtydligar att ”rigorous analysis is marked by transparency regarding the process of sorting, choosing and organizing the data”. Tracy (2010), s. 842.

5.2.5 Den komparativa ansatsen

Fördelen på ett generellt plan med en komparativ ansats är att kontraster framträder, vilket möjliggör nya perspektiv och en ökad förståelse, men samtidigt måste man vara medveten om de fallgropar som den komparativa ansatsen för med sig. Grant behandlar dessa svårigheter i samband med externa prov. Han menar dels att externa prov är ett vitt begrepp som omfattar många olika slags prov, dels att de kontextuella omständigheterna i olika studier försvårar en jämförande ansats.²²⁵

Föreliggande studie har en uttalad temporal komparativ ansats. HLP-lärarna intervjuades 2009 och KP-lärarna 2016, vilket har möjliggjort en undersökning av förändring över tid. Komparationen är dock problematisk, framför allt i två avseenden. Det första problemet handlar om urvalet av lärare. Initialt var tanken att jag i den uppföljande intervjuerien 2016 på nytt skulle vända mig till de HLP-lärare som hade intervjuats i licentiatuppsatsen. Det visade sig dock att dessa lärare inte hade använt sig av KP i sin bedömningspraktik och jag behövde därför göra nya överväganden gällande urvalet av KP-lärare. Det andra problemet är att jag i den uppföljande intervjuerien utgick från en delvis omarbetad intervjuguide. I linje med Marianne Sjölands resonemang kring metodologiska dilemman i samband den komparativa ansatsen behöver jag därför beakta ”problem kring eventuell asymmetri mellan objekten som kan leda till att den ena enheten utgör ett slags måttstock på den andra”.²²⁶ På samma sätt som Sjöland ställs jag inför denna problematik eftersom min licentiatuppsats kan utgöra en sådan måttstock i relation till avhandlingen. Jag delar Sjölands uppfattning att problematiken delvis kan hanteras genom att eftersträva metodologisk transparens.²²⁷

5.2.6 Generaliserbarhet

Som jag skrev inledningsvis kan relationen mellan mikro- och makrostudier kopplas till begreppet generaliserbarhet. Med generaliserbarhet avses i metodologiska sammanhang huruvida analysresultaten är giltiga i andra sammanhang än i den specifika undersökningen. Begreppet har utgjort en

²²⁵ Grant (2003), s. 130.

²²⁶ Sjöland (2016), s. 46.

²²⁷ För ett mer elaborerat resonemang kring den komparativa ansatsen i en historiekulturellt förankrad studie, se Sjöland (2016), s. 45–47.

klassisk akilleshäl för studier baserade på kvalitativ metod. Ingrid Carlgren lyfter frågan om generaliserbarhet genom att belysa motsättningen mellan å ena sidan forskningsansatser med en dragning åt det generella, universella och abstrakta och å andra sidan forskningsansatser med dragning åt det specifika, partikulära och konkreta. I de förra genereras kunskaper genom generalisering och abstraktion och i de senare genom specificering och konkretion. Hon menar att syftet med den senare ”partikularistiska” ansatsen ofta är att utgöra ett första steg som sedan kan och bör följas av andra och större studier med högre generaliseringsambitioner. Carlgren menar vidare att en partikularistisk ansats är särskilt väl lämpad för forskning kopplad till professionella verksamheter, till exempel lärarprofessionen.

Jag delar Carlgrens uppfattning vilket innebär att jag i min forskning inte eftersträvar kontextfria generaliseringar, istället måste läsarna (och lärarna) få stöd i att göra sina egna försök att se hur resultaten kan överföras till sin egen kontext.²²⁸ Detta resonemang ligger i linje med Thomsson som menar att forskare med en kvalitativ ansats bör eftersträva en resonerande, argumenterande och teoretiserande forskning som gör generaliseringen möjlig att anta.²²⁹

²²⁸ Carlgren (2011), s. 73.

²²⁹ Thomsson (2002), s. 33. För ett mer elaborerat resonemang kring generaliserbarhet i en historiekulturellt förankrad studie, se Rudnert (2019), s. 97–100.

6. KONSTRUKTIONEN

I det här inledande empiriska kapitlet analyseras historielärarnas uppfattningar om de frivilliga externa provens **konstruktion**. Kapitlet innehåller två resultatredovisande delar. I den första delen redovisas HLP-resultaten och i den andra delen KP-resultaten. För att kunna sätta in historielärarnas uppfattningar om konstruktionen i ett historie- och bedömningskulturellt sammanhang inleds varje resultatredovisande del med en kortfattad beskrivning av den kursplan som respektive prov skulle förhålla sig till. Därefter positioneras provens bedömningsverktyg i relation till de bedömnings- och historiekulturella traditioner som har redovisats tidigare. Den bedömningskulturella positioneringen görs genom att analytiskt fokus riktas mot bedömningsverktygens format och den historiekulturella positioneringen görs genom att analytiskt fokus riktas mot det konstrukt som ligger till grund för dessa bedömningsverktyg. Genom att sedan relatera *positioneringen av proven* till historielärarnas *uppfattningar om proven* framträder såväl bedömningskulturella som historiekulturella diskrepanser. I kapitlets avslutande resultatdiskuterande del analyseras dessa diskrepanser i ljuset av den fjärde kvadrantens dilemma.

6.1. HLP

6.1.1 Kursplanen 1994/2000

HLP konstruerades för att användas som ett slutprov för A-kursen i historia. Kursplanen för denna kurs formulerades initialt 1994 och genomgick sedan en mindre revidering sex år senare. Det finns en hel del tidigare forskning som på olika sätt har haft ambitionen att positionera domänlogiken i 1994/2000 års kursplan. Även om det inte råder någon fullständig konsensus inom fältet är den samlade bilden att kursplanens domänlogik har en tydlig HT-traditionell förankring.²³⁰ Vad gäller kursplanens pro-

²³⁰ Se till exempel Rosenlund (2016) som i sin analys av 1994/2000 års kursplan beskriver den som övergången från en mer innehållsbaserad tradition till en kontinental HT-tradition. Se även Eliasson (2009); Persson

gressionslogik är forskningsläget inte lika omfattande men även här tenderar tidigare forskning att positionera logiken som HT-traditionell.²³¹ Kursplanen saknar en explicit formulerad emfaslogik och det är en logik som inte har uppmärksammats i tidigare forskning.²³² I vilken utsträckning HLP:s bedömningsverktyg sedan är valida eller inte i relation till kursplanens konstrukt kan diskuteras, men det är en frågeställning som ligger utanför detta forskningsprojekt.

6.1.2 Observationsverktygen

Observationsverktygens formatlogik

HLP konstruerades och genomfördes under en tioårsperiod mellan åren 2003-2013. Provet bestod av ett antal delprov och varje delprov innehöll i sin tur antingen en större uppgift eller ett antal kortare deluppgifter. Antalet delprov förändrades dock över tid. Under de fyra första åren bestod provet av fyra delprov (A, B, C och D) och under de sex avslutande åren bestod provet av tre delprov (A, B och C).²³³ I tabellen nedan redovisas uppgifternas format i respektive delprov under de fyra första åren.

(2018). Åström Elmersjö (2017) argumenterar dock för att övergången från MH-traditionen till HT-traditionen kan spåras i tidigare styrdokument.

²³¹ Eliasson konstaterar att progressionslogiken i 1994 års betygskriterier var tvådelad. Logiken byggde på ett görande (ett verb) som sedan hanterades dikotomt (gör/gör icke verbet). Eliasson (2014), s. 279. Rosenlund uppmärksammar att progressionslogiken i kursplanen bar prägel av en splittrad ad hoc-lösning eftersom kriterierna för de lägre och högre betygen bygger på olika kunskapssyner. De förra på en mer objektivistisk kunskapssyn och de senare på en mer konstruktivistisk kunskapssyn. Han menar att detta är problematiskt, men i linje med de brittiska progressionsmodellerna. (Rosenlund (2019).

²³² Rosenlund (2016) uppmärksammar i första hand emfaslogiken i historielärarnas egen bedömningspraktik.

²³³ Under det första året bestod HLP av fem delprov. Sekvenseringen av delproven varierade över tid, vilket innebär att bokstavsbenämningen på respektive delprov varierade. För att tydliggöra framställningen utgår jag från benämningen i 2009 års prov, se webbplatsen <http://www.historieprov.se/prov%202009.pdf>

Tabell 1: Uppgifternas format under åren 2003-2006

Delprov A	A-delens uppgifter var till formatet mättraditionella. Delprovet innehöll en eller två slutna uppgifter med fokus på referenskunskaper där eleven skulle placera ett antal bilder på rätt plats på en tidslinje.
Delprov B	B-delens uppgifter var till formatet tolkningstraditionella. Delprovet innehöll 10-15 öppna essäuppgifter varav tre skulle besvaras. ²³⁴ Ungefär hälften av dessa uppgifter var traditionella rubrikuppgifter med fokus på referenskunskaper medan den andra hälften var dokumentbaserade uppgifter med fokus på källhantering.
Delprov C	C-delens uppgift var till formatet tolkningstraditionell. Delprovet innehöll en omfattande öppen dokumentbaserad essäuppgift med fokus på källhantering. Till skillnad från B-delens dokumentbaserade uppgifter behövde eleven hantera flera olika källor i denna uppgift.
Delprov D	D-delens uppgift var till formatet tolkningstraditionell. Delprovet innehöll en omfattande öppen dokumentbaserad uppgift med fokus på temporal orientering.

Som framgår av tabellen innehöll HLP under de fyra första åren en blandning av slutna och öppna uppgiftsformat, dock med en tydlig övervikt av det öppna, tolkningstraditionella, formatet. Över tid sker dock två viktiga förändringar som påverkade balansen mellan uppgiftsformaten: 1) A-delens slutna uppgift, den bildbaserade tidslinjen, ersattes med tjugo slutna flervalsuppgifter, varav eleven skulle besvara tio och 2) D-delens öppna essäuppgift togs bort, vilket innebar att antalet delprov minskade från fyra till tre. Båda dessa formatlogiska förändringar innebar att balansen mellan de mättraditionella och tolkningstraditionella uppgiftsformaten i HLP försköts i en mättraditionell riktning.

²³⁴ B-delens uppgifter hämtades från det finländska ämnesrealprovet (se not 3). Ofta återfinns uppgifter i delprov B med exakt samma ordalydelse som i den finländska förlagan.

Observationsverktygens domänlogik

Efter att ha positionerat uppgifternas formatlogik i relation till de bedömningskulturella traditionerna är det nu dags att positionera uppgifternas domänlogik i relation till de historiekulturella traditionerna.

HLP bestod som sagt av flera olika delprov som i sin tur innehöll en eller flera deluppgifter. Under de första åren bestod A-delen av bildbaserade tidslinjer som prövade elevens förmåga att rekonstruera ett givet narrativ genom att identifiera ett antal fenomen och sedan placera dem i rätt kronologisk ordning. Tidslinjerna ersattes därefter av flervalsuppgifter. Nedan ses en sådan flervalsuppgift hämtad från 2009 års A-del.

9. Vad var en kogg?

- a. En eftertraktad exotisk krydda.
- b. En speciell yrkesgrupp i medeltida städer.
- c. Ett handelsfartyg.
- d. Ett bredbladigt svärd.
- e. En färgglad huvudbonad från 1300-talet.

Dessa flervalsuppgifter prövade elevens förmåga att återge detaljerade sakhistoriska kunskaper förankrade i ett antal givna narrativ. För både tidslinjen och flervalsuppgifterna kan således den övergripande tankeoperationen kopplas till narrativ rekonstruktion, vilket ligger i linje med den MH-traditionella domänlogiken. Övergången från tidslinjen till flervalsuppgifterna innebar dock att denna MH-traditionella förankring förstärktes eftersom den sakhistoriska rekonstruktionen i högre utsträckning än bildbaserad kronologiska rekonstruktionen efterfrågar rena minneskunskaper.

Att domänlogiskt positionera B-delens essäuppgifter är inte helt enkelt. I delprovet finns renodlat MH-traditionella uppgifter med fokus på rekonstruktion, men det finns även renodlat HT-traditionella uppgifter med fokus på såväl konstruktion som dekonstruktion. Något som gör positioneringen än mer komplex är att flera av uppgifterna i B-delen kan beskrivas som flerdimensionella, det vill säga att uppgiften prövar fler än en förmåga. Uppgift 2 i 2009 års prov är ett exempel på en sådan flerdimensionell uppgift.

2. Det moderna Europa har övertagit ett omfattande kulturarv från antikens Grekland och Rom. Presentera ett område av den antika kulturen och bedöm dess betydelse.

Uppgiften är tvådelad och i den första delen provas elevens förmåga att beskriva ett område av den antika kulturen. Att deskriptivt rekonstruera en del av ett givet narrativ är en förmåga som är förankrad i den MH-traditionella domänlogiken. I uppgiftens andra del provas elevens förmåga att konstruera ett narrativ genom att etablera ett samband mellan dåtid och nutid. Det är en förmåga som är förankrad i den HT-traditionella domänlogiken. Sammanfattningsvis kan HLP:s B-del beskrivas som ett smörgåsbord där eleven kan välja mellan uppgifter förankrade i såväl den MH-traditionella som den HT-traditionella domänlogiken.²³⁵

C-delens dokumentbaserade essäuppgift är däremot tydligt förankrad i den HT-traditionella domänlogiken. I 2009 års prov hade C-uppgiften följande uppgiftsformulering:

Läs och jämför de fem texterna och fundera över deras trovärdighet. Genomför sedan en skriftlig diskussion om hur de ska bedömas utifrån de källkritiska begreppen – samtidighet, beroende och tendens – och ta sedan ställning till vad som hände Wilhelm Jönsson på Mårtensafton 1926.

I uppgiften provas dels förmågan att dekonstruera narrativen i ett antal källor genom att tillämpa källkritisk metod, dels förmågan att konstruera ett narrativ genom att formulera en slutsats med stöd av dessa källor. Båda dessa förmågor är förankrade i den HT-traditionella domänlogiken.

Även D-uppgiften, det vill säga den uppgift som togs bort, är tydligt förankrad i HT-traditionen. Nedanstående D-uppgift är hämtad från 2006 års prov.

²³⁵ Att HLP:s B-del utformades på detta sätt ska ses i ljuset av det saknades tydliga anvisningar för det centrala ämnesinnehållet i 1994/2000 års kursplan.

Obemärkt?

På en hemsida för släktforskare i januari 2006 hittar vi följande svar på en fråga från Maria om vad förkortningen O.B.R. betyder i en kyrkobok.

”Hej Maria

En gissning: O.B.R. betyder obemärkt. Flickan föddes inte på hemorten utan på något annat ställe, där ingen visste vem hon var. Kanske på något ställe där det fanns särskilda hus för ändamålet.

mvh

Rune”

Annonserna nedan kommer från veckotidningen *Husmodern* januari 1945. Sådana annonser fanns i alla veckotidningar på 1940- och 1950-talen. När de försvann vet vi inte eftersom ingen historiker undersökt det. Läs igenom annonserna och besvara frågorna nedan.

Godkänd

Fråga: Vilket problem har kvinnorna som satt in annonserna?

Väl godkänd

Fråga: Vad säger annonserna om unga kvinnors situation i Sverige för ungefär 60 år sedan?

Mycket väl godkänd

Fråga: Hur och varför har situationen för unga kvinnor i Sverige förändrats?

Vad som provas i D-uppgifterna varierar över tid, men genomgående i dessa ofta flerdimensionella uppgifter, är att de i något avseende provar elevens förmåga att hantera temporala samband. I den här uppgiften är det ett tvådelat temporalt samband som provas, det vill säga en jämförelse

I uppgiften provas explicit elevens förmåga att konstruera narrativ genom att etablera samband mellan dåtid, nutid och framtid. Den här förmågan att hantera tre tidsdimensioner i relation till varandra utgör en central delkomponent i den kontinentala HT-traditionens domänlogik.

För att sammanfatta den historiekulturella positioneringen av HLP-uppgifternas domänlogik kan man konstatera att de flesta av provets uppgifter var HT-traditionellt förankrade, men att det även fanns uppgifter med MH-traditionell förankring. Över tid förstärks denna MH-traditionella förankring genom A-delens revidering och D-delens exkludering.

6.1.3 Uppfattningar om observationsverktygen

Uppfattningar om observationsverktygens format

HLP-lärarna uppmärksammar inte uppgifternas format i någon större utsträckning. Henrik resonerar dock kring de flervalsfrågor som fanns med i de senare årens versioner av HLP:

Jag har haft sådana här kryssfrågor länge dels för att det ju skönt att rätta. Sen finns det ju ett absolut rätt svar då och sen tycker jag också att det är bra för dem som har lite läs och skrivsvårigheter. Då behöver de inte skriva utan då bara kryssar de. Så jag har jobbat med det sättet att fråga och man kan ju göra frågorna rätt kluriga om man vill och man kan ha flera svarsalternativ som är rätt och så vidare, så det var inget nytt för mig.²³⁶

Henrik uppfattar att flervalsuppgifterna i HLP ligger i linje med hans eget sätt att konstruera uppgifter och han lyfter fram flera fördelar med detta slutna uppgiftsformat. En av dessa fördelar är att det finns ”ett absolut rätt svar”, men samtidigt lyfter han även fram möjligheten att ”göra frågorna rätt kluriga om man vill”. Att konstruera flervalsuppgifter med ”flera svarsalternativ som är rätt” skulle kunna tolkas som ett sätt för Henrik att hantera den fjärde kvadrantens dilemma, det vill säga att konstruera slutna uppgifter förankrade i en HT-traditionell domänlogik.

²³⁶ Intervju med Henrik, 2009-11-26

Uppfattningar om observationsverktygens domänlogik

HLP-lärarna uppmärksammar uppgifternas domänlogik i högre utsträckning än vad de uppmärksammar uppgifternas format. Det är framför allt domänlogiken i C- och D-uppgifterna som uppmärksammas, det vill säga de två uppgifter i HLP vars domänlogik har den tydligaste HT-traditionella förankringen. Den källkritiska C-uppgiften uppfattas av HLP-lärarna som en självklar del av historieämnets domänlogik. Håkan uttrycker sig på följande sätt:

Den sista tredjedelen är ju den källkritiska och den har vi ju inte tummat på i och med att det är ju ... för många är det ju ämnets kärna ... och det som oftast ... Alltså om man läser i kriterierna vad som ska ingå så är ju källkritik en given sak. Det är väl förintelsen och källkritik som är oavvisligt och då kan man ju inte ta bort den och hävda liksom att det här är ett viktigt prov.²³⁷

Till skillnad från den källkritiska C-uppgiften uttrycker HLP-lärarna sig i kritiska ordalag om D-uppgiften, den uppgift som exkluderades. Över huvud taget råder det en viss oklarhet om vilken förmåga uppgiften avser att pröva. Henrik uttrycker sig på följande sätt om 2006 års D-uppgift:

Alltså, där är ju en fråga som nu lite grann har försvunnit. Jag vet inte riktigt vad de har kallat den för, kanske den allmänna frågan. Vi hade ett år om en liten annons som talade om att man sökte en fosterfamilj, lite så kryptiskt. Man skulle nog komma dit som utsatt för lönskaläge och så skulle man vara där några månader och föda och sen skulle man återvända till familjen såsom om man hade varit på resa. Det var från 1940-talet tror jag så det var ju inte jättelänge sedan, det kunde ha varit deras mormor. Så skulle de liksom förklara det här. Vad är detta för någonting? Och det var väldigt få som klarade det, men i bedömningsmatrikeln så kunde man ändå få poäng om man kunde liksom ringa in tidsandan. Vad är detta för någonting som händer liksom, utan att man kunde förstå steg för steg vad som blev följden av den här annonsen.²³⁸

²³⁷ Intervju med Håkan, 2010-01-19

²³⁸ Intervju med Henrik, 2009-11-26

Efter att ha beskrivit uppgiften i allmänna ordalag uttrycker Henrik en osäkerhet om uppgiftens domänlogik. Eleverna skulle "liksom förklara den här. Vad är detta?" Han konstaterar vidare att eleverna kunde få poäng om de kunde "liksom ringa in tidsandan". Henrik konstaterar i sitt resonemang att det var väldigt få elever som klarade uppgiften. Det är i sammanhanget intressant att notera hur Henrik, som sitter med i konstruktionsgruppen för HLP, distanserar sig från uppgiften genom sitt val av pronomen när han kommenterar hur den ska benämnas: "Jag vet inte riktigt vad de har kallat den för, kanske den allmänna frågan".

Helena, som tydligt kopplar D-uppgiften till begreppet historiemedvetande, är den av HLP-lärarna som uttrycker starkast kritik mot uppgiften:

Helena: Det där historiemedvetandet tyckte jag var ett rent intelligenstest.

Jag: Det där tycker jag är lite intressant. Kan du utveckla den tankegången?

Helena: Ja, du ska i princip läsa någonting och så ska du förstå vad du läser och sedan dra lite slutsatser. [...] Du ska liksom vara lite smart och kunna tänka självständigt.²³⁹

När Helena lite längre fram under intervjun för ett resonemang där hon ställer uppgifterna i de olika delproven mot varandra förtydligar hon sin uppfattning att D-uppgiften intar en domänlogisk särställning i negativ bemärkelse.

Efter varje prov så säger vi att det här skulle inte vara med och det här skulle inte vara med och så där. Jag var egentligen nästan mest nöjd med det första provet, när vi tog de finska frågorna rätt av liksom. En del av de historieuppfattnings ... historiemedvetande och sånt har jag mer upplevt som någon slags intelligenstest. Om man tänker att de går i ettan ... jag vet inte om jag har tyckt att de hörde hemma på A-kursen. Nu har vi tagit bort det. Och ibland så har man tyckt att källkritiken har varit för svår svenska ... ur språket ... föråldrat språk och så vidare. Och sist så hade jag nån synpunkt på kryssfrågorna också.²⁴⁰

²³⁹ Intervju med Helena, 2009-11-25

²⁴⁰ Intervju med Helena, 2009-11-25

Helena beskriver att hon är mest nöjd med uppgifterna i B-delen, eller de "finska frågorna" som hon benämner dem. A-delens flervalsuppgifter och C-delens källkritiska uppgifter ifrågasätts i viss mån, men inte på domänlogiska grunder. Det är endast D-uppgifterna, de uppgifter som enligt Helena prövar "historieuppfattning ... historiemedvetande och sånt" som hon ifrågasätter domänlogiskt, det vill säga om den verkligen utgör en del av A-kursen: "jag vet inte om jag har tyckt att de hörde hemma på A-kursen".

Den här skillnaden i hur HLP-lärarna uppfattar de två stora HT-förankrade uppgifterna i HLP indikerar att begreppet "källkritik" tycks vara ett väl etablerat begrepp och en självklar del av HLP-lärarnas domänlogik medan begreppet "historiemedvetande", som förvisso har funnits med som ett begrepp i kursplanen sedan 1994, inte tycks ha fått något större genomslag i deras domänlogik.²⁴¹ Håkan kopplar begreppet till A-delens bildbaserade tidslinje och i sitt resonemang antyder han en osäkerhet kring vad som menas med historiemedvetande.

Jag: Du var ju inne på det här begreppet historiemedvetande förut.

Tycker du att provet har haft ambitionen att mäta det?

Håkan: Jag tänker lite den här tidslinjen ... att där är det ändå ... alltså att ha någon relation till tiden. Om det är historiemedvetande ... jag vet inte riktigt vad historiemedvetande är men ...

Jag: Det är ingen diskussion ni har tagit i gruppen?

Håkan: Alltså han, den gode Eliasson, nämner väl historiemedvetande ibland när han dyker upp ... han är inte med alltid.²⁴²

Håkans resonemang kring begreppet förstärker bilden av att det finns en diskrepans mellan HLP-lärarnas domänlogik och den domänlogik som manifesteras i HLP. Den här diskrepansen kommer att behandlas mer utförligt i kapitlets slutdiskussion.

²⁴¹ Gunnemyr (2010), s. 235–236.

²⁴² Intervju med Håkan, 2010-01-19

6.1.4 Värderingsverktygen

Värderingsverktygens formatlogik

Inledningsvis är det viktigt att klargöra att HLP var ett kriterierelaterat prov, vilket innebar att elevprestationerna skulle relateras till i förväg fastställda kunskapskriterier. Dessa kriterier fanns angivna i ett värderingsverktyg som benämndes ”Bedömningsgrunder”.²⁴³ Eftersom HLP bestod av ett antal delprov med olika slags uppgifter innehöll bedömningsgrunderna flera olika slags värderingsverktyg. I tabellen nedan redovisas bedömningsgrundernas format för respektive delprov under de fyra första åren.

²⁴³ Bedömningsgrunderna konstruerades av konstruktionsgruppen och skickades ut samtidigt som provet.

Tabell 2: Bedömningsgrundernas format under perioden 2003-2006

Delprov A	Bedömningsgrunderna för A-delens bildbaserade tidslinjer bestod av ett dikotomt facit. Lärarens tolkningsutrymme var således reducerat till noll, vilket ligger i linje med mättraditionen.
Delprov B	Bedömningsgrunderna för B-delens uppgifter bestod av poängbaserade matriser till respektive uppgift. I matrisen angavs kriterier för poängintervallen 1-4 respektive 5-6. Vad som särskilde kvaliteten inom respektive poängintervall preciserades inte och det fanns inte heller några elevexempel eller värderingskommentarer i anslutning till matrisen. ²⁴⁴ Allt sammantaget gavs läraren ett förhållandevis stort tolkningsutrymme, vilket ligger i linje med tolkningstraditionen.
Delprov C	Bedömningsgrunderna för C-delens uppgift bestod av en betygsbaserad matris samt ett svarsexempel för de olika betygsnivåerna. I matrisen angavs kriterier för betygsnivåerna Godkänd, Väl godkänd och Mycket väl godkänd. Lärarens tolkningsutrymme var således förhållandevis begränsat vilket ligger i linje med mättraditionen.
Delprov D	Bedömningsgrunderna för D-delen bestod av samma delar som för C-delen, det vill säga en betygsbaserad matris kompletterad med svarsexempel för de tre betygsnivåerna. Även D-delens bedömningsgrunder var således till formatet mättraditionellt förankrade.

Som framgår av tabellen var HLP:s bedömningsgrunder i huvudsak mättraditionellt förankrade. Den här mättraditionella förankringen förändrades inte över tid. Varken revideringen av A-delen eller D-delens exkludering innebar att lärarens tolkningsutrymme begränsades ytterligare.

²⁴⁴ Matriserna i B-delen är i de flesta fall identiska med matriserna i de finländska ämnesrealproven.

Värderingsverktygens progressionslogik

Efter att ha positionerat bedömningsgrundernas format i relation till de bedömningskulturella traditionerna är det nu dags att positionera deras progressionslogik i relation till de historiekulturella traditionerna.

I A-delens facit avgörs elevprestationens kvalitet av huruvida svaret är korrekt eller inte. Svarsalternativen i flervalsfrågorna är antingen sanna eller falska, vilket ligger i linje med den objektivistiska kunskapssyn som genomsyrar den MH-traditionella progressionslogiken.²⁴⁵

Progressionslogiken i B-delens matriser är däremot HT-traditionellt förankrad. I exemplet nedan ses matrisen till uppgift 2 i 2009 års B-del.

2. Det moderna Europa har övertagit ett omfattande kulturarv från antikens Grekland och Rom. Presentera ett område av den antika kulturen och bedöm dess betydelse.

1-4: Uppgiften förutsätter att skribenten har kännedom om den antika kulturen och beskriver något enskilt kulturområde med exempel och hänvisar till senare tider då fenomenet förekommit.

5-6: Fördjupad behandling av ifrågavarande kulturområde och dito om antiken och övriga tidsperioden som behandlas.

Progressionsindikatorerna som ligger till grund för minimikravet, det vill säga skillnaden mellan 0 och 1 poäng, kan kategoriseras som kvalitativa. Antalet exempel som krävs för att beskriva det enskilda kulturområdet kvantifieras till exempel inte. Inte heller preciseras kraven för de samband som eleven ska etablera mellan antiken och ”senare tider”. Progressionsindikatorerna som ligger till grund för optimumnormen, det vill säga skillnaden mellan 4 och 5 poäng kan också kategoriseras som kvalitativa. Skillnaden mellan de två poängintervallen bygger på kvalitativa värdeord och handlar om huruvida svaret ger uttryck för en ”fördjupad” behandling eller inte. Vad som menas med ”fördjupad behandling” preciseras inte. Detta sammantaget innebär att progressionslogiken i B-delens matris baseras på kvalitativa indikatorer, vilket ligger i linje med den HT-traditionella progressionslogiken.

²⁴⁵ Jämför med VanSledrights (2014) viktade svarsalternativ och efterföljande förhandling.

Progressionslogiken i C-delens matris är däremot i högre grad mättraditionellt förankrad. I exemplet nedan ses matrisen till 2009 års C-del.

Läs och jämför de fem texterna och fundera över deras trovärdighet. Genomför sedan en skriftlig diskussion om hur de ska bedömas utifrån de källkritiska begreppen – samtidighet, beroende och tendens – och ta sedan ställning till vad som hände Wilhelm Jönsson på Mårtensafton 1926.

Godkänd

Svarsexempel: /.../

Kriterier: Kan använda alla tre källkritiska begrepp på någon källa och vet vad de innebär.

Väl godkänd

Svarsexempel: /.../

Kriterier: Kan diskutera kring alla källor med relevant användning av källkritiska begrepp.

Mycket väl godkänd:

Svarsexempel: /.../

Kriterier: Diskuterar kring alla källor utifrån alla tre begreppen. Identifierar skillnader och likheter i källorna. Tar ställning på ett korrekt sätt.

Progressionsindikatorerna som ligger till grund för minimikravet kvantifieras dels genom att antalet källkritiska begrepp som ska användas preciseras (*tre*), dels genom att antalet källor som ska hanteras preciseras (*någon källa*). Progressionsindikatorerna för optimumnormerna är delvis kvantifierade. Skillnaden mellan G-nivån och VG-nivån inrymmer en kvantitativ indikator som handlar om att antalet källor har betydelse för svarets kvalitet (*någon källa/alla källor*), men det finns även en additiv indikator eftersom kravet för VG-nivån omfattar, utöver användningen av kriterierna, även någon form av diskussion. Skillnaden mellan VG-nivån och MVG-nivån inrymmer på motsvarande sätt både kvantitativa och additiva progressionsindikatorer. Den kvantitativa indikatorn kommer till

uttryck genom att antalet källkritiska begrepp i relation till respektive källa preciseras (*alla källor utifrån alla tre begreppen*). Den additiva indikatorn kommer till uttryck i kravet på att svaret ska inrymma en tydlig jämförelse *samt* ett tydligt ställningstagande. Detta sammantaget innebär att C-delens progressionslogik förvisso är splittrad men att den i högre utsträckning än B-delens bygger på kvantitativa indikatorer, vilket ligger i linje med den MH-traditionella progressionslogiken.

Progressionslogiken i D-delens matriser är däremot renodlat HT-traditionell. I exemplet nedan ses matrisen för 2007 års D-del.

Godkänd

Studera tabellen över bränsleproduktionen från början av 1800-talet fram till 1990. Vilka är de största förändringarna och vad beror de på?

Svarsexempel

/.../

Bedömningskriterier

Jämföra diakront - grundläggande kunskaper om industrialisering, befolkningsökning och transportrevolution

Väl godkänd

Vad har inte förändrats så mycket och vad beror det på? Vilka slutsatser kan du dra av det?

Svarsexempel

/.../

Bedömningskriterier

Jämföra diakront och synkront - grundläggande kunskap om energianvändning i u-världen

Mycket väl godkänd

Studera artikeln från Sydsvenskan 7 februari 2007!

Hur ser sambandet ut mellan de åtgärder som diskuteras för framtiden i Malmö och de förändringar som tabellen till föregående frågor visar? Beskriv en alternativ framtid!

Svarsexempel

/.../

Bedömningskriterier

Kunna använda historisk kunskap för att förstå dagens och framtidens förändringar.

Analysera olikheter i dagens värld i ett framtidsperspektiv.

Till skillnad från C-delen baseras progressionslogiken i D-delens matris helt på kvalitativa och additiva indikatorer. Minimikravet kopplas till deluppgift 1 och innebär att eleven i sitt svar ska konstruera ett narrativ ge-

nom att etablera ett samband mellan dåtid och nutid. I minimikravet inryms en kvalitativ indikator i form av ett värdeord (*grundläggande*). Optimumnormerna, som kopplas till deluppgifterna 2 och 3, är i första hand additiva. Skillnaden mellan G- och VG-nivån är att elevprestationen inrymmer *såväl en diakron som en synkron jämförelse*. Skillnaden mellan VG-nivån och MVG-nivån innebär att elevsvaret även ska inrymma ett temporalt resonemang med fokus på framtiden. Detta sammantaget innebär att D-delens progressionslogik kan kategoriseras som HT-traditionellt förankrad.

Sammanfattningsvis vad gäller progressionslogiken i HLP:s bedömningsgrunder framträder en splittrad bild. A-delens facit är MH-traditionellt förankrat medan B- och D-delens matriser är HT-traditionellt förankrade. C-delens matriser är svåra att positionera eftersom de inrymmer såväl kvantitativa som additiva progressionsindikatorer. D-delens exkludering innebar således (indirekt) att provets progressionslogiska förankring försköts i en MH-traditionell riktning.

6.1.5 Uppfattningar om värderingsverktygen

Värderingsverktygens format- och progressionslogik

HLP-lärarna uppmärksammar inte HLP:s bedömningsgrunder i samma utsträckning som de uppmärksammar HLP:s uppgifter, och i den mån de ger uttryck för några uppfattningar om bedömningsgrunderna är dessa inte samstämmiga. Håkan uttrycker sig i positiva ordalag om bedömningsgrunderna:

Inget problem att använda mig av de kriterierna ... helt överens. Alltså de kriterierna att bedöma hur man ska poängsätta och bedöma de här proven är jag helt överens om. Det är inget ... det är inte fel ... det är rätt ... så skulle jag också göra facit, så det är inget problem.²⁴⁶

Huruvida Håkans positiva omdöme avser bedömningsgrundernas formatlogik och/eller progressionslogik framkommer inte. Till skillnad från Håkan är Hanna mer kritisk i sin beskrivning av HLP:s bedömningsgrunder:

²⁴⁶ Intervju med Håkan, 2010-01-19

Betygskriterierna har ju inte varit så välformulerade ... eller de har varit svårtolkade tycker jag. Vi har diskuterat det mycket i gruppen också när vi har haft rättningmöte. /... / När vi har diskuterat bedömningen så har det ju varit svårt ibland för att vi har uppfattat att man ska bedöma på olika sätt eller att ... det efterfrågar olika saker. Vi har uppfattat att det efterfrågas olika saker helt enkelt och eleverna har ofta missuppfattat vad som efterfrågas ... och då kan det vara svårt att göra en bedömning. Och sen är det vissa som tycker att här kan man kanske vara snäll och ge godkänt ändå medan någon annan säger ... nä, men det står ju faktiskt inte här. Så det gäller att finslipa det och hitta en tydlighet i instruktionerna till eleverna och i bedömningsmallen.²⁴⁷

Hanna beskriver att det inom kollegiet ofta har funnits olika uppfattningar om hur en elevprestation ska värderas och hon efterfrågar därför en ökad ”tydlighet” i bedömningsgrunderna, vilket ligger i linje med det mättraditionella idealet att begränsa lärarens tolkningsutrymme. Huruvida Hannas önskemål om en ökad tydlighet avser bedömningsgrundernas formatlogik eller progressionslogik kan diskuteras. Om det är formatlogiken som åsyftas så kan tydligheten öka genom flera svarsexempel och/eller bedömningskommentarer. Om det är progressionslogiken som åsyftas så efterfrågar hon däremot en ökad tydlighet i de matriser som finns i bedömningsgrunderna.

6.2.6 Sammanvägningsverktygen

HLP saknade ett explicit sammanvägningsverktyg. Det fanns således inte några riktlinjer för hur prestationsomdömena på respektive delprov skulle vägas samman till ett samlat prestationsomdöme. Frånvaron av ett sammanvägningsverktyg och därmed någon form av emfaslogik ger HLP-lärarna ett stort tolkningsutrymme, vilket ligger i linje med tolkningstraditionen. Det är den enskilde läraren som avgör om något delprovresultat väger tyngre än de andra.

²⁴⁷ Intervju med Hanna, 2009-11-19

6.1.7 Uppfattningar om sammanvägningsverktygen

HLP-lärarna uppmärksammar inte frånvaron av ett sammanvägningsverktyg och därmed frånvaron av en emfaslogik. Att HLP-lärarna inte efterfrågar ett sådant verktyg indikerar att de uppfattar prestationsomdömena på de enskilda delproven som meningsfulla i sig själva. Vid betygs-sättningen kan de till exempel använda sig av C-delens prestationsomdöme som ett underlag för att värdera elevens förmåga att hantera källor. Det är även möjligt att tolka frånvaron som en indikation på att de vill tona ner HLP:s summativa syfte. En sammanvägning av de enskilda prestationsomdömena till ett samlat prestationsomdöme i form av ett provbetyg skulle underlätta en summativ jämförelse på såväl elevnivå som på lärar- och skolnivå. Det här resonemanget om hur HLP-lärarna använder sig av resultaten kommer att fördjupas i kapitel 8.

6.1.8 HLP – en sammanfattning

Observationsverktygen

HLP:s uppgifter innehöll under de fyra första åren en blandning av uppgifter med såväl slutna som öppna format, dock med en tydlig övervikt av det öppna, tolkningstraditionella, formatet. Domänlogiskt var dessa uppgifter huvudsakligen HT-traditionellt förankrade, men det fanns även uppgifter med MH-traditionell förankring. Över tid innebar dock revideringen av A-delen och exkluderingen av D-delen att formatlogiken försköts i en mer mättraditionell riktning och att domänlogiken försköts i en mer MH-traditionell riktning.

HLP-lärarna uppmärksammar uppgifternas formatlogik i begränsad omfattning medan däremot uppgifternas domänlogik uppmärksammas i hög omfattning. Det är framför allt domänlogiken i C- och D-uppgifterna som uppmärksammas och i deras kritik av D-uppgiften framkommer att det finns en domänlogisk diskrepans gällande den förmåga som prövas i denna uppgift. Lite förenklat handlar diskrepansen om huruvida begreppet historiemedvetande är en del av domänlogiken eller inte.

Värderingsverktygen

HLP:s bedömningsgrunder var till formatet i huvudsak mättraditionellt förankrade. Progressionslogiskt framträder dock en splittrad bild eftersom A-delens facit och delvis C-delens matris var MH-traditionellt förankrade, medan B- och D-delens matriser var HT-traditionellt förankrade.

D-delens exkludering innebar att provets progressionslogiska förankring försköts i en MH-traditionell riktning.

HLP-lärarna uppmärksammar bedömningsgrunderna i begränsad omfattning och det gäller såväl formatlogiken som progressionslogiken. Den domänlogiska diskrepans som framträder i deras uppfattningar om HLP:s uppgifter motsvaras således inte av någon motsvarande progressionslogisk diskrepans gällande HLP:s bedömningsgrunder.

Sammanvägningsverktygen

HLP saknade ett explicit sammanvägningsverktyg och därmed riktlinjer för provets emfaslogik. HLP-lärarna uppmärksammar inte frånvaron av ett sådant verktyg och de efterfrågar inte heller riktlinjer för provets emfaslogik.

6.2 KP

6.2.1 2011 års kursplan

KP konstruerades för att kunna användas som ett slutprov för kursen Historia 1b. Domänlogiken som ligger till grund för denna kurs skrivs fram i 2011 års kursplan. Tidigare forskning tenderar att positionera domänlogiken i denna kursplan som HT-traditionell.²⁴⁸ Vad gäller kursplanens progressionslogik har den uppmärksamrats i begränsad uppfattning av tidigare forskning, men även den beskrivs som HT-traditionell.²⁴⁹ Kursplanens emfaslogik gällande den substantiella dimensionen av domänlogiken är inte särskild framträdande. I anvisningarna gällande det centrala innehållet finns dock ett antal ”ska”-satser som markerar vilka ämnesområden som väger tyngst. Vad gäller den procedurala dimensionen framkommer att de olika förmågorna ska väga lika tungt. Någon tidigare forskning kring emfaslogiken i 2011 års kursplan har jag inte hittat. I vilken utsträckning KP:s bedömningsverktyg är valida eller inte i

²⁴⁸ Den övergripande frågeställningen inom fältet gällande 2011 års kursplan har i huvudsak handlat om balansen mellan den kontinentala respektive anglosaxiska tolkningen av HT-traditionen. Jarhall (2020) skriver i sin forskningsöversikt att 2011 års kursplan för grundskolan kan ses som ett närmande mellan traditionerna. Rosenlund (2016) menar att det sker en förskjutning i jämförelse med 1994/2000 års kursplan för gymnasiet i den bemärkelsen att 2011 års kursplan i högre utsträckning betonar historieämnets konstruktivism (historia som konstruktion). Anders Person skriver i sin avhandling att 2011 års styrdokument har fokus på metaperspektiven snarare än historia per se. Persson (2017), s.136.

²⁴⁹ Eliasson argumenterar för att verben i 1994 års styrdokument har ersatts av adverb (2014), s. 279. Rosenlund (2019) uppmärksammar att de kunskapskrav som formuleras i 2011 års kursplan vilar tungt på bruket av värdeord. Han argumenterar sedan för att dessa värdeord inte är förankrade i den historiedidaktiska forskningen.

relation till denna kursplan är som sagt en frågeställning som ligger utanför detta forskningsprojekt.²⁵⁰

6.2.2 Observationsverktygen

Observationsverktygens formatlogik

KP konstruerades och genomfördes under åren 2015–2018. Provet bestod under hela perioden av två delprov. I tabellen nedan redovisas uppgiftsformaten i respektive delprov under det första året.

Tabell 3. Uppgifternas format 2015

Delprov 1	Delprov 1 innehöll det första året nio uppgifter: fem öppna essäuppgifter, en öppen kortsvarsuppgift och tre slutna flervalsuppgifter. De flesta av uppgifterna var således till formatet tolkningstraditionellt förankrade, men det fanns även uppgifter med en mättraditionell förankring.
Delprov 2	Delprov 2 innehöll två uppgifter med fokus på källhantering: en vanlig dokumentbaserad essäuppgift och en längre mer krävande essäuppgift som kan liknas vid ett dokumentbaserat case. Båda uppgifterna var således till formatet tolkningstraditionellt förankrade.

Som framgår av tabellen var KP:s uppgifter till formatet i huvudsak tolkningstraditionellt förankrade. Uppgiftsformaten i delprov 2 förändrades inte över tid, men i delprov 1 skedde två tydliga förändringar från 2015 till 2016, förändringar som sedan blev bestående under de tre därpå följande åren. Förändringarna innebar dels att det totala antalet provuppgifter blev något färre, dels att de slutna flervalsuppgifterna helt togs bort. Över tid sker således en formatförskjutning i tolkningstraditionell riktning

Flervalsfrågornas exkludering är en viktig skillnad mellan HLP och KP. I HLP fanns det initialt inte några flervalsfrågor, men de introduceras efter hand, medan det i KP initialt fanns ett antal flervalsfrågor som sedan

²⁵⁰ Ammert & Eliasson (2019) har granskat KP:s validitet i relation till styrdokumentet. Deras uppfattning är att provets konstrukt ligger i linje med styrdokumentets konstrukt.

lyftes bort. Den här formatförskjutningen i motsatta riktningar är något som jag kommer att återkomma till i det här kapitlets slutdiskussion.

Observationsverktygens domänlogik

Efter att ha positionerat uppgifternas formatlogik i relation till de bedömningskulturella traditionerna är det nu dags att positionera uppgifternas domänlogik i relation till de historiekulturella traditionerna. Som framgått ovan förändrades KP:s formatlogik över tid, men vad gäller uppgifternas domänlogik framträder en mer statisk bild. Under de fyra år som KP konstruerades prövades samtliga förmågor som skrivs fram i kursplanen, det vill säga referenskunskaper, temporal orientering, källhantering, begreppshantering och historieanvändning.

I delprov 1, som efter 2015 bestod av åtta uppgifter, prövades samtliga förmågor utom källhantering.²⁵¹ De enskilda uppgifterna var till övervägande del endimensionella, det vill säga de prövade endast en förmåga, men det fanns även uppgifter som var tvådimensionella. Uppgift 3 från 2016 års delprov 1 är ett exempel på en sådan tvådimensionell uppgift:

²⁵¹ I det första provet från 2015 prövades samtliga förmågor i delprov 1. I de efterföljande proven utgår dock källhanteringen från delprov 1 och koncentreras till delprov 2.

Uppgift 3. Vilken betydelse har den enskilda personen?

I rutan ser du exempel på personer som kan kopplas till viktiga historiska händelser eller processer. Du ska välja att arbeta med en av dessa personer.

Uppgift

<u>Person</u>	<u>Händelse/process</u>
• Jeanne d'Arc	Frankrikes framgångar under hundraårskriget.
• Gustav II Adolf	Sveriges framgångar under trettioåriga kriget.
• Adolf Hitler	Att nazisterna tog makten i Tyskland 1933.
• Elise Ottesen-Jensen	Att svenska kvinnor skulle få rätt till abort.
• Michail Gorbatjov	Berlinmurens fall.
• Valfri person	En händelse där personen kan ha haft betydelse.

Välj **en** av personerna i rutan.

Här ser du två tolkningar av personens betydelse för händelsen/processen:

- Personen hade stor betydelse för händelsen/processen.
 - Personen hade liten betydelse för händelsen/processen. Det fanns annat som var viktigare.
1. Hitta stöd för båda tolkningarna genom att ge historiska exempel.
 2. Jämför de båda tolkningarna och diskutera deras styrkor och svagheter när det gäller att förklara den historiska händelsen.

Använd begreppen aktör och struktur när du besvarar uppgiften.

I uppgiften provas två olika förmågor. I deluppgift 1 provas elevens förmåga att med stöd av substantiella kunskaper underbygga rekonstruktionen av *två olika tolkningar* av ett narrativ. Den här formen av konstruktivistisk rekonstruktion är en tankeoperation som ligger i linje med den HT-traditionella domänlogiken och den är utmärkande för samtliga referensramsuppgifter i KP efter 2015. I deluppgift 2 provas förmågan att föra ett resonemang *om* historia med stöd av metahistoriska begrepp, vilket även det ligger i linje med den HT-traditionella domänlogiken.

Ovanstående öppna essäuppgift kan med fördel jämföras med en av de slutna uppgifter som fanns med i 2015 års prov.

<p>Den historiska frågan: Varför ändrades Europas karta under 1800-talet och i början av 1900-talet?</p>	1. Välj det begrepp (a eller b) som är rimligast när man besvarar den historiska frågan.	2. Välj det alternativ(I, II eller III) som är det rimligaste svaret på den historiska frågan och som dessutom passar ihop med begreppet du valt.
	<p>a. Konservatismen</p> <p>b. Nationalismen</p>	I. Många ville behålla maktfördelningen och ståndssamhället som det var.
		II. Många ville öka handeln mellan länderna genom att sänka tullavgifterna.
		III. Många ville skapa länder där de flesta hade samma språk eller kände sig som medborgare i samma land.

Uppgiften är på samma sätt som den öppna essäuppgiften tvådelad. Eleven ska inte bara besvara den historiska frågan, det vill säga rekonstruera ett givet narrativ genom att välja korrekt svarsalternativ, utan dessutom välja det historiska begrepp som är relevant i sammanhanget. Uppgiften i sin helhet kräver således något mer utöver rekonstruktionen och kan därför inte kategoriseras som renodlat MH-traditionell. Uppgiften kan inte heller beskrivas som renodlat HT-traditionell eftersom den inte kräver att eleven ska hantera två olika tolkningar av narrativet. Att det slutna uppgiftsformatet föll bort efter 2015 års prov kan tolkas som att KP-konstruktörerna initialt gjorde ett försök att hantera den fjärde kvadrantens di-

lemma, det vill säga konstruera slutna uppgifter förankrade i en HT-traditionell domänlogik, men att de sedan gjorde bedömningen att dessa uppgifter inte var valida i relation till den HT-traditionella domänlogiken.

Om man lämnar referensramskunskaperna och begreppshanteringen fanns det i delprov 1 även uppgifter som gjorde anspråk på att pröva temporal orientering. Nedanstående exempel är hämtat från 2016 års prov.

Uppgift 7. Ett tredje världskrig?

Uppgift

Under 1900-talet utkämpades två världskrig. Använd dina kunskaper om världskrigen och resonera om vad som talar för och vad som talar emot att ett tredje världskrig kommer att inträffa i framtiden.

Jämför tidsperioden för världskrigen under 1900-talet(orsaker/konsekvenser) med situationen i världen idag när du resonerar om framtiden.

Uppgiften kan jämföras med D-uppgiften från 2007 i HLP, som behandlades under analysen av HLP. I båda uppgifterna prövas elevens förmåga att konstruera ett narrativ genom att etablera samband mellan dåtid, nutid och framtid. Den här förmågan att hantera de tre tidsdimensionerna i relation till varandra utgör en central delkomponent i den kontinentala HT-traditionens domänlogik

Slutligen fanns i delprov 1 uppgifter som gjorde anspråk på att pröva elevens förmåga att resonera om historieanvändning. Nedanstående uppgift är hämtad från 2016 års prov.

Uppgift 6. Att använda historia om vikingatiden

Den här affischen användes i Norge av nazisterna för att värva norska soldater. På affischen syns ett vikingaskepp. På affischen står det: "Med Waffen-SS* och den norska legionen mot den gemensamma fienden ... mot bolsjevismen**".

*Waffen-SS var en del av den nazistyska krigsorganisationen.

**Bolsjevismen var ett namn på den sovjetiska kommunismen.

Uppgift

Vilket syfte kan nazisterna ha haft när de valde att använda historia om just vikingatiden för att värva soldater?

Jag tror att nazisternas syfte med att använda just vikingatiden för att värva norska soldater var...

I uppgiften provas elevens förmåga att dra en slutsats om varför ett givet narrativ, i det här fallet historia om vikingatiden, används i ett specifikt syfte. Den här förmågan att resonera om hur historia används är på samma sätt som den tredelade temporala tankeoperationen domänlogiskt förankrad i den kontinentala tolkningen av HT-traditionen.

Om man lämnar delprov 1 och riktar blicken mot delprov 2 kan man konstatera att formatlogiken i detta delprov, till skillnad från delprov 1,

inte förändrades över tid. Delprov 2 innehöll under hela perioden två öppna essäuppgifter med fokus på källhantering. Skillnaden mellan dessa båda uppgifter är att den första är mer begränsad (färre källor) och endast prövade den dekonstruerande aspekten av den källhanterande förmågan, medan den andra uppgiften var mer omfattande (fler källor) och även prövade den konstruerande aspekten av den källhanterande förmågan. Båda uppgifterna i delprov 2 kan således kategoriseras som HT-traditionellt förankrade, men förankringen är särskilt tydligt i den andra, mer omfattande deluppgiften.

Sammanfattningsvis kan KP-uppgifternas domänlogik beskrivas som HT-traditionellt förankrad, vilket kan jämföras med HLP-uppgifternas mera splittrade domänlogiska förankring. Över tid innebar exkluderingen av KP:s slutna uppgifter att KP:s HT-traditionella förankring stärktes ytterligare, vilket kan jämföras med HLP, där inkluderingen av de slutna uppgifterna innebar en förskjutning i MH-traditionell riktning.

6.2.3 Uppfattningar om observationsverktygen

Uppfattningar om observationsverktygens formatlogik

KP-lärarna uppmärksammar uppgifternas format i något större utsträckning än HLP-lärarna. Karl uttrycker sig på följande sätt när han resonerar om uppgiftsformatet i KP:

Sättet dom är uppbyggda på och bildmaterialet eller nånting annat som är kopplat till frågan. Till exempel det var nån fråga nu om varför tyska ockupationsmakten i Norge använde historia om vikingatiden på ett visst sätt och så var det en bild på en viking och en tysk waffen SS-soldat. Jag menar ... att hitta detta är ju enormt tidskrävande ju ... så här ger ju Skolverket oss ett gratismaterial med fantastiskt hög kvalitet.²⁵²

Huruvida det är uppgifternas format- eller domänlogik som Karl avser när han beskriver att "Skolverket ger oss ett gratismaterial med fantastisk hög kvalitet" är oklart, men överlag uttrycker sig Karl och de övriga KP-lärarna i positiva ordalag när de resonerar om uppgifterna i KP. Det går dock inte att bortse från att KP-lärarnas uppfattningar i det avseendet kan

²⁵² Intervju med Karl, 2016-12-06.

påverkas av intervjusituationen, det vill säga det faktum att intervjuaren ingick i den grupp som konstruerade uppgifterna i KP.

Uppfattningar om observationsverktygens domänlogik

På samma sätt som HLP-lärarna uppmärksammar KP-lärarna uppgifternas domänlogik i betydligt högre utsträckning än deras formatlogik. Ett tydligt mönster hos HLP-lärarna var kritiken mot D-uppgiftens domänlogik, det vill säga den HLP-uppgift som hade den tydligaste förankringen i den kontinentala HT-traditionen. Även KP-lärarna uppmärksammar i första hand de KP-uppgifter som är förankrade i den kontinentala HT-traditionen. Särskilt de uppgifter som prövade elevens förmåga att etablera temporala samband mellan dåtid, nutid och framtid tar stort utrymme. Kristina uttrycker sig på följande sätt om framtidsperspektivet i den tredelade temporala tankeoperationen:

Jag tycker det är jätteskumt ... jag har jättesvårt att förlika mig med de här framtidsidéerna. Alltså jag förstår det här med historiemedvetandet, jag är fullständigt med på det / ... / och återigen kan man känna så här, ja men det har vi väl alltid gjort, men några sierskor har vi inte varit så mycket innan. Jag tycker det här är svårt, men jag förstår, för grejen ligger ju inte i att tala om hur det ska bli, utan det ligger i att kunna motivera och se kontinuitet och förändring ... och att kunna se vad är det som har hänt förut, hur blev det då. Jag är fullständigt med på det, men för eleverna tror jag att det är lite knepigt.²⁵³

På samma sätt som HLP-lärarna kopplar Kristina sitt resonemang om uppgiften till begreppet historiemedvetande. Hon betonar att hon förstår tankeoperationen men hon tydliggör samtidigt att hon inte uppfattar att framtidsperspektivet utgör en del av hennes egen domänlogik. Även de övriga KP-lärarna uppmärksammar framtidsperspektivet, men till skillnad från Kristina uppfattar de i högre utsträckning perspektivet som en del av sin egen domänlogik. Karin uttrycker sig på följande sätt:

Det står ju i de centrala målen ... att man ... och det tycker jag är jätteviktigt. Jag brukar alltid försöka få dom att tänka på hur kommer det

²⁵³ Intervju med Kristina, 2016-11-21.

att se ut ... med tanke på hur det har (betoning) sett ut ... om de kan dra nån röd tråd ... genom historien ... så att ... framtiden är ju viktig att ha med tycker jag ... för att få dem historiemedvetna ... och kunna diskutera. /.../ Jag brukar ha en fråga ibland när jag har haft ett vanligt prov på det här 1850-1930 i svensk historia, så har jag haft en fråga hur tror du att det kommer att se ut i Sverige om 50 år baserat på det du har lärt dig nu om hur det såg ut i Sverige under den här tidsperioden ... men det är en svår fråga.²⁵⁴

Karin resonerar kring uppgiftens domänlogik och hon beskriver att framtidsperspektivet som viktigt men svårt. Hon beskriver även att hon konstruerar egna frågor som inbegriper framtidsperspektivet. Kristers resonemang om framtidsperspektivet, som utgår från en specifik KP-uppgift, följer samma mönster:

Där hade vi uppgiften om bland annat tredje världskriget, som jag vet... en av mina elever tänkte, men ska vi prata om framtiden där ... nä, men det handlar ju om dåtiden, men du ska se framtiden an ... ja ja, ok men då förstår jag att man ska se lite på vad det kan bli av världen ... men jag tror en del elever har svårt att se det här, det temporala, att här har vi dået, nu har vi nuet, och sen har vi framtiden, för dom ser gärna att det handlar om dået mest ... och jag som lärare ska jag nog bli lite bättre kanske, det har jag insett, att jobba mer med framtiden, för att vi befinner oss här och nu, men så finns det ett då och så finns det en framtid ... men jag tyckte om den uppgiften om tredje världskriget, den var en viktig uppgift som många elever tyckte var väldigt intressant att skriva om.²⁵⁵

På samma sätt som Karin uppfattar Krister framtidsperspektivet som en viktig men för eleverna komplex tankeoperation och han ser behovet av att arbeta mer med detta. I KP-lärarnas resonemang framkommer dock att det råder delade meningar inom kollegiet huruvida framtidsperspektivet hör hemma inom historieämnets domänlogik. Karl uttrycker sig på följande sätt:

²⁵⁴ Intervju med Karin, 2016-11-08.

²⁵⁵ Intervju med Krister, 2016-11-24.

Det är kunskapskravet vi har pratat mest om ... (Imiterar upprörd kollega) "Vi kan inte säga någonting om framtiden ... hur ska dom kunna göra det?" Du fattar inte hur infekterat detta har varit. /.../ Jag har flera kollegor som tycker att såna här frågor är rena skitet. "Ordet framtiden hör inte hemma i historia" ... jag har flera kollegor som säger så ... rakt ut.²⁵⁶

Det här kollegiala motståndet mot framtidsperspektivet kan jämföras med HLP-lärarna motstånd mot D-uppgiften. Ett sätt att beskriva skillnaden är att det hos HLP-lärarna handlar om en *vertikal* förhandling, det vill säga en förhandling mellan formuleringsarenan och realiseringsarenan, medan det hos KP-lärarna handlar om en *kollegial* förhandling på realiseringsarenan, en förhandling som i förlängningen leder till en kollegial polarisering inom ämneslaget. Den här kollegiala polariseringen kommer jag att återkomma till i kapitel 9.

6.2.4 Värderingsverktygen

Värderingsverktygens format

KP:s värderingsverktyg benämns som "bedömningsanvisningar" och de är generellt sett mer elaborerade jämfört med HLP:s "bedömningsgrunder". I tabellen nedan redovisas bedömningsanvisningarnas format för respektive delprov 2015.

²⁵⁶ Intervju med Karl, 2016-12-06.

Tabell 4: Bedömningsanvisningarnas format 2015

Delprov 1	KP:s bedömningsanvisningar var till formatet mättraditionellt förankrade. För de tre slutna flervalsuppgifterna bestod av anvisningarna ett dikotomt facit, vilket reducerade lärarens tolkningsutrymme till noll. Bedömningsanvisningarna för de öppna uppgifterna var i grunden konstruerade på samma sätt som bedömningsgrunderna för HLP:s C- och D-delar, det vill säga betygsbaserade matriser som kompletterades med autentiska svarsexempel. Den stora skillnaden är att KP:s bedömningsanvisningar även inrymde ”bedömningsaspekter” och ”bedömningskommentarer”. I bedömningsaspekterna angavs ”de aspekter av svaren som ska bedömas” och i bedömningskommentarerna beskrevs ”hur progressionen i matrisen kommer till uttryck i svaret”. Såväl bedömningsaspekterna som bedömningskommentarerna innebar att lärarens tolkningsutrymme i hög grad begränsades, vilket ligger i linje med mättraditionen.
Delprov 2	Bedömningsanvisningarna för de två öppna uppgifterna i delprov 2 var konstruerade på samma sätt som anvisningarna för de öppna uppgifterna i delprov 1, och således mättraditionellt förankrade.

Som framgår av tabellen var KP:s bedömningsanvisningar mättraditionellt förankrade. Efter 2015 lyftes de slutna flervalsuppgifterna bort, vilket innebar att det inte längre fanns något dikotomt facit i KP:s bedömningsanvisningar. Denna förändring innebar dock inte någon betydande fokusförskjutning över tid. Helhetsintrycket, om man jämför KP:s bedömningsanvisningar med HLP:s bedömningsgrunder, är således att båda till formatet var mättraditionellt förankrade, men att denna förankring var tydligare i KP.

Värderingsverktygens progressionslogik

Efter att ha positionerat bedömningsanvisningarnas format i relation till de bedömningskulturella traditionerna är det nu dags att positionera deras progressionslogik i relation till de historiekulturella traditionerna.

Progressionslogiken i KP:s bedömningsanvisningar bygger på den progressionslogik som finns framskriven i kursplanens kunskapskrav för historia 1b. I kunskapskraven finns exempel på såväl kvantitativa som additiva progressionsindikatorer, men den grundläggande logiken är att progressionen skrivs fram genom kvalitativa indikatorer i form av värdeord. Värdeorden varierar mellan förmågorna och inom förmågorna, men den mest frekventa progressionskedjan uttrycks med adverbena *översiktlig – utförlig – utförlig och nyanserad*.²⁵⁷ I exemplet nedan ses matrisen till den tidigare nämnda uppgift 3 i delprov 1 från 2016 års B-del.

Belägg för nivå E: eleven redogör översiktligt	Belägg för nivå C: eleven redogör utförligt och motiverar med enkla omdömen	Belägg för nivå A: eleven redogör utförligt och nyanserat och motiverar med nyanserade omdömen
I svaret hittar eleven stöd för båda tolkningarna genom att ge något rimligt historiskt exempel för varje tolkning.	I svaret hittar eleven stöd för båda tolkningarna genom att ge några rimliga historiska exempel. De historiska exemplen används för att skapa historiska sammanhang. Eleven diskuterar om den ena tolkningen är rimligare än den andra.	Utöver kraven för C De historiska exemplen visar på olika perspektiv, t.ex. ekonomiska, politiska, sociala, långsiktiga/kortsiktiga och nationella/internationella. Eleven väger relevanta styrkor och svagheter hos tolkningarna mot varandra.

I den övre skuggade delen av matrisen redovisas de två progressionskedjor som finns angivna i kursplanens kunskapskrav. I den nedre, ofärgade delen finns KP-konstruktörernas uppgiftsspecifika tolkning av dessa progressionskedjor. Som framgår av exemplet är minimikravets progressionsindikatorer kvantifierade, dels genom att antalet tolkningar som ska

²⁵⁷ Hur dessa värdeord kan tolkas på en generell nivå anges i Skolverkets stödmaterial, men i detta sammanhang är det den uppgiftsspecifika tolkning som kommer till uttryck i KP:s matriser som ligger till grund för analysen.

behandlas i resonemanget har preciserats (*båda*), dels genom att antalet historiska exempel som ska underbygga resonemanget har preciserats (*något rimligt historiskt exempel*). Även optimumnormernas progressionsindikatorer är kvantifierade. Skillnaden mellan E-nivån och C-nivån skrivs fram genom en kvantitativ progressionsindikator som handlar om att antalet historiska exempel har betydelse för resonemangets kvalitet (*några rimliga historiska exempel*). På C-nivån finns även additiva progressionsindikatorer som handlar om krav på kontextualisering av de historiska exemplen samt krav på att svaret ska innehålla ett resonemang om tolkningsalternativens rimlighet. Skillnaden mellan C-nivån och A-nivån inrymmer inte några kvantitativa progressionsindikatorer, men däremot såväl additiva som kvalitativa indikatorer. Den additiva aspekten kommer till uttryck i kravet på att resonemanget även ska inrymma ett perspektivbyte. Den kvalitativa aspekten handlar implicit om att det metakognitiva resonemanget gällande tolkningsalternativens rimlighet ska föras utifrån en konstruktivistisk kunskapssyn.

Progressionslogiken i samtliga uppgiftsspecifika KP-matriser är på motsvarande sätt uppbyggda av en kombination av kvantitativa, additiva och kvalitativa progressionsindikatorer. Progressionslogiken i KP:s matriser kan således, på samma sätt som progressionslogiken i HLP:s matriser, beskrivas som splittrad, med inslag av såväl MH-traditionella som HT-traditionella indikatorer. Denna splittrade bild ligger i linje med den utbredda uppfattningen inom fältet att progressionslogiken utgör ett efterfatt område.²⁵⁸

6.2.5 Uppfattningar om värderingsverktygen

Värderingsverktygens format och progressionslogik

KP-lärarnas uppfattningar om bedömningsanvisningarna är inte lika omfattande och tydliga som deras uppfattningar om provets uppgifter. Ett mönster som går att urskilja är dock att deras uppfattningar om bedömningsanvisningarna inte är entydigt positiva i samma utsträckning som deras uppfattningar om uppgifterna. Krister är den av KP-lärarna som är mest positivt inställd till bedömningsanvisningarna:

²⁵⁸ Se kapitel 3.

Jag tyckte också bedömningsanvisningarna var tydliga, väldigt bra exempel också, elevexempel, när det gäller hur man skulle tolka som lärare dom här svaren som man sen fick, bedömningsaspekter kring referenskunskap och kring att använda begrepp, och dom var till stor nytta dom här /.../, så jag tyckte om dom här exemplen som ni hade här i bedömningsanvisningarna, dom var till stor nytta ... det tyckte både jag och min kollega när vi såg på detta.²⁵⁹

Krister efterfrågar tydlighet och han uppfattar att KP:s bedömningsanvisningar uppfyller denna önskan. Han betonar särskilt elevexemplens betydelse för att skapa denna tydlighet. Kristers resonemang kan jämföras med HLP-läraren Hannas uppfattningar om den bristande tydligheten i HLP:s bedömningsgrunder.

Kristina delar i många avseenden Kristers positiva uppfattning om bedömningsanvisningarna i KP:

Det är ju så välarbetat detta så att det är ju väldigt svårt tycker jag att på nått sätt säga att nämen det här var inte rätt, eller det här var svårt att använda eller så, för att det är ju så många som har testat det. Jag tänker att det fungerar bra och man får stöd för det man ska göra och exemplen är bra /.../ Och det står ju också då att bedömningen utgår från följande kunskapskrav, man får det i början, och det är också viktigt att man kan koppla det dit.²⁶⁰

Samtidigt som Kristina beskriver uppgifterna som välarbetade så betonar hon att uppgiftskonstruktörernas tolkning är svår att ifrågasätta. Den här normativa undertonen som kommer fram i många av KP-lärarnas resonemang är inte alls lika framträdande hos HLP-lärarna. Det är en viktig skillnad som kommer att uppmärksammas i kapitlets slutdiskussion.

Till skillnad från Krister är Kristina däremot mer tveksam till bedömningsanvisningarnas format.

Kristina: Jag är ju oerhört kluven inför det här med matriser.

Jag: Ja, det är det som jag tycker är spännande

Kristina: Jag har följt den diskussionen hela vägen och jag vägrar som

²⁵⁹ Intervju med Krister, 2016-11-24.

²⁶⁰ Intervju med Kristina, 2016-11-21.

erfaren lärare att tvingas in i ett hörn av fyrkanter ... för jag tycker inte det är rätt ... men som formativ bedömning, som en del av den formativa bedömningen, så är det pedagogiskt.²⁶¹

Ett sätt att tolka Kristina kritik av matriser är att hon uppfattar att de begränsar hennes tolkningsfrihet i alltför stor utsträckning. Hon tar med andra ord avstånd från den mättraditionella ambitionen att begränsa lärarens tolkningsfrihet. Även Karin uttrycker ett visst motstånd mot matriserna i KP. Karin uttrycker sig på följande sätt om matrisen till en av uppgifterna i proven.

Några har jag väl tyckt varit lite komplicerade ... men nu kommer jag inte ihåg riktigt, var det ... och det tyckte jag den här uppgift 9 här /.../ Den tyckte jag ... alltså jag gav nästan upp ... jag tyckte det blev så komplicerat /.../Ja, det var väl att det var med alla de här tre (förmågorna) och att det blev ... hur ska jag hålla isär det när jag rättar det ... alltså vad ska jag se vilket som är vad ... och det är ju ... och det måste jag ju kunna kanske.²⁶²

Huruvida det är matrisens format eller progressionslogik som ligger till grund för Karins motstånd är något oklart, men i grunden handlar hennes kritik om att bedömningsanvisningarna är svåra att hantera. Den mättraditionella ambitionen att begränsa lärarens tolkningsfrihet i samband med värderingen av öppna uppgifter har resulterat i ett alltför ”komplicerat” värderingsverktyg.

Karl är den av KP-lärarna som uppehåller sig längst vid KP:s bedömningsanvisningar. Han är också den av KP-lärarna som är mest kritisk till dem.

Jag tycker det är svårt att rätta provet ... hade jag bara kunnat rätta det själv på nått sätt. Jag läser det och tittar på kunskapskraven ... denna fråga behandlar detta kunskapskravet ... och så har jag ämnesplanen framför mig. Så läser jag elevsvaret och så tittar jag på kunskapskraven och så kanske jag har något elevexempel som (ohörbart) Skolverket till stöd ... det hade räckt för mig. /.../ om jag hade tittat tillräckligt

²⁶¹ Intervju med Kristina, 2016-11-21.

²⁶² Intervju med Karin, 2016-11-08.

många gånger på ett antal elevsvar så jag hade förstått hur Skolverket vill att jag ska tänka, förstår du vad jag menar? Är där för mycket information i matrisen då kan det vara svårt också. Då har det nästan varit enklare att försöka jämföra elevsvaret med exempelsvaren.²⁶³

På samma sätt som Karin uppfattar Karl att KP:s matriser är svåra att hantera. Han menar att de innehåller för mycket information och han skulle föredra att endast använda elevexemplen som stöd för värderingen.

Sammanfattningsvis framträder ett mönster där flertalet av KP-lärarna uttrycker någon form av kritik mot KP:s matriser. En möjlig tolkning av KP-lärarnas avståndstagande från matriserna skulle kunna vara att det finns en metodologisk diskrepans mellan deras egen formatlogik och den formatlogik som kommer till uttryck i KP:s matriser. I sitt motstånd mot den mättraditionella KP-matrisen positionerar sig KP-lärarna i linje med den tolkningstraditionella ambitionen att konstruera värderingsverktyg som ger läraren ett stort tolkningsutrymme. De uppfattar KP-matriserna som alltför mättraditionella och förordar i stället en mer tolkningstraditionell värdering baserad på elevsvaren. I grunden handlar deras motstånd om tillit till professionens beprövade erfarenhet. KP-lärarna vill inte låta sig automatiseras och avprofessionaliseras. De värjer sig mot ”att man ska sitta och räkna belägg och så”.²⁶⁴

6.2.6 Sammanvägningsverktygen

Till skillnad från HLP så innehöll KP ett sammanvägningsverktyg. I ett formulär benämnt ”Sammanställning av elevresultat” skulle läraren fylla i ett prestationsomdöme (”belägg”) på varje enskild uppgift i de båda delproven. Formuläret begränsar lärarens tolkningsutrymme och är således till formatet tydligt mättraditionellt förankrat.

Vid sammanvägningen till ett samlat prestationsomdöme (”provbedömning”) skulle läraren ta hänsyn till två faktorer. Den första var elevens totala antal belägg och den andra var ”breddvillkoren”, det vill säga riktlinjer för provets emfaslogik. Den emfatiska logiken tydliggör dels att samtliga procedurala förmågor väger lika tungt, dels att det på E-nivån inte

²⁶³ Intervju med Karl, 2016-12-06.

²⁶⁴ Intervju med Karl, 2016-12-06.

lämnas utrymme för någon form av kompensatorisk sammanvägning. Eftersom KP:s emfaslogik endast omfattar domänlogikens procedurala dimension kan den kategoriseras som HT-traditionellt förankrad.

6.2.7 Uppfattningar om sammanvägningsverktygen

KP-lärarna uppmärksammar inte KP:s sammanvägningsverktyg. Det framkommer således inte huruvida de uppfattar att verktygets formatlogik och/eller emfaslogik ligger i linje med deras egen bedömningspraktik eller inte. Att KP-lärarna, på samma sätt som HLP-lärarna, inte uppmärksammar sammanvägningen kan som tidigare framkommit tolkas på två olika sätt. Det indikerar dels att det enskilda prestationsomdömet är meningsfullt i sig, till exempel vid betygssättning, dels att betydelsen av summativa jämförelser mellan elever, lärare och skolor tonas ner. Resonemanget om hur KP-lärarna använder sig av resultaten kommer att fördjupas i kapitel 8.

6.2.8 KP-sammanfattning

Observationsverktygen

KP innehöll under det första året en blandning av slutna och öppna uppgiftsformat, dock med en tydlig övervikt av det öppna, tolkningstraditionella, formatet. Domänlogiskt var dessa uppgifter huvudsakligen HT-traditionellt förankrade, men det fanns även uppgifter med en mer MH-traditionell förankring. KP förändrades över tid och på samma sätt som för HLP omfattar denna förändring såväl formatlogiken som domänlogiken, men det är en förändring som sker i motsatta riktningar jämfört med HLP. För KP innebar exkluderingen av de slutna uppgifterna en formatlogisk förskjutning i tolkningstraditionell riktning och en domänlogisk förskjutning i HT-traditionell riktning.

På samma sätt som HLP-lärarna uppmärksammar KP-lärarna i första hand observationsverktygens domänlogik. Även KP-lärarna uppfattar att det finns en diskrepans mellan deras egen domänlogik och den domänlogik som kommer till uttryck i provets uppgifter. Denna diskrepans framträder tydligast i de uppgifter som prövar förmågan att etablera temporala samband mellan dåtid, nutid och framtid. Det är framför allt framtidsperspektivet som uppmärksammas och som föranleder mer elaborerade resonemang.

Värderingsverktygen

KP:s bedömningsanvisningar var till formatet mättraditionellt förankrade och jämfört med HLP:s bedömningsgrunder var denna förankring än mer uttalad. Vad gäller KP:s progressionslogik är bilden inte lika tydlig. Progressionsindikatorerna i KP:s matriser är såväl kvantitativa, additiva som kvalitativa. Det är således, på samma sätt som för HLP:s bedömningsgrunder, inte möjligt att förankra KP:s progressionslogik i någon genomgående specifik historiekulturell tradition.

Till skillnad från HLP-lärarna framkommer det hos KP-lärarna ett tydligt mönster i deras uppfattningar om KP:s värderingsverktyg. Flera av KP-lärarna uttrycker ett motstånd mot KP:s mättraditionella matriser. KP-lärarnas motstånd kan tolkas som en form av formatlogisk diskrepans.

Sammanvägningsverktygen

KP:s sammanvägningsverktyg var till formatet mättraditionellt förankrat och dess emfaslogik kan kategoriseras som HT-traditionell. KP-lärarna uppmärksammar inte sammanvägningsverktyget, vilket innebär att det inte framkommer några eventuella diskrepanser.

6.3 Resultatdiskussion

Resultatdiskussion består av tre delar. I den första delen behandlas provkonstruktörerna och formuleringsarenan. I den andra delen behandlas historielärarna och realiseringsarenan. I den tredje delen presenteras en övergripande tentativ slutsats som kommer att återkomma och utvecklas i de efterföljande empiriska kapitlen.

6.3.1 Provkonstruktörerna och formuleringsarenan

I analysen framkommer att HLP-konstruktörerna inte gjorde något försök att konstruera mättraditionella observationsverktyg med anspråk på att mäta HT-traditionella kunskaper. De försökte istället hantera den fjärde kvadrantens dilemma genom att konstruera mättraditionella värderingsverktyg som gjorde anspråk på att mäta kvaliteten i elevprestationer genererade av tolkningstraditionella observationsverktyg. De försköt således hanteringen av dilemmat från observationsprocessen till värderingsprocessen genom att konstruera värderingsverktyg som begränsade lärens tolkningsutrymme.

Till skillnad från HLP-konstruktörerna gjorde däremot KP-konstruktörerna initialt ett försök att konstruera mättraditionella observationsverktyg med anspråk på att mäta HT-traditionella kunskaper. I 2015 års KP fanns det tre slutna flervalsuppgifter som gjorde anspråk på att vara förankrade i en HT-traditionell domänlogik. I de övriga sex uppgifterna hanterade KP-konstruktörerna den fjärde kvadrantens dilemma på samma sätt som HLP-konstruktörerna, det vill säga de försköt dilemmat till värderingsprocessen genom att utforma värderingsverktyg som begränsade lärarnas tolkningsutrymme.

I analysen framkommer vidare att både HLP-konstruktörerna och KP-konstruktörerna förändrade provens konstruktion över tid. HLP:s konstruktion förändrades såväl historiekulturellt som bedömningskulturellt. Historiekulturellt försköts HLP:s domänlogik i en MH-traditionell riktning genom att provets D-del exkluderades. Bedömningskulturellt försköts HLP:s formatlogik i en mättraditionell riktning genom att flervalsuppgifter inkluderades i provets A-del. KP:s konstruktion förändrades däremot endast bedömningskulturellt och till skillnad från HLP försköts formatlogiken i en tolkningstraditionell riktning genom att provets flervalsuppgifter exkluderades. Den här skillnaden mellan provens förändring över tid kan uppfattas som att provkonstruktörerna hanterar den fjärde kvadrantens dilemma på olika sätt.

HLP-konstruktörernas sätt att hantera den fjärde kvadrantens dilemma kan tolkas som en HT-traditionell reträtt, det vill säga att de låter den HT-traditionella domänlogiken delvis ge vika för den mättraditionella formatlogiken. Att provkonstruktörer på det här sättet tenderar att retirera epistemologiskt när de ställs inför den fjärde kvadrantens dilemma har uppmärksammats i tidigare forskning. VanSledright beskriver samma fenomen när han argumenterar för att det finns en risk att provkonstruktörer som egentligen är HT-traditionellt förankrade tenderar att "sacrifice construct validity to achieve adequate test reliability coefficients".²⁶⁵ KP-konstruktörernas sätt att hantera den fjärde kvadrantens dilemma kan däremot tolkas som en mättraditionell reträtt, det vill säga de låter de mättraditionella kraven på reliabilitet ge vika för en HT-traditionell konstruktvaliditet. För att parafrasera VanSledright så offrade KP-konstruktörerna reliabili-

²⁶⁵ VanSledright (2014), s. 8. Jfr med Wineburgs (2004) resonemang om att konstruktörer av externa prov tenderar att förskjuta balansen i mättraditionell riktning.

teten på validitetens altare när de gav upp försöket att konstruera mättraditionella observationsverktyg med anspråk på att mäta HT-traditionella kunskaper.

6.3.2 Historielärarna och realiseringsarenan

I studien framkommer att både HLP-lärarna och KP-lärarna uppfattar att det finns diskrepanser mellan deras egen domänlogiska uppfattning och den domänlogik som kommer till uttryck i provens uppgifter. HLP-lärarnas domänlogiska diskrepans manifesteras framför allt i deras uppfattningar om D-uppgiften, det vill säga i den uppgift som på olika sätt prövar elevens förmåga att resonera om tredelade temporala samband. När HLP-beskriver den tankeoperation som prövas i D-uppgiften använder de begreppet "historiemedvetande". KP-lärarnas domänlogiska diskrepans är mera omfattande, men även hos dem manifesteras den framför allt i deras uppfattningar om de uppgifter som handlar om tredelade temporala samband. De beskriver dessa uppgifter genom att fokusera på framtidsperspektivet, det vill säga det tredje ledet i tankeoperationen. Hos båda lärarkollektiven manifesteras således diskrepansen i uppgifter som har en tydlig förankring i den HT-kontinentala domänlogiken.

Historielärarnas sätt att hantera denna domänlogiska diskrepans skiljer sig dock åt. HLP-lärarna gör i högre grad motstånd mot provets domänlogik medan KP-lärarna i högre utsträckning är lojala med provets domänlogik och uppfattar den som svår, men "rätt". KP-lärarna uppfattar således i högre utsträckning KP:s domänlogiska budskap som normativa. I KP-lärarnas utsagor framkommer dock att det inte är alla inom historielärare ämneslaget som delar deras acceptans av provets domänlogik.

Dessa skillnader kan tolkas som att lärarkollektiven involveras i olika historiekulturella processer. HLP-lärarna involveras framför allt i en vertikal historiekulturell *förhandling om* budskapen, en förhandling som i förlängningen innebar att den kontinentala tolkningen av HT-traditionen, manifesterad i begreppet historiemedvetande, tonas ner i HLP:s domänlogik. KP-lärarna involveras i högre grad i en historiekulturell *förmedling av* budskapen, en förmedling som i förlängningen innebar att den kontinentala tolkningen av HT-traditionen, manifesterad i framtidsperspektivet, betonas i KP-lärarnas egen domänlogik.

Huruvida HLP:s domänlogiska förändring över tid, det vill säga D-uppgiftens exkludering, ska tolkas som en följd av en historiekulturellt

förankrad epistemologisk förhandling eller som en bedömningskulturellt förankrad metodologisk reträtt till följd av den fjärde kvadrantens dilemma är en öppen fråga. Den ena tolkningen behöver dock inte utesluta den andra.

KP:s formatlogiska förändring över tid, det vill säga flervälsfrågornas sorti, kan däremot tolkas som en bedömningskulturell reträtt till följd av den fjärde kvadrantens dilemma. Det finns ingenting i det empiriska materialet som talar för att formatförskjutningen skulle vara en följd av en vertikal budskapsförhandling kring uppgifternas formatlogik.

6.3.3 En tentativ slutsats

Både HLP-lärarna och KP-lärarna beskriver en diskrepans mellan provets domänlogik och deras egen domänlogik, en diskrepans som framförallt manifesteras i de HT-traditionella uppgifter som prövar elevens förmåga att hantera tredelade temporala samband. Hos HLP-lärarna framkommer ett tydligt motstånd mot dessa uppgifter vilket föranleder en historiekulturell förhandling om provets domänlogik som resulterar i att uppgiften exkluderas från HLP. KP-lärarna är däremot lojala med provets domänlogik och de involveras i en historiekulturell budskapsförmedling. En tentativ slutsats baserad på denna skillnad är att HLP-lärarna i första hand uppfattar HLP som ett resultatgenererande *bedömningskulturellt* verktyg medan KP-lärarna i första hand uppfattar KP som ett budskapsförmedlande *historiekulturellt* verktyg.

7. GENOMFÖRANDET

I det här empiriska kapitlet behandlas historielärarnas beskrivningar av hur de **genomför** de frivilliga externa proven. Inledningsvis positioneras provens processreglerande direktiv i relation till de bedömningskulturella traditioner som har redovisats tidigare. Därefter positioneras historielärarnas beskrivningar av hur de i realiteten genomför bedömningen, vilket innebär att bedömningskulturella skillnader mellan HLP-lärarna och KP-lärarna framträder. I kapitlets avslutande resultatdiskussion kommer dessa bedömningskulturella skillnader att sättas in i ett historiekulturellt sammanhang samt relateras till den fjärde kvadrantens dilemma.

7.1 HLP

7.1.1 De processreglerande direktiven

HLP:s direktiv för hur observationen skulle genomföras var kortfattade, men ändå tydligt förankrade i de mättraditionella kraven på standardisering. I direktiven framgick till exempel vilket datum provet skulle skrivas samt den skrivtid som eleverna hade till sitt förfogande. Indirekt framgick även att provet skulle skrivas under kontrollerade former och på en bestämd skrivplats. Värderingen och sammanvägningen av elevprestationerna skedde lokalt ute på skolorna, men hur dessa processer skulle genomföras reglerades inte i några explicita direktiv. Frånvaron av sådana direktiv innebar att ämneslagen eller den enskilda läraren kunde situationanpassa dessa delar av genomförandet, vilket kan uppfattas som en tolkningstraditionell förankring.²⁶⁶

7.1.2 Observationsprocessen

Samtliga HLP-lärare uttrycker att de genomför observationen i linje med HLP:s mättraditionella observationsdirektiv. Kraven på standardisering ifrågasätts således inte, men några av lärarna uppfattar ändå att det finns

²⁶⁶ Från och med 2010 innehöll HLP en lärarinstruktion där observationsdirektiven preciserades och där vissa värderingsdirektiv fanns angivna.

ett visst handlingsutrymme. Ett exempel är Håkan som beskriver en i huvudsak standardiserad process, men där lärarna ges ett visst handlingsutrymme gällande skrivdatum:

Jag har det som en särskild punkt till på måndag när vi har ämneskonferens. Där ska vi se vilka som ska vara med i årets prov och hur ska vi lägga det. Ska vi lägga det tillsammans eller ska vi ge något alternativ, för man behöver inte göra samtidigt som de andra. Och då ska jag ha klart för mig hur många grupper som kommer att göra det denna gången, alltså sammanlagt.²⁶⁷

Ett återkommande inslag i HLP-lärarnas beskrivningar av observationen är att ett gemensamt skrivdatum och en gemensam skrivplats som en del av standardiseringen inte sker per automatik. Håkan uttrycker sig på följande sätt:

Det är jävligt jobbigt ... speciellt att få till det här med ... alltså på den egna skolan ... att alla ska skriva samtidigt och göra skrivschema och bråka med språklärare och mattelärare som inte vill släppa ifrån sig lektioner.²⁶⁸

Trots dessa praktiska svårigheter väljer ändå Håkan och de övriga HLP-lärarna att genomföra observationen under standardiserade former. Värdet av att standardisera processen tycks vara tungt vägande.

Den ende av HLP-lärarna som i viss mån avviker från mönstret är Henrik. Även Henrik genomför observationen under standardiserade former, men till skillnad från de andra HLP-lärarna väljer han ut vilka undervisningsgrupper som ska skriva provet. Under de år som Henrik har använt sig av HLP har han låtit "sex utvalda grupper" skriva provet och tendensen över tid har varit att det blivit allt mer sällan. Han menar att det är undervisningsgruppens sammansättning som avgör om han ska använda sig av provet eller inte. Henrik undervisar huvudsakligen elever på yrkesförberedande program. De här eleverna har av olika skäl valt att läsa A-kursen i historia som valbar kurs eller som individuellt val. Vanligtvis

²⁶⁷ Intervju med Håkan, 2010-01-19.

²⁶⁸ Intervju med Håkan, 2010-01-19.

låter han inte dessa elever skriva HLP då han menar att han inte vill utsätta sina elever för ”ännu ett misslyckande”.

Om jag har en grupp framför mig där jag vet att jag skulle döda allt intresse genom att fokusera på ett slutprov i maj så struntar jag i det. Då tänker jag: vad är kursmålen, hur ska jag nå till kursmålen? [...] Jag har till exempel en grupp nu som jag för död och pina inte skulle utsätta för det där provet.²⁶⁹

Provet är enligt Henrik anpassat för elever som läser teoretiska program och han menar att provet förutsätter en ”traditionell undervisning” och att det är konstruerat efter en ”mall” som inte passar eleverna på de yrkesförberedande programmen.²⁷⁰ Henrik uppfattar sin undervisning i dessa grupper som elevcentrerat och projektinriktat med många kreativa inslag. Det är intressant att följa Henriks tankar om vad som händer med hans undervisningspraktik när han tar ställning till om han ska låta en undervisningsgrupp skriva provet eller inte.

Jag: Menar du att den här typen av undervisning inte lämpar sig att avslutas med ett stort prov?

Henrik: Jag menar inte att det inte lämpar sig att avslutas med ett stort prov. Jag kan avsluta med ett stort prov, men med ett stort prov som är anpassat till det arbetssättet jag har haft. [...] Om vi skulle få Skåneprovet som obligatorisk så hade jag naturligtvis till viss del fått lägga om min undervisning. Jag kan till exempel inte lägga ett par månader på ett projekt.²⁷¹

I de undervisningsgrupper där Henrik bestämmer sig för att använda sig av HLP väljer han således bort det projektinriktade arbetssätt som vanligtvis präglar hans undervisning. Det här skulle kunna indikera att HLP befäster mera traditionella undervisningsmönster.

²⁶⁹ Intervju med Henrik, 2009-11-26.

²⁷⁰ Intervju med Henrik, 2009-11-26.

²⁷¹ Intervju med Henrik, 2009-11-26.

7.1.3 Värderingsprocessen

HLP saknade explicita värderingsdirektiv, vilket indirekt ligger i linje med tolkningstraditionen. Trots detta betonar flera HLP-lärare ändå vikten av att värderingen av elevprestationerna sker under mättraditionella former. Helena beskriver värderingsprocessen på sin skola på följande sätt:

Då byter vi prov med varandra och sen försöker vi vara väldigt noga med att vi rättar helt enligt facit /.../ Där hade vi till en början lite problem, men jag tror att det har blivit bättre ... att vi liksom verkligen rättar efter facit och inte ger poäng för något som inte står i facit, då blir det ju helt fel. Och sen så byter vi och tittar lite grann på hur vi har rättat samma fråga. Vi tittar ... man frågar nån kollega ... tycker du det är ok ... tycker du detta är en ok rättning ... nej, jag hade inte gett så många poäng för det svaret eller han hade nog fått en halv till och så där.²⁷²

Som framgår av utsagan beskriver Helena en standardiserad värderingsprocess där lärarna inte värderar den egna klassens prestationer och där värderingen kompletteras med en efterföljande kollegial samvärdering där särskilt svårvärderade prestationer tas upp till diskussion.

Även Hannas beskrivning av värderingsprocessen ligger i linje med de mättraditionella kraven på standardisering. Hon beskriver värderingsprocessen på följande sätt:

Vi samrättar mellan klasserna, alltså man rättar inte sin egen klass prov utan man byter med varandra eller mellan skolorna /.../ och det är eleverna väl medvetna om ... och jag tycker det är viktigt också att de ser att det är någon sorts ... att vi rättar efter samma kriterier och att det påminner om ett nationellt prov.²⁷³

På samma sätt som Hanna beskriver Helena en systematiserad kollegial samvärdering där den enskilde läraren inte värderar sina egna elevers provsvar. Hon betonar vikten av en standardiserad värderingsprocess genom att dra paralleller till de nationella proven.

²⁷² Intervju med Helena, 2009-11-25.

²⁷³ Intervju med Hanna, 2009-11-19.

Håkan beskriver en värderingsprocess som delvis avviker från Helenas och Hannas. Han beskriver att det i regel är den undervisande läraren som värderar elevprestationerna, men han uttrycker samtidigt att det inom kollegiet finns en ambition att värdera elevprestationerna gemensamt när han berättar att de medverkande lärarna vid ett tillfälle genomförde en samvärdering utav en av provets delar (C-delen). Man skulle kunna beskriva denna värderingsprocess som en partiell kollegial samvärdering.

Henrik avviker på nytt från de övriga HLP-lärarna. På hans skola värderar den undervisande läraren sina egna prov. Henrik förklarar att denna tolkningstraditionella värderingsprocess beror på att en mer standardiserad värdering i form av en systematisk kollegial samvärdering skulle vara alltför tidskrävande, ”det skulle dra ut på tiden ännu mer”.²⁷⁴

7.1.4 Sammanvägningsprocessen

Som tidigare framgått saknade HLP sammanvägningsverktyg och det fanns således inte några direktiv för hur sammanvägningen skulle genomföras. HLP-lärarna uppmärksammar inte frånvaron av sammanvägningsdirektiv och det finns inte heller några utsagor där HLP-lärarna beskriver att de genomför en sammanvägning av provets olika delar.²⁷⁵

7.1.5 HLP-sammanfattning

Observationen

HLP:s observationsdirektiv reglerade observationsprocessen i linje med de mättraditionella kraven på standardisering. Trots att HLP-lärarna uppfattar att det kan vara problematiskt att efterleva dessa ideal väljer ändå de flesta att vara lojala med direktiven. Det innebär ett merarbete att genomföra observationen under standardiserade former, men det är ett pris man är beredd att betala. Henrik, som undervisar på en skola med i huvudsak yrkesförberedande program, ger dock uttryck för ett friare förhållningsätt gentemot observationsdirektiven. När han väljer ut vilka undervisningsgrupper som ska genomföra provet utmanar han de mättraditionella kraven på standardisering och rör sig i riktning mot tolkningstraditionens mer situerade, klassrumsnära och autentiska ideal.

²⁷⁴ Intervju med Henrik, 2009-11-26.

²⁷⁵ Jfr avsnitt 6.1.7.

Värderingen

HLP saknade explicita direktiv gällande värderingsprocessen, något som indirekt ligger i linje med tolkningstraditionens betoning på situering. Trots frånvaron av värderingsdirektiv är det endast en av HLP-lärarna som värderar sina egna prov. Två av de andra genomför en kollegial samvärdering och den tredje uttrycker en ambition att genomföra en sådan gemensam samvärdering. Den kollegiala samvärdering som HLP-lärarna beskriver är systematiserad och syftar i första hand till att reducera den enskilda lärarens betydelse under värderingsprocessen. Det innebär att HLP-lärarna beskriver en värderingsprocess som i hög utsträckning är mättraditionellt förankrad.

Sammanvägningen

HLP saknade sammanvägningsdirektiv och därmed riktlinjer för hur sammanvägningen skulle genomföras. HLP-lärarna efterfrågar inte några sådana riktlinjer och det finns inte heller några utsagor där det framkommer att de genomför en sammanvägning av provets olika delar.

7.2 KP

7.2.1 De processreglerande direktiven

KP:s observationsdirektiv var mer omfattande än HLP:s. Där framkom att provets två delar skulle genomföras vid skilda tillfällen. Vidare framkom att observationen kunde göras vid de tidpunkter som skolan fann lämpliga, så till skillnad från HLP fanns det inget specificerat datum för observationen. Provtiden reglerades för respektive delprov, 90 respektive 80 minuter. Till skillnad från HLP reglerades även vilka slags anpassningar som kunde göras vid observationen, till exempel ytterligare skrivtid för elever med läs- och skrivsvårigheter samt tillgång till lexikon för elever med annat modersmål än svenska. Det framkom att det är skolan som ansvarar för dessa anpassningar, men samtidigt betonades att det är viktigt att skolan gör dessa anpassningar ”så att provet fortfarande prövar de kunskaper och förmågor som gäller för provet”.²⁷⁶ Ytterligare en skillnad jämfört med HLP är de sekretessregler som fanns angivna för hanteringen av provet. Allt sammantaget visar att KP var ett prov vars observationsdirektiv tydligt var förankrade i mättraditionen.

²⁷⁶ Värt att notera är att verbet ”får” användes i samband med skrivningarna om anpassning. Direktiven reglerade således vilka anpassningar som var tillåtna att göra.

På samma sätt som HLP saknade KP processreglerande direktiv gällande värderingen och sammanvägningen. Båda processerna genomfördes lokalt ute på skolorna, men hur de skulle genomföras reglerades inte. Lärarnas handlingsutrymme var således stort, vilket ligger i linje med det tolkningsstraditionella idealet att situera bedömningsprocessen.

7.2.2 Observationsprocessen

KP-lärarna är till skillnad från HLP-lärarna inte samstämmiga när de beskriver hur de genomför observationen, det vill säga hur de operationaliserar KP:s observationsdirektiv. Krister beskriver en standardiserad process i linje med KP:s mättraditionella direktiv:

Principen är att alla som läser historia 1b ska göra provet och för närvarande [...] är det ungefär inalles hundra elever, alla ska försöka göra det, alla ska göra det. Provet görs oftast i en stor sal, allihopa sitter där, vi är två lärare som bevakar dom. Om man har dyslexi eller om det finns speciella omständigheter så skriver man det på nån slags, vi kallar det studioteket eller resurscentrum. [...] Vi gör det också naturligtvis på en och samma gång, tillsammans allihopa.²⁷⁷

Krister är som framgår fullt ut lojal med KP:s direktiv. Han betonar på samma sätt som HLP-lärarna att alla ska göra provet och att alla ska göra det samtidigt. Även hans resonemang om anpassningar ligger nära KP:s direktiv. När han beskriver hur två lärare ”bevakar” eleverna understryks Kristers lojalitet med KP:s direktiv och hans processuella förankring i mättraditionen.

De tre övriga KP-lärarna uppvisar inte samma lojalitet med KP:s observationsdirektiv. I deras utsagor framträder en processlogisk diskrepans. Karin är den lärare som ligger längst ifrån Krister och HLP-lärarna.

Karin: En gång, när jag var med och provade ut så använde jag hela provet /.../ men annars så är det ju att ha i delmoment.

Jag: Kan du utveckla det lite mer?

Karin: Jo, men jag har gjort det till exempel när vi har läst om Sverige under slutet av 1800-talet, då med migrationen, så har jag använt ett

²⁷⁷ Intervju med Krister, 2016-11-24.

delmoment som har varit med ... det kanske inte finns längre kvar om aktör och struktur för svenskarna då ... vad det var som styrde att de flyttade utomlands, mest USA ju, så då har jag använt det.²⁷⁸

Karin beskriver att hon inte använder provet i sin helhet utan att hon hämtar enskilda uppgifter från provet, uppgifter som hon sedan integrerar i sin egen bedömningspraktik som en form av examinerande avslutning på ett delmoment under kursen. Att på det här sättet välja ut specifika uppgifter från KP är ett tydligt exempel på en tolkningstraditionell situation av observationen. När Karin beskriver vilka delmoment som hon brukar avsluta med uppgifter hämtade från KP, ger hon ytterligare ett exempel, ett exempel som handlar om källor och historiebruk.

Jag brukar köra det ganska tidigt ... källkritik och historiebruk i ettan. /.../ Jag brukar gå igenom kriterierna, de här källkritiska kriterierna. och, ja kvarlevor och berättande källor och sånt där och sen så fick dom den här uppgiften och sen fick de ju också en om historiebruk, det tog jag ju också ... om Thermopyle.²⁷⁹

Det är svårt att se något tydligt mönster hos Karin när det gäller vilka uppgifter hon väljer ut. Den första uppgiften om den svenska migrationen under 1800-talet prövar referenskunskaper och begreppshantering, medan den andra handlar om källhantering och historiebruk.

Även Karl beskriver att han inte använder provet i sin helhet: "Nån gång har jag gjort det, annars har jag plockat ut".²⁸⁰ När Karl har använt sig av enskilda uppgifter i sin bedömningspraktik har han framför allt använt sig av de frisläppta exempeluppgifterna.

Jag har printat ut allt material så att jag har det utskrivet och så använder jag exempeluppgifterna. "Hur man kan använda historia om kors-tågen", det var en uppgift där Jan Guillou och George Bush använder historia, eller "Blå tåget och Willy Brandt använder historia om kalla kriget".²⁸¹

²⁷⁸ Intervju med Karin, 2016-11-08.

²⁷⁹ Intervju med Karin, 2016-11-08.

²⁸⁰ Intervju med Karl, 2016-12-06.

²⁸¹ Intervju med Karl, 2016-12-06.

När Karl ger exempel på vilka uppgifter han har valt ut är det som framgår två historiebruksuppgifter han nämner. Att han lyfter fram just historiebruksuppgifterna är intressant och kan jämföras med Kristina.

Kristina berättar att hennes sätt att genomföra observationen har förändrats över tid. 2013 (sic!) genomförde hon en utprovningssversion av KP i sin helhet i två av sina undervisningsgrupper. 2014 använde hon inte provet alls. 2015 genomfördes provet på nytt, men under nya former.

Vi kunde genomföra det nu, då fick vi ihop 80 minuter som vi hade pusslat själva med att flytta lektioner och be kollegor skjuta lite på och så där ... så vi kunde faktiskt ... vi hade faktiskt tur, vi kunde få ihop samtidigt 80 minuter.²⁸²

Kristina beskriver 2015 års observation som delvis standardiserad och på samma sätt som HLP-lärarna uppfattar hon att det är problematiskt att skapa utrymme i schemat för ett gemensamt provtillfälle. Hon beskriver längre fram under intervjun att stödet från skolledningen var begränsat men att hon tillsammans med två kollegor lyckades genomföra en gemensam observation. I beskrivningen framkommer att Kristina och hennes kollegor använde KP som ett slags slutprov, men på samma sätt som Karin och Karl situerade de observationen genom att välja ut delar av provet. 2015 valde de ut två uppgifter från KP och som ämnesansvarig uppfattar Kristina att hon var drivande i valet av uppgifter.

Kristina: Vi genomförde två frågor av förra årets prov. Det kunde jag genomdriva.

Jag: Ok, vilka delar var det ni valde att ... ?

Kristina: Vi valde, det är dom två som jag har valt ut här sen, och det är historiebruk och källhantering ... och det är väl dom två där jag ... det var inte svårt att komma överens om ... historiebruk är, känner många sig väldigt osäkra på vad det är och hur man ska kunna bedöma det, och när det gäller källhantering så är det svårt att hitta ett bra material för att kunna göra det, det finns ju ett färdigt bra material här, så att dom två frågorna var vi ... det var inte svårt att komma överens om ... men däremot ett helt prov var dom inte beredda, i sina klasser

²⁸² Intervju med Kristina, 2016-11-21.

att genomdriva.²⁸³

Kristina motiverar valet av uppgifter på olika sätt. Historiebruksuppgiften valdes ut där för att de inom ämneslaget var osäkra på vad historiebruk är och hur man ska bedöma det, medan källhanteringen valdes ut för att det är tidskrävande att konstruera en sådan uppgift. Kristinas resonemang om historiebruksuppgiften kan jämföras med HLP-lärarnas resonemang om D-uppgiften och ”historiemedvetande”. I båda fallen uppfattas uppgiftens domänlogik som svårtolkad, men till skillnad från HLP-lärarna, som tar avstånd och bortförhandlar den svårtolkade uppgiften, så väljer Kristina och hennes kollegor att använda uppgiften i deras gemensamma observation. Den här skillnaden kommer att uppmärksammas och behandlas utförligare i kapitlets slutdiskussion.

På samma sätt som Kristina involverar Karin sina ämneskollegor i sitt resonemang kring observationen. Hon berättar att hon tillsammans med en kollega bestämde sig för att välja ut en KP-uppgift och sedan använda den i flera klasser.

Han och jag känner att nu måste vi få lite nystart med det här med historia, så vi har dragit igång ... och det har vi gjort ... egentligen var det ju då ... det är en del utav provet som är på historia 1a1 om källkritik och det bestämde vi att det ska vi ... det provet ska vi använda i alla grupper ... historiegrupper ... från 50-poängskursen till fördjupningskursen och se liksom vad får vi för resultat.²⁸⁴

Till skillnad från Kristina var inte observationen standardiserad i den bemärkelsen att alla skulle genomföra uppgiften vid samma tidpunkt. Karin berättar vidare att de övriga fyra historielärarna i ämneslaget anslutit sig till projektet, men att hon är osäker på om alla verkligen har genomfört observationen: ”Ja, fast jag vet ju inte, hon som är då lite ... på tvären. För vi träffades precis innan lovet och då så sa hon att hon inte hunnit haft det där provet än”.²⁸⁵

²⁸³ Intervju med Kristina, 2016-11-21.

²⁸⁴ Intervju med Karin, 2016-11-08.

²⁸⁵ Intervju med Karin, 2016-11-08.

Det kan avslutningsvis noteras att de tre KP-lärare som beskriver att de endast använder delar av provet samtliga uttrycker en ambition att i framtiden genomföra provet i sin helhet under standardiserade former. De tre KP-lärarna anger olika skäl till varför de inte genomför provet i sin helhet, men samtliga hänvisar i grunden till det kollegiala motstånd som berördes som hastigast i det föregående kapitlet. Det här kollegiala motståndet kommer som sagt att belysas utförligare i det avslutande empirikapitlet, se avsnitt 9.5.3.

7.3.2 Värderingsprocessen

På samma sätt som i HLP så saknas det i KP explicita direktiv gällande värderingsprocessen. Lärarnas handlingsutrymme är således stort, vilket ligger i linje med tolkningstraditionens logik. Trots detta uttrycker samtliga KP-lärare en strävan efter att standardisera värderingen. De beskriver på samma sätt som HLP-läraren Håkan en partiell kollegial samvärdering där den undervisande läraren initialt värderar elevsvaren för att därefter välja ut några elevsvar som värderas kollegialt. Denna samvärdering kan vara mer eller mindre systematisk. Karl beskriver en relativt osystematisk samvärdering:

Jag: Rättar du dina egna prov?

Karl: Ja, jag har gjort det, men sen har jag ... Vi är några kollegor som har tittat på varandras prov ... enskilda frågor eller enskilda prov.

Jag: Är det särskilt utvalda frågor då som ni tittar på?

Karl: Där vi upplever att vi vill få nån annans ögon på det. Men det har vi gjort ganska mycket i och för sig ... tittat på prov ... tittat på elevsvar, tillsammans ... i kollegiet.²⁸⁶

Till skillnad från Karl beskriver Kristina en mer systematisk process:

Sambedömning gjorde vi så här att vi rättade våra prov, och sen så valde vi ut fem stycken elevsvar var som vi av någon anledning ville diskutera. Då tar man ju såna som man tycker ligger på gränsen åt nått håll så att säga, eller där man har svårt att bestämma sig ... och så kopierade vi upp de här femton svaren, vi var tre lärare då, satt vi och

²⁸⁶ Intervju med Karl, 2016-12-06.

diskuterade, och det var väldigt givande.²⁸⁷

Krister är den av KP-lärarna som har drivit den partiella samvärderingsprocessen längst i en mättraditionell riktning. Han beskriver en systematisk process där en tredjedel av svaren i form av ”110 buntar” väljs ut för samvärdering tillsammans med en kollega.

Jag tror vi börjar med att vi läser igenom ... förra året var det två lärare, vi läste igenom våra egna och så läste vi igenom ungefär hälften av den andres provskörd och sen satt vi oss ned och rättade våra egna, men vi valde ut ungefär en tredjedel för att ha en sambedomning /.../ så först så ser vi på alla prov tillsammans och så rättar jag som undervisande lärare, och sen så har vi en diskussion kring dom ungefär en tredjedel som vi kanske är tveksamma om eller som vi vill ha en diskussion om.²⁸⁸

Att det är just Krister som drivit den partiella samvärderingen längst i mättraditionell riktning kan kopplas till att han är den ende av KP-lärarna som låter sina elever genomföra provet i sin helhet under standardiserade former.

Karins beskrivning av hur hon genomför värderingen avviker från de tre andra KP-lärarna.

Ja, och då har jag ju bedömt ensam ... nu har vi ju då försökt ... först nu då kommit igång och sambedomt ... nu ska vi på den här uppgiften om ... vad hette den .. Vad hände egentligen på skeppet ... Little George var det väl ... där ska vi ju ... jag har rättat mina och alla har nog rättat, men sen så ska vi sätta oss och gå igenom hur vi har rättat och sambedöma lite mer ... men innan, alltså när jag har gjort mina egna uppgifter, eller kört då från portalen så har jag ju rättat ensam.²⁸⁹

²⁸⁷ Intervju med Kristina, 2016-11-21.

²⁸⁸ Intervju med Krister, 2016-11-24.

²⁸⁹ Intervju med Karin, 2016-11-08.

Karin beskriver att hon vanligtvis värderar svaren på egen hand, men att hon försöker att komma igång med någon form av partiell kollegial samvärdering i anslutning till den KP-uppgift som ämneslaget har beslutat att samtliga elevgrupper ska genomföra.

7.3.3 Sammanvägningsprocessen

Som tidigare framgått saknade KP direktiv för hur sammanvägningen skulle genomföras. På samma sätt som HLP-lärarna uppmärksammar inte KP-lärarna frånvaron av sammanvägningsdirektiv och det finns inte heller några utsagor där KP-lärarna beskriver att de genomför en sammanvägning av provets olika delar.²⁹⁰

7.3.4 Sammanfattning

Observationen

På samma sätt som i HLP är KP:s observationsdirektiv mättraditionellt förankrade, vilket innebär höga krav på standardisering. Till skillnad från HLP-lärarna är KP-lärarna inte lojala med dessa direktiv när de genomför observationen. Den här processlogiska diskrepansen manifesteras framför allt när KP-lärarna situerar observationen genom att välja ut uppgifter från KP som de sedan på olika sätt integrerar i sin egen bedömningspraktik.

Värderingen

Varken HLP eller KP har några processreglerande direktiv gällande värderingen, vilket indirekt ligger i linje med tolkningstraditionen. Detta till trots tillämpar båda lärargrupperna någon form av partiell kollegial samvärdering och/eller kollegial extern värdering. Skillnaden mellan lärargrupperna är att KP-lärarna i något högre utsträckning standardiserar denna samvärdering. Denna processlogiska skillnad indikerar att den mättraditionella betoningen på reliabilitet och mätsäkerhet väger tyngre hos HLP-lärarna än hos KP-lärarna.²⁹¹

²⁹⁰ Jfr avsnitt 6.1.7.

²⁹¹ Behovet hos HLP-lärarna att involvera kollegorna i värderingen skulle kunna spegla en osäkerhet gällande tolkningen av kursplanens progressionslogik.

Sammanvägningen

KP saknar sammanvägningsdirektiv och därmed riktlinjer för hur sammanvägningen skulle genomföras. KP-lärarna efterfrågar inte några sådana riktlinjer och det finns inte heller några utsagor där det framkommer att de genomför en sammanvägning av provets olika delar.

7.3 Resultatdiskussion

7.3.1 Att standardisera eller situera genomförandet

I studien framkommer att historielärarna **genomför** frivilliga externa prov i historia på olika sätt. Denna skillnad framträder tydligast i deras sätt att genomföra observationen.²⁹² HLP-lärarna beskriver att de i linje med HLP:s mättraditionella direktiv genomför observationen under standardiserade former. Även KP:s observationsdirektiv är mättraditionellt förankrade, men till skillnad från HLP-lärarna så frigör sig KP-lärarna i hög utsträckning från dessa direktiv. De uttrycker förvisso en ambition att genomföra observationen under standardiserade former, men i realiteten är det endast en av dem som faktiskt gör detta. De andra KP-lärarna väljer i stället att genomföra observationen i linje med den tolkningstraditionella betoningen på situering framför standardisering. Denna processlogiska diskrepans kommer till uttryck på flera olika sätt, men en viktig del av situeringen är att KP-lärarna inte använder provet i sin helhet utan att de väljer ut vilka uppgifter som de vill använda. Att tre av KP-lärarna väljer ut uppgifter från KP, antingen för att integrera dem i sin egen bedömningspraktik, eller för att tillsammans med kollegor genomföra dem under mer standardiserade former, är ett viktigt resultat i sig, men utifrån ett historiekulturellt perspektiv är det kanske ännu viktigare att analysera vilka uppgifter KP-lärarna väljer ut.

7.3.2 Att begränsa eller bredda domänlogiken

När KP-lärarna väljer ut uppgifter väljer de i första hand uppgifter som prövar källhantering och historiebruk, det vill säga två förmågor som är förankrade i den HT-traditionella domänlogiken. I det föregående kapitlet framkom att historiebruk är en förmåga som KP-lärarna uppfattar som en

²⁹² De övergripande resultat som framkommer gällande värderingsprocessen är inte lika tydliga som resultaten gällande observationsprocessen. Det framträder inte heller samma tydliga mönster, det vill säga att HLP-lärarna är lojala med HLP:s observationsdirektiv, medan KP-lärarna tenderar att distansera sig från KP:s observationsdirektiv.

ny, svårtolkad och svårbedömd del av historieämnets domänlogik. Till skillnad från HLP-lärarna, som valde bort den uppgift som prövade en förmåga som inte låg i linje med deras egen domänlogik, så väljer KP-lärarna aktivt ut en uppgift som prövar en förmåga som inte ligger i linje med deras egen domänlogik. HLP-lärarna och KP-lärarna hanterar således uppgifterna i respektive prov på diametralt olika sätt.

Lärarkollektivens olika sätt att hantera uppgifterna i respektive prov kan relateras till ett välkänt fenomen inom det bedömningsteoretiska fältet, ett fenomen som handlar om att mättraditionella externa prov tenderar att begränsa komplexa konstrukt ("narrowing the construct").²⁹³ Att begränsa ett konstrukt kan innebära att komplexa delkomponenter helt lyfts bort, men det kan också innebära att komplexa delkomponenter förskjuts epistemologiskt. Det är möjligt att tolka HLP-lärarnas förhandling om D-uppgiften som ett exempel på att provets domänlogik begränsas. KP-lärarnas selektiva hantering av HLP-uppgifterna handlar däremot inte om att begränsa provets domänlogik utan snarare om att bredda deras egen domänlogik, det vill säga att använda uppgifterna i KP som ett stöd för att integrera de "kontinentala" delarna av kursplanens HT-traditionella domänlogik i sin egen undervisnings- och bedömningspraktik.

KP-lärarnas sätt att hantera uppgifterna kan även relateras till den fjärde kvadrantens dilemma. Som framkom i det föregående empiriska kapitlet hanterade såväl HLP- som KP-konstruktörerna dilemmat genom att förskjuta det från observationen till värderingen. Att KP-lärarna väljer ut just historiebruksuppgifterna skulle kunna tolkas som att de inte bara behövde ett stöd för att integrera uppgiftens domänlogik i sin egen undervisnings- och bedömningspraktik, utan att de även hade ett behov av ett värderingsverktyg som tydliggjorde hur prestationerna skulle värderas. KP:s bedömningsanvisningar i form av mättraditionellt förankrade matri-ser, i kombination med autentiska elevexempel och värderingskommentarer, gav dem ett värderingsverktyg som underlättade värderingen av denna svårbedömda förmåga.

I ljuset av den fjärde kvadrantens dilemma kan skillnaden också tolkas som två olika reträtter. HLP-lärarnas lojalitet med de mättraditionella processreglerande direktiven innebär att de retirerar epistemologiskt. De ef-

²⁹³ Se t.ex. Grant (2003).

terfrågar mättraditionella verktyg som genererar mätsäkra resultat samtidigt som de distanserar sig från delar av 1994/2000 års HT-traditionella domänlogik. KP-lärarna retirerar metodologiskt när de distanserar sig från KP:s mättraditionella processreglerande direktiv. De är däremot lojala mot 2011 års HT-traditionella domänlogik och de efterfrågar bedömningsverktyg som kan utgöra ett stöd för hur denna domänlogik kan tolkas och operationaliseras.

7.3.3 Den tentativa slutsatsen

HLP-lärarna beskriver en standardiserad observation i linje med HLP:s mättraditionella direktiv. Resultatet stärker den tentativa slutsatsen att HLP-lärarna i första hand uppfattar HLP som ett resultatgenererande bedömningskulturellt verktyg och inte som ett budskapsförmedlande historiekulturellt verktyg.

Till skillnad från HLP-lärarna beskriver tre av KP-lärarna en situerad observation som inte ligger i linje med KP:s mättraditionella direktiv. De situerar observationen genom att välja ut enskilda uppgifter från KP som de sedan integrerar i sin egen bedömningspraktik. Valet av enskilda uppgifter indikerar att de fokuserar på förmågor som inte ligger i linje med deras egen domänlogik, företrädesvis uppgifter som handlar om historiebruk. Resultatet stärker den tentativa slutsatsen att KP-lärarna i första hand uppfattar KP som ett budskapsförmedlande historiekulturellt verktyg och inte som ett resultatgenererande bedömningskulturellt verktyg.

8. ANVÄNDNINGEN

I det här empiriska kapitlet behandlas historielärarnas beskrivningar av hur de **använder** de frivilliga externa proven. Hur ett externt prov är tänkt att användas regleras i provets syftesreglerande direktiv. Kapitlets disposition är i stora drag densamma som i föregående kapitel. Inledningsvis positioneras provens syftesreglerande direktiv i relation till de bedömningskulturella traditioner som har redovisats tidigare. Därefter positioneras historielärarnas beskrivningar av hur de i realiteten använder proven, vilket innebär att skillnader mellan HLP-lärarna och KP-lärarna framträder. I kapitlets avslutande resultatdiskussion kommer dessa bedömningskulturella skillnader att sättas i ett historiekulturellt sammanhang samt relateras till den fjärde kvadrantens dilemma.

8.1 HLP

8.1.1 De syftesreglerande direktiven

HLP saknade explicita syftesreglerande direktiv. Det framkommer därför inte om provets huvudsakliga syfte i första hand var summativt eller formativt. I praktiken var det således ämneslagen eller den enskilde läraren som avgjorde hur provet skulle användas. Frånvaron av syftesreglerande direktiv skulle kunna uppfattas som att provet var tolkningstraditionellt förankrat, men det skulle även kunna vara en följd av att syftet uppfattades som givet och därför inte behövdes tydliggöras i en explicit syftesformulering.

Eftersom HLP saknade en explicit syftesformulering inleds redovisningen av HLP-lärarnas utsagor med ett avsnitt om deras uppfattningar om provets intentionella syfte, det vill säga vilket syfte de uppfattar att HLP-konstruktörerna har haft när de konstruerade provet.

8.1.2 Uppfattningar om provets intentionella syfte

HLP-lärarna uppfattningar om provets intentionella syfte varierar. Helena uttrycker sig på följande sätt:

Syftet uppfattade jag att vi skulle få på något vis någon slags likvärdighet i betygen huvudsakligen ... inte så mycket konsensus kring något innehåll, men en slags likvärdig bedömning.²⁹⁴

Helena betonar att det intentionella syftet inte handlade så mycket om ”konsensus kring något innehåll” utan att det var likvärdigheten vid betygssättningen som var det huvudsakliga syftet. Hennes uppfattning kan jämföras med Håkan som uttrycker sig på följande sätt om HLP:s intentionella syfte:

Jag: Hur uppfattade du bakgrunden, intentionerna, till att det blev ett prov?

Håkan: Man ville ha ... alltså historia har länge liksom famlat i mörkret ... olika lärare har haft med sig olika käpphästar ... några har varit jätteduktiga på första världskriget, andra på amerikanska befrielsekriget och så vidare och så vidare./.../ och det har inte funnits någon riktig kanon över vad är det egentligen som vi ska peka på /.../ ändå så har folk gått ut med sina MVG eller sina femmor och ingen har vetat vad de har kunnat och inte kunnat ... och det växer fram då ett intresse och veta vad vi sysslar med /.../ och då har det växt fram alltså ... finns det ... skulle vi kunna kanske skapa instrument för att kolla det här ... att vi verkligen ... om vi åtminstone håller en minimum ... eller är någorlunda överens.²⁹⁵

Till skillnad från Helena uppfattar Håkan att det just var en strävan efter konsensus kring innehållet som var det huvudsakliga intentionella syftet med HLP. Håkans tankar kring frånvaron av en kanon som ger historielärare tydliga riktlinjer vad gäller valet av narrativ kan relateras till utformningen av kursplanen för Historia A. Den anger, till skillnad från den nuvarande kursplanen för Historia 1b, inget centralt innehåll, vilket skulle kunna förklara varför valet av narrativ uppfattas som problematiskt av Håkan. I detta sammanhang kan det även vara intressant att notera att HLP:s B-del, en del av provet som prövar specifika substantiella kunskaper, har avdramatiserats genom att eleven endast ska besvara tre av de femton uppgifterna i denna del. Om syftet med provet hade varit att skapa

²⁹⁴ Intervju med Helena, 2009-11-25.

²⁹⁵ Intervju med Håkan, 2010-01-19.

en ökad samsyn bland lärarna gällande vilka narrativ som de ska betona i sin undervisning, så motverkar provets utformning detta syfte. Om eleven hade behövt besvara samtliga uppgifter hade det tvingat fram en diskussion om en gemensam substantiell kanon, och i förlängningen en ökad substantiell samsyn. Provets utformning indikerar således att det intentionella syftet med provet snarare handlar om att skapa en procedural kanon, det vill säga vilka förmågor som ska betonas.

Hanna uppfattar att det intentionella syftet var tvådelat, att det omfattade både likvärdighet och innehåll:

Jag uppfattade att det dels handlade om att man skulle kunna ha det som utgångspunkt för att diskutera likvärdig bedömning och betygs-sättning ... och att man kanske skulle enas om nån sorts stoff som ska ingå på A-kursen, att det finns ju nån form av informell kanon ... och att man kanske skulle erbjuda en likvärdig undervisning, att man ger eleverna samma typ av allmänbildning.²⁹⁶

Vid sidan av likvärdighet och gemensam kanon framkommer ytterligare ett syfte. Henrik uttrycker sig på följande sätt:

Mm ... ja, jag känner jag var lite i utkanten där i början ... men alltså en tanke var väl att höja ämnets status ... och sen var det ju det här ... alltså rent ... att man ville ha någon form av rättvisare betyg ... alltså den diskussionen som kom ... kanske för hela gymnasieskolan och då blev detta ett sätt.²⁹⁷

Henrik lyfter fram ett tredje intentionellt syfte som handlar om historieämnets status. Den underliggande tanken är att genomförandet av ett externt prov i historia under standardiserade former ger ämnet ”större tyngd”, ett starkare signalvärde. Den här variationen gällande uppfattningar om provets intentionella syfte är inte lika uttalad när HLP-lärarnas sedan beskriver hur de själva använder HLP.

²⁹⁶ Intervju med Hanna, 2009-11-19.

²⁹⁷ Intervju med Henrik, 2009-11-26.

8.1.3 Det primära användningsområdet

Samtliga HLP-lärare beskriver att de använder HLP för att generera ett provresultat som de sedan kan använda på olika sätt. Ett genomgående mönster är att det primära användningsområdet för detta provresultat är att använda det som ett underlag vid betygssättningen, det vill säga en summativ användning på elevnivå. När Håkan förklarar för sina elever hur provresultatet kommer att användas vid betygssättningen uttrycker han sig på följande sätt:

Resultatet på detta prov kommer inte påverka ert betyg negativt utan det enda som kan hända är att ni har ytterligare en möjlighet att visa för mig att ni är värda ett högre betyg ... men aldrig så att om ni ligger på ett MVG ska hamna ner på ett VG eller G bara för att ni har skrivit ner er på det provet.²⁹⁸

Håkans sätt att använda provresultatet kan beskrivas som en slags enkelriktad kalibrering av det befintliga betygssunderlaget. I de fall där det finns en positiv diskrepans mellan elevens resultat på HLP och det övriga betygssunderlaget, det vill säga när provbetyget är högre än det tänkta slutbetyget, använder Håkan provresultatet summativt för att motivera det högre betyget. I de fall där det finns en negativ diskrepans bortser han däremot från provresultatet i sin betygssättning. Håkans beskrivning av hur han använder provresultatet vid betygssättningen kan jämföras med Hannas beskrivning:

I år så har jag medvetet liksom tagit med det här provet i planeringen och visat upp det för eleverna så de vet att det kommer till våren. Jag har sagt till dem att det är ett prov som alla andra, att det kan både höja och sänka, men det är viktigt prov. Så att det blir ungefär som man använder nationella provet i svenska till exempel, att det är ett viktigt prov, men det är inte det viktigaste ... och det är inget slutprov heller.²⁹⁹

Till skillnad från Håkan så betonar Hanna att provresultatet kan påverka elevernas betyg i båda riktningarna. I sitt resonemang jämför hon sitt sätt

²⁹⁸ Intervju med Håkan, 2010-01-19.

²⁹⁹ Intervju med Hanna, 2009-11-19.

att använda resultatet med de nationella proven. I de dåvarande användningsdirektiven för de nationella proven betonades att provresultatet skulle användas som ett stöd vid betygssättningen. Hannas resonemang ligger dock närmre de nuvarande direktiven som innebär att de nationella provens resultat särskilt ska beaktas vid betygssättningen.

8.1.4 Multifunktionalitet

Vid sidan av den summativa resultat användningen på elevnivå beskriver HLP-lärarna ytterligare användningsområden för provresultatet. Bilden av HLP som ett multifunktionellt prov är tydlig. Håkan är den av HLP-lärarna som är mest kluven inför att använda provresultatet summativt på en elevnivå:

Det är ett frivilligt prov som de som vill gör /.../ om man frivilligt går in i detta så hoppas jag att mina kollegor också tänker så här ... att vi inte använder det för att värdera elevernas prestationer utan värderar huruvida vårt arbetssätt gentemot det här provet är lyckat eller inte lyckat ... alltså så ser jag det.³⁰⁰

Håkan beskriver hur han vill använda provresultaten formativt på lärarnivå, det vill säga som resultatbaserad återkoppling för att förbättra sin egen undervisning. När Håkan förklarar för sina elever varför han använder HLP uttrycker han sig på följande sätt:

Håkan: Alltså, det är mest en utvärdering av min undervisning än vad det är en utvärdering av er själva ... säger jag till dem.

Jag: Är det så du uppfattar provets syfte i första hand?

B: Ja ... tills vidare så anser jag att ..., det är ett frivilligt prov som de som vill gör, alltså de lärare som vill befatta sig med detta. Det behandlar alltså sådana allmänna saker som hur kursen är upplagd och vilka kvaliteter som man ska bedöma ... men det är inte alls ... det tar inte alls hänsyn till gruppen och hur man har arbetat tidigare ... så då är det viktigt att man tar det som ett diagnostiskt prov för sig själv.³⁰¹

³⁰⁰ Intervju med Håkan, 2010-01-19.

³⁰¹ Intervju med Håkan, 2010-01-19.

Genom att använda provresultatet på detta sätt utvärderar Håkan såväl sin undervisningspraktik som sin bedömningspraktik. Om elevernas resultat på HLP är höga indikerar det att hans undervisningspraktik fungerar och om elevernas resultat ligger i linje med hans egna interna resultatutfall indikerar det att hans bedömningspraktik fungerar. Det sistnämnda är viktigt för Håkan:

Det är mycket intressant ... för resultatet ... det totala resultatet förstår du ... gav vid handen att mina elever hade presterat precis vad de hade presterat tidigare ... så jag menar ... här i det här fallet var det inte så att det skiljde sig väldigt mycket från vad jag hade tidigare.³⁰²

Håkan poängterar att det finns en tydlig korrelation mellan HLP-resultaten och hans egna interna resultat, vilket bekräftar att hans bedömningspraktik är väl fungerande.

HLP:s multifunktionalitet framkommer även i Helenas beskrivning av hur hon använder HLP-resultaten i ett summativt syfte på en lärarnivå:

Jag tycker det är intressant att se vad ... har mina elever lärt sig tillräckligt i förhållande till andra elever här på skolan? Det känns lite grann som en kvalitetsmätare ... eller har jag inte lyckats med de här eleverna fullt ut och så. Jag gör det inte för att se om mina elever uppnår någon slags topprestationer, men det känns bra om de fixar det här provet för då känner jag liksom att då har de nog fått med sig det som nog är viktigt i kursen.³⁰³

Genom att jämföra sina egna elevgruppers resultat med andra lärares elevgrupper får Helena en bekräftelse på sin egen undervisningskompetens: ”Det känns lite grann som en kvalitetsmätare”. Även Håkan beskriver att han och hans kollegor initialt använde HLP-resultaten som ett underlag för att skapa sig en uppfattning om hur deras egna elevers resultat förhöll sig till elevresultaten på andra skolor:

Vad är det för betyg vi sätter ... sätter vi betyg över eller under ... kräver vi mer av våra elever än vad de kräver på andra skolor ... det är det

³⁰² Intervju med Henrik, 2009-11-26.

³⁰³ Intervju med Helena, 2009-11-25.

... det naturliga som man går och svävar i okunnighet när det inte finns någon standard att utgå ifrån /.../ även om det är väldigt subjektivt, men i alla fall ... något mått alltså ... man kan säga ... ja, okej, mina elever skrev si och så ... och det stämmer ganska bra med din bild eller det stämmer inte alls eller ... man kan diskutera kring något resultat, men ... vi hade inget ... så bättre än inget är detta ... detta är bättre än inget.³⁰⁴

Håkan beskriver att det har funnits en ambition på hans skola att använda resultaten summativt på såväl lärarnivå som skolnivå, men att bristen på tid har inneburit att det inte har blivit av. Han beskriver sedan ett exempel där en lärare från en annan skola i samma kommun hade satt MVG på alla elever utom en och att detta hade föranlett en diskussion inom ämneslaget.

Det tillät jag mig liksom att betvivla ... alltså ... och då sa jag helt framt så här att ... just den där klassen hade varit väldigt intressant att jämföra med andra klasser ... i ett sådant prov. Är det verkligen så ... finns det verkligen 29 MVG och en VG i en klass även om det är en NV-klass på skola X? Det blev aldrig så att den klassen eller den läraren som sätter de betygen var med på de här proven, så vi fick aldrig veta ... men det är de grejerna som vi var intresserade av.³⁰⁵

Håkan ifrågasätter bedömningskompetensen hos den här läraren och han vill använda HLP-resultaten för att kontrollera rimligheten i hans betygsättning.

8.1.5 Sammanfattning

HLP saknar syftesreglerande direktiv och HLP-lärarnas uppfattningar om provets intentionella syfte varierar. När de beskriver hur de använder HLP blir bilden mer samstämmig. Det primära användningsområdet är att de använder provresultaten som ett betygsunderlag vid betygssättningen, det vill säga en summativ användning på elevnivå, vilket ligger i linje med den mättraditionella syfteslogiken. I HLP-lärarnas utsagor blir det dock

³⁰⁴ Intervju med Håkan, 2010-01-19.

³⁰⁵ Intervju med Håkan, 2010-01-19.

tydligt att provet är multifunktionellt. Vid sidan av det primära användningsområdet framkommer även sekundära användningsområden som i första hand handlar om resultatbaserad formativ återkoppling på lärarnivå.

8.2 KP

8.2.1 De syftesreglerande direktiven

Till skillnad från HLP fanns det i KP explicita syftesreglerande direktiv. I dessa framkom att KP:s syfte var att ”stödja en likvärdig och rättvis bedömning och betygssättning”. Direktivet klargör att syftet med KP var att generera ett provresultat som sedan skulle användas summativt på en elevnivå. KP:s syftesreglerande direktiv var således tydligt mättraditionellt förankrade.³⁰⁶

8.2.2 Det primära användningsområdet

Trots att det i KP:s direktiv explicit framkommer att provets syfte var att generera provresultat som skulle användas summativt på en elevnivå framkommer det hos tre av KP-lärarna att deras primära användningsområde för KP är något annat. Karin motiverar varför hon använder sig av KP på följande sätt:

Ibland blir man ju sådär osäker på, gör jag rätt, undervisar jag om rätt saker och bedömer jag rätt. Och då tycker jag att proven har blivit en stor trygghet för mig, att nu vet jag att jag har hamnat rätt. Ja, det var till exempel det första provet jag var med och provade ut. Det handlade ju om kalla kriget och då var det en naturklass jag hade där och dom var väl lite så här när jag sa de till dom att de skulle få ett nationellt prov. Men sen när jag fick det, jag visste ju inte alls vad det skulle handla om så kände jag det att ... ja, men det var ju precis rätt ju.³⁰⁷

Karins motivering bygger på hennes reaktion när hon ser provet, inte på det efterföljande resultatutfallet. Provets resultatgenererande syfte är således underordnad provets budskapsförmedlande syfte. Det som utmärker

³⁰⁶ KP:s syftesformulering, så som den skrivs fram i Skolverkets direktiv till provkonstruktörerna, är däremot multifunktionell. Provet ska vara ett summativt stöd för betygssättningen, men det ska även fylla formativa funktioner.

³⁰⁷ Intervju med Karin, 2016-11-08.

det budskapsförmedlande syftet är att resonemangen är giltiga oavsett om provet genomförs eller inte. Karins användning av KP kan beskrivas som en verktygsbaserad formativ användning. Ett annat sätt att uttrycka denna tankegång är att hon använder proven som ett verktyg för att kvalitets-säkra sin tolkning av konstruktet.

I utsagan refererar Karin till en specifik uppgift om kalla kriget som prövade elevens förmåga att hantera olika tolkningar av detta narrativ. Hon uppfattar att hennes val av centralt innehåll och hennes sätt att historiedidaktiskt tolka och hantera detta innehåll har varit "rätt". Karins sätt att referera till KP som "rätt" indikerar att KP:s budskapsförmedlande roll får en normativ prägel för Karin. Hon får en bekräftelse på att hennes tolkning av kursplanens konstrukt är korrekt. Till skillnad från HLP-lärarna så använder sig Karin av provets verktyg mer som en mall än som ett diskussionsunderlag.

Jag har nog blivit, som jag sa innan, blivit tydligare med efter proven ... jag har liksom känt mig bekräftad utav det ... har fått hjälp att konkretisera målen har skrivit här ... så att ... ja, det är en trygghet för mig /.../ Det är lättare att tolka de centrala målen, eller, ja det som Skolverket står i ... när jag har använt mig utav era prov.³⁰⁸

På samma sätt som Karin betonar Krister att det inte är den resultatgenererande funktionen hos KP som i första ligger till grund för hans användning av KP.

Det främsta värdet har ju varit det här didaktiska samtalet, man lär sig av ... hur man ska lägga upp det ... med betoning på begrepp till exempel, kanske mer modern internationell historia, jobba med användning och källkritik på det sättet som det finns bland annat på kursprovet, så att för mig har ju det här varit en fantastisk fortbildning kan man säga nästan ju, att undervisa utifrån Gyll och utifrån kunskaps ... utifrån kursprovet ... fortbildning som lärare som man har ... tycker att jag har haft stor nytta av.³⁰⁹

³⁰⁸ Intervju med Karin, 2016-11-08.

³⁰⁹ Intervju med Krister, 2016-11-24.

Krister uppfattar KP som en möjlighet till fortbildning och att KP:s bedömningsverktyg kan användas för att lära sig ”hur man lägger upp det”. Återigen framkommer KP:s budskapsförmedlande funktion. Det finns hos Krister en oro för att något kan saknas i hans tolkning av kursplanens domänlogik.

Jag vill egentligen ha en ... det låter så fint ... jag vill att de ska lära sig mer, att dom ska få mer kunskap om historia, utöka sitt historiemedvetande, ökad måluppfyllelse ... alltså inte så att jag och mina kollegor, kan prata för mig själv, från 2011 - 2015 då inte har lyckats eller att dom inte lärt sig nånting, men det kanske hela tiden fått en känsla, jag personligen, att det är nånting vi har missat till exempel ... har jag haft för stor betoning på nåt speciellt område, för mig har det kunnat varit tydligare koppling till central innehåll och kunskapskrav och själva grunden för ämnet när man jobbar med kursprovet, det har betytt mycket att jag kan knyta upp mina uppgifter, mina olika block eller studiedelar, kurser som jag har då, till det centrala innehållet, det har blivit en tydlighet i detta.³¹⁰

Det finns hos Krister en ängslan att han har ”missat” någonting. Krister exemplifierar genom att lyfta fram emfaslogiken i det centrala innehållet när han oroar sig för om han har haft ”för stor betoning på nåt speciellt område”.

Även Karl betonar KP:s budskapsförmedlande funktion.

Genom att jag läser frågorna så förstår jag också hur jag ska tänka ... och sen så har vi ett underlag som vi sen kan diskutera likvärdighet och bedömning utifrån, för det har varit det svåra tidigare. Så det har ju egentligen väldigt mycket handlat om lärarnivå. /.../ Då kursprovet kom 2015 så var vi flera stycken som tände direkt ... och på ämnesträffar så ... vi tog några ämnesträffar och gick igenom ... vi diskuterade fråga för fråga ... som var på kursprovet 2015 ... vi plockade ut nåt elevsvar och vi satt och gnuggade det.³¹¹

³¹⁰ Intervju med Krister, 2016-11-24.

³¹¹ Intervju med Karl, 2016-12-06.

Karl är tydlig när säger att uppgifterna i KP ger honom en förståelse för ”hur jag ska tänka”. Till skillnad från Karin och Krister är han noga med att detta gäller både domänlogiken och progressionslogiken. I sitt fortsatta resonemang konkretiserar han detta genom att fokusera på elevsvaren:

Om jag hade tittat tillräckligt många gånger på ett antal elevsvar ... så jag hade förstätt hur ... Skolverket vill att jag ska tänka ... förstår du vad jag menar.³¹²

På samma sätt som de övriga KP-lärarna resonerar Karl om provet som fortbildning. Det underlättar tolkningen av vad Skolverket vill. Även Karl betonar således provets budskapsförmedlande syfte och ikläder sig rollen som budskapsmottagare.

8.2.3 Multifunktionalitet

Även hos KP-lärarna framkommer det externa provets multifunktionalitet. Tydligast framkommer det hos Kristina som betonar att KP fyller multipla syften, både summativa och formativa. Hon urskiljer tre olika användningsområden. Till skillnad från de övriga KP-lärarna så är den summativa användningen grundläggande för Kristina:

Jag vet att vi sätter olika betyg på olika saker, det är väldigt stor bredd ... det vet jag ... och jag tycker inte att det är rätt ... så att det är ju rättvisepatoset ligger väl i botten där ... men sen visste jag ju också att det skulle ge dom andra effekterna.³¹³

De två andra ”effekter” som Kristina resonerar kring benämner hon som samarbete och fortbildning:

Man skulle kunna, så att säga, tvinga oss kollegor inom historieämnet att faktiskt samarbeta kring en gemensam uppgift, jag försökte väldigt länge, innan kursprovet fanns, att hitta en gemensam uppgift, men det var ingenting som jag kunde lyckas med /.../ Jag visste om vi skulle kunna samarbeta kring gemensamma uppgift, och titta vad eleverna

³¹² Intervju med Karl, 2016-12-06.

³¹³ Intervju med Kristina, 2016-11-21.

och se på resultatet, så skulle det ofel ..., alltså ofelbart bli lite fortbildning på köpet, och det var också ett mål, så ... rättvis bedömning, det skulle skapa samarbete och det skulle skapa fortbildning ... det var, det var som var min tanke.³¹⁴

Samarbetet har en mer metodologisk dimension medan fortbildningen har en mer epistemologisk dimension.

Även de tre lärare som betonar den verktygsbaserade användningen beskriver att de i någon form använder KP-resultaten i sin undervisningspraktik. Flera av KP-lärarna uttrycker att de använder KP-resultat som ett underlag vid betygssättningen, det vill säga i ett summativt syfte på elevnivå.

Krister, den ende av KP-lärarna som använder provet i sin helhet, uttrycker sig på följande sätt när han förklarar för sina elever hur han använder resultaten från KP:

Det här är ett kursprov, det sammanfattar hela kursen och skulle ni misslyckas kapitalt, du har legat kanske på C så är det ju en tankeställare onekligen från min sida, så att det här kommer att vara en viktig del av betygssättningen, men inte det som avgör. /.../ Det här sänker inte dig, det kan egentligen bara lyfta dig, men det är ett mycket viktigt prov, det kan ända visa fingervisning om hur det kommer att gå på kursbetyget.³¹⁵

Krister betonar att KP är ett viktigt prov, men på samma sätt som HLP-läraren Henrik så använder han provresultatet enkelriktat. Karin, som genomförde delar av provet som en examinerande avslutning på ett internt utformat delmoment, noterar att fanns en positiv diskrepans mellan resultatutfallet på KP-uppgifterna och resultaten från hennes internt utformade uppgifter. Hennes sätt att använda KP-resultaten blir därför per automatik enkelriktat.

Ja, jag tror att de har fått högre betyg på dom här proven än de får i slutbetyg i regel ... men ... alltså dom, jag vet inte, jag har väl varit ... känt, jag har fått lätta lite kanske på mina krav efter proven, för att de

³¹⁴ Intervju med Kristina, 2016-11-21.

³¹⁵ Intervju med Krister, 2016-11-24.

når ganska höga betyg på proven och sen så har jag känt att ja men så här höga har jag ju inte satt innan, men ... och satt ... det har nog påverkat mer till elevernas fördel.³¹⁶

Som framkommit tidigare så använder även Kristina utvalda uppgifter från KP som en slags slutexamination. Hon beskriver att hon 2015 använde sig av två uppgifter från KP som prövade källhantering respektive historiebruk. I sin kommunikation med eleverna tydliggör hon att det är utfallet på dessa uppgifter som är direkt avgörande för deras betyg när det gäller dessa båda centrala förmågor. Examinationens summativa syfte är uttalat och insatserna är höga (high stake), vilket ligger i linje med de mättraditionella idealen.

Även Karl uppfattar att han använder KP i ett summativt syfte. Karl konkretiserar sitt resonemang med ett exempel där han använder KP-resultat summativt, som ett stöd vid något som kan liknas vid en betygsförhandling.

Jag hade en elev som gjorde väldigt bra ifrån sig vid nått tillfälle och sen skrev väldigt dåligt vid flera av dom här ... [KP-uppgifterna]. Det var en uppgift som handlade om den sovjetiska kollektiviseringen där man mötte ett antal källor som handlade om vilka konsekvenser den sovjetiska kollektiviseringen fick och sen var det den här med Blå tåget och Willy Brandt använder historia ... och hon gjorde ju riktigt kasst där ... och hon ville ju ha ett A i historia ... och då kom jag ihåg att ... du får inget A i historia. Alltså det fanns flera tillfällen i samband med de här exempeluppgifterna ... (ohörbart) ... riktigt svagt ifrån sig ... väldigt verbal tjej, kan prata mycket och tala för sig i klassrummet. I dom här sammanhangen så visade hon ingenting som var på en A-nivå ... så jag kommer inte att sätta ... du fick ett C i historia. Det var ett exempel och så modig hade jag nog inte vågat vara annars kanske. Det gav mig lite råg i ryggen just att det var Skolverkets material.³¹⁷

Notera att Karl låter KP-resultaten väga tungt vid betygssättningen och att de ger honom ”lite råg i ryggen” att stå på sig i förhandlingen eftersom de är resultat genererade från verktyg framtagna av Skolverket.

³¹⁶ Intervju med Karin, 2016-11-08.

³¹⁷ Intervju med Karl 2016-12-06.

Även KP-lärarna använder således provresultaten i ett summativt syfte på elevnivå, vilket ligger i linje med mättraditionen, men till skillnad från HLP-lärarna är dock detta syfte sekundärt i relation till det formativa syftet. KP-lärarna beskriver på samma sätt som HLP-lärarna att provresultaten används summativt vid betygssättningen, men i betydligt högre grad än HLP-lärarna betonar KP-lärarna KP:s formativa syfte. Ett exempel är Krister som uttrycker sig på följande sätt:

Det ska vara nånting som ger ett stöd till mig som lärare, rent summativt. Här har vi Katrin: ”Krister, jag vill ha ett A i betyg”. Jag: ”Titta, du fick ett E på kursprovet”. Men det ska ändå kunna finnas nån form av formativt inslag. Jag vill inte att det ska bli ... näst sista lektionen, prov, och sen bara visa resultatet. Då blir det för trubbigt.³¹⁸

Krister använder resultatet summativt, men han betonar att resultatet även används i ett formativt syfte. Flera av KP-lärarna beskriver hur de använder provresultat som ett underlag för återkoppling. Genom återkopplingen tydliggörs relationen mellan befintliga kunskaper/förmågor som de kommer till uttryck i resultatet och eftersträfvansvärda kunskaper/förmågor som de beskrivs i konstruktet

I mötet med empirin framkom ytterligare ett syfte. Krister resonerar på följande sätt:

Legitimiteten har ökat faktiskt [...] för det jag säger ... titta här Christian ... du fick ett E på detta ... här ser du, så här, ... det är lite som ett nationellt prov på nått sätt, vilket jag aldrig varit med om, men att ... eleverna har varit väldigt positiva till det, för det har blivit nån sorts objektivitetsaspekt ... nu är det inte bara Krister som hittar på utan nu är det faktiskt Skolverkets prov här.³¹⁹

Krister använder KP för att legitimera sin egen bedömningspraktik. Han gör detta genom att relatera KP-resultaten till tidigare, internt genererade, resultat: ”Nu är det inte bara Krister som hittar på utan nu är det faktiskt

³¹⁸ Intervju med Krister, 2016-11-24.

³¹⁹ Intervju med Krister, 2016-11-24.

Skolverkets prov här”. Den här legitimerande funktionen har uppmärksamats tidigare av forskare som intresserat sig för utvärderingar.³²⁰

8.2.4 Sammanfattning

KP:s syftesreglerande direktiv var mättraditionellt förankrade och klargör tydligt att syftet med provet var att generera ett provresultat som skulle användas summativt på elevnivå. Detta till trots beskriver tre av KP-lärarna att de i första hand använder provet i ett formativt syfte. Det är dock inte ett resultatbaserat återkopplande formativt syfte som betonas utan ett verktygsbaserat budskapsförmedlande formativt syfte. KP-lärarna är således i högre grad än HLP-lärarna fokuserade på verktygen i sig än på de resultat som dessa verktyg genererar. Även hos KP-lärarna framkommer dock att provet är multifunktionellt. Vid sidan av det primära verktygsbaserade användningsområdet framkommer även sekundära resultatbaserade användningsområden.

8.3 Resultatdiskussion

8.3.1 Summativ eller formativ användning?

Historielärarna i studien anger genomgående flera olika anledningar till varför de använder sig av ett frivilligt externt prov. Proven är således multifunktionella. Den grundläggande skillnaden mellan HLP-lärarna och KP-lärarna är dock att de förra i första hand beskriver en resultatbaserad summativ användning på elevnivå, medan de senare i första hand beskriver en verktygsbaserad formativ användning på lärarnivå.

Utgångspunkten för det fortsatta resonemanget är att lärarkollektiven använder sig av respektive prov för att hantera olika problem. HLP-lärarna använder HLP-resultaten summativt på elevnivå för att hantera ett likvärdighetsproblem medan KP-lärarna använder KP-verktygen formativt på en lärarnivå för att hantera ett alignmentproblem.

8.3.2 Från likvärdighetsproblem till alignmentproblem

Likvärdighetsproblemet

Likvärdighetsproblemet är i grunden ett bedömningskulturellt problem som handlar om att lärares värdering av kunskaper och färdigheter inte

³²⁰ Se t.ex. Vedung, som särskiljer sex olika syften med utvärderingar, däribland det legitimerande syftet.

sker på likvärdigt sätt, det vill säga att en och samma elevprestation värderas olika beroende på vem som värderar prestationen.

Även om likvärdighetsproblemet i grunden är ett bedömningskulturellt problem kan det uppfattas och hanteras på olika sätt. Om det uppfattas som ett metodologiskt problem ligger fokus på värderingsverktygens format och de processreglerande värderingsdirektiven. Om det däremot uppfattas som ett epistemologiskt problem riktas fokus mot konstruktet i allmänhet och mot progressionslogiken i synnerhet.

I spåren av likvärdighetsproblematiken följer därför både en bedömningskulturell metodologisk och en historiekulturell epistemologisk förhandling. Den bedömningskulturella metodologiska diskussionen kan till exempel handla om behovet av en ökad användning av mätsäkra observationsverktyg i form av flervalsuppgifter eller behovet av externa standardiserade värderingsprocesser. Den historiekulturella epistemologiska förhandlingen kan till exempel handla om behovet av en kunskapstaxonomisk tankereda eller nödvändigheten av att kvantifiera progressionsindikatorerna. Oavsett om likvärdighetsproblematiken uppfattas som ett metodologiskt eller epistemologiskt problem innebär hanteringen av problemet att på olika sätt begränsa bedömarens tolkningsutrymme i samband med bedömningen.

Alignmentproblemet

Alignmentproblemet är i grunden ett historiekulturellt problem som handlar om att det finns en diskrepans mellan historielärarnas egen domänlogik och den domänlogik som genomsyrar skolans styrdokument.

När alignmentproblemet utgör det behov som det externa provet svarar mot och ges tolkningsföreträdare tonas provens datagenerande funktion ner och proven får en tydligare budskapsförmedlande funktion. Provens bedömningsverktyg konkretiserar det konstrukt som ligger till grund för styrdokumentet och de tjänar som ett mer eller mindre normativt diskussionsunderlag för lärarna. Det externa provet blir en del av lärarnas professionsutveckling, vilket innebär att det är provets budskapsförmedlande funktion som hamnar i fokus.

I spåren av alignmentproblematiken följer i första hand en historiekulturell förhandling eller förmedling om konstruktets domänlogik. Det kan handla om behovet av en substantiell kanon (till exempel antikens vara

eller inte vara), men det kan också handla om behovet av en procedural vägledning (till exempel vad förmågan att använda historia innebär).

Förändring över tid i ljust av den fjärde kvadrantens dilemma

HLP-lärarna använde HLP-resultaten för att säkerställa en likvärdig bedömning. KP-lärarna använde KP:s bedömningsverktyg för att säkerställa att deras tolkningar av domänlogiken i Gyll låg i linje med Skolverkets intentioner. Det sker således en fokusförskjutning från likvärdighet till alignment. I ljust av den fjärde kvadrantens dilemma kan denna fokusförskjutning tolkas som en fråga om problemformuleringsprivilegium och tolkningsföreträde. HLP-lärarnas fokus på bedömningens metodologiska dimension, manifesterad i likvärdighetsproblemet, innebär att reliabiliteten ges tolkningsföreträde framför validiteten. KP-lärarnas fokus på bedömningens epistemologiska dimension, manifesterad i alignmentproblemet, innebär att validiteten ges tolkningsföreträde framför reliabiliteten. Den här fokusförskjutningen kommer att behandlas mer utförligt i slutdiskussionen.

8.3.3 Den tentativa slutsatsen

Både HLP-lärarna och KP-lärarna uppfattar att provet fyller multipla syften, men HLP-lärarna tenderar att i första hand använda HLP-resultat i ett summativt syfte för att hantera ett likvärdighetsproblem. Det stärker den tentativa slutsatsen att HLP-lärarna i första hand uppfattar HLP som ett resultatgenererande bedömningskulturellt verktyg och inte som ett budskapsförmedlande historiekulturellt verktyg.

KP-lärarna tenderar att i första hand använda KP-verktyg i ett formativt syfte för att hantera ett alignment-problem. Resultatet stärker den tentativa slutsatsen att KP-lärarna i första hand uppfattar KP som ett budskapsförmedlande historiekulturellt verktyg och inte som ett resultatgenererande bedömningskulturellt verktyg.

9. PÅVERKAN

9.1 Påverkansområden

I det här avslutande empiriska kapitlet redovisas hur historielärarna uppfattar att respektive prov **påverkar** deras bedömningspraktik. Kapitlet inrymmer på samma sätt som de tre andra empiriska kapitlen en resultatredovisande del och en resultatdiskuterande del där resultaten analyseras och relateras till den fjärde kvadrantens dilemma. Utöver dessa delar innehåller kapitlet även ett avsnitt där tidigare forskning med särskild relevans för studien redovisas.

I den resultatredovisande delen särskiljer jag mellan epistemologisk påverkan, metodologisk påverkan och kollegial påverkan. Med epistemologisk påverkan avses påverkan på bedömningspraktikens *vad*-dimension, det vill säga konstruktet i bred bemärkelse. Med metodologisk påverkan avses påverkan på bedömningspraktikens *hur*-dimension, det vill säga bedömningens format- och processlogik. Med kollegial påverkan avses påverkan på den kollegiala interaktionen, det vill säga de samtal och samarbeten kring bedömning som äger rum ämneskollegor emellan.³²¹

Under analysarbetet har det framträtt fem dominerande teman eller påverkansområden: tre epistemologiska, ett metodologiskt och ett kollegialt. Dessa presenteras i tabellform nedan (se tabell 5).

³²¹ Ett närbesläktat, men bredare begrepp, är lärarkultur. Se Letitia Fickel (2006), s. 76. I ett svenskt sammanhang utgör ämneslaget den kollektiva enheten för dessa samtal.

Tabell 5: Påverkansområden

Epistemologisk påverkan	Metodologisk påverkan	Kollegial påverkan
Konstrukt	Format	Interaktion
Domänlogikens substantiella dimension	Observationsverktygets format	Kollegiala samtal och samarbeten
Domänlogikens procedurala dimension		
Progressionslogiken		

Som framkommer i tabellen omfattar de tre epistemologiska påverkansområdena den substantiella dimensionen av domänlogiken, det vill säga huruvida det externa provet påverkar historielärarnas val av narrativ (stoffurval); den procedurala dimensionen, det vill säga påverkan på vilka förmågor som betonas i undervisning och bedömning; och progressionslogiken, det vill säga huruvida det externa provet påverkar historielärarnas uppfattningar om hur kunskaper i historia ska värderas.

Det metodologiska påverkansområdet omfattar påverkan på observationsverktygens format, det vill säga lärarnas sätt att konstruera uppgifter.

Det femte påverkansområdet, kollegial interaktion, kan brytas ner i en teoretisk och en praktisk dimension. Lärare kan interagera på ett teoretiskt plan genom att samtala om bedömning, men de kan också interagera på ett praktiskt plan genom att samarbeta i samband med bedömning, till exempel genom att samvärdera eller genom att konstruera ett prov tillsammans. I resultatredovisningen kommer jag dock inte att särskilja mellan kollegiala samtal och kollegiala samarbeten.³²²

9.2 Tidigare forskning

I det här avsnittet kommer jag att uppmärksamma tidigare forskning som kan kopplas till påverkansområdena. I redovisningen av dessa studier

³²² Den kollegiala interaktionen inrymmer dels en kvantitativ aspekt som handlar om hur ofta och hur länge lärare interagerar, dels en kvalitativ aspekt som handlar om interaktions innehåll, det vill säga vilka aspekter av bedömningspraktiken som interaktionen handlar om. Det empiriska materialet ger mig främst underlag för att behandla den kvalitativa aspekten av provens påverkan på den kollegiala interaktionen.

finns det genomgående en gråzon huruvida det externa provets påverkanseffekt avser bedömningspraktiken eller undervisningspraktiken. Denna gråzon beror på att de flesta av studierna inte explicit särskiljer dessa båda praktiker från varandra. Som framkommit tidigare i det övergripande kapitlet om tidigare forskning så är bedömningspraktiken i många empiriska studier ofta endast en av flera aspekter i en större undersökning med fokus på undervisningspraktiken.

9.1.1 Epistemologisk påverkan

Påverkan på domänlogiken

Tidigare forskning har i hög grad uppmärksammat att externa prov i historia tenderar att påverka den substantiella dimensionen av historielärares domänlogik. S. G. Grant sammanfattar utifrån ett amerikanskt perspektiv sina intryck på följande sätt:

Although untangling the relative influence of state curricula from state exams can be difficult, the principal pedagogical effect of state history tests seems to be on teachers' content decisions. Most of the teachers profiled in these chapters report making a range of small to large changes in the subject matters ideas they teach.³²³

Grants egen forskning nyanserar dock bilden av att de externa proven i huvudsak påverkar den substantiella dimensionen av historielärares domänlogik. I en undersökning med fokus på New York Regents Exams (NYRE) kommenterar Paula, en ung historielärare, hur hon uppfattar att NYRE påverkar hennes sätt att använda sig av dokument i sin undervisning.

[...] The way I use documents is for kids to make interpretations of them, of the documents, based on their knowledge of the topic. And that means that there may be a very wide variety of answers to whatever the question is. When I use them how I think the state wants to use them, there is probably not a wide variety of answers to the question. There's probably a pretty specific answer to the question that goes with the document that I'm presenting them.³²⁴

³²³ S. G. Grant (2003), s. 307.

³²⁴ Grant (2006), s. 40.

Grant menar att Paulas sätt att använda dokument i sin undervisningspraktik inte ligger i linje med provets intentioner och att hon därför påverkas till att justera sin undervisningspraktik.³²⁵

Paula has created, in effect, two teaching approaches: One is explicitly state curriculum and test-directed, while the other is explicitly self-directed. [...] Instead of crafting an instructional plan that coherently promotes ambitious teaching and learning, then, Paula constructs a wobbly mix of the rich and the dull, the engaging and the pedantic.³²⁶

Grants analys av Paulas erfarenheter är intressanta. Han konstaterar att oerfarna historielärare tenderar att justera sin undervisningspraktik genom ett slags tudelning. Han menar dock att även oerfarna lärare med en tydlig ämnesuppfattning kan undvika att kapitulera inför ett externt prov. I en tidigare studie använder han sig av begreppet *ambitious teachers* för att beskriva hur lärare med en djupt rotad ämnessyn fortsätter att undervisa på sitt eget sätt trots att provet ger andra budskap.³²⁷

Även Jill Gradwell har i sin forskning intresserat sig för de externa provens påverkan på den procedurala dimensionen av historielärares domänlogik. Till skillnad från Grant har hon i dessa studier inte funnit någon tydlig påverkan. Samtidigt konstaterar hon dock att andra undersökningar har funnit en påverkan till det ”bättre”, det vill säga att historielärare uppfattar att de externa proven påverkar dem att i högre utsträckning uppmärksamma mer komplexa procedurala förmågor.³²⁸

Påverkan på progressionslogiken

I min licentiat-uppsats uppmärksammar jag den ”dubbla värderingsprocessen” i samband med värderingen av de finländska ämnesrealproven. Den dubbla värderingsprocessen innebär att den undervisande läraren gör en initial värdering och att externa censorer sedan gör en slutlig värdering. I studien framkommer att de finländska historielärarna uppfattar det som

³²⁵ I praktiken innebär Paulas resonemang en fokusförskjutning från en ämnesuppfattning förankrad i en tolkningsbaserad konstruktivism mot en mer objektivistiskt förankrad ämnesuppfattning.

³²⁶ Grant (2006), s. 40–41.

³²⁷ Grant (2003).

³²⁸ Gradwell (2006), s. 158.

viktigt att deras initiala värdering av elevernas prestationer på ämnesrealprovet ligger i linje med censorernas slutliga värdering.³²⁹

Det finländska exemplet visar betydelsen av att involvera lärarna i värderingsprocessen. Den dubbla värderingsprocessen skapar förutsättningar för en fortlöpande dialog mellan de externa aktörerna och historielärarna. I en kanadensisk undersökning framkommer att en sådan interaktion i samband med värdering av provsvar kan få många positiva följder.³³⁰ Lärarna i undersökningen uppfattade att deras involvering i värderingsprocessen av ett externt prov bidrog till att klargöra deras pedagogiska grunduppfattning, bekräfta eller förbättra deras undervisnings- och bedömningspraktik samt stärka deras uppfattning av läraryrket som en profession.³³¹

Caroline Gipps har i en engelsk kontext uppmärksammat vad hon benämner som group moderation, det vill säga när en grupp lärare, med eller utan externa experter, kommer samman för att diskutera bedömning. Hon menar att det är en process som i högsta grad stärker ämnesutvecklingen och som dessutom får återverkningar för lärares faktiska undervisning.³³²

9.2.2 Metodologisk påverkan

Påverkan på observationsverktygets format

I min licentiat-uppsats uppmärksammar jag även det finländska ämnesrealprovets påverkanseffekt på de finländska lärarnas sätt att konstruera observationsverktyg.³³³ Studien visar att provet i hög grad påverkar deras sätt att konstruera provuppgifter. I övrigt har jag inte hittat några studier som berör detta påverkansområde. Frånvaron av relevanta forskningsresultat ligger i linje med den forskningsöversikt som redovisats tidigare. Tidigare forskningsresultat visar i stort sett samstämmigt att det framför allt är den epistemologiska dimensionen av historielärares undervisningspraktik som påverkas av externa prov, det vill säga *vad* historielärarna

³²⁹ Gunnemyr (2011).

³³⁰ I Kanada är det lärare som konstruerar proven och värderar provsvaren. I t.ex. British Columbia är det mer än 300 lärare som utvecklar provfrågorna och fler än 1600 lärare som är involverade i värderingen av provsvaren varje år. Den kanadensiska modellen skiljer sig både från den amerikanska modellen där proven i huvudsak konstrueras och värderas av kommersiella aktörer utanför det politiska systemet och från den engelska modellen där inflytandet ligger hos universiteten. Se Gambell & Hunter (2004), s. 699.

³³¹ Gambell & Hunter (2004), s. 697.

³³² Gipps (1994), s. 117–122.

³³³ Gunnemyr (2011).

undervisar om. Hon menar att väldigt få undersökningar har visat på någon påverkan inom den metodologiska dimensionen, det vill säga *hur* lärare undervisar.³³⁴

9.2.3 Kollegial påverkan

Påverkan på den kollegiala interaktionen

Generellt sett kan tidigare forskning om den kollegiala interaktionen kopplas till begreppet lärarsocialisation. Det finns en omfattande forskning som visar att för lärare utgör ämneskollegorna och ämneslaget den viktigaste organisatoriska påverkansfaktorn i lärares identitetsbygge.³³⁵ Susan Stodolsky och Pamela Grossman beskriver framväxten av ämnesspecifika subkulturer inom ett ämneslag. Lärare identifierar sig med sina ämneskollegor och tar till sig deras uppfattningar om vad som är kunskap och hur man på bästa sätt ska undervisa om dessa kunskaper. Den här starka identifikationen med ämneskollegor kan förklaras av att lärare utbildas ämnesvis, att de organiseras i ämneslag och att de engagerar sig i ämnesföreningar.³³⁶

I en amerikansk kontext visar tidigare forskning att ämnesspecifik kollegial interaktion har betydligt större potential att leda till att lärares bedömnings- och undervisningspraktik utvecklas jämfört med en mer ämnesövergripande kollegial interaktion. Lärares bedömnings- och undervisningspraktik utvecklas när de ges möjlighet att tillsammans med ämneskollegor utgå från problem som de uppfattar som relevanta och som de på ett direkt sätt kan relatera till sin egen verksamhet. Vidare tenderar de förändringar som följer av ämnesspecifik kollegial interaktion att både bli mer djupgående, det vill säga att lärares ämnes- och kunskapssyn förändras i grunden, och bli mer bestående över tid. Förändringarna blir något mer än bara metodologiska tips och handgrepp.³³⁷

Tidigare forskning gällande hur externa prov påverkar historielärares kollegiala interaktion är begränsad. Den enda vetenskapliga forskning jag har påträffat, som uttryckligen undersöker hur den kollegiala interaktionen hos historielärare påverkas när ett externt prov implementeras, är

³³⁴ Cimbricz (2002), s. 1–21.

³³⁵ För en forskningsöversikt, se Fickel (2006), s. 77.

³³⁶ Grossman & Stodolsky (1995), s. 5–11.

³³⁷ För en forskningsöversikt, se Quinlan & Åkerlind (2000), s. 25–27.

en studie av Letitia Fickel.³³⁸ Hennes studie beskriver ett ämneslag i Kentucky. Det är ett väl fungerande ämneslag med en samsyn kring syftet med historieundervisningen. I studien framkommer att den kollegiala interaktionen initialt intensifierades i samband med att nya styrdokument och ett nytt externt prov implementerades i Kentucky. Interaktionen medförde bland annat att historielärarna samordnade stoffurvalet i högre utsträckning än tidigare och att lärarna utarbetade en helt ny kurs för att deras elever skulle komma bättre förberedda till provet. De första åren efter implementeringen visade eleverna goda resultat och kollegiets arbetssätt bekräftades. Sedan sker en markant försämring av elevernas resultat och då förändrades kollegiets strategi. Fokus flyttas från styrdokumenten till provresultatet; från undervisningspraktiken till bedömningspraktiken, vilket bland annat medförde att lärarna i högre utsträckning än tidigare började samtala och samarbeta om hur provfrågor skulle konstrueras och hur provsvarens kvalitet skulle värderas. Huvudfrågan blev hur man ska lära eleverna att svara på de uppgifter som man trodde att provet skulle innehålla. Det handlar om att vänja eleverna vid en viss slags uppgifter och i förlängningen om en utveckling mot konformitet gällande värderingen. Fickels studie visar att det externa provets betydelse för hur den kollegiala interaktionen påverkas kan förändras över tid. Den kollegiala interaktionen intensifierades i samband med att ett externt prov implementerades.³³⁹

9.2.4 Tidigare forskning i relation till föreliggande studie

I redovisningen av tidigare forskning är det påtagligt hur gråzonen mellan undervisningspraktik- och bedömningspraktik framträder. Som tidigare nämnts är det i första hand en forskningsmetodologisk konsekvens, men gråzonen kan även belysas utifrån ett teoretiskt perspektiv. I det teoretiska ramverket driver jag tesen att domänlogiken utgör progressionslogikens fundament, vilket i praktiken innebär att den epistemologiskt baserade gränsdragningen mellan undervisning och bedömning är flytande.

Av redovisningen framgår vidare att det framför allt är påverkanseffekten på domänlogiken som har uppmärksamrats inom fältet. Av särskilt intresse för min forskning är dock Fickels studie gällande påverkan

³³⁸ Fickel (2006).

³³⁹ Fickel (2006), s. 75–100.

på den kollegiala interaktionen. Studien inrymmer liksom min en temporal dimension och hennes resultat att den kollegiala interaktionen förändras över tid i samband med att elevresultaten försämrats har en direkt koppling till mitt syfte.

En mera övergripande diskussion där jag ställer mina resultat mot tidigare forskning kommer att föras i det avslutande kapitlet, se avsnitt 10.3.

9.3 HLP

9.3.1 Epistemologisk påverkan

Påverkan på domänlogiken – den substantiella dimensionen

HLP-lärarnas uppfattningar om huruvida deras val av narrativ påverkas av HLP är inte samstämmiga. Henrik uppfattar att hans val av narrativ inte påverkas över huvud taget. Två av de andra HLP-lärarna, Helena och Hanna, uppfattar att deras val av narrativ påverkas i viss utsträckning. Helena uttrycker sig på följande sätt:

Lite gör det ju det på något vis. Man tänker alltid liksom att man kan inte lika mycket göra som man själv vill liksom, utan det är ju på något vis att man känner att får de nu det här provet så kan de ju inte känna att det här har vi inte gått igenom.³⁴⁰

Håkan uppfattar, till skillnad från de övriga tre HLP-lärarna, att hans val av narrativ påverkas i hög utsträckning.

Ja, jag påverkas naturligtvis eftersom jag, inte vill ge mindre chanser till mina elever än vad andra elever får och då måste de få chansen. Och då berättar jag klart och tydligt för dem att de och de områdena måste vi gå igenom. Så även om vi tidigare sa att vi kanske kursivt skulle gå över det så gör vi inte det nu. Vi måste verkligen lära oss de här områdena, det är viktigt. För jag menar om vi ska kunna göra rätt för oss i slutet på ett prov så måste ni i alla fall ha läst igenom det.³⁴¹

Håkans uppfattning kan relateras till att han, till skillnad från de övriga HLP-lärarna, väljer ut vilka undervisningsgrupper som ska genomföra

³⁴⁰ Intervju med Helena, 2009-11-25.

³⁴¹ Intervju med Håkan, 2010-01-19.

provet. För att eleverna i dessa grupper ska lyckas på provet ändrar Håkan sitt vanliga val av narrativ.

Påverkan på domänlogiken – den procedurala dimensionen

Den variation som framkom gällande den substantiella dimensionen återkommer i HLP-lärarnas uppfattningar om HLP:s påverkan på den procedurala dimensionen av deras domänlogik. Henrik uppfattar som tidigare att han inte påverkas över huvud taget, medan Helena och Hanna uppfattar att HLP har påverkat dem i viss utsträckning. Båda uppfattar att det framför allt är deras sätt att hantera den källkritiska förmågan som har påverkats. Hanna uttrycker sig på följande sätt:

Den källkritiska delen lägger jag ganska mycket tid på och det finns ju många lärare som inte gör det alls nästan, som inte tar upp det. Men jag gör det, dels för att jag tycker det är någonting som är så himla centralt i historieämnet, men också för att, ja, då har jag faktiskt det här provet i bakhuvudet, att det ska de behärska liksom. Men de ska också känna att de behärskar det när de har läst kursen. Det ska inte bara vara någonting som man tar upp i början av kursen och sen glömer man bort det.³⁴²

Hanna uppfattar att den källkritiska uppgiften i HLP har påverkat den procedurala dimensionen av hennes domänlogik. Till skillnad från den källkritiska uppgiften uttrycker ingen av HLP-lärarna någon uppfattning att D-uppgifterna, det till säga de uppgifter som prövade elevernas förmåga till temporal orientering, har påverkat dem. Den här skillnaden i hur HLP-lärarna uppfattar de två stora dokumentbaserade uppgifterna i HLP indikerar att lärarna i viss utsträckning väljer vad de låter sig påverkas av.

Påverkan på progressionslogiken

Även vad gäller påverkan på progressionslogiken framkommer en variation av uppfattningar. Helena uppfattar i allmänna ordalag att hon har blivit ”lite bättre på att bedöma tack vare det här provet”³⁴³ medan Henrik är tydlig i sin uppfattning att han inte påverkas alls. Hanna uttrycker ingen bestämd uppfattning huruvida hennes sätt att värdera provsvar påverkas

³⁴² Intervju med Hanna, 2009-11-19.

³⁴³ Intervju med Helena, 2009-11-25.

av HLP, men hon uttrycker en viss kritik mot HLP:s bedömningsgrunder då hon uppfattar att de ”inte varit så välformulerade eller de har varit svårtolkade”.³⁴⁴ Hannas negativa uppfattning skulle kunna innebära att hon saknar incitament för att låta sig påverkas. Håkan har en avvikande uppfattning om bedömningsgrunderna: ”Jag har inget problem att använda mig av de kriterierna [...] så skulle jag också göra facit, så det är inget problem”.³⁴⁵ Han uppfattar dock att värderingsverktyget inte är ett funktionellt värderingsverktyg för de elever som han arbetar med.

Jag tror inte att jag har hundra procent nytta av att använda mig av provet för att säga nu kan jag med säkerhet sätta det där IG:et eller det där G:et. För att fortfarande är det så det ger mig inte bedömningsgrunder på det sättet som jag behöver för att mäta de här eleverna som inte är mätbara enligt traditionella parametrar.³⁴⁶

Håkan uppfattar att hans huvudsakliga värderingsproblem gäller gränsdragningen mellan IG och G hos elever med språksvårigheter och det är ett problem som han inte kan lösa med HLP. På samma sätt som Hanna saknar Håkan därför incitament att låta sig påverkas i någon större utsträckning.

Ytterligare en aspekt i HLP-lärarnas utsagor gällande HLP:s påverkan på deras progressionslogik handlar explicit om interbedömarreliabiliteten inom ämneslaget. Det finns en samstämmighet hos HLP-lärarna i deras uppfattningar att det uppkommer ett likvärdighetsproblem i samband med värderingsprocessen. Elevernas kunskaper i historia värderas på olika sätt av olika lärare. Helena uttrycker sig på följande sätt:

Det är jätteviktigt att vi visar varandra och liksom tittar och försöker prata med kollegor. Hur bedömer man det här svaret tycker du? Är det här ett MVG-svar eller är det ett VG-svar? Jag tycker det är intressant det här med kollegor. Att en del kollegor rättar strängt och andra rättar mindre strängt och vi har nog blivit lite mera uppmärksamma på hur vi rättar, att där är en viss skillnad ändå, trots att vi har ett gemensamt

³⁴⁴ Intervju med Hanna, 2009-11-19.

³⁴⁵ Intervju med Håkan, 2010-01-19.

³⁴⁶ Intervju med Håkan, 2010-01-19.

prov, trots att svaren finns, att facit finns.³⁴⁷

Helena uppmärksammar att det finns progressionslogiska skillnader mellan lärarna inom ämneslaget som innebär att ”en del kollegor rättar strängt och andra rättar mindre strängt”. Hon uppfattar att HLP har inneburit en ökad medvetenhet om problematiken, men understryker samtidigt att det fortfarande kvarstår vissa brister gällande den kollegiala interbedömarreliabiliteten.

9.3.2 Metodologisk påverkan

Påverkan på observationsverktygens format

Även HLP-lärarnas uppfattningar om hur HLP påverkar deras sätt att konstruera uppgifter varierar. På ett individuellt plan går det att känna igen mönstret från den epistemologiska dimensionen. Henrik uppfattar på nytt att han inte påverkas så mycket av HLP. Håkan uppfattar att han inte påverkas av HLP eftersom han anser att uppgiftsformaten i HLP har legat i linje med hans vanliga sätt att konstruera uppgifter i traditionella skriftliga prov. Som framkommit tidigare uppfattar Håkan dock att han har förändrat sitt sätt att konstruera prov och provfrågor i takt med att hans elevunderlag har förändrats. Han menar därför att det är svårt att förena HLP:s uppgiftsformat med hans nuvarande, mer empatiska, sätt att tänka kring undervisning och uppgiftskonstruktion.

Helena uppfattar på samma sätt som tidigare att hon påverkas i viss utsträckning:

Något, tror jag, något tror jag att de har det faktiskt. Jag har funderat mer kring formuleringar och lite mer att jag har funderat på kring mål och så. Förr så gjorde jag det kanske lite mer lättvindigt, att det här är stoffet och det blev lite mer detaljfrågor och så.³⁴⁸

Helena för ett allmänt hållet resonemang om att hon har blivit mer medveten och reflekterande i sitt sätt att tänka kring uppgiftskonstruktion. Till skillnad från Helena och de andra två HLP-lärarna uppfattar Hanna att hennes sätt att konstruera uppgifter har påverkats i ganska hög grad.

³⁴⁷ Intervju med Helena, 2009-11-25.

³⁴⁸ Intervju med Helena, 2009-11-25.

Generellt sett så har jag varit ganska negativt inställd till prov. Jag har inte haft så mycket prov i mina kurser över huvud taget. Jag har haft mer essäliknande prov där de till och med har fått förbereda frågorna hemma. Eller att de har fått tre stora frågor på provet som jag har satt ett betyg på istället eller gett ett omdöme på. Men jag har faktiskt, sen jag blev engagerad i detta, så har jag kanske omvärderat lite därför att det behöver inte vara så att ett prov liksom gör anspråk på att mäta på ett naturvetenskapligt sätt, utan man kan komma åt analysen och reflektionen där också, det beror helt på hur det är utformat. Så att jag har blivit ganska intresserad av det här hur man formulerar prov och så efter detta, det är ganska kul. Så jag ska faktiskt prova nästa vecka, i den klass som ska göra det här provet till våren, så ska jag utforma ett prov på medeltiden på samma vis, i tre delar, så att en del är kryssfrågor och en del är analysfrågor och en del är källkritisk del, men jag kommer inte göra så att de får välja utan alla ska svara på alla frågor.³⁴⁹

Hanna uppfattar att hon före HLP hade en mer avvaktande inställning till skriftliga prov, men att HLP har påverkat henne att tänka i andra banor. Hon berättar vidare att hon i ett kommande internt prov kommer att tillämpa den tredelade provstrukturen från HLP. Hannas fokusförskjutning gällande uppgiftsformatet kan tolkas som en påverkan i mättraditionell riktning.

9.3.3 Kollegial påverkan

Påverkan på den kollegiala interaktionen

Till skillnad från de övriga påverkansområdena är HLP-lärarna samstämmiga i sin uppfattningar att HLP påverkar den kollegiala interaktionen. Allra tydligast i sin uppfattning är Hanna som uttrycker sig på följande sätt:

Jag tycker det är jättepositivt för att det bidrar till att vi diskuterar stoff i kursen. Det bidrar till att vi diskuterar bedömning, det bidrar till att vi diskuterar likheter och skillnader mellan programmen, att vi sam-

³⁴⁹ Intervju med Hanna, 2009-11-19.

bedömer, så jag tycker att det är jätteviktigt för vår ämnesgrupp faktiskt. Men det skulle lika gärna kunna vara ett gemensamt projekt till exempel som vi jobbar med, så det behöver inte vara ett prov i den här bemärkelsen.³⁵⁰

Hanna värdesätter de kollegiala samtal och samarbeten som användningen av HLP genererar. Notera dock att Hanna betonar att det inte behöver vara ett prov, det kan lika gärna vara något annat projekt. Det väsentliga är att den kollegiala interaktionen ökar. I praktiken handlar den kollegiala interaktionen om att främja och stimulera såväl samtal som samarbeten kring bedömningsprocessens alla delar, vilket i förlängningen kan ses som ett led i en generell professionsutveckling.

Samtliga HLP-lärare uppfattar att de använder HLP som ett underlag för kollegial interaktion. Det finns dock skillnader mellan lärarna vad i HLP som ligger till grund för denna interaktion. Henrik uppfattar att interaktionen huvudsakligen kretsar kring provuppgifternas domänlogik och format, medan Håkan uppfattar att samtalen framför allt handlar om hur provresultaten ska användas. Helena för ett mer generellt resonemang när hon uppfattar att HLP har lett till en ökad benägenhet att samtala om bedömning och att ta hjälp av varandra i bedömningsfrågor.

9.3.4 HLP – en sammanfattning

Epistemologisk påverkan

HLP-lärarnas uppfattningar om huruvida HLP påverkar den substantiella dimensionen av deras domänlogik varierar. En av dem uppfattar ingen påverkan alls medan en annan uppfattar att provet påverkar hans val av narrativ i högsta grad. Ett liknande mönster framkommer när det gäller den procedurala dimensionen av domänlogiken. En av lärarna uppfattar ingen påverkan medan två av lärarna uppfattar att HLP i viss utsträckning har påverkat dem att betona den källkritiska förmågan i sin undervisning och bedömning. Även när det gäller progressionslogiken återkommer samma variationsmönster, men här framkommer även ett visst motstånd mot HLP:s bedömningsgrunder och att de därför saknar incitament att låta sig påverkas. När det gäller progressionslogiken framkommer även en

³⁵⁰ Intervju med Hanna, 2009-11-19.

återkommande uppfattning att den kollegiala interbedömarreliabiliteten har stärkts, men att det kvarstår vissa brister.

Metodologisk påverkan

Den variation som framkommer vad gäller epistemologisk påverkan gäller även HLP-lärarnas uppfattningar om provets påverkan på deras egen formatlogik gällande observationsverktygen. På nytt uppfattar en av lärarna att han inte påverkas alls i sitt sätt att utforma prov och provuppgifter medan en annan uppfattar att hon påverkas i ganska hög utsträckning. Den senare uppfattar att hon använder skriftliga format i högre utsträckning än tidigare i sin bedömningspraktik och hon berättar att hon har för avsikt att använda sig av HLP:s tredelade konstruktion i ett internt moment.

Kollegial påverkan

Det är först i uppfattningarna gällande den kollegiala interaktionen som det framträder ett tydligt mönster hos HLP-lärarna. Det finns en klar samstämmighet i deras uppfattningar att HLP har utgjort ett viktigt discussionsunderlag och en katalysator för den kollegiala interaktionen inom det egna ämneslaget.

9.4 KP

9.4.1 Epistemologisk påverkan

Påverkan på domänlogiken – den substantiella dimensionen

Krister är den ende av KP-lärarna som ger uttryck för att KP har påverkat den substantiella dimensionen av hans domänlogik.

Jag tror att jag har tagit upp lite mer faktiskt om ekonomisk, social historia ändå, det har ju varit, om man ser på kursprovet, frågor om till exempel migration, om social utveckling, om ekonomisk utveckling, det har jag nog inte tagit upp, om jag ska vara ärlig, så mycket före 2015, det har jag tagit upp mer nu.³⁵¹

Att det är just Krister som uttrycker denna uppfattning är kanske inte så överraskande eftersom han är den ende av KP-lärarna som genomför KP

³⁵¹ Intervju med Krister, 2016-11-24.

i sin helhet under standardiserade former. Kristers uppfattning att provet i viss utsträckning påverkat hans val av narrativ (centralt innehåll) ligger i linje med några av HLP-lärarna. Den substantiella dimensionen av domänlogiken får större betydelse om ett prov används i sin helhet och om provresultatet sedan ska användas i ett summativt syfte.

Kristers uppfattning kan jämföras med Karins resonemang kring hur KP har påverkat hennes val av narrativ:

Jag: Om jag har förstått dig rätt så använder du framför allt enskilda uppgifter i anslutning till undervisningsmoment.

Karin: Ja, precis, så är det ju ... och då styr det ju kanske lite vad jag har för moment om jag vet att det finns en bra uppgift.

Jag: Menar du att ... vad kommer först där - hönan eller ägget?

Karin: Nej, det är nog ... jag har ju bestämt att det här måste vi gå igenom och sen kollar jag då ... finns det någon bra uppgift som jag kan rätta mig efter lite grann ... och ibland gör det ju det och ibland gör det ju inte det, för det är ju inte hundra procent heltäckande.

Jag: Absolut inte ... men du menar att själva ditt urval av centralt innehåll, det kommer först och sen så ser du om du hittar någon uppgift som du skulle kunna använda dig av.

Karin: Ja.³⁵²

Som framkommer i utsagan uppfattar Karin att KP:s betydelse för hennes val av narrativ är begränsat. Hennes val av narrativ kommer först och därefter väljer hon ut någon uppgift från KP som hon kan använda i anslutning till detta narrativ.

Påverkan på domänlogiken – den procedurala dimensionen

Till skillnad från den substantiella dimensionen så uppfattar samtliga KP-lärare att KP har påverkat den procedurala dimensionen av deras domänlogik. Karin är den av KP-lärarna som tydligast beskriver en fokusförskjutning från en snävare MH-förankrad domänlogik med ett tydligt fokus på rekonstruktion till en bredare HT-förankrad domänlogik. I nedanstående resonemang tar hon avstamp i de uppgiftsmatriser i form av rutor som fanns infogade i KP-uppgifterna.

³⁵² Intervju med Karin, 2016-11-08.

De här rutorna tycker jag [...] man vet att man använder och når kursmålen på ett riktigt sätt och det tycker jag är väldigt bra för innan Gyll var det ju mer fokus på historiska kunskaper för eleverna, men nu att man får med det här med begrepp och källor och användning ... man vet vad man ska titta på ... så det tycker jag är väldigt bra.³⁵³

Lite längre fram i intervjun förtydligar Karin sitt resonemang om hur hon påverkats av KP. Hon uppfattar att det har blivit lättare att tolka de centrala målen i 2011 års kursplaner i historia. Det är framför allt förmågan att hantera begrepp som har blivit tydligare för Karin:

Det här med det kan man ju komma in på då ... med begrepp ... det tycker jag har blivit mycket tydligare för mig efter de här proven kom ... för jag hade nog mycket snävare syn på vad begreppen var som man skulle ta upp ... nu tycker jag det är ... ja, det står om begrepp då med källkritiken ... det är ju då äkthet och närhet och tiden och ... och det hade jag kanske inte tänkt att ... är det begrepp ... ja, men, och epoken är begrepp och aktör och struktur och så här ... jag känner, det är bra att jag har fått det tydligare för mig.³⁵⁴

Karin uppfattar att användningen av KP har inneburit att hon har fått en tydligare uppfattning om vad som menas med begrepp och vilka begrepp som är relevanta att behandla i undervisningen. Hon menar också att hennes sätt att uppfatta den grundläggande tankeoperationen i samband med begreppshantering har påverkats:

Jag har nog gått från mer att ... bara kunskaperna, till att de ... att jag ser att de kan använda sig utav det dom har lärt sig på ett annat sätt, alltså mer konkret liksom ... vad dom förstår, så att dom ... hur ska jag uttrycka mig ... från alltså bara skriva ner massa kunskap till att dom kan använda sig utav begreppen och förstår begreppen ... sätta in det i mer, liksom, praktisk betydelse.³⁵⁵

³⁵³ Intervju med Karin, 2016-11-08.

³⁵⁴ Intervju med Karin, 2016-11-08.

³⁵⁵ Intervju med Karin, 2016-11-08.

Karin beskriver en förändring där hon tidigare prövade elevernas kunskaper *om begreppen* till att nu pröva deras förmåga att *använda begreppen* i ett resonemang – en förändring från ordförståelse till arbetsredskap. Denna förändring tydliggör en fokusförskjutning från den MH-traditionella betoningen på rekonstruktion och refererat till den HT-traditionella betoningen på (de)konstruktion och resonemang. Karin konkretiserar sedan sitt resonemang när hon beskriver hur eleverna hanterade en KP-uppgift om historiebruk:

Men som det här de gjorde nu om historiebruk. Då tyckte jag liksom att de kunde ju se att det här var ett politiskt historiebruk eller det var bara kommersiellt. Dom fattade det på ett annat sätt än att de bara skulle ha skrivit upp ... ”skriv hur många historiebruk det finns”. För där kunde jag ju haft ett prov då ... ”skriv fem historiebruk” ... men här då fick de ju praktiskt använda sig av sina kunskaper och då känner de ju också liksom, ja, men så här ... ”ja, men nu fattar jag ju vad det här handlar om på ett annat sätt”.³⁵⁶

Karin beskriver att hon tidigare hade utformat uppgifter där eleverna skulle rekonstruera, det vill säga minnas någon form av typologi.³⁵⁷ Hon menar att användningen av KP-uppgifter har inneburit att eleverna i högre utsträckning tillämpar typologin i ett resonemang.

Även Kristina uppfattar att KP har påverkat hennes sätt att tolka 2011 års kursplaner.

När jag kom att använda kursprovet i historia så gav det mig personligen mer än jag trodde. Det öppnade ögonen mer på mig än vad jag trodde för att komma närmare kursplanen ... jag förstod lite bättre en del saker. Det här med begrepp till exempel hade jag bara kopplat till referenskunskaper. Jag hade inte kopplat det till tolkningen, så för mig var, då 2013 (sic) när jag fick vara med och prova ut det så hade det ... så gav det mig mer än, det gav mig mer än jag trodde för att komma närmare kursplanen.³⁵⁸

³⁵⁶ Intervju med Karin, 2016-11-08.

³⁵⁷ Troligen den typologi som har sitt ursprung i Klas-Göran Karlssons teoribygge. Se Karlsson (2014).

³⁵⁸ Intervju med Kristina, 2016-11-21.

På samma sätt som Karin lyfter hon fram just begreppshantering som ett exempel på hur KP har påverkat henne. Hon förtydligar sedan sitt resonemang:

Det här med begrepp, alltså ordet begrepp över huvud taget, begreppskunskapskravet hade jag liksom, okej dom ska kunna feodalism, ja jag förstår det, det var jag helt med på ... men de begrepp som tillhör tolkning, exempelvis aktör och struktur, eller ... eller dom andra sakerna som har med tolkningen att göra, dom hade jag inte fått med alls ... och det är fortfarande svårt att gå långt där, känner jag, när hinner man alltså, det är fortfarande knepigt ... men det var en aha-upplevelse för mig.³⁵⁹

Tidigare uppfattade Kristina begreppen som i första hand innehållsbegrepp med koppling till referenskunskap. Hon uppfattar att KP har påverkat henne att betona de metahistoriska begreppens betydelse, särskilt i relationen till ”tolkningen”. Över huvud taget är just tolkningen av narrativ ett återkommande tema i KP-lärarnas uppfattningar om KP:s påverkan. Kristina vidareutvecklar sitt resonemang om hur KP påverkat hennes uppfattningar om tolkningens betydelse:

När man tittar på hur kursprovet är uppbyggt så handlar det ju väldigt mycket om att försöka, så mycket det någonsin går, att komma fram till tolkning ... alltså, det är så oerhört lätt att man bara stannar på nått sätt vid det man måste ha för att tolka. De måste ju ... man kan ju inte tolka orsakerna till franska revolutionen utan att veta vad franska revolutionen var ... och då hamnar man ju ofta att man har fullt upp med att undervisa om franska revolutionen. Så att kunna integrera så att säga referenskunskaperna med tolkning av dem, det tyckte jag att jag fick hjälp med.³⁶⁰

Kristina beskriver en domänlogisk förskjutning från *kunskaper om narrativet* till *tolkningen av narrativet*. Det är en förskjutning som belyser en grundläggande skillnad mellan den MH-traditionella domänlogiken och HT-traditionella domänlogiken. Det är relationen mellan narrativet i sig

³⁵⁹ Intervju med Kristina, 2016-11-21.

³⁶⁰ Intervju med Kristina, 2016-11-21.

(minneskunskap) och tolkningen av narrativet (historiskt tänkande) som Kristina uppfattar som betydelsefull.

Även Karl resonerar om tolkningens betydelse när han beskriver sina tankegångar i samband med värderingen av en elevprestation:

Nu ska vi se ... vad är det nu som är det primära här. Ok, hitta stöd för båda tolkningarna, det är huvudgrejen då, alltså. Jämför de båda tolkningarna, ok. Ja, här var det många elever som hade svårt att ... röra sig här. Så det är något som jag har lärt mig i alla fall, det här ordet "tolkningar", det är viktigt ... hitta stöd för den tolkningen och den. Så återigen, provet ger oss lärare väldigt mycket stöd hur vi ska uttrycka oss och lägga upp vår undervisning.³⁶¹

Karl uppfattar att KP har påverkat hans sätt att uppfatta ämnet som ett tolkningsämne i linje med HT-traditionen. Uppgiften handlar inte om att rekonstruera ett etablerat narrativ, det vill säga en MH-traditionell tankeoperation baserad på en objektivistisk kunskapssyn. Uppgiften kräver en HT-traditionell tankeoperation, baserad på en konstruktivistisk kunskapsyn, där det är dekonstruktionen av narrativen som är det primära.

Påverkan på progressionslogiken

När det gäller KP-lärarnas uppfattningar om hur KP har påverkat deras progressionslogik framträder den HT-traditionella fokusförskjutningen inte lika tydligt som gällande domänlogiken. Karin uttrycker sig på följande sätt:

Jag har blivit mildare i min bedömning [...] säkrare på att värdera elevsvaren ... och det går fortare att rätta för jag vet på ett bättre sätt vad jag ska titta på utifrån de här betygsstegen, det blir mer konkret. Jag vet vad man ska titta på för ett E, för C och ett A [...] och då vet, då ser jag ju också i bedömningsanvisningarna ... "du diskuterar om den ena tolkningen är rimligare än den andra" ... "och använder båda begreppen på ett rimligt sätt" ... det blir väldigt tydligt ... det går fortare.³⁶²

³⁶¹ Intervju med Karl, 2016-12-06.

³⁶² Intervju med Karin, 2016-11-08.

Karin uppfattar att hennes användning av KP:s bedömningsanvisningar har gjort att hon har blivit ”mildare” i sin värdering av elevprestationer. Användningen har även inneburit att hon har blivit tydligare och säkrare i sin tolkning av kunskapskravens olika betygssteg. Det här är ett återkommande tema hos KP-lärarna när de beskriver KP:s påverkan på deras progressionslogik. Krister uttrycker sig på följande sätt:

Jag tror faktiskt att jag har blivit lite tuffare när det gäller kunskapskraven, jag tror nästan det. Jag har, känns det, lite mer belägg ... lite mer på fötterna. Jag är lite tuffare i kunskapskraven för dom högre betygen nu än vad jag hade för några år sen. Jag känner att jag vet med en tydlighet exakt vad jag ska begära från en elev. Det kvittar egentligen vilket kunskapskrav det är, vilken uppgift det är ... jag kan ändå väldigt tydligt förklara för eleven att ... om vi ser till matrisen här så måste du ha detta ... nyansering innebär problematisering, jämförelser, se på likheter och skillnader, se på vilka grupper som påverkades av detta. /.../ Så jag känner att ... njae, snarare att jag har blivit säkrare. Därför tycker jag att kursprovet har varit ett fantastiskt, egentligen en enda lång fortbildning för lärarna, för mig då, för mina kollegor ... att man kan vara tydligare i vad man vill ha och jag känner mig tryggare också i att förmedla det till eleverna.³⁶³

Till skillnad från Karin uppfattar Krister att han blivit ”tuffare” i sitt sätt att tolka kunskapskraven, men även han uppfattar att KP har inneburit att han har blivit säkrare och tydligare i sitt sätt att värdera elevsvar. Krister förtydligar sitt resonemang:

Jag tycker helt klart att uppgifterna här med matriserna och kunskapskraven har förbättrat mitt eget sätt att förstå de här, som du säger, märkes- eller nyckel-, eller honnörsorden, det som är det fetstilade, som alltid har varit lite ... inte svårt, men ... en utmaning för läraren är att försöka begripliggöra för eleverna.³⁶⁴

Krister uppfattar att KP har varit ett stöd för honom att tolka värdeorden i kunskapskraven, det vill säga de kvalitativa progressionsindikatorerna.

³⁶³ Intervju med Krister, 2016-11-24.

³⁶⁴ Intervju med Krister, 2016-11-24.

Huruvida det har inneburit en progressionslogisk förskjutning i HT-traditionell riktning framkommer dock inte.

Även Karl tycker sig ha fått stöd av KP och det stödmaterial som vuxit fram i spåren av KP:

Samtidigt har jag fått mycket, mycket stöd i hur ... man kan tänka ... så att jag tycker att det är mycket ... det är enklare att bedöma ... men om jag har blivit mer generös eller hårdare eller så, det ... det vet jag egentligen inte ... jag kanske bedömer saker på ett annat sätt än vad jag gjorde förr.³⁶⁵

Karl tar inte ställning till om användningen av KP har inneburit att han har blivit ”mer generös eller hårdare” i sin värdering, men han uppfattar att bedömer saker ”på ett annat sätt” än vad han gjorde förr. Han konkretiserar sin tankegång i följande resonemang:

Det här "historiska exempel" och "historiska sammanhang" ... det var mycket viktigt. Ja, alltså, vad är kvalitativa aspekter i ett prov, det är inte helt självklart egentligen /.../ en elev kan skriva väldigt mycket, men alltså vad är det som /.../ Det diskuterade vi också ganska mycket på en ämnesträff, alltså vad är historiska exempel och vad är historiska sammanhang.³⁶⁶

Karl för ett progressionslogiskt resonemang kring ”de kvalitativa aspekterna” i ett elevsvar och lutar sig då mot KP-formuleringarna ”historiska exempel” och ”historiska sammanhang”. Hans resonemang indikerar en progressionslogisk förskjutning i HT-traditionell riktning, en förskjutning från kvantitativa till kvalitativa progressionsindikatorer. Karl antyder att den här förskjutningen inte omfattar alla inom ämneslaget:

Mitt sätt att bedöma elever är väl ... jag vet inte ... allsidigare skulle jag säga ... jag vet inte om jag är strängare, jag kanske ... jag inser ju att ... men det insåg jag för länge sedan ... men jag har ju kollegor här som i princip bara mäter referenskunskap ... i form av prov med poäng ... och hela det här materialet har ju ... det har ju varit en resa in i att

³⁶⁵ Intervju med Karl, 2016-12-06.

³⁶⁶ Intervju med Karl, 2016-12-06.

jag förstår ... det är nånting mycket mer komplext.³⁶⁷

Karl jämför sin egen bedömningspraktik med kollegor som betonar referenskunskap och som kvantifierar prestationsomdömena. Han menar att han gör något mer komplext. Den här polariseringen inom ämneslaget kommer att behandlas mer utförligt i avsnitt 9.4.3.

9.4.2 Metodologisk påverkan

Påverkan på observationsverktygens format

Till skillnad från HLP-lärarna är KP-lärarna överlag positivt inställda till de uppgiftsformat som återfinns i KP. De uppfattar också att de har blivit bättre på att konstruera uppgifter. Huruvida denna påverkan är epistemologisk eller metodologisk kan dock diskuteras. Karin uttrycker sig på följande sätt:

Jag tycker att era frågor är bättre utformade ... jag känner att dom får ju en verkligen tänka till. Det kanske var mer det här lite att "rapa upp" när jag gjorde ett prov innan. Här måste dom verkligen använda sina kunskaper på ett bättre sätt.³⁶⁸

Karin uppfattar att hon tidigare hade en mer minnesbaserad, MH-traditionell, kunskapsuppfattning när hon konstruerade uppgifter. Även Karl uppfattar att han har blivit en bättre uppgiftskonstruktör efter KP:

Och detta har ju helt, det har ju öppnat upp ögonen enormt mycket att jag måste formulera en fråga som är kopplad till en förmåga som ska mätas ... i ett kunskapskrav. Det har ökat min förmåga att göra det väldigt mycket ... så jag har blivit mycket, mycket skickligare i mitt sätt att formulera frågor /.../ Jag har förstått, genom att titta på det här materialet, att jag själv kan skapa sådana uppgifter.³⁶⁹

Notera att Karl använder samma formulering som många andra KP-lärare: att KP har "öppnat upp ögonen". Karl tycker att han har blivit bättre på att formulera uppgifter som ger eleven möjlighet att visa kunskaper i

³⁶⁷ Intervju med Karl, 2016-12-06.

³⁶⁸ Intervju med Karin, 2016-11-08.

³⁶⁹ Intervju med Karl, 2016-12-06.

historia. Han uppfattar att han på ett bättre sätt formulerar uppgifter som är kopplade till kursplanens domänlogik (förmågor) och lite längre fram i intervjun framkommer även att han uppfattar att han har blivit duktigare på att formulera uppgifter som diskriminerar. ”Jag har också blivit bättre på att formulera frågor som verkligen sällar ut de riktigt avgörande kvaliteterna”.³⁷⁰

Även Kristina uppfattar att KP har påverkat henne sätt att konstruera uppgifter:

En del är samma ... (ohörbart) favoriter, men jag tänker på ett annat sätt och jag gör dem på ett annat sätt. Det har ju med kunskapskraven att göra för ska man försöka täcka upp kunskapskraven och ge eleverna chans att testa på historiebruk mer än en gång, då får man ligga i. Då kan man inte harva på framåt i historien, det går ju inte, utan jag måste ju kombinera. Jag måste ju ha minst två saker samtidigt så att säga. Jag kan inte ta ett kunskapskrav i taget, utan jag måste ju hela tiden ha med fler. /.../ Jag kan liksom inte ... jag måste inkludera historiebruk eller källhantering på varje, eller helst båda, på varje epok ... på de moment jag jobbar med.³⁷¹

Kristina uppfattar att KP har påverkat hennes sätt att konstruera uppgifter. Hon uppfattar att KP:s flerdimensionella uppgifter har påverkat och underlättat hennes arbete med att konstruera uppgifter som täcker flera delar av domänlogiken. Hon beskriver att hon att i varje uppgift av tidsskäl, utöver referenskunskaper, måste pröva ytterligare någon förmåga.

9.4.3 Kollegial påverkan

Kollegial interaktion

I linje med HLP-lärarna uppfattar KP-lärarna att det externa provet har haft en stor betydelse för den kollegiala interaktionen. Krister uttrycker sig på följande sätt:

Det har betytt väldigt mycket. Det har varit någonting som vi har kunna fästa vårt tänkande kring, och vårt arbete för att utveckla ämnet.

³⁷⁰ Intervju med Karl, 2016-12-06

³⁷¹ Intervju med Kristina, 2016-11-21.

Tidigare så har det varit väldigt givande att utveckla ämnet, men fördelen med att ha kursprovet är att man kan ha någonting att bygga upp det på, man kan ha någonting att falla tillbaka på.³⁷²

Krister uppfattar att KP har bidragit till en ökad interaktion inom ämneslaget och att det har varit ”någonting att bygga upp det på, man kan ha någonting att falla tillbaka på.” Även Kristina uppfattar att KP har varit ett stöd för de kollegiala samtal som hon som ämnesföreträdare har velat föra på ämnesmötena:

Jag tog över 2007 och jag tror att från 2009 och framåt så har det här varit en av mina saker som jag velat driva ... någon form av gemensam bedömning, och sen i samband med, som jag sa, med 2011 så kändes det som att det lättade, alltså ... det fanns, när kursproven började komma, så fanns det ju någonting konkret att ta på.³⁷³

På samma sätt som Krister beskriver Kristina KP som ett underlag att utgå ifrån som ”någonting konkret att ta på”. Karins uppfattning ligger i linje med de övriga KP-lärarnas uppfattning:

Jag tycker det skulle underlätta, för att då har vi ju någonting gemensamt att jobba mot. Och då om man har bestämt sig för att nu ska vi göra det här och vi ska diskutera det här så borde det ju underlätta. Jag tycker att alla borde känna en trygghet i det och jag tror att det kommer att bli så. Vi har en ny kollega som är jättetacksam för dom här diskussionerna vi har fört nu.³⁷⁴

Karin uppfattar att KP skulle underlätta den kollegiala interaktionen eftersom det utgör något som de tillsammans kan arbeta mot. Samtidigt så beskriver Karin att det inte är alla inom ämneslaget som delar hennes uppfattning:

Det är ju då framför allt en kollega som ju inte vill samarbeta, hon har den andra samhällsettan nu och jag vet inte vad hon gör och det är lite

³⁷² Intervju med Krister, 2016-11-24.

³⁷³ Intervju med Kristina, 2016-11-21.

³⁷⁴ Intervju med Karin, 2016-11-08.

skrämmande.³⁷⁵

Karins beskriver att det är en kollega i ämneslaget som inte velat samarbeta kring KP. Den här tendensen till kollegial polarisering framkommer även hos två av de andra KP-lärarna.³⁷⁶

Kollegial polarisering

Till skillnad från HLP-lärarna beskriver flera av KP-lärarna att det har funnits ett motstånd inom ämneskollegiet mot det externa provet. Kristina beskriver motståndet mot KP på följande sätt:

Det finns ett motstånd mot provet, nu tänker jag ... det spelar ingen roll om det är en eller flera eller så, men motståndet är att vi blir styrda av provet, vi kommer bara att undervisa utifrån det som provet kommer att ta upp, som vi tror att provet kommer att ta upp ... vi blir begränsade och det vill man inte acceptera ... det är en största argumentet emot ... och så jämför man med USA och lite andra länder där man bara undervisar mot examensprovet ... och det andra är att ... ett känt citat från en kollega: "vi kan inte utsätta våra ungdomar för sånt" ... och det tredje är då det här jag redogjorde för förut att, och den är mer outtalad. att man är orolig för att provet ska ge ett resultat som, som inte stämmer överens med det man själv har innan, och därmed så hamnar man i nån form av knipa.³⁷⁷

Kristina resonerar kring tre olika anledningar till det kollegiala motståndet mot KP. Den första anledningen handlar om att provet får en alltför styrande funktion på deras epistemologiska vägval; den andra att provets mättraditionella direktiv inte ligger i linje med deras egen mer tolknings-traditionella uppfattning om bedömning; och den tredje om att provresultatet kan indikera en diskrepans mellan provresultat och betygssättning, vilket indirekt skulle kunna indikera en bristfällig bedömningskompetens.

³⁷⁵ Intervju med Karin, 2016-11-08.

³⁷⁶ Kristers ämneslag består endast av två personer. Båda använder KP och ingenting i Kristers utsagor tyder på att provet legat till grund för någon form av kollegial polarisering.

³⁷⁷ Intervju med Kristina, 2016-11-21.

Det här motståndet mot KP har bidragit till att den kollegiala interaktionen har haft en baksida i form av en kollegial polarisering inom ämneslaget. Karl beskriver interaktionens dubbla ansikten på följande sätt:

Då kursprovet kom 2015 så var vi flera stycken som tände direkt och på ämnesträffar så ... Vi tog några ämnesträffar och gick igenom ... vi diskuterade fråga för fråga som var på kursprovet 2015. Vi plockade ut något elevsvar och vi satt och gnuggade det. Men det är en väldigt kluven bild. /.../ Vi är några lärare som är väldigt medvetna och använder detta materialet och sen är det några lärare som i princip, eller av olika skäl, inte använder sig av bedömningsstödet./.../ För så är det, här är dom som inte ens har öppnat detta material. Dom har nog knappt ens inte tittat i kursplanen, eller ämnesplanen. /.../ Så jag vet inte vad som kommer att hända nästa läsår. Hälften av oss kommer att ha utvecklats något enormt, vissa är kvar på ruta ett.³⁷⁸

Karl beskriver hur KP har legat till grund för ämnesutvecklande diskussioner inom ämneslaget, men samtidigt beskriver han hur provet har polariserat ämneslaget. Han funderar över hur denna polarisering kan påverka det framtida samarbetet inom kollegiet. Karl tycker att frivilligheten är ett bekymmer och att rektor borde göra KP obligatoriskt. Han konkretiserar sedan konsekvensen av den kollegiala polariseringen genom ett exempel:

I en klass, det var så i fjor då, en av mina kollegor, en yngre lärare, han använde kursprovet i en NV3:a. Medan hans kollega hade ett helt annat prov ... alltså, "vilket år dog Napoleon?" ... "när bröt första världskriget ut?" De var ganska irriterade därför att det var ju på en helt annan nivå. Det Martin gjorde var ju så högkvalitativt och han hade sådana prov hela läsåret, medan den andre läraren hade prov som ... Det beror på vad man ... bra eller dåliga, men de mätte ju inte någonting av det som ska mätas. De mätte ju enstaka faktakunskaper och så. Det behöver inte vara fel på det, men det är ju bara en del av ...³⁷⁹

Karls sätt att konkretisera den kollegiala polariseringen är ett tydligt exempel på hur två olika historiekulturella traditioner möts. Han beskriver,

³⁷⁸ Intervju med Karl, 2016-12-06.

³⁷⁹ Intervju med Karl, 2016-12-06.

mer eller mindre medvetet, sig själv och sin yngre kollega som representeranter för HT-traditionen och den andra kollegan som en representant för MH-traditionen.

9.4.4 KP – en sammanfattning

Epistemologisk påverkan

Till skillnad från HLP-lärarna uppfattar KP-lärarna att provet har påverkat deras domänlogik i hög utsträckning. Det är en bred HT-traditionell påverkan som i första hand gäller den procedurala dimensionen av deras domänlogik och som omfattar såväl synen på narrativ som synen på grundläggande begrepp.

Det framkommer även att KP-lärarna uppfattar att deras progressionslogik har påverkats. Det handlar dock inte om någon epistemologisk förskjutning, utan snarare om att de uppfattar att de har blivit säkrare i att tolka kunskapskravens progressionsindikatorer. Till skillnad från HLP-lärarna handlar den progressionslogiska påverkanseffekten inte i första hand om att säkerställa kollegial interbedömarreliabilitet, utan om att tolka kunskapskravens värdeord på ett ”rätt sätt”, det vill säga en fråga om alignment.

Metodologisk påverkan

Till skillnad från HLP-lärarna är KP-lärarna överlag positivt inställda till de uppgiftsformat som återfinns i KP. KP-lärarna uppfattar även i högre utsträckning att de har förändrat sitt sätt att konstruera uppgifter. De betonar att de har blivit bättre uppgiftskonstruktörer. Det är dock svårt att avgöra om den denna uppfattning kan räknas som en metodologisk eller epistemologisk påverkan, det vill säga om det är uppgiftens format eller domänlogik som ligger till grund för deras resonemang.

Kollegial påverkan

KP-lärarna uppfattar att KP har bidragit till en ökad kollegial interaktion. Denna interaktion har dock skapat en polarisering inom ämneslaget mellan de som använder KP och de som inte gör det. KP-lärarna uppfattar att ”icke-användarna” eller ”de andra” är förankrade i en MH-traditionell domänlogik, medan de själva, med KP som ett tolkningsverktyg, har blivit allt mer säkra i sin HT-traditionella förankring. Någon motsvarande kollegial polarisering framkommer inte hos HLP-lärarna.

9.5 Resultatdiskussion

9.5.1 Påverkanseffekt

I resultatdelen framkommer att historielärarna uppfattar att externa prov påverkar deras bedömningspraktik på något sätt. På vilka olika sätt och i vilken omfattning som lärarna uppfattar att provet påverkar dem, varierar både inom och mellan de båda lärarkollektiven. Det är dock möjligt att urskilja två övergripande komparativa mönster: 1) KP-lärarna uppfattar i högre grad och mera samstämmigt än HLP-lärarna att de i något avseende påverkas, det vill säga påverkanseffekten är högre och mer homogen hos KP-lärarna. 2) HLP-lärarna uppfattar att det i första hand är den kollegiala interaktionen som påverkas, medan KP-lärarna uppfattar en bredare påverkanseffekt som utöver den kollegiala påverkanseffekten även omfattar såväl en epistemologisk som en metodologisk påverkan.

Skillnaden i påverkanseffekt kan relateras till hur de båda lärarkollektiven använder proven. HLP-lärarna använder HLP i första hand som ett resultatgenererande mätverktyg för att lösa ett metodologiskt likvärdighetsproblem, vilket innebär att provets epistemologiska påverkanseffekt blir begränsad. KP-lärarna använder däremot KP i första hand som ett tolkningsverktyg för att lösa ett epistemologiskt alignmentproblem, vilket innebär att KP:s epistemologiska påverkanseffekt blir betydande.

I ljuset av den fjärde kvadrantens dilemma kan skillnaden i påverkanseffekt dessutom relateras till hur de båda lärarkollektiven uppfattar och genomför proven. Båda lärarkollektiven uppfattar att det finns en diskrepans mellan provens domänlogik och deras egen domänlogik. HLP-lärarna tenderar att vara lojala med HLP:s mättraditionella processreglerande direktiv och att välja bort (bortförhandla) de delar av HLP:s HT-traditionella domänlogik som inte ligger i linje med deras egen domänlogik. Det här sättet att hantera den fjärde kvadrantens dilemma innebär att provet får en begränsad epistemologisk påverkanseffekt. KP-lärarna tenderar att välja ut de delar av KP:s HT-traditionella domänlogik som inte ligger i linje med deras domänlogik och att distansera sig från KP:s mättraditionella processreglerande direktiv. Det här sättet att hantera den fjärde kvadrantens dilemma innebär en omfattande epistemologisk påverkanseffekt.

Sammanfattningsvis indikerar resultatet att provets påverkanseffekt är beroende av den bedömningskulturella kontexten. I en tolkningstraditionell

bedömningskultur är det historiekulturella förhandlingsutrymmet stort och påverkans effekten låg. I en mättraditionell bedömningskultur är det historiekulturella förhandlingsutrymmet begränsat och påverkans effekten hög. Den här skillnaden kommer att behandlas mer utförligt i slutkapitlet.

9.5.2 Kollegial eller vertikal polarisering

Båda lärarkollektiven uppfattar att den kollegiala interaktionen påverkas av proven, men KP-lärarna betonar i högre grad att den kollegiala interaktionen inrymmer en baksida i form av en kollegial polarisering. Den kollegiala polarisering som KP-lärarna beskriver äger rum på realiseringsarenan och handlar i grunden om en kollegial förhandling om domänlogiken. KP-lärarna uppfattar att de behöver stöd för att hantera den specifika tolkning av HT-traditionen som präglar 2011 års domänlogik. När KP-lärarnas HT-traditionella tolkning av domänlogiken påverkas av KP involveras de i en kollegial förhandling med de lärare i ämneslaget som inte använder sig av KP. I förlängningen innebär förhandlingen en polarisering inom ämneslaget, en polarisering mellan historielärare som accepterar 2011 års HT-traditionella domänlogik och de som gör motstånd.

Ett delvis annat sätt att tolka den kollegiala polariseringen är att uppfatta implementeringen av Gyll och KP som att en outtalad informell historiekultur utmanas av en uttalad formell historiekultur. Diskrepansen mellan den formella och den informella historiekulturen skapar en osäkerhet hos professionen. Historielärare måste förhålla sig till detta spänningsfält och här kan man se att historielärare positionerar sig på olika sätt. De lärare som ser ett värde i att ligga nära den formella historiekulturen, som den kommer till uttryck i kursplanen, ser KP som ett stöd för att tolka kursplanen på ”rätt” sätt. De lärare som av någon anledning vill distansera sig från den formella historiekulturen och hålla fast vid den informella historiekulturen, ser tvärtom ett behov att distansera sig från KP.

Den kollegiala polariseringen i samband med KP kan också jämföras med den vertikala förhandlingen i samband med HLP. Som framkom i kapitel 6 var det en förhandling som föranledde en domänlogisk reträtt på formuleringsarenan. HLP-lärarnas bortförhandling av de delar av HLP:s

HT-traditionella domänlogik som inte låg i linje med deras egen domänlogik skapade således en vertikal polarisering mellan formuleringsarenan och realiseringsarenan.

9.5.3 Den tentativa slutsatsen

I kapitlet framkommer att HLP-lärarna uppfattar en begränsad epistemologisk påverkan i samband med HLP. Det är ett resultat som stärker den tentativa slutsatsen att HLP-lärarna inte uppfattar HLP som ett budskapsförmedlande historiekulturellt verktyg utan som ett resultatgenererande bedömningskulturellt verktyg.

KP-lärarna beskriver däremot en omfattande epistemologisk påverkan i samband med KP. Det är ett resultat som stärker den tentativa slutsatsen att KP-lärarna uppfattar KP som ett budskapsförmedlande historiekulturellt verktyg och inte som ett resultatgenererande bedömningskulturellt verktyg.

10. SAMMANFATTNING OCH SLUTDISKUSSION

Det övergripande syftet med forskningsprojektet har varit att undersöka historiekulturell förändring över tid med avseende på skolämnet historia, och då särskilt beakta vilken betydelse kunskapsbedömning i allmänhet och externa prov i synnerhet kan ha för denna förändring. Den temporala brytpunkten för denna förändringsprocess har varit 2011 års skolpolitiska reformer manifesterad i läroplanen Gy11.

Avhandlingens övergripande syfte har operationaliserats i en mikrostudie där jag har undersökt hur historielärare på gymnasiet uppfattar frivilliga externa prov i historia. I studien undersöks historielärares uppfattningar och beskrivningar av 1) provens **konstruktion**, 2) provens **genomförande**, 3) provens **användning** och 4) provens **påverkan** på deras bedömningspraktik.

I studien har åtta historielärare intervjuats varav fyra har använt sig av Historielärarnas prov (HLP) och fyra av Kursprovet i historia (KP). HLP-lärarna intervjuades före 2011 års skolpolitiska reformer medan KP-lärarna intervjuades efter 2011 års skolpolitiska reformer.

De empiriska resultat som studien genererat har genomgående analyserats i ljuset av den fjärde kvadrantens dilemma, det vill säga hypotesen att relationen mellan den historiekulturella HT-traditionen och den bedömningskulturella mättraditionen inrymmer ett latent spänningsförhållande.

10.1 De empiriska resultaten

I studien framkommer att historielärarnas uppfattningar om provens **konstruktion** inrymmer såväl likheter som skillnader. Både HLP-lärarna och KP-lärarna uppfattar att det finns en diskrepans mellan deras egen domänlogiska uppfattning och den domänlogik som kommer till uttryck i provens konstrukt, men deras sätt att hantera denna diskrepans skiljer sig åt. HLP-lärarna gör i högre grad motstånd mot provets domänlogik medan

KP-lärarna i högre grad uppfattar provets domänlogik som den ”rätta” domänlogiken. KP-lärarna uppfattar således i högre utsträckning provets domänlogik som ett normativt underlag.

Studien visar vidare att historielärarnas beskrivningar av hur de **genomför** frivilliga externa prov i historia skiljer sig åt. HLP-lärarna beskriver i högre grad att de är lojala gentemot HLP:s mättraditionella direktiv och att de genomför HLP under standardiserade former. KP-lärarna beskriver däremot i högre grad att de frigör sig från KP:s mättraditionella direktiv och att de genomför provet under mer situerade former.

Studien visar även att historielärarna beskriver att de **använder** frivilliga externa prov i historia på olika sätt. HLP-lärarna beskriver i högre grad att de använder HLP:s *resultat* summativt för att lösa ett likvärdighetsproblem medan KP-lärarna i högre grad beskriver att de använder KP:s *verktyg* formativt för att lösa ett alignmentproblem.

Studien visar slutligen att historielärarna uppfattar att deras bedömningspraktik **påverkas** på olika sätt av frivilliga externa prov i historia. HLP-lärarna uppfattar att deras bedömningspraktik påverkas i begränsad omfattning och att denna påverkan i första hand omfattar den kollegiala interaktionen. KP-lärarna uppfattar däremot att deras bedömningspraktik påverkas i hög utsträckning och att denna påverkan, utöver den kollegiala interaktionen, även omfattar såväl den epistemologiska som den metodologiska dimensionen av deras bedömningspraktik.

Studien i sin helhet visar således att det finns stora skillnader mellan HLP-lärarnas och KP-lärarnas sätt att uppfatta och hantera ett frivilligt externt prov i historia. För att möjliggöra en analys av dessa skillnader har resultaten förankrats i ett teoretiskt ramverk.

10.2 Det teoretiska ramverket

Avhandlingens teoretiska ramverk är omfattande och kan i vissa avseenden beskrivas som teoriutvecklande. Det är framför allt den uttalade ambitionen att fläta samman historiedidaktisk och bedömningsteoretisk teoriutveckling som ligger till grund för denna teoriutveckling. Teoriutvecklingen manifesteras framför allt i hypotesen om den fjärde kvadrantens dilemma, men även i ett antal konceptuella modeller och begreppsutredande avsnitt.

I praktiken innebär de teoriutvecklande delarna i avhandlingen att jag synliggör sambandet mellan bedömningens vad- och hur-dimension på ett

tydligt sätt. Ett exempel är korrelationen mellan det tredelade konstrukt-begreppet och den tredelade processen i samband med bedömningens genomförande. Jag kan utifrån denna teoribaserade grund argumentera för att domänlogiken i första hand ska kopplas till observationen och konstruktionen av observationsverktyg; att progressionslogiken ska kopplas till värderingen och konstruktionen av värderingsverktyg; och att emfas-logiken ska kopplas till sammanvägningen och konstruktionen av sammanvägningsverktyg. Jag vill hävda att den här tredelade korrelationen underlättar synliggörandet av de komplexa epistemologiska och metodologiska vägval som provkonstruktörer och historielärare ställs inför när de ska bedöma elevernas kunskaper i historia. I föreliggande avhandling har den tredelade korrelationen i högsta grad underlättat analysen av de epistemologiska och metodologiska dimensionerna i såväl provens bedömningsverktyg som i lärarnas uppfattningar om dessa verktyg.

Avhandlingens primära portalbegrepp är historiekultur. Historiekulturen, som den kommer till uttryck i skolan, definieras i avhandlingen som *det kommunikativa sammanhang inom vilket olika sätt att uppfatta det förflutna lyfts fram och ges mening*. Detta kommunikativa sammanhang tolkas som en dubbel process, dels som en ständigt pågående *förhandling om* hur historieämnet ska uppfattas, dels som en ständigt pågående *förmedling av* dessa olika uppfattningar. Dessa båda processer är sammanvävda och utgör tillsammans grunden för en specifik historiekultur.

Ett externt prov i historia definieras i avhandlingen som en historiekulturell produkt som konstrueras av provkonstruktörer på en historiekulturell formuleringsarena och som implementeras av en historielärare på en historiekulturell realiseringsarena. Vid konstruktionen och implementeringen utgör det externa provet ett underlag för såväl historiekulturell förhandling som historiekulturell förmedling.

Skolans historiekultur är dock inget isolerat fenomen utan en tråd bland flera i den komplexa kulturella väv som formar både skolan och samhället i stort. Denna kulturella väv inrymmer en mängd olika subkulturella trådar, däribland den bedömningskulturella tråden. I föreliggande projekt görs antagandet att det finns ett betydelsefullt samband mellan den historiekulturella och bedömningskulturella tråden, ett samband som innebär att skolans historiekultur påverkas/omformas av den bedömningskulturella kontexten.

I projektet görs även antagandet att det uppstår ett särskilt spänningsförhållande mellan skolans historiekultur och bedömningskultur under de perioder då skolans historiekultur är HT-traditionellt förankrad samtidigt som skolans bedömningskultur är mättraditionellt förankrad, ett spänningsförhållande som benämns som den fjärde kvadrantens dilemma. I relation till ett externt prov operationaliseras dilemmat på olika sätt beroende på vilken arena man befinner sig. På formuleringsarenan operationaliseras dilemmat vid konstruktionen av provet och på realiseringsarenan vid implementeringen av provet.

Med stöd av det teoretiska ramverket kan de empiriska resultat som redovisats förankras kontextuellt i en förändringsprocess där 2011 års skolpolitiska reformer manifesterade i Gyll utgör den temporala brytpunkten.

10.3 Empiri möter teori

I studien framkommer att historielärarna uppfattar och hanterar det externa provet på olika sätt. En möjlig tolkning av denna skillnad är att proven svarar mot olika behov eller problem. I föreliggande studie har två sådana problemområden utkristalliserats: likvärdighetsproblemet och alignmentproblemet. HLP-lärarna använder provets *resultat* i ett summativt syfte för att hantera ett likvärdighetsproblem medan KP-lärarna använder provets *verktyg* i ett formativt syfte för att hantera ett alignmentproblem. En slutsats, baserad på denna skillnad, som tagit form under analysarbetet är att HLP-lärarna i första hand uppfattar HLP som ett resultatgenererande *bedömningskulturellt* verktyg medan KP-lärarna i första hand uppfattar KP som ett budskapsförmedlande *historiekulturellt* verktyg.

Med stöd av det teoretiska ramverket kan den här problemförskjutningen kontextualiseras och beskrivas som två olika narrativ.

10.3.1 HLP och likvärdighetsproblemet

HLP-narrativet

Under efterkrigstiden präglades historieämnet i Sverige av en gradvis nedmontering som når sin kulmen i samband med de stora skolpolitiska reformerna 1994. Under denna nedmontering fortlever ett historieämne som till stora delar är MH-traditionellt förankrat. 1994 års skolpolitiska reformer innebär dock att ett HT-traditionellt historieämne etableras på

den historiekulturella formuleringsarenan. 1994 års skolpolitiska reformer innebar även en övergång från ett normrelaterat till ett målrelaterat bedömningssystem. Den här övergången kan beskrivas som en övergång från en mättraditionell till en tolkningstraditionell bedömningskultur. Likvärdigheten i det målrelaterade bedömningssystemet skulle garanteras genom att nationellt uppsatta kunskapskriterier tolkades lokalt och att dessa tolkningar sedan mynnade ut i en kollegial samsyn. Tilliten till professionen var initialt mycket hög och lärarna fick till och med i uppdrag att skriva fram egna lokala betygskriterier. Efterhand kom dock allt fler signaler om en stigande betygsinflation, vilket gav upphov till en diskussion huruvida lärarnas värdering av elevernas kunskaper verkligen var likvärdig.

I den här situationen går HLP-lärarna i bräsch och skapar en kollegial formuleringsarena där ett antal kollegor inom Skåneregionen tillsammans utformar ett externt prov som kan lösa likvärdighetsproblemet. När HLP-konstruktörerna tog sig an uppdraget ställdes de inför den fjärde kvadrantens dilemma, det vill säga att konstruera ett prov som hanterade ett HT-traditionellt konstrukt samtidigt som det genererade mätsäkra provresultat. De försökte hantera den fjärde kvadrantens dilemma genom att konstruera mättraditionella värderingsverktyg som gjorde anspråk på att mäta kvaliteten i elevprestationer genererade av tolkningstraditionella observationsverktyg.

På realiseringsarenan manifesteras den fjärde kvadrantens dilemma i samband med provets implementering. För HLP-lärarna handlar dilemmat om i vilken utsträckning de förmår vara lojala med det prov som HLP-konstruktörerna hade konstruerat. När HLP-lärarna beskriver hur de uppfattar, genomför och använder sig av HLP framkommer att de i hög grad är lojala med HLP:s mättraditionella format och direktiv. De uttrycker inte något större motstånd mot de mättraditionella värderingsverktygen. De genomför provet under standardiserade former och de använder provets resultat i ett summativt syfte. Motsvarande lojalitet gäller dock inte HLP:s HT-traditionella domänlogik. När HLP-lärarna beskriver hur de uppfattar provets konstruktion framkommer att det finns diskrepanser mellan deras egen domänlogik och den domänlogik som ligger till grund för HLP. Denna diskrepans manifesteras framför allt i deras uppfattningar om provets D-uppgift, det vill säga den uppgift som har den

tydligaste förankringen i en kontinentalt förankrad HT-traditionell domänlogik.

HLP-narrativet och historiekulturens dubbla processer

När HLP-konstruktörerna exkluderar D-uppgiften från HLP kan deras agerande tolkas som ett resultat av en historiekulturell förhandling, en *förhandling* om de kontinentala delkomponenterna inom den HT-traditionella domänlogiken. Att HLP-lärarna lyckas bortförhandla D-uppgiften indikerar att deras förhandlingsposition var stark. Denna starka förhandlingsposition kan relateras till de tolkningstraditionella ideal som präglar bedömningskulturen under perioden. Idealet att bejaka professionens tolkningsutrymme gav HLP-lärarna stora möjligheter att omtolka eller bortse ifrån delar av den domänlogik som skrivs fram i styrdokumentet.

HLP-lärarnas uppfattningar att deras bedömningspraktik endast i begränsad omfattning påverkas av HLP kan på motsvarande sätt tolkas som att provets budskapsförmedlande funktion är förhållandevis svag. HLP-lärarna uppfattar HLP i första hand som ett resultatgenererande bedömningskulturellt mätverktyg för att lösa ett likvärdighetsproblem och inte som ett budskapsförmedlande historiekulturellt tolkningsverktyg för att tolka den nya kursplanen. Likvärdighetsproblemet överskuggar alignmentproblemet och policy-influenser i form av styrdokument och externa prov uppfattas inte som ett normativa underlag utan som diskussionsunderlag. Provets historiekulturella påverkanseffekt blir därför begränsad och den domänlogiska diskrepansen mellan formuleringsarenan och realiseringsarenan består.

HLP-narrativet i relation till tidigare forskning

Av HLP-narrativet framgår att det vid implementeringen av HLP fanns domänlogiska diskrepanser mellan den HT-traditionella domänlogiken i historieämnets kursplaner 1994/2000 och HLP-lärarnas egen domänlogik. HLP-lärarna i min studie beskriver inte explicit sin egen bedömningspraktik, men i deras uppfattningar om HLP:s domänlogik framgår hur HLP-lärarna framförallt brottas med begreppet ”historiemedvetande”, manifesterad i de HLP-uppgifter som gjorde anspråk på att pröva en tredelad temporal tankeoperation.

Den domänlogiska diskrepansen som framkommer i HLP-narrativet ligger delvis i linje med tidigare alignment-baserade studier inom fältet. I

Rosenlunds studie framkommer att historielärare verksamma på gymnasiet under den aktuella perioden i hög grad fortsatte att konstruera bedömningsuppgifter som endast mätte innehållsrelaterade minneskunskaper, det vill säga uppgifter med en tydlig MH-traditionell förankring.³⁸⁰ Fribergs studie förstärker bilden av en MH-traditionell bedömningspraktik då han i sin studie konstaterar att det i första hand tycks vara innehållskunskaper som utgör den primära basen för lärarnas bedömningar i klassrummet: "Innehåll först, förmågor sedan".³⁸¹ Rosenlund och Fribergs studier indikerar en mer omfattande domänlogisk diskrepans än den som framkommer i föreliggande studie.

Diskrepansen som framkommer i HLP-narrativet ligger mer i linje med resultaten i Samuelssons studie. Till skillnad från Rosenlund och Friberg, beskriver han en större spännvidd mellan lärarnas bedömningspraktik. Tre av de åtta lärarna i hans studie avviker förvisso från de HT-traditionella intentionerna i 1994/2000 års kursplan och Samuelsson beskriver deras bedömningspraktik som "deklarativ allmändidaktisk praxis" där fokus ligger på att eleverna ska "skriva av rätt" och "minnas". Denna deklarativa kunskapssyn med fokus på "reproduktion av det förflutna" motsvarar i hög grad den MH-traditionella domänlogiken. Två av de åtta lärarna i hans studie tillämpar dock en mer "processuell historiedidaktisk praxis", i linje med den HT-traditionella domänlogiken.³⁸² HLP-lärarnas uppfattningar att den källkritiska uppgiften i HLP prövar en självklar och viktig del av historieämnets domänlogik indikerar att även deras domänlogik ligger i linje med den HT-traditionella domänlogiken och att den domänlogiska diskrepansen som framkommer i HLP-narrativet i första hand omfattar de kontinentala delarna av den HT-traditionella domänlogiken.

10.3.2 KP och alignmentproblematiken

KP-narrativet

2011 års skolpolitiska reformer manifesterade i Gyll innebar att historieämnet uppgraderades och att det återigen blev ett ämne som alla elever på gymnasiet skulle läsa. Det historieämne som nu etableras på formuleringsarenan och som skrivs fram i 2011 års kursplaner är på samma sätt som historieämnet i 1994/2000 års kursplaner HT-traditionellt förankrat.

³⁸⁰ Rosenlund (2016).

³⁸¹ Friberg (2017).

³⁸² Samuelsson (2015).

Den nya kursplanen, som till skillnad från den föregående kursplanen formulerades av historiedidaktiker, hade en tydligare teoriförankring, vilket innebar att historielärarna behövde hantera domänlogiska delkomponenter som inte hade varit lika framträdande i 1994/2000 års kursplaner. Dessa delkomponenter handlar i första hand om hur historia kan användas och om betydelsen av det framtidsinriktade perspektivet i tolkningen av historiska förändringsprocesser.

På den bedömningskulturella formuleringsarenan sker en tydlig förskjutning i samband med 2011 års skolpolitiska reformer, en förskjutning som innebar att mättraditionella aktörer flyttade fram sina positioner. Det likvärdighetsproblem som HLP-lärarna försökte hantera på en regional formuleringsarena lyftes i samband med 2011 års reformer upp till den nationella formuleringsarenan. Betygsinflationen skulle hejdas genom tydligare styrdokument och fler externa prov.

I efterdyningarna av 2011 års mättraditionella förskjutning gav Skolverket KP-konstruktörerna i uppdrag att konstruera mätsäkra prov som prövade det HT-traditionella konstrukt som skrevs fram i 2011 års kursplaner. Det innebar att KP-konstruktörerna, på samma sätt som HLP-konstruktörerna, ställdes inför den fjärde kvadrantens dilemma. Initialt försöker KP-konstruktörerna att hantera dilemmat genom att konstruera ett prov som inrymde en blandning av mättraditionella observations- och värderingsverktyg. Efterhand övergår de dock till att hantera dilemmat på samma sätt som HLP-konstruktörerna, det vill säga genom att konstruera mättraditionella värderingsverktyg som gjorde anspråk på att mäta kvaliteten i elevprestationer genererade av tolkningstraditionella observationsverktyg.

På motsvarande sätt som för HLP-lärarna innebar implementeringen av KP att KP-lärarna ställdes inför den fjärde kvadrantens dilemma. När KP-lärarna beskriver hur de uppfattar KP:s konstruktion framkommer på samma sätt som hos HLP-lärarna att det finns domänlogiska diskrepanser mellan deras egen domänlogik och den domänlogik som kommer till uttryck i provet. Denna diskrepans manifesterades framförallt i de uppgifter som betonar betydelsen av det framtidsinriktade perspektivet i tolkningen av historiska förändringsprocesser. Till skillnad från HLP-lärarna är dock KP-lärarna lojala med provets HT-traditionella domänlogik, som de uppfattar som svår men ”rätt”. Diskrepansen uppfattas således i högre grad som ett alignmentproblem. KP-lärarnas lojalitet omfattar däremot inte

KP:s mättraditionella direktiv. KP-lärarna tonar ner de mättraditionella ambitionerna som KP inrymmer och de beskriver en situerad observation där de väljer ut KP-uppgifter som prövar svårtolkade delkomponenter i det HT-traditionella konstruktet.

KP-narrativet och historiekulturens dubbla processer

Till skillnad från HLP-lärarna väljer KP-lärarna att vara lojala med den HT-traditionella domänlogiken. Den domänlogiska diskrepansen ligger därför inte till grund för någon *förhandling* om de kontinentala delkomponenterna inom den HT-traditionella domänlogiken. Att KP-lärarna, till skillnad från HLP-lärarna, inte bortförhandlar de uppgifter som inte ligger i linje med deras egen domänlogik indikerar att deras förhandlingsposition är svag. Denna svaga förhandlingsposition kan relateras till de mättraditionella ideal som präglar periodens bedömningskultur. Till skillnad från HLP-perioden är det centralisering och kontroll av professionen som kännetecknar periodens bedömningskulturella credo. Policy-influenser i form av styrdokument och externa prov uppfattas som ett normativt underlag. Idealet att begränsa professionens tolkningsutrymme gav KP-lärarna små möjligheter att omtolka eller bortse ifrån den domänlogik som skrivs fram i 2011 års kursplan.

När KP-lärarna beskriver att de använder KP som ett formativt tolkningsverktyg och att deras bedömningspraktik påverkas i hög utsträckning kan det uppfattas som att provet ligger till grund för en historiekulturell budskapsförmedling, en *förmedling* som i första hand omfattar de budskap som handlar om de kontinentala delkomponenterna inom den HT-traditionella domänlogiken. KP-lärarna använder KP som ett budskapsförmedlande historiekulturellt verktyg för att förtydliga domänlogiken i 2011 års kursplan. Noteras kan att KP-lärarna uppmärksammar domänlogiken i högre utsträckning än progressionslogiken. En möjlig tolkning av denna prioritering är att KP-lärarna behövde förankra domänlogiken i sin egen ämnesuppfattning innan de var redo att ta sig an progressionslogiken. De måste försäkra sig om att de mäter rätt innan de tar sig an uppdraget att mäta rättvist.

I studien framkommer dock att inte alla historielärare uppfattar KP som ett budskapsförmedlande historiekulturellt verktyg. KP-lärarna beskriver en kollegial polarisering inom ämneslaget, en polarisering mellan lärare

som bortförhandlar de nya delarna av ämneslogiken och lärare som använder provet för att införliva de nya komponenterna i sin egen domänlogik.

KP-narrativet i relation till tidigare forskning

Av KP-narrativet framgår att det vid implementeringen av KP fanns domänlogiska diskrepanser mellan den HT-traditionella domänlogiken i historieämnets kursplaner 2011 och KP-lärarnas egen domänlogik. Narrativet tydliggör även att KP-lärarnas bedömningspraktik i hög utsträckning påverkas av KP och att denna påverkan i första hand handlar om domänlogiken och de kontinentala delkomponenterna inom den HT-traditionella domänlogiken. Narrativet visar slutligen att denna påverkan förleder en kollegial polarisering inom KP-lärarnas ämneslag.

Den påverkansseffekt som framkommer i KP-narrativet ligger delvis i linje med tidigare effekt-baserad forskning. I Jarhalls studie intervjuas historielärare på grundskolan före och efter Lgr11.³⁸³ Ett av studiens resultat är att historielärare verksamma efter Lgr 11 uppfattar att de nationella prov i historia som introduceras 2013 påverkar såväl deras val av centralt innehåll som det sätt som de konstruerar uppgifter och bedömer elevernas kunskaper. I studien framkommer bland annat att de uppfattar att de nationella proven har påverkat dem att i högre utsträckning betona källhantering och historiebruk i sin undervisning, det vill säga en förskjutning mot en mer HT-traditionell domänlogik.

I jämförelse med KP-narrativet framkommer två viktiga skillnader. Det ena är att påverkansseffekten hos grundskolelärarna tycks vara högre och omfatta även det den substantiella dimensionen av domänlogiken (det centrala innehållet) och det andra är att den polarisering som framkommer hos KP-lärarna på gymnasiet uteblir hos grundskolelärarna. Båda dessa skillnader kan relateras till bedömningskulturen och till skillnaderna mellan de båda externa proven. Det nationella provet är obligatoriskt och provresultaten ska vara ett stöd vid betygssättningen. KP var däremot ett frivilligt prov och det är läraren själv som avgör på vilket sätt resultaten ska användas. Förekomsten av ett obligatoriskt betygstödjande externt prov kan tolkas som att grundskolans bedömningskultur har en tydligare

³⁸³ Jarhall (2020).

mättraditionell förankring och att den historiekulturella budskapsförmedlingen blir mer uttalad. Även den kollegiala polariseringen kan ses i ljuset av den bedömningskulturella dimensionen. Avsaknaden av ett obligatoriskt prov på gymnasiet innebär att bedömningskulturens mättraditionella förankring inte är lika uttalad, vilket ger KP-lärarnas kollegor ett större historiekulturellt förhandlingsutrymme.

Även Fickels studie belyser bedömningskulturens betydelse för påverkanseffekten. I hennes studie om hur interaktionen inom ett ämneslag i Kentucky påverkas av ett externt prov framkommer hur påverkanseffekten förändras när elevernas resultat försämrats. Hon beskriver hur fokus flyttas från styrdokumentet till provresultatet; från undervisningspraktiken till bedömningspraktiken, vilket bland annat medförde att lärarna i högre utsträckning än tidigare började samtala och samarbeta om hur provfrågor skulle konstrueras och hur provsvarens kvalitet skulle värderas.³⁸⁴

Fickels studie visar i linje med föreliggande studie att externa prov har en hög påverkanseffekt i en mättraditionell bedömningskultur. Den visar dock tydligare än föreliggande studie att konsekvenserna av denna påverkanseffekt kan variera över tid. Den initiala påverkanseffekten hos Kentucky-lärarna med fokus på styrdokument och ämnesutveckling påminner i hög grad om KP-lärarnas sätt att uppfatta det externa provet som ett budskapsförmedlande historiekulturellt verktyg till stöd för att tolka domänlogiken i det konstrukt som kommer till uttryck i styrdokumentet. Den senare påverkanseffekten hos Kentucky-lärarna är fortfarande hög men fokus förflyttas i högre grad till konstruktets progressionslogik och till den konkreta konstruktionen av värderingsverktyg. Skillnaden kan återigen tillskrivas frivillighetsaspekten i KP-provet, vilket som sagt innebär att bedömningskulturens mättraditionella förankring inte är lika uttalad, vilket ger KP-lärarnas kollegor ett större historiekulturellt förhandlingsutrymme.

Avslutningsvis vill jag kortfattat ställa frivillighetsaspekten och resultaten i föreliggande studie i relation till de effekt-baserade studier som i första hand belyser de negativa effekterna av externa prov. Som framgår av Berg & Perssons forskningsöversikt utgör de ”policy-kritiska” studierna en huvudfåra inom fältet och tidigare forskning har uppmärksammat

³⁸⁴ Fickel (2006).

en bred palett av negativa effekter som omfattar allt från att proven påverkar historielärare att begränsa sin egen domänlogik ("narrowing the construct") till den tudelning av undervisningen som framkommer i Grants studie.³⁸⁵ I föreliggande studie där frivilliga externa prov utgör analysobjekt är de negativa effekterna begränsade, tvärtom uppfattar såväl HLP-lärarna som KP-lärarna de externa proven som ett stöd för ämnesutvecklingen. Frivillighetsaspekten är viktig att beakta och kommer att lyftas fram i slutdiskussionens avslutande del.

10.4 Slutdiskussion

10.4.1 Den fjärde kvadrantens dilemma

Ett "dilemma" innebär enligt SAOL att ställas inför en valsituation och oavsett vad du väljer kommer ditt val att medföra negativa konsekvenser. Valsituationen i den fjärde kvadrantens dilemma innebär att ställas inför valet att vara lojal med de mättraditionella kraven på standardisering och mätsäkerhet eller att vara lojal med det HT-traditionella kravet på en konstruktivistisk ämnesuppfattning. Om man som historielärare eller provkonstruktör väljer att vara lojal med de mättraditionella kraven blir den negativa konsekvensen en alignment-problematik, det vill säga problem med att konstruera bedömningsverktyg som uppfattas som valida i förhållande till det HT-traditionella konstruktet. Om man å andra sidan väljer att vara lojal med de HT-traditionella kraven blir den negativa konsekvensen en likvärdighetsproblematik, det vill säga problem med att säkerställa reliabla bedömningsresultat.

10.4.2 Den heliga graal

I studien framkommer att historielärare och provkonstruktörer som ställs inför den fjärde kvadrantens dilemma har tre alternativ. Det första alternativet är att retirera metodologiskt, det vill säga att tona ner kraven på mätsäkerhet, till exempel genom att använda öppna essäuppgifter. En metodologisk reträtt, det vill säga en fokusförskjutning i tolkningstraditionell riktning, innebär att bruket av öppna verktyg och direktiv ökar, något som skapar ett ökat tolknings- och handlingsutrymme, vilket i sin tur kan leda till en betygsinflation. Det andra alternativet är att retirera epistemologiskt, det vill säga att förskjuta bedömningens fokus i riktning mot mer

³⁸⁵ Berg & Persson (2020). Grant (2003).

minnesbaserade kunskaper. En sådan reträtt får dock betydande konsekvenser. En epistemologisk reträtt, det vill säga en fokusförskjutning i MH-traditionell kan få såväl domänlogiska konsekvenser ("narrowing the construct") som progressionslogiska konsekvenser i form av kvantifiering. Det tredje alternativet är att stanna kvar och ta sig an uppdraget att lösa dilemmat genom att konstruera mättraditionella bedömningsverktyg som gör anspråk på att pröva kunskaper förankrade i en HT-traditionell epistemologi. Detta tredje alternativ kan beskrivas som forskningsfältets motsvarighet till sökandet efter den heliga graal.

Föreställningen att det är möjligt att verkligen finna den heliga graal, det vill säga "lösa" den fjärde kvadrantens dilemma, är dock som sagt problematisk. Fokus på likvärdighet, det vill säga en vertikal reträtt, kommer att skapa ett alignmentproblem medan fokus på alignment, det vill säga en horisontell reträtt, kommer att skapa ett likvärdighetsproblem. Istället för att hitta en lösning på dilemmat handlar det snarare om att i den uppkomna valsituationen sträva efter en balans mellan två negativa konsekvenser. Den heliga graal ska därför inte uppfattas som en lösning på dilemmat, utan som en hantering av detsamma.

Därmed inte sagt att de som är verksamma inom fältet saknar valmöjligheter i sin strävan att konstruera bedömningsverktyg som skapar balans mellan de mättraditionella och de HT-traditionella idealen. I studien ser man hur provkonstruktörerna för HLP och KP försöker hantera dilemmat på olika sätt. KP-konstruktörerna gör initialt ett försök att konstruera mättraditionella observationsverktyg som gör anspråk på att vara valida i relation till det HT-traditionella konstruktet. Efterhand överges detta försök och de försöker istället, på samma sätt som HLP-konstruktörerna, att förskjuta dilemmat från observationen till värderingen genom att konstruera mättraditionella värderingsverktyg som gör anspråk på att särskilja kvaliteter i elevsvar som genererats av HT-traditionella essäuppgifter. Huruvida detta sätt att förskjuta dilemmat till värderingen är en framkomlig väg för att hantera den fjärde kvadrantens dilemma kan diskuteras. Kanske är en epistemologisk eller metodologisk reträtt i någondera riktningen nödvändig för att skapa en solid bedömning?

10.4.3 Den dubbla pendelrörelsen

I studien framkommer vidare att den fjärde kvadrantens dilemma i grunden är ett latent och mer djupgående dilemma, med omfattande historiekulturella och historiedidaktiska konsekvenser. I ett längre perspektiv är det också tydligt att dess relevans har varierat över tid. En vanligt förekommande metafor för att beskriva förändring över tid är pendelrörelsen. Eftersom den fjärde kvadrantens dilemma är en tvådimensionell tankemodell krävs dock en metafor i form av en dubbel pendelrörelse som omfattar såväl den historiekulturella som den bedömningskulturella dimensionen. Metaforens komplexitet stegras ytterligare om man dessutom beaktar att dessa dubbla pendelrörelser kan analyseras på olika arenor. I studien framkommer att formuleringsarenans pendelrörelser inte alltid korrelerar med realiseringsarenans pendelrörelser.

Under perioden 1994–2011, det vill säga perioden före Gyll, har formuleringsarenans historiekulturella pendel en tydlig HT-traditionell förankring. På realiseringsarenan är denna förankring inte lika tydlig. Förhandlingen i samband med implementeringen av de nya HT-traditionella kursplanerna manifesterar en historiekulturell pendelrörelse i otakt, en slags eftersläpning på realiseringsarenan. Även den bedömningskulturella pendelrörelsen hamnar under motsvarande period i otakt. På formuleringsarenan utarbetades en ny läroplan som genomsyrades av den tolkningstraditionella betoningen på tolkningsfrihet. På realiseringsarenan är den tolkningstraditionella förankringen inte lika tydlig. HLP-lärarnas behov av ett mättraditionellt externt prov för att hantera likvärdighetsproblemet manifesterar denna eftersläpning.

Under perioden efter Gyll förblir formuleringsarenans historiekulturella pendel HT-traditionellt förankrad. På realiseringsarenan ser man en polarisering mellan lärare som i samband med implementeringen anammar den nya HT-traditionella kursplanen och de lärare som gör motstånd. Vad gäller den bedömningskulturella pendeln sker en förskjutning i mättraditionell riktning manifesterad i styrdokumentet och KP. HLP är en regional förlöpare och Gyll den slutliga manifestationen av denna mättraditionella pendelrörelse.

Om man sätter dessa pendelrörelser i relation till varandra kan man se att den fjärde kvadrantens dilemma före Gyll förvisso är ett relevant problem för HLP-konstruktörerna, men eftersom den bedömningskulturella pendeln på den nationella formuleringsarenan är tolkningstraditionellt

förankrad blir dilemmat inte lika relevant för HLP-lärarna i deras egen bedömningspraktik. Efter Gyll är dilemmat i högsta grad relevant för såväl provkonstruktörer som historielärare. Hur dessa båda pendelrörelser kommer att förhålla sig till varandra framgent är dock i högsta grad en öppen fråga.

10.4.4 I kristallkulan

Det finns signaler om att den bedömningskulturella pendeln på formuleringsarenan fortsätter att röra sig i mättraditionell riktning. Tydligast syns dessa signaler i grundskolan och då särskilt i samband med utformningen och hanteringen av de nationella proven. Sedan 2018 ska de nationella provens resultat särskilt beaktas vid betygssättningen och med början 2024 kommer dessutom de externa proven att digitaliseras. I samband med digitalisering kommer både bedömningsdirektiven och bedömningsverktygen att få en mer mättraditionell prägel. Vad händer med skolans historiekultur om dessa mättraditionella vindar blir bestående och den bedömningskulturella pendeln kvarstannar i sitt mättraditionella läge? I denna avslutande del av slutdiskussion tar jag mig friheten att blicka in i kristallkulan och där tycker jag mig se två tänkbara kontrasterande historiekulturella scenarion.

I det första scenariot har den historiekulturella pendeln förskjutits i en MH-traditionell riktning. I hanteringen av den fjärde kvadrantens dilemma har de mättraditionella kraven på reliabilitet inneburit en epistemologisk förskjutning som bland annat medfört en omfattande revidering av kursplanens konstrukt. Det är en revidering som omfattar samtliga delar av konstruktet, det vill säga såväl domänlogiken som emfas- och progressionslogiken. Domänlogiskt innebär revideringen att de procedurala delarna av domänlogiken begränsas och det är framför allt de svårbedömda delkomponenterna av domänlogiken som exkluderas, i första hand historiebruk och framtidsperspektivet i den tredelade temporal tankeoperationen. Den substantiella dimensionen av domänlogiken breddas där emot, flera narrativ inryms nu i det centrala innehållet. Emfaslogiskt innebär revideringen att förmågan att kunna rekonstruera ett givet narrativ (kunskaper om historia) betonas medan förmågan att (de)konstruera narrativ har tonas ned. De innebär i sin tur att innehållsbegreppen vinner terräng och att de metahistoriska begreppen tappar mark. Progressions-

logiskt innebär revideringen att kunskapskravens optimumnormer kvantifieras, vilket indirekt innebär att substantiella breddkunskaperna värderas högre än procedurala djupkunskaper. Vissa svårbedömda delar av domänlogiken blir dessutom progressionslösa, det vill säga helt sakna optimumnormer. I klarspråk innebär det här första historiekulturella scenariot att "faktakunskapernas" egenvärde lyfts fram. Historieläraren uppdrag blir mer inriktat på MH-traditionell allmänbildning än på samtidsförståelse och framtidsorientering

I det andra scenariot kvarstannar den historiekulturella pendeln i sitt HT-traditionella läge och det genomförs inte några större revideringar av de nuvarande HT-traditionella kursplanerna. Det här är ett scenario som kommer att ställa höga krav på utveckling av historielärarnas bedömningskompetens eftersom de behöver hantera den fjärde kvadrantens dilemma i sin egen bedömningspraktik. Stöd för en sådan kompetensutveckling skulle kunna ske på olika sätt. En fortbildningsinsats riktad mot strategiska delar av HT-traditionen, likt den som lanseras av Skolverket i samband med implementeringen av Lgr22, kan vara en möjlighet. Etablerandet av en provbank med frivilliga externa prov är en annan. Dessa prov kan som framgår av studien användas av historielärarna för att tydliggöra den HT-traditionella domänlogiken. Det vilar dock ett tungt ansvar på provkonstruktörerna att utforma sådana externa prov som inrymmer en rimlig balans mellan öppna och slutna observationsverktyg. Det innebär att även fortsättningsvis försöka konstruera slutna uppgifter som prövar den HT-traditionella domänlogiken, men sådana prov behöver också inrymma öppna uppgifter vilka på ett mer transparent sätt tydliggör domänlogiken. Emfaslogiken i de proven bör vara tydligt formulerad och utgöra ett stöd för historielärarna när de ska avgöra vilken tyngd de procedurala delarna av domänlogiken ska ha.

Den största utmaningen i det här scenariot handlar dock om progressionslogiken. I föreliggande studie framkom att både HLP-konstruktörerna och KP-konstruktörerna försökte hantera den fjärde kvadrantens dilemma genom att förskjuta det till värderingsprocessen. Om historielärare i sin kompetensutveckling inte får tillgång till värderingsverktyg som ger tydliga riktlinjer om hur elevernas förväntade kognitiva utveckling från nybörjare till expert manifesteras riskerar värderingen att baseras på naiva vardagliga kognitiva teorier.

Därför är behovet av framtida historiedidaktisk forskning som fokuserar på progressionslogikens optimumnormer stort. Det är dock viktigt att denna forskning har en uttalat praktisknära ansats och att den utgör en brygga mellan grundforskning och lärares professionella kompetensutveckling och yrkesutövning. Tidigare forskning med en sådan inriktning inom fältet har haft en metodologisk slagsida mot intervjubaserade studier. Föreliggande studie förstärker bilden av denna slagsida. Behovet av framtida studier inom fältet med andra metodologiska ansatser är därför stort, till exempel observationsstudier och/eller text- och dokumentbaserade studier. Jag ser även ett behov av att mikrostudier i linje med föreliggande studie följs upp av andra och större studier med högre generaliseringsambitioner.

Oavsett vilka möjligheter som finns i detta scenario för historieämnet att hantera det jag kallar den fjärde kvadrantens dilemma, så är det grundläggande villkoret att det historiekulturella och bedömningskulturella dilemmat tydliggörs. Det är min ambition att föreliggande studie ska kunna bidra till det.

Avslutningsvis skulle jag vilja återkoppla till avhandlingens inledning och det som där sägs om studiens relevans samt till Ingrid Carlgren. Hon beskriver att mycket av den forskning som bedrivs inom utbildningsområdet utgörs av forskning om och på lärare och hon efterfrågar i högre utsträckning forskning för lärare. Hon menar vidare att utgångspunkten för en sådan forskningsansats bör tas i ett uppfattat problem i verksamheten eller i en fråga som har anknytning till ett uppfattat problem. I föreliggande avhandling har den fjärde kvadrantens dilemma utgjort detta problem och min förhoppning är att den här avhandlingen kan vara till stöd för historielärare när de tar sig an framtida utmaningar som handlar om kunskapsbedömning i en föränderlig historiekultur.

SUMMARY

Introduction

Is it possible to assess knowledge of history in a measurable way without affecting our way of perceiving what is meant by historical knowledge? This question is the starting point for the overall purpose of this thesis that is to investigate how historical culture, as it is expressed in the school subject of history, changes over time. The analytical focus for the study has been to explore the significance of knowledge assessment in general, and of voluntary external tests in particular, for this change. The temporal focus for the study has been the 2011 school policy reforms manifested in the curriculum for upper-secondary school (Gy11).

The overall purpose has been operationalized in a micro-study where I have investigated how history teachers in upper-secondary school perceive voluntary external tests in history. In the study I examine history teachers' perceptions of 1) the design of the tests, 2) the implementation of the tests, 3) the use of the tests and 4) the impact of the tests on their assessment practice.

Theory

The theoretical framework is extensive and can in some respects be described as theory-developing. It is above all the ambition to combine theory of history didactics and theory of assessment that forms the basis for this theory development.

Historical culture is the central concept of the dissertation. The concept is defined as the communicative context in which different ways of perceiving the past are highlighted and given meaning. This communicative context is interpreted as a double process, partly as an ongoing *negotiation* of how the subject of history should be perceived, and partly as an ongoing *communication* of these different perceptions. These two processes are intertwined and together form the basis of a specific historical culture.

However, historical culture is not an isolated phenomenon but a thread among several in the complex cultural fabric that shapes both the school and society at large. This cultural fabric contains a variety of subcultural threads, including the cultural thread of assessment. An assumption is made that there is a significant connection between the two threads of historical culture and assessment, a connection which means that the school's history culture is influenced and transformed by the school's assessment-cultural context and vice versa.

In line with the theory-developing ambition, I intend to test the overall hypothesis that the relationship between the school's historical-cultural dimension and the assessment-cultural dimension contains a latent tension referred to as the dilemma of the fourth quadrant. The dilemma is based on the basic assumption that there is always a connection between *what* is assessed and *how* something is assessed, in this context between the historical-cultural what-dimension and the assessment cultural how-dimension.

The historical-cultural what-dimension is multifaceted, but analytically it is possible to distinguish two broad epistemological traditions: the *Memory History tradition* (MH tradition) and the *Historical Thinking tradition* (HT tradition). The former is rooted in a more objectivist and the latter in a more constructivist view of historical knowledge.

The assessment-cultural how-dimension is also complex, but in the field of assessment theory it is common to distinguish between a psychometrically anchored *measurement-tradition* and pedagogically anchored *interpretation-tradition*. Slightly simplified, it can be said that the former to a greater extent methodologically limits the assessor's room for interpretation, while the latter to a greater extent affirms the assessor's room for interpretation.

As stated, the hypothesis of the dilemma of the fourth quadrant is based on the basic assumption that there is always a connection between the school's historical-cultural dimension and the assessment-cultural dimension. A second assumption is that this connection contains a latent special tension between the HT-tradition and the measurement-tradition. In other words, it is problematic to combine HT-traditional epistemology with measurement-traditional methodology.

Method

The empirical material was collected from qualitative interviews with eight history teachers, of whom four has used the History Teachers' Test (HLP) and four has used the Course Test in History (KP). The HLP teachers were interviewed before Gyll, while the KP teachers were interviewed after Gyll.

The empirical material was collected from qualitative interviews with eight history teachers. The first sub-study included four teachers during the year 2009 who had used a voluntary external test named the History Teachers' Test (HLP). The second sub study included four teachers during the year 2016 four who had used a voluntary external test named the Course Test in History (KP).

In a qualitative survey, it is often impossible to strive for a representative sample in a more strict sense, but I have had an ambition that the sample should be as varied as possible. Since the tests are voluntary, the basic variable is that the teacher has used one of the tests. I also had the ambition to let gender and experience of teaching constitute variables, but especially the latter variable was in practice difficult to live up to. The interviewed teachers are employed at eight different upper secondary schools located in southern Sweden.

The interviews were recorded, transcribed and interpreted using an abductive principle of shifting between theory and empiricism. I began the analysis work by tracing statements that can be described as key statements, that is specific pieces of data that is empirically rich and contains analytical possibilities. Then I designed a theory-based coding of teachers' statements linked to the research questions that form the basis of the study. However, many of the teachers perceived the impact of the tests on their assessment practice in a way that differed from the initial coding. These perceptions therefore, in line with the abductive principle, came to constitute a specific coding.

Results

The study as a whole shows that there are major differences between the HLP teachers' and KP teachers' ways of perceiving a voluntary external test in history.

The study shows that the history teachers' perceptions of the **design of the tests** contain both similarities and differences. Both the HLP teachers

and the KP teachers perceive that there is a discrepancy between their own view of the subject of history and the view that is expressed in the test construct, but their ways of dealing with this discrepancy differ. The HLP teachers to a greater extent oppose the view expressed in the test, while the KP teachers to a greater extent perceive the view in the test as the “correct” view. The KP teachers thus to a greater extent perceive the view of the subject of history that is expressed in the test as a normative basis.

The study further shows that the history teachers' perceptions of the **implementation of the tests** differ. HLP teachers describe to a greater extent that they are loyal to HLP's measurement traditional directives and that they implement HLP in standardized forms. The KP teachers, on the other hand, describe to a greater extent that they free themselves from the KP's measurement traditional directives and that they implement the test in more situational forms.

The study also shows that history teachers perceive the **use of the tests** in different ways. The HLP teachers describe to a greater extent that they use HLP's *results* in a summative way to solve an equivalence problem, while the KP teachers to a greater extent perceive that they use KP's *tools* formatively to solve an alignment problem.

Finally, the study shows that history teachers perceive **the impact of the tests** on their assessment practice in different ways. The HLP teachers perceive that their assessment practice is affected to a limited extent and that this influence primarily consists of the collegial interaction. The KP teachers, on the other hand, perceive that their assessment practice is influenced to a great extent and that this influence, in addition to the collegial interaction, also includes both the epistemological and the methodological dimension of their assessment practice.

Discussion

The study shows that history teachers perceive and handle the external tests in different ways. One possible interpretation of this difference is that the tests respond to different needs or problems. In the study, two such problem areas have emerged, the equivalence problem and the alignment problem. A conclusion based on the empirical results is that the HLP

teachers use the test results for a summative purpose to deal with an equivalence problem, while the KP teachers use the test tools for a formative purpose to deal with an alignment problem.

In line with the above reasoning, the HLP teachers' perceptions that their assessment practice only to a limited extent is affected by HLP can be interpreted that the test's communicative function is relatively weak. An overall conclusion that follows is that the HLP teachers primarily perceive HLP as a result-generating assessment-cultural tool for solving an equivalence problem. The equivalence problem overshadows the alignment problem.

The KP teachers' perceptions that their assessment practice is affected to a great extent by KP can in a similar way be interpreted as the test's communicative function being relatively strong. An overall conclusion that follows is that the KP teachers primarily perceive KP as a historical-cultural message-conveying tool for interpreting the syllabus. The alignment problem overshadows the equivalence problem.

The problem shift from an equivalence problem to an alignment problem can be related to the overall purpose of the thesis, that is to investigate how and why historical culture, as it is expressed in the school subject of history, changes over time. A common metaphor for describing change over time is the pendulum motion. However, since the dilemma of the fourth quadrant is a two-dimensional model of thought, a metaphor is required in the form of a double pendulum movement that encompasses both the historical cultural and the assessment cultural dimension. The complexity of the metaphor is further increased if one also considers that these double pendulum movements can be analyzed in different arenas. The study shows that the pendulum movements in the formulation arena does not always correlate with the pendulum movements in the realization arena.

During the period 1994–2011, i.e. the period before Gy11, the historical cultural pendulum of the formulation arena, manifested in the 1994/2000 course syllabus, has a clear HT-traditional anchoring. The HLP teachers' perception that there is a discrepancy between their own view of the subject of history and the view that is expressed in the construct of HLP, indicates that this anchoring is not as clear in the realization

arena. The negotiation in connection with the implementation of HLP indicates a historical-cultural pendulum movement out of sync a kind of lag in the realization arena. The assessment cultural pendulum movement also ends up out of sync during the corresponding period. In the formulation arena, the new curriculum that was drawn up 1994 was anchored in the interpretation-tradition and its emphasis on freedom of interpretation. The HLP teachers' need for an external test anchored in the measurement-tradition to deal with the problem of equivalence indicates a kind of lag in the formulation arena.

During the period after Gy11, the historical cultural pendulum of the formulation arena, manifested in the 2011 course syllabys, remains HT-traditionally anchored. In the same way as the HLP teachers, the KP teachers' perception that there is a discrepancy between their own view of the subject of history and the view that is expressed in the test construct indicates that this anchoring is not as clear in the realization arena. However, unlike the HLP teachers, the KP teachers are not involved in a negotiation in connection to the implementation of KP. On the contrary, they use KP as a message-conveying tool for interpreting the new HT traditional syllabus. This difference between the two teacher collectives can be related to the difference concerning the assessment cultural pendulum. With regards to this pendulum, it is possible to see a shift after Gy11 in a measurement traditional direction, manifested in both the Curriculum and in the increased use of external tests. This assessment cultural shift limits the KP teachers' room for negotiation.

To sum up, the study indicates that the implementation of Gy11 has resulted in a change in the historical culture, as it is expressed in history teachers' view of the school subject of history. The study also indicates that there is a connection between the school's historical cultural change and the school's assessment cultural change, manifested in the history teachers' different ways of perceiving and using external tests.

REFERENSER

Muntliga källor

Intervjuer (materialet finns i författarens ägo.)

Intervju med Hanna, 2009-11-19

Intervju med Helena, 2009-11-25

Intervju med Henrik, 2009-11-26

Intervju med Håkan, 2010-01-19

Intervju med Karin, 2016-11-08

Intervju med Kristina, 2016-11-21

Intervju med Krister, 2016-11-24

Intervju med Karl, 2016-12-06

Litteratur

Ahonen, Sirkka, (2004). "Vad mäter uppgifterna i realprovet?", i Till, Jan-Erik (red.), *Historia och samhällslära som en del av realprovet åren 2001–2004*, Helsingfors.

Allal, Linda (2013). Teachers' Professional Judgement in Assessment: A Cognitive Act and a Socially Situated Practice. *Assessment in Education: Principles, Policy & Practice*, 20.

Alvén, Fredrik (2017). *Tänka rätt och tycka lämpligt: historieämnet i skärningspunkten mellan att fostra kulturbärare och förbereda kulturbbyggare*. Diss. Malmö: Malmö högskola.

Alvesson, Mats & Sköldberg, Kaj (2008). *Tolkning och reflektion: vetenskapsfilosofi och kvalitativ metod*. (2., [uppdaterade] uppl.) Lund: Studentlitteratur.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational, & Psychological Testing (US). (2014). *Standards for educational and psychological testing*. American Educational Research Assn.

- Ammert, Niklas (2008). *Det osamtidigas samtidighet. Historiemedvetande i svenska historieläroböcker under hundra år*. Diss. Lund: Lunds universitet.
- Ammert, Niklas och Eliasson, Per (2019). *Externa historieprov : en studie av validitetsfrågor*. Malmö universitet, Fakulteten för lärande och samhälle (Skrifter med historiska perspektiv: 19).
- Andersson Hult, Lars (2021). "Varför skall eleverna ägna sig åt själlöst minnesplugg? Om tilliten till historielärarna under moderniteten och de nationella proven införande i det senmoderna samhället" i Ludvigsson, David & Andersson Hult, Lars (red.) Att bedöma i historieämnet: perspektiv på historisk kunskap, bedömning och prov. [Jönköping]: Historielärarnas förening.
- Arensmeier, Cecilia et al. (2016). "De nationella proven och deras effekter i årskurs 6 och 9: En intervjustudie med elever, lärare och skolledare". Bilaga i Nationella proven i grundskolans årskurs 6 och 9: En uppföljning av lärares och elevers uppfattningar om proven. Skolverket. Rapport 447. Stockholm:2016.
- Aronsson, Peter (2004). *Historiebruk: att använda det förflutna*. Lund: Studentlitteratur.
- Barton, Keith & Levstik, Linda (2004). *Teaching History for the Common Good*, Mahwah, N.J.
- Berg, Mikael & Persson, Anders (2020). Graderande granskning och förklarade glapp [Elektronisk resurs] Svensk historie- och samhällskunskapsdidaktisk forskning om lärares bedömningspraktik 2009–2019. Nordidactica. 1, 18-44.
- Brookhart, Susan M. (2015). "Assessment of Historical Thinking in Practice" i *New directions in assessing historical thinking*. Kadriye E. & Seixas, P. (red.). New York, NY: Routledge.
- Carlgren, Ingrid (2011). "Forskning ja, men i vilket syfte och om vad? Om avsaknaden och behovet av en "klinisk" mellanrumsforskning", i Eklund, S. (red.), *Lärare som praktiker och forskare om praxisnära forskningsmodeller*. Stockholm: Stiftelsen SAF i samarbete med Lärarförbundet.
- Carlgren, Ingrid & Forsberg, Eva & Lindberg, Viveca (2009). "Diskussion" i Carlgren, Ingrid & Forsberg, Eva & Lindberg, Viveca (red.), *Perspektiv på den svenska skolans kunskapsdiskussion*, Stockholm.

- Cimbricz, Sandra (2002). "State-Mandated Testing and Teachers' Beliefs and Practice", *Education Policy Analysis Archives*, 2002:10(2).
- Clark, Anna & Grever, Maria (2018). "Historical Consciousness: Conceptualizations and Educational Applications" i Metzger, Scott Alan & Harris, Lauren McArthur (red.) *The Wiley international handbook of history teaching and learning*. New York: Wiley-Blackwell.
- Crooks, Terry. J., Kane, Michael. T., & Cohen, Allan. S. (1996). *Threats to the valid use of assessments*. Assessment in Education: Principles, Policy & Practice, 3(3).
- Cuban, Larry. (2009). *Hugging the middle: how teachers teach in an era of testing and accountability*. New York, NY: Teachers College, Columbia University.
- Danielsson Malmros, Ingmarie (2012). *Det var en gång ett land:- berättelser om svenskhet i historieläroböcker och elevers föreställningsvärldar*. Diss. Lund : Lunds universitet.
- Denscombe, Martyn (2000). *Forskningshandboken – för småskaliga forskningsprojekt inom samhällsvetenskaperna*, Lund: Studentlitteratur.
- Duquette, Catherine (2015). "Relating Historical Consciousness to Historical Thinking Through Assessment" i Ercikan, Kadriye & Seixas, Peter (red.), *New directions in assessing historical thinking*. Florence: Routledge Ltd.
- Eliasson, Per (2009). "Kan ett historiemedvetande fördjupas?" i Karlsson, Klas-Göran & Zander, Ulf (red.) (2009). *Historien är nu: en introduktion till historiedidaktiken*. Lund: Studentlitteratur.
- Eliasson (2014). "Historieämnet i de nya läroplanerna" i Karlsson, Klas-Göran & Zander, Ulf (red.) (2014). *Historien är närvarande: historiedidaktik som teori och tillämpning*. 1. uppl. Lund: Studentlitteratur.
- Emerson, Robert M. (2004). Working with "key incidents" i C. Seale, G.Giampietro, J. F. Gubrium & D. Silverman (red.) *Qualitative research practice* (s. 457–472. Thousand Oak, CA: Sage.
- Englund, Tomas (1986). *Läroplanens och skolkunskapens politiska dimension*. Göteborg: Daidalos.
- Ercikan, Kadriye, & Seixas, Peter (red.) (2015). *New directions in assessing historical thinking*. Florence: Routledge Ltd.
- Evans, Ronald W. (1989) "Teacher Conceptions of History", *Theory & Research in Social Education* 1989:3.

- Evans, Ronald W. (1990). "Teacher Conceptions of History Revisited: Ideology, Curriculum, and Student Belief", *Theory & Research in Social Education*, 1990:2.
- Fickel, Letitia (2006). "Paradox of Practice: Expanding and Contracting Curriculum in a High-Stakes Climate", i S. G. Grant, S. G (red.), *Measuring history: Cases of high-stakes testing across the US*, Greenwich, CT.
- Forsberg, Eva (2010) *Svensk forskning om bedömning: en kartläggning*, Stockholm: Vetenskapsrådet.
- Forsberg, Eva & Lundahl, Christian (2006). "Kunskapsbedömningar som styrmedia", *Utbildning & Demokrati*, 2006:15.
- Forsberg, Eva & Lundahl, Christian (2009). "Skolans interna och externa kunskapsbedömningar", i Lindblad, Rita Foss & Lander Rolf (red.), *Att säkra det osäkra*, Lund: Studentlitteratur.
- Friberg, Henrik (2017). Historielärares klassrumsbedömningar: innehåll och roll i ett historiedidaktiskt perspektiv. Licentiatavhandling Karlstad : Karlstads universitet.
- Gambell, Trevor & Hunter, Darryl (2004). "Teacher scoring of large-scale assessment: professional development or debilitation?", *Journal of Curriculum Studies*, Vol. 36 No. 6.
- Gibbs, Graham (2007). *Analysing qualitative data*. Los Angeles, [Calif.]: SAGE.
- Gipps, Caroline V. (1994). *Beyond testing: towards a theory of educational assessment*. London: Falmer.
- Glaser, Joakim (2015). *Fotboll från Mielke till Merkel: kontinuitet, brott och förändring i supporterkultur i östra Tyskland*. Diss. Malmö: Malmö högskola.
- Gradwell, Jill (2006). "Teaching in Spite of, Rather than Because of, the Test: A Case of Ambitious History Teaching in New York State i S. G. Grant, S. G (red.), *Measuring history: Cases of high-stakes testing across the US*, Greenwich, CT.
- Grant, S. G. (2003). *History lessons: teaching, learning and testing in U. S. high school classrooms*, New York: Routledge.
- Grant, S. G. (2006). "Research on History Tests" i S. G. Grant, S. G (red.), *Measuring history: Cases of high-stakes testing across the US*, Greenwich, CT.

- Grant, S.G. (2018). "Teaching Practices in History Education" i Metzger, Scott Alan & Harris, Lauren McArthur (red.) *The Wiley international handbook of history teaching and learning*. New York: Wiley-Blackwell.
- Grever, Maria & Adriaansen Robbert-Jan (2017). "Historical Culture: A Concept Revisited" i Carretero Maria, Berger Stefan & Grever, Maria (red). *Palgrave Handbook of Research in Historical Culture and Education*. Palgrave Macmillan UK.
- Grossman, Pamela L. och Stodolsky, Susan (1995) "The Role of School Subjects in Secondary School Teaching", *Educational Researcher*, Vol. 24, No. 8.
- Gullberg, Tom (2006) "Historia är intressant – men hur skall historia undervisas?", i Oker-Blom, Gun & Sandvik, Mia (red.), *Historia i nuet*, Helsingfors: Utbildningsstyrelsen.
- Gunnemyr, Per (2010). "I huvudet på en finländsk provkonstruktör", i Per Eliasson, Klas-Göran Karlsson, Henrik Rosengren och Charlotte Tornbjær (red.), *Historia på väg mot framtiden*, Lunds universitet.
- Gunnemyr, Per (2011). *Likvärdighet till priset av likformighet. En studie av hur och varför svenska och finländska historielärare på gymnasiet uppfattar att de påverkas av externa prov i historia*, Lund: Lunds universitet och Malmö högskola.
- Gustafsson, Bengt, Hermerén, Göran & Petersson, Bo (2005). *Vad är god forskningssed?: synpunkter, riktlinjer och exempel*. Stockholm: Vetenskapsrådet.
- Jarhall, Jessica (2020). *Historia från kursplan till klassrum: perspektiv på lärares historieundervisning från Lpo 94 till Lgr 11*. Diss. Malmö : Malmö universitet.
- Jönsson, Anders (2013). *Lärande bedömning*, Malmö: Gleerups.
- Jönsson, Anders (2017[2018]). *Prov eller bedömning?: Att tolka och använda provresultat och omdömen*. Första upplagan Malmö: Gleerups
- Jönsson, Anders, Rosenlund, David & Alvé, Fredrik (2017). "Complement or contamination: A study of the validity of multiple-choice items when assessing reasoning skills in physics". *Front. Educ.* 2:48.
- Tillgänglig på Internet: <http://urn.kb.se/resolve?urn=urn:nbn:se:hkr:diva-17254>
- Karlegård, Christer & Karlsson, Klas-Göran (1997) *Historiedidaktik*, Lund: Studentlitteratur.

- Karlsson, Klas-Göran (2009). "Historiedidaktik: begrepp, teori och analys", i Karlsson, Klas-Göran & Zander, Ulf (red.) (2009). *Historien är nu: en introduktion till historiedidaktiken*. 2., [uppdaterade och bearbetade] uppl. Lund: Studentlitteratur.
- Karlsson, Klas-Göran (2010). *Europeiska möten med historien: historiekulturella perspektiv på andra världskriget, förintelsen och den kommunistiska terrorn/ Klas-Göran Karlsson*. Stockholm: Atlantis.
- Karlsson, Klas-Göran (2014). Historia, historiedidaktik, och historiekultur - teori och perspektiv" i Klas-Göran Karlsson & Zander, Ulf (red.). *Historien är närvarande: historiedidaktik som teori och tillämpning*. Lund: Studentlitteratur.
- Karlsson, Per-Arne (2014). Undervisning och lärande i historia: ett kreativt rum för narrativ kompetens. Stockholm: Acta Universitatis Stockholmiensis.
- Kayser Nielsen, Nils (2010). Historiens förvandlingar: historiebrug fra monumenter till oplevelseøkonomi. Aarhus: Aarhus Universitetsforlag.
- Kristensson Uggla, Bengt (2002). Slaget om verkligheten: filosofi, omvärldsanalys, tolkning. Eslöv: B. Östlings bokförl. Symposion.
- Kvande, Lise & Naastad, Nils (2015). *Hva ska vi med historie? Historiedidaktik i teori og praksis*. Oslo: Universitetsforlaget.
- Kölbl, Carlos & Konrad, Lisa (2015). "Historical Consciousness in Germany: Concept, Implementation, Assessment" i Ercikan, K., & Seixas, P. (red), *New directions in assessing historical thinking*. Florence: Routledge Ltd.
- Larsson, Hans Albin (2001). *Barnet kastades ut med badvattnet. Historien om hur skolans historieundervisning närmast blev historia. Aktuellt om Historia 2001:2*, Bromma: Historielärarnas förening.
- Lee, Peter & Ashby, Rosalyn (2000). "Progression in historical understanding among students ages 7-14" i Stearns, Peter N., Seixas, Peter C & Wineburg, Samuel S (red.). *Knowing, teaching, and learning history: national and international perspectives*. New York: New York University Press.
- Lévesque, Stéphane (2008). *Thinking Historically. Educating Students for the Twenty-First Century*, Toronto: Toronto University Press.
- Lévesque, Stéphane & Clark, Penney (2018). "Historical Thinking: Definitions and Educational Applications" i Metzger, Scott Alan & Harris,

- Lauren McArthur (red.) *The Wiley international handbook of history teaching and learning*. New York: Wiley-Blackwell.
- Lifmark, David (2004). Skolforskarens etiska dilemman och relationer till sina informanter. *Etiska lärare–moraliska barn*, Ingår i: Forskningsarbete pågår: Nationella forskarskolan i pedagogiskt arbete (NaPA) / [ed] Per-Olof Erixon, Umeå: Umeå universitet.
- Lindberg (2009). ”Kunskapsuppfattningar och inter/nationella redskap för bedömning i Carlgren, Ingrid, Forsberg, Eva & Lindberg, Viveca (red.) *Perspektiv på den svenska skolans kunskapsdiskussion*. Stockholm: Stockholms universitets förlag.
- Lindensjö, Bo & Lundgren, Ulf P (2000). *Utbildningsreformer och politisk styrning*, Stockholm: HLS förl.
- Lipsky, Michael (2010). *Street-level bureaucracy: dilemmas of the individual in public services*. 30th anniversary expanded ed. New York: Russell Sage Foundation.
- Ludvigsson, David & Andersson Hult, Lars (red.) (2021). Att bedöma i historieämnet: perspektiv på historisk kunskap, bedömning och prov. [Jönköping]: Historielärarnas förening.
- Lundahl, Christian (2006). *Viljan att veta vad andra vet: kunskapsbedömning i tidigmodern, modern och senmodern skola*. Diss. Uppsala : Uppsala universitet.
- Lundahl, Christian (2009). *Varför Nationella prov? – framväxt, dilemma, möjligheter*, Lund: Studentlitteratur.
- Lundahl, Christian och Tveit, Sverre (2014). “Att legitimera nationella prov i Sverige och i Norge – en fråga om profession och tradition, *Pedagogisk forskning i Sverige*, årg 19 Nr 4.5.
- Löfström, Jan & Virta, Arja & Van den Berg, Marko (2010). “Who Actually Sets the Criteria for Social Studies Literacy? The National Core Curricula and the Matriculation Examination as Guidelines for Social Studies Teaching in Finland in the 2000’s”, *Journal of Social Science Education*, Vol. 9, No. 4.
- Madaus, George. F. (1988). “The influence of testing the curriculum”, i Tanner, Laurel (ed.), *Eighty-seventh Yearbook of the National Society for the Study of Education. Part I*, Chicago, Illinois.
- Maxwell, Alexander (2010). “Assessment Strategies for a History Exam, or, Why Short-Answer Questions are Better than In-Class Essays”, *The History Teacher*, Vol. 43 No. 2.

- Monte-Sano, C., & Reisman, A. (2015). "Studying historical understanding" in Corno L., & Anderman E. M. (Eds), *Handbook of educational psychology*. New York: Routledge.
- Morawski, Jan (2010). *Mellan frihet och kontroll: om läroplanskonstruktioner i svensk skola*. Diss. Örebro : Örebro universitet.
- Nordgren, Kenneth (2016). "How to Do Things With History: Use of History as a Link Between Historical Consciousness and Historical Culture", *Theory & Research in Social Education*, Volume 44.
- Nygren, Thomas (2011). *History in the service of mankind: international guidelines and history education in upper secondary schools in Sweden, 1927-2002*. Diss. (sammanfattning) Umeå: Umeå universitet.
- Nyström, Peter (2004). *Rätt mått på prov: om validering av bedömningar i skolan*, Umeå: Univ.
- Pellegrino, James. W., Chudowsky, Naomi & Glaser, Robert (red.) (2001). *Knowing what students know: the science and design of educational assessment*, National Academy Press, Washington, D.C.
- Persson, Anders (2017). *Lärartillvaro och historieundervisning: innebörder av ett nytt uppdrag i de mätbara resultatens tid*, Institutionen för idé- och samhällsstudier, Umeå universitet, Diss. Umeå: Umeå universitet.
- Persson, Helén (2018). *Historia i futurum: om progression i historia i läroböcker och styrdokument 1919-2012*, Historiska institutionen, Lunds universitet, Diss. Lund : Lunds universitet.
- Portelli, Alessandra (1991). *The death of Luigi Trastulli, and other stories: form and meaning in oral history*. Albany, N.Y.: State Univ. of New York Press.
- Prop. 2013/14:1 Utgiftsområde 16 (2013/14).
- Quinlan, Kathleen M. & Åkerlind, Gerlese S. (2000). "Factors affecting departmental peer collaboration for faculty development: Two cases in context", *Higher Education* Vol. 40.
- Quinlan, Mary Kay (2011). i Ritchie, Donald A. (red.) *The Oxford handbook of oral history*. New York: Oxford University Press.
- Radinsky et al. (2015). "Historical Thinking: In Search of Conceptual and Practical Guidance for the Design and Use of Assessments of Student Competence" i Ercikan, Kadriye, & Seixas, Peter (red), *New directions in assessing historical thinking*. Florence: Routledge Ltd.

- Reisman (2019). Reisman, Abby Brimsek, Emily & Hollywood, Claire (2019). "Assessment of Historical Analysis and Argumentation (AHAA): A New Measure of Document-Based Historical Thinking", *Cognition and Instruction*, 37:4.
- Rennstam, Jens & Wästerfors, David (2015). *Från stoff till studie: om analysarbete i kvalitativ forskning*. (1. uppl.) Lund: Studentlitteratur.
- Ritchie, Donald A. (2015) *Doing Oral History*, New York, NY: Twayne Publishers.
- Rosenlund, David (2010). "Är en förklaring alltid en förklaring", i Per Eliasson, Klas-Göran Karlsson, Henrik Rosengren och Charlotte Tornbjær (red.), *Historia på väg mot framtiden*, Lunds universitet.
- Rosenlund, David (2016). *History education as content, methods or orientation?: a study of curriculum prescriptions, teacher-made tasks and student strategies*. Diss. Malmö: Malmö högskola.
- Rosenlund, David (2019). "Internal Consistency in a Swedish History Curriculum: A Study of Vertical Knowledge Discourses in Aims, Content and Level Descriptors". *Journal of Curriculum Studies*. 2019:51(1).
- Rudnert, Joel (2019). *Bland stenyxor och Tv-spel: om barn, historisk tid och när unga blir delaktiga i historiekulturen*. Diss. Malmö: Malmö universitet.
- Rudnert, Joel & Bruér, Mikael (2021). "Flervalsuppgifter i historia" i Ludvigsson, David & Andersson Hult, Lars (red.) Att bedöma i historieämnet: perspektiv på historisk kunskap, bedömning och prov. [Jönköping]: Historielärarnas förening.
- Rüsen, Jörn (2004). *Berättande och förnuft*, Göteborg: Daidalos.
- Samuelsson, Johan (2010). "Censorer och inspektörer: från kanonkoll till reformkontroll i ämnet historia 1960–1982", opublicerat föredrag presenterat vid *Fjärde årliga konferensen inom nationella nätverket för historiedidaktisk forskning*, Malmö Högskola 2010-03-05.
- Samuelsson, Johan (2015). "Grundskollärares ämnesdidaktiska bedömningspraxis", *Pedagogisk forskning i Sverige*, Vol. 20, nr 1-2.
- Samuelsson, Johan (2021). "Läroverkslärares bedömningspraktik, ett aktualiserbart yrkesarv för framtiden" i Ludvigsson, David & Andersson Hult, Lars (red.) Att bedöma i historieämnet: perspektiv på historisk kunskap, bedömning och prov. [Jönköping]: Historielärarnas förening.

- Schüllerqvist, Bengt (2005). *Svensk historiedidaktisk forskning*. Stockholm: Vetenskapsrådet.
- Schwab, Joseph J. (1978). *Science, curriculum, and liberal education: selected essays*. Chicago: Univ. of Chicago P.
- Scully, Darina (2017). "Constructing multiple-choice items to measure higher order thinking, Practical Assessment", *Research and Evaluation*, Vol. 22 (4).
- Seixas, Peter (2008). "Foreword" i Lévesque, Stéphane (2008). *Thinking historically: educating students for the twenty-first century*. Toronto: University of Toronto Press.
- Seixas, Peter et al. (2015). "A Design Process for Assessing Historical thinking: The Case of a One-hour Test" i Ercikan, Kadriye, & Seixas, Peter (red), *New directions in assessing historical thinking*. Florence: Routledge Ltd.
- Shemilt, Dennis (2009). "Drinking an Ocean and Pissing a Cupful: How Adolescents Make Sense of History", i Symcox, Linda & Wilschut, Arie (red.) (2009). *National history standards: the problem of the canon and the future of teaching history*. Charlotte, NC: Information Age Pub.
- Shemilt, Denis (2018). "Assessment of Learning in History Education: Past, Present and Possible Futures" i Metzger, Scott Alan & Harris, Lauren McArthur (red.) *The Wiley international handbook of history teaching and learning*. New York: Wiley-Blackwell.
- Sjöland, Marianne (2016). *Historia från tidskriftsredaktionen: en komparativ studie av Populär historias och History today's historieskrivning*. Diss. Lund : Lunds universitet.
- Smith, Mark, Breakstone, Joel & Wineburg, Sam (2019). "History Assessments of Thinking: A Validity Study", *Cognition and Instruction*, 37:1, 118-144.
- Skolinspektionen (2010). Dnr: U2009/4877/G, *Kontrollrättning av nationella prov i grundskolan och gymnasieskolan – Redovisning av regeringsuppdrag*, Stockholm 2010.
- Skolinspektionen (2015). *Undervisningen i historia*, Stockholm: Skolinspektionen.
- Skolverket (2003). *Det nationella provsystemet – vad varför och varthän*, Stockholm: Skolverket.
- Skolverket (2005). *Nationella prov i gymnasieskolan – ett stöd för likvärdig betygssättning*, Stockholm: Skolverket.

- Skolverket (2008). Dnr 64-2008:258, *Central rättning av nationella prov*. Stockholm: Skolverket.
- Skolverket (2009). *Likvärdig betygssättning i gymnasieskolan. En analys av sambandet mellan nationella prov och kursbetyg*, Stockholm: Skolverket.
- Skolverket (2009). Dnr: 2009:337, *Redovisning av regeringsuppdrag att ge förslag på hur det nationella provsystemet bör utvecklas och utformas*. Stockholm: Skolverket.
- SOU 2007:28, *Tydliga mål och kunskapskrav i grundskolan. Förslag till nytt mål- och uppföljningssystem*.
- Stenlås, Niklas (2009). *En kår i kläm: läraryrket mellan professionella ideal och statliga reformideologier*. Stockholm: Finansdepartementet, Regeringskansliet.
- Strannegård, Lars (2019). Utan antikens historia i läroplanen riskeras Sveriges konkurrenskraft. *Dagens Nyheter*. 1 oktober.
<https://www.dn.se/debatt/utan-antikens-historia-i-laroplanen-riskeras-sveriges-konkurrenskraft/> (Hämtad 2021-11-06).
- Taras, Maddalena (2007). "Summative assessment: the missing link for formative assessment", *Journal of Further and Higher Education*, 33:1.
- Thomsson, Heléne (2002). *Reflexiva intervjuer*. Lund: Studentlitteratur
- Thompson, Paul (1980). *Det förgångnas röst: den muntliga historieforskningens grunder*. Stockholm: Gidlund.
- Thornberg, Pia (2014). Kan sambedömning leda till en mer likvärdig bedömning och betygssättning? (Dissertation). Hämtad från <http://urn.kb.se/resolve?urn=urn:nbn:se:hkr:diva-11848>.
- Thorp, Robert (2013). "Vad är ett historiemedvetande egentligen och varför är det viktigt" i David Ludvigsson (red.), *Kritiska perspektiv på historiedidaktiken*, Aktuellt om historia, 2013:2.
- Tracy, Sarah J. (2010). Qualitative Quality: Eight "Big-Tent" Criteria for Excellent Qualitative Research. *Qualitative Inquiry*, 2010:16.
- Trost, Jan (2005). *Kvalitativa intervjuer*. 3. uppl. Lund: Studentlitteratur.
- Van Boxtel, Carla & Van Drie, Jannet (2018). "Conceptualizations and Educational Applications" i Metzger, Scott Alan & Harris, Lauren McArthur (red.) *The Wiley international handbook of history teaching and learning*. New York: Wiley-Blackwell.

- VanSledright, Bruce & Limon, M. (2006). "Learning and teaching social studies: A review of cognitive research in history and geography" i Alexander, P. A., & Winne P. H., (red.), *Handbook of educational psychology (2nd ed.)* Mahwah, N.J.: Erlbaum.
- VanSledright, Bruce (2014). *Assessing historical thinking and understanding: innovative designs for new standards*. New York, NY: Routledge.
- VanSledright, Bruce (2015). "Assessing for Learning i History Classroom" i Ercikan, Kadriye & Seixas, Peter C. (red.) (2015). *New directions in assessing historical thinking*. New York, NY: Routledge.
- Vedung, Evert (2002). "Utvärdering och de sex användningarna", i Rombach, Björn & Sahlin, Kerstin (red.), *Från sanningssökande till styrmedel: moderna utvärderingar i offentlig sektor*, 2. uppl., Santérus, Stockholm.
- Virta, Arja (1997). "Evaluering, kunskap och historieuppfattning", i Karlegård, Christer & Karlsson, Klas-Göran (red.), *Historiedidaktik*, Lund: Studentlitteratur.
- Vogler, Kenneth E. & Virtue David (2007). "Just the Facts, Ma'am": Teaching Social Studies in the Era of Standards and High-Stakes Testing", *The Social Studies*, March/April.
- Waldis, Monika et al (2015) i "Material-Based and Open-Ended Writing tasks for Assessing Narrative Competence Among Students" i Ercikan, Kadriye & Seixas, Peter C. (red.) (2015). *New directions in assessing historical thinking*. New York, NY: Routledge.
- Widén, Pär (2010). *Bedömningsmakten: berättelser om stat, lärare och elev, 1960-1995*. Diss. Linköping : Linköpings universitet.
- Wikström, Christina (2013). *Konsten att göra bra prov: vad lärare behöver veta om kunskapsmätning* . 1. utg. Stockholm: Natur & Kultur.
- Wineburg, Sam (2004). "Crazy for History" i *Journal of American History*, 90(4).
- Zancanella, Don (1992). "The Influence of State-Mandated Testing on Teachers of Literature", *Educational Evaluation and Policy Analysis*, Vol. 14 No. 3.
- Zander, Ulf (2001). *Fornstora dagar, moderna tider: bruk av och debatter om svensk historia från sekelskifte till sekelskifte*. Diss. Lund : Univ.

- Åström Elmersjö, Henrik (2017). *En av staten godkänd historia: förhandsgranskning av svenska läromedel och omförhandlingen av historieämnet 1938-1991*. Lund: Nordic Academic Press.
- Östling, Johan (2008). *Nazismens sensmoral: svenska erfarenheter i andra världskrigets efterdyning*. Diss. Lund : Lunds universitet.

BILAGOR

Bilaga 1

Intervjuguide för intervjuserie 1 – HLP-lärarna

Introduktion

Jag har ju redan i vår mailkontakt beskrivit syftet med den här intervjun och vad den ska handla om. (Mitt forskningsprojekt syftar till att undersöka hur gymnasielärare i historia uppfattar externa prov.)

Jag hoppas att den här intervjun kan bli mer av ett samtal där du öppet berättar för mig om hur du uppfattar Historielärarnas prov. Det är viktigt för mig att du också framför eventuella kritiska synpunkter!

För att vara säkra på att jag inte missar viktig information och för att kunna kontrollera informationen under analysens gång, vill jag gärna spela in intervjun på band. Är det OK?

Jag har ett antal frågeområden som jag vill ta upp under samtalets gång och jag beräknar att intervjun tar ca en timme.

Har du några frågor innan vi sätter igång?

Fråga 1

”Allra först vill jag bara kolla vissa bakgrundsuppgifter om dig”

- Hur lång erfarenhet har du av att undervisa i historia?
- Vilka undervisningsgrupper arbetar du huvudsakligen med?
- Hur många undervisningsgrupper har du ”förberett” för provet?

Fråga 2

”Vill du berätta om vilka syften du uppfattar låg bakom provets tillkomst 2005/2006?”

- a. Varför tillkom provet? På vems/vilkas initiativ? Vilken/vilka funktioner skulle provet fylla?
- b. Har det skett någon förändring vad gäller provets syften sedan 2005?
- c. Uppfattar du att provet används i linje med de tänkta intentionerna?
- d. Var har du valt att använda dig av Skåneprovet?

Fråga 3

”Har du varit delaktig i utformandet/konstruktionen av provet?”

Om delaktig

”Vill du berätta hur du uppfattar arbetsprocessen i samband med provkonstruktionen?”

- a. Vem/vilka avgör vilka personer som ska delta i arbetet med provet? Har det skett någon förändring sedan 2005/2006?
- b. Varför har du valt att vara med i provkonstruktörsgruppen?
- c. Vilka konkreta förutsättningar har du haft för arbetet med provet? Har det skett någon förändring sedan 2005/2006? (Tid/Plats)
- d. Vem/vilka personer deltar i arbetet med provet? Har det skett någon förändring sedan 2005/2006? (Lärare i ämnet? På andra skolor? Elever?)
- e. Kan du beskriva arbetsfördelningen mellan de personer som deltar i arbetet med provet? Är det någon som har huvudansvaret?

- f. Uppfattar du att du kan påverka provets utformning? Har det skett någon förändring sedan 2005/2006? (Relationer)

Om ej delaktig

”Vill du berätta hur du uppfattar arbetsprocessen i samband med provkonstruktionen?”

- a. Vem/vilka avgör vilka personer som ska delta i arbetet med provet? Har det skett någon förändring sedan 2005/2006?
- b. Varför har du valt att inte vara med i provkonstruktörsgruppen?
- c. Uppfattar du att du kan påverka provets utformning? Har det skett någon förändring sedan 2005/2006? Har det skett någon förändring om du jämför med det gamla realprovet?

Fråga 4

”Vill du berätta hur du uppfattar vilka kunskaper provet ska mäta?

- a. Vilket innehåll? Vilka kompetenser/förmågor
- b. Vilka olika slags frågetyper innehåller provet?

Fråga 5

”Vill du berätta hur du uppfattar provets betydelse för din egen undervisning?”

Planering

- a. Vilken betydelse har provet för planeringen av din undervisning? (Vilket innehåll? Vilka förmågor/kompetenser? Undervisningsmetoder? Har det skett någon förändring sedan 2006?

- b. Vilken betydelse har det material, utöver provfrågorna, som medföljer provet? (Lärarinstruktion/Elevinstruktion/Bedömningsmall)
- c. Vilken betydelse har styrdokumentet för planeringen av din undervisning?
- d. Vilken betydelse har läroböckerna för planeringen av din undervisning?
- e. Vilken betydelse har elevernas inställning till provet för planeringen av din undervisning?

Genomförande

- a. Använder du dig av (tidigare utgåvor av) provet i din undervisning?

Bedömning

- a. Kan du beskriva hur det går till när du bedömer provsvaren? (Svårigheter? Sambedömning?)
- b. Vilken betydelse har provet haft för din syn på bedömning? (Kan man mäta kunskaper i historia?) (Summativ/Formativ)
- c. Vilken betydelse har provet för utformningen av dina egna interna prov?
- d. Vilken betydelse har provet för din betygsättning?

Fråga 6

”Vill du berätta hur du uppfattar att provet påverkar samarbetet och ämnesutvecklingen på din skola?”

- a. Förs det några uttolkande diskussioner om provet i andra forum, utöver den provkonstruerande gruppen? Vad tas upp i dessa forum? Har det skett någon förändring sedan 2005/2006? (nationellt, lokalt, skola, arbetslag, ämneslag)
- b. Hur uppfattar du elevernas reaktioner på provet?
- c. Hur uppfattar du lärarnas reaktioner på provet?
- d. Hur uppfattar du skolledningens reaktioner på provet?

Fråga 7

”Vill du berätta hur du uppfattar att provet uppfattas i samhället i stort?”

- a. Finns det en medial beskrivning av provet? Har beskrivningen förändrats sedan 2006?

Fråga 8

”Om vi nu mot slutet av intervjun skulle summera och avrunda det här: vilka för- och nackdelar anser du att provet i dess nuvarande utformning har?”

- a. Fördelar och nackdelar?
- b. Tankar kring förändringen sedan 2005/2006?
- c. Tankar inför framtiden?

Avslutning

”Hur förberedde du dig för intervjun? Hur har det påverkat hur du svarade?”

”Finns det något du vill tillägga innan vi avslutar intervjun?”

Stort tack för hjälpen!

Bilaga 2

Intervjuguide för intervjuserie 2

Information inför intervjun

Vid intervjutillfället kommer vi att samtala om gymnasiets kursprov för historia 1b. Jag har bifogat 2016 års version av provet och till intervjun ber jag dig att välja ut en eller ett par uppgifter samt en eller ett par bedömningsanvisningar som du som du skulle vilja ta upp till diskussion. Intervjun beräknas ta 60-90 minuter i anspråk. Med ditt godkännande bandas samtalet för att möjliggöra transkribering i efterhand. *Stort tack för din medverkan!*

Bakgrundsfaktorer

- Utbildning
 - Grundutbildning (omfattning, när)
 - Fortbildning (allmänt, specifikt om bedömning och betygssättning)
- Lärarerfarenhet (ämneskombination, undervisningsgrupper, särskilda uppdrag, karriärtjänst)

Frågeområde 1

- Här samtalar vi om *varför* du använder kursprovet i historia
 - Syfte på elevnivå?
 - Syfte på lärarnivå?

Frågeområde 2

- Här samtalar vi om *hur du gör* när du använder kursprovet i historia
 - Genomförandet av provet (t.ex. vilka undervisningsgrupper, eventuella anpassningar)
 - Värderingen av elevsvaren (t.ex. sambedömning)
 - Betygssättningen i relation till provresultatet

Frågeområde 3

- Här samtalar vi om de uppgifter och de bedömningsanvisningar som du har valt ut från 2016 års version av kursprovet.
 - Varför har du valt ut just dessa uppgifter?
 - Varför har du valt ut just dessa bedömningsanvisningar?

Frågeområde 4

- Här samtalar vi om *vilken betydelse* du uppfattar att kursprovet har haft för
 - ditt sätt att konstruera bedömningsuppgifter
 - ditt sätt att värdera elevsvar
 - din betygssättning
 - det kollegiala samarbetet på din skola

Är det möjligt att bedöma kunskaper i historia på ett mätsäkert sätt utan att det påverkar vårt sätt att uppfatta vad som menas med historisk kunskap? Den här frågeställningen utgör utgångspunkten för avhandlingens övergripande syfte som är att undersöka historiekulturell förändring över tid med avseende på skolämnet historia, och då särskilt beakta vilken betydelse kunskapsbedömning i allmänhet och frivilliga externa prov i synnerhet kan ha för denna förändring. Undersökningens temporala nav utgörs av läroplansreformen Gyll.

I en intervjubaserad studie undersöks hur historielärare på gymnasiet uppfattade och hanterade två frivilliga externa prov: Historielärarnas prov och det frivilliga Kursprovet för historia 1b. De empiriska resultat som studien genererat har genomgående analyserats i ljuset av den fjärde kvadrantens dilemma. Dilemmat bygger på antagandet att det alltid finns ett samband mellan vad som bedöms och hur någonting bedöms och att det är särskilt problematiskt att förena historia som tolkningsämne med höga krav på mätsäkerhet.

Per Gunnemyr är anställd vid Malmö Universitet. Den fjärde kvadrantens dilemma är hans doktorsavhandling i historia.

ISBN 978-91-7877-234-6 (print)

ISBN 978-91-7877-235-3 (pdf)

ISSN 1651-4513

MALMÖ UNIVERSITY
205 06 MALMÖ, SWEDEN
WWW.MAU.SE