

# Att mäta historiskt tänkande

ARJA VIRTÄ\*

Åbo universitet

Per Gunnemyr, *Den fjärde kvadrantens dilemma: Kunskapsbedömning i en föränderlig historiekultur* (Malmö: Malmö universitet 2021). 251 s.

”Den fjärde kvadrantens dilemma”, huvudrubriken i Per Gunnemyrs avhandling, väcker nyfikenhet och tankarna riktas kanske inte direkt till historien eller historiedidaktik. Underrubriken avslöjar ändå att studien handlar om historieundervisning.

Den fjärde kvadranten betyder en del i fyrfältsmatrisen som formas av två axlar. Den historiekulturella dimensionen (vad) utgör i Gunnemyrs avhandling den vertikala axeln, medan den horisontella axeln motsvarar bedömningskultur (hur-dimensionen). Även om fyrfältet är en förenkling, är det ett praktiskt verktyg för att åskådliggöra mångfacetterade och abstrakta företeelser.

Den vertikala historiekulturella vad-dimensionen symboliserar olika sätt att uppfatta historisk kunskap. Gunnemyr skriver om två motsatta epistemologiska traditioner som närmast angår historieundervisning, nämligen *Memory History*-traditionen (MH), som i detta sammanhang innebär historiskt innehåll som bör minnas, och *Historical Thinking*-traditionen (HT), som betonar historiskt tänkande, alltså mer krävande saker än minneskunskaper. Enligt den sistnämnda traditionen skulle läraren och bedömaren sträva efter och förutsätta att eleverna lär sig tänka historiskt, granska källor kritiskt, behärska begreppen historiemedvetenhet och historiebruk och så vidare. Dessa två traditioner presenteras i historiedidaktisk litteratur ofta som motpoler med tillägg att det i praktisk undervisning givetvis finns många mellanformer.

Den horisontella axeln innebär olika traditioner i bedömning. Tolknings-tradition betyder att läraren eller bedömaren har tolkningsutrymme och elevsvaren är öppna, som i traditionella essäuppgifter. Motpolen är den psy-

\* Professor emerita i historie- och samhällskunskapsdidaktik; fakultetsopponent

kometriskt orienterade mättraditionen då bedömarens tolkningsutrymme begränsas och elevsvaren är entydiga att värdera och bedöma, som till exempel flersvarsfrågor.

Den fjärde kvadranten ligger alltså i den del som begränsas mellan dimensionerna HT (historiskt tänkande) och den exakta mätningen. Enligt *Historical Thinking*-traditionen skulle man i undervisningen förutsätta tankefärdigheter som är typiska för historieämnet, men ändå skulle läraren bedöma elevernas svar på ett sätt som kan kallas mätning, med så vattentäta kriterier som möjligt. Här formas ett dilemma – den fjärde kvadrantens dilemma – och i sin avhandling vill Per Gunnemyr lösa dilemmat, eller åtminstone undersöka hur historielärare uppfattar och hanterar det här ofta latenta spänningsfältet.

Avhandlingens syfte är ”att undersöka historiekulturell förändring över tid med avseende på skolämnet historia, och då särskilt att beakta vilken betydelse kunskapsbedömning i allmänhet och externa prov i synnerhet har för denna förändring” (s. 23). Fyra forskningsfrågor som operationaliserar huvudsyftet handlar alla om historielärares uppfattningar eller verksamhet. Hur uppfattar lärare de frivilliga externa provens konstruktion; hur beskriver de att de genomför proven; hur beskriver de att de använder proven; hur uppfattar de provens påverkan på sin bedömningspraktik?

Frågeställningarna är tydligt avgränsade. Man kan ställa frågan om avgränsningen till och med var för smalt fokuserad. Kunde man även ha frågat vilken ämnesuppfattning eller historiesyn lärare hade och hur den var kopplad till deras bedömning – eller kunde man ha frågat om deras bedömning mer allmänt? Gunnemyrs avgränsning är ändå motiverade och studien blir en välstrukturerad helhet.

Metodologiskt och till sin struktur är avhandlingen en typisk utbildningsvetenskaplig studie. Den börjar med teoretiska delar som presenterar tidigare forskning samt teorier och begrepp. Därefter följer metodavsnittet, fyra kapitel som innehåller redogörelse samt tolkning av resultat och till slut sammanfattning och slutsatser.

Gunnemyrs avhandling är historiedidaktisk och har sålunda en tvärvetenskaplig karaktär, och den kan placeras i utrymmet mellan historieämnet och utbildningsvetenskap, närmare bestämt bedömningsforskning. Författaren betonar också att arbetet har en praktisknära ansats. Den handlar om frågor som är aktuella för historieundervisningen och lärares verksamhet. Studien handlar om ett mer generellt epistemologiskt problem, om det egentligen är möjligt att *mäta* historiska kunskaper eller göra dem mätbara så att man inte fastnar vid irrelevanta detaljer som är enkla att räkna.

Gunnemyrs avhandling har ett teoriutvecklande mål genom att skapa en kombination av historiedidaktiska teorier och bedömningsteorier. Tvärve-

tenskaplighet uttrycks i de huvudbegrepp som enligt forskaren bygger upp avhandlingens teoretiska grund, alltså historiekultur och bedömningskultur. Gunnemyr anser att historiekultur är primär och bedömningskultur sekundär som huvudbegrepp. Ordningen kunde också ha varit annan, eftersom vikten verkar ha varit minst lika mycket i bedömning och prov som i historieämnet eller historiekultur. Till exempel i forskningsfrågorna nämns historiekultur inte alls utan alla de fyra frågorna syftar till bedömning – lärares sätt att förhålla sig till externa frivilliga prov och använda dem. Dessutom är det fråga om lärares uppfattningar och erfarenheter, och därför kunde även dessa två begrepp ha förtjänat något utrymme. Ett begrepp som även hade kunnat få mer uppmärksamhet i den här avhandlingen är *assessment literacy*. Det definieras i korthet som lärarens förmåga att konstruera och använda bedömningsverktyg samt bedöma vad eleverna kan och vet.

Historiekultur används här som ett brett och övergripande begrepp, men fokus är hela tiden i skolans historiekultur och i synnerhet i de externa frivilliga proven som historiekulturella produkter. Historiekultur är därför ett ganska stort paraplybegrepp över den här företeelsen, även om skolan är en historiekulturell arena och proven onekligen på ett konkret sätt återspeglar skolans historiekultur och lärarens historiesyn.

Den historiekulturella förändring som nämns i avhandlingens underrubrik och syfte, hänvisar till de ändringar som följde i historieämnet i gymnasiet efter skolpolitiska reformer, framför allt 2011. Redan 1994 års kursplan ändrade ämnet genom att betona historiskt tänkande och lyfta fram begreppet historiemedvetande som följd av internationell historiedidaktisk utveckling, även om historia annars blev ett tämligen litet ämne. En tydlig profilmförhöjning för historieämnet följde med 2011 års skolpolitiska reformer och förnyade kursplaner, inte minst därför att historia då fick status som ett kärnämne. Åtminstone enligt styrdokumentet skulle vikten under de senaste två årtiondena ha flyttats från den gamla dammiga ”*memory history*” med faktainnehållsbetoning till ett mer krävande sätt att undervisa med vikt vid historiska färdigheter – analysförmåga, källkritik, historiemedvetande och historiebruk. Dessa färdigheter är krävande att lära sig, undervisa och bedöma.

Bedömning är alltid en stor utmaning för elever, men även för lärare. Den handlar om ett likvärdigt och rättvist sätt att bedöma olika elever och deras kunskaper, men också motsvarighet (*alignment*) till de mål som ställs i kursplanerna. Ett sätt att lösa bedömningsproblem har varit att använda externa prov som konstrueras av någon annan än klassens egen lärare. I svenska gymnasier används inte obligatoriska externa prov i historieämnet, men man har genomfört kortvariga projekt i vilka man har experimenterat med frivilliga prov. I Skåne skapade en grupp historielärare ett frivilligt

externt prov för gymnasier i början av 2000-talet, och efter 2011 års reform konstruerades på uppdrag av Skolverket kursprov, men även det sistnämnda projektet var ganska kortvarigt. I avhandlingen används förkortningar HLP för Skåne-provet och KP för det senare kursprovet.

Per Gunnemyr själv arbetade som provkonstruktör för kursprovet. Hans empiriska material består av semistrukturerade individuella intervjuer med åtta gymnasielärare som hade tillräcklig erfarenhet av frivilliga externa prov. De lärare som Gunnemyr intervjuade 2009 redan för sin licentiatavhandling, hade tillämpat det tidigare historieläraryrket (HLP-lärare). Den senare intervjugruppen (2016) hade använt kursprovet (KP-lärare). Genom att intervjua gymnasielärare undersökte han inte bara lärarnas uppfattningar och erfarenheter, utan också eventuella förändringar i deras sätt att uppfatta eller använda proven.

Antalet informanter är tämligen litet av förstäligen skäl: antalet gymnasielärare som tillämpade dessa externa prov var inte särskilt stort. Det begränsade antalet har gett forskaren en möjlighet att genomföra detaljerad analys, som han kallar intervjubaserad mikrostudie.

Gunnemyr karakteriserar avhandlingens analytiska form som abduktiv pendelrörelse mellan teori och empiri. Han har skapat teoribaserad kodning för att analysera det empiriska materialet. Koderna bygger på olika kategoriseringar av historiekultur och bedömningskultur. Analysen har genomförts på ett mycket logiskt och systematiskt sätt i linje med forskningsfrågorna. Det här är en tydlig styrka i avhandlingen. Ändå kan man fråga sig om analysen likaväl kunde ha kallats deduktiv på grund av dess starka teorianknytning. Rapporteringen är också stringent, men delvis svårläst, därför att begreppsapparaten bakom kodningen och tolkningarna är mångbottnad. Rapporteringstekniskt har det säkert varit tydligt och exakt att använda förkortningar – sådana som HLP, KP, HT och MH – men för läsaren kan detta vara en utmaning.

Gunnemyr positionerar sin avhandling i relation till tidigare forskning i samma tematik. Forskning kring historiedidaktik generellt och kring historieundervisning i synnerhet har utvecklats starkt i Sverige under de senaste två årtiondena. Det finns ett antal intressanta studier som handlar om bedömning i historieämnet. Av dessa kan nämnas till exempel David Rosenlunds avhandling (2016) och Johan Samuelssons studie (2015) om *alignment*-problematik mellan gymnasielärares bedömning och kursplan 1994/2000.<sup>1</sup> Jessica Jarhall har i sin avhandling (2020) forskat i nationella provens påver-

1. David Rosenlund, *History Education as Content, Methods or Orientation? A Study of Curriculum Prescriptions, Teacher-made Tasks and Student Strategies* (Malmö 2016); Johan Samuelsson, "Grundskollärares ämnesdidaktiska bedömningspraxis: Exemplet historia", *Pedagogisk forskning i Sverige*, 20:1–2 (2015) s. 54–78.

kan i grundskollärares undervisning.<sup>2</sup> Anders Perssons avhandling (2017) fokuserade på grundskollärares uppfattningar om bedömning, uppdrag som fick större vikt efter 2011 års skolpolitiska reform.<sup>3</sup> Gunnemyrs avhandling bidrar till denna forskningstematik med fokus på gymnasielärare och externa prov.

Som nämndes i början används historiekultur och bedömningskultur som huvudbegrepp i avhandlingen. Teoriavsnitten visar att Gunnemyr känner väl till båda teoriområden, utifrån vilka han bygger upp den analytiska apparaten för att behandla sitt empiriska material. Det historiekulturella analysverktyget – som egentligen innebär ett sätt att uppfatta historieundervisning – bygger på Seixas och Ercikans dimensioner av historieämnet.<sup>4</sup> Dessa är ämnets domänlogik, emfaslogik och progressionslogik. Närmare bestämt handlar det om vilka kunskaper och färdigheter som krävs i ämnet, vilka av dessa som är de viktigaste, och hur man skiljer mellan olika kvalitetsnivåer i kunskaper. Bedömningsteoretiska kategorier formas av observation, sammanvägning och värdering. Av dessa formas koderna till analysering av vad lärare har berättat om sina uppfattningar om proven och deras sätt att genomföra och använda proven.

Resultaten är intressanta och väcker många tankar om bedömningen och kursplanernas utveckling. Enligt Gunnemyrs analys ändrades också själva proven under forskningsperioden. I HLP-provet kunde man se den bedömningstradition som betonade tolkning, men i KP-provet fanns i stället en tydlig tendens mot mer strukturerad mätning. Det finns likheter men även några tydliga skillnader mellan de två grupperna som intervjuades i linje med de prov de har använt. HLP-lärare tenderade att använda de frivilliga proven som ett bedömningsinstrument medan KP-lärare såg dem mer som ett historiekulturellt medel.

Enligt resultaten menade lärare att det fanns några diskrepanser mellan deras egna uppfattningar om ämnet (undervisningens domänlogik) och de externa provens domänlogik. Det fanns ändå skillnader mellan de två grupperna. HLP-lärare var mer kritiska mot provens domänlogik, medan KP-lärare sade att riktningen var rätt men svår. Det verkar som att lärare hade vant sig tämligen väl vid till exempel källkritiska dokumentbaserade uppgifter, som alltså är i linje med HT-traditionen, men det som kändes främmande var att hantera uppgifter med betoning på begreppet historiemedvetande.

2. Jessica Jarhall, *Historia från kursplan till klassrum: Perspektiv på lärares historieundervisning från Lpo 94 till Lgr 11* (Malmö 2020).

3. Anders Persson, *Lärartillvaro och historieundervisning: Innebörder av ett nytt uppdrag i de mätbara resultatens tid* (Umeå 2017).

4. Peter Seixas & Kadriye Ercikan, "A Design Process for Assessing Historical Thinking: The Case of a One-hour Test", i Kadriye Ercikan & Peter Seixas (red.), *New Directions in Assessing Historical Thinking* (New York 2015).

De tre dimensionerna då – nu – framtiden var svåra att koppla ihop i bedömningen, särskilt enligt HLP-gruppen. Kursplanerna 2011 innehöll begreppet historiebruk och KP-lärare upplevde det som svårt att hantera i bedömningen. Så hände det ofta att lärare valde bort uppgifter som handlade om historiemedvetande och historiebruk. Dessa två nyckelbegrepp har kanske inte blivit ordentligt etablerade i undervisningen även om de erkändes som viktiga i historiens domänlogik. Särskilt i den nyare intervjugruppen blev det tydligt att de nya proven hade ökat kollegialt samarbete kring bedömning, men även kollegialt motstånd och polarisering i uppfattningar förekom.

Kvalitativa studier som ofta fokuserar på ett begränsat antal informanter bidrar sällan med direkt generaliserbara resultat. Det har inte heller varit avsikten med Gunnemyrs avhandling, men den lyfter i alla fall fram drag och förhållningssätt som är relevanta för bedömning och historieundervisning generellt.

Det är oavgjort hur man kan lösa den fjärde kvadrantens dilemma som i Gunnemyrs ord är "ett latent och mer djupgående dilemma, med omfattande historiekulturella och historiedidaktiska konsekvenser" (s. 222). Förhållningssättet till dilemmat har varierat i tid som i en ständig pendelrörelse. I slutkapitlet skriver Gunnemyr om kristallkulan med tanke på att bedömningen i fortsättningen kan betona mättraditionen som följd av bland annat digitaliseringen av proven. En fråga som uppstår är vad som då händer med skolans historiekultur. I detta sammanhang kan man betona lärarnas professionalitet och deras medvetenhet om dilemmat. Det här omfattar lärares kunskapsyn och ämnesuppfattningar, som har mycket att göra med deras sätt att undervisa och bedöma. Bedömning är en del av lärarens yrkesskicklighet och professionalitet.

Sammanfattningsvis kan jag konstatera att Per Gunnemyrs avhandling är ett välkommet bidrag till historiedidaktisk forskning, särskilt bedömningens område. Forskaren har genomfört sitt forskningsuppdrag logiskt och noggrant enligt vetenskapliga normer och följt forskningsetiska och forskarens principer. Den empiriska analysen är starkt knuten till teorin, och valen motiveras på ett övertygande sätt. I disputationen som ordnades digitalt via Zoom, försvarade han sin avhandling på ett övertygande sätt och bevisade sin förmåga för vetenskaplig diskussion.