

# Samtal om skrivbedömning

## Lärares normer, beslut och samstämmighet

Per Blomqvist

Akademisk avhandling för avläggande av filosofie doktorsexamen i språkdidaktik vid Stockholms universitet som offentligen kommer att försvaras fredagen den 8 juni 2018 kl. 10.00 i William Olsson-salen, Geovetenskapens hus, Svante Arrhenius väg 14.

### Abstract

*Teachers' conversations about writing assessment. Norms, pedagogical decisions and grading*

This thesis reports and discusses results from a qualitative study of Swedish teachers' writing assessment in upper secondary school. Based on teacher group discussions, the study investigates teachers' interactions, expressions of norms and decisions when assessing and grading students' writing in the subject of Swedish.

The aim of the study is to describe and analyze *how* teachers interact and *what* they pay attention to when discussing writing assessment. In relation to this aim the following three research questions were posed: (1) What characterizes the teachers' conversations about writing assessment? (2) What assessment norms do the teachers express? (3) What decisions do they make?

The theoretical perspectives are dialogical and pedagogical. From a dialogical point of view, the assessment conversations are considered as social acts where teachers in interaction collectively create meaning. Furthermore, the assessment conversations are considered as pedagogical, as a part of teachers' reflective work to discuss how to assess students' work and how to make decisions about further instructions and grading. The methods used to describe and analyze interactional patterns, expressions of norms and decisions in these assessment conversations are topic analysis and initiative-response-analysis.

The empirical data consists of audio and video recordings from three teacher groups' conversations about writing assessment, comprising a total of 17 Swedish teachers from three upper secondary schools. Data was also collected via a questionnaire with the individual teacher's grading of the students' writing and evaluation of the assessment conversations.

The results show that teachers' assessments of students' writing focus on much more, and partly other, criteria than their pedagogical decisions. The quality standards that teachers express in the discussions about student texts focus mostly on communicative quality, language style and text structure but also on content and the use of sources. The pedagogical decisions, on the other hand, almost exclusively focus on text structure and the use of sources.

The results also show that shortcomings in the students' texts are crucial for teachers' summative assessments. Meanwhile, the teachers also express that students' age, their writing development in the course and the national test must be considered. These assessment norms can be compensatory and have a substantial impact on these teachers' decisions on summative assessments. The teacher groups show considerable variation in the basis for their decisions regarding summative assessment of students' writing.

Finally, the teachers demonstrate a high degree of consistency within each teacher group when discussing summative assessment of students' writing. However, the assessments that individual teachers make after the discussions comply only to some extent with the group's decisions. The least degree of consistency was shown by the teacher group who, in the discussion, seemed the most consistent. Decision-making processes in these discussions are characterized by low intensity, where few assessment alternatives are being considered. The overall assessment consistency between the teacher groups was found to be low.

**Keywords:** *Writing assessment, social moderation, assessment norm, pedagogical decision, decision-making process, summative assessment.*

Stockholm 2018

<http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-154769>

ISBN 978-91-7797-236-5  
ISBN 978-91-7797-237-2

Institutionen för språkdidaktik

Stockholms universitet, 106 91 Stockholm



Stockholms  
universitet



SAMTAL OM SKRIVBEDÖMNING

Per Blomqvist





Stockholms  
universitet

# Samtal om skrivbedömning

Lärares normer, beslut och samstämmighet

Per Blomqvist

©Per Blomqvist, Stockholms universitet 2018

ISBN tryckt 978-91-7797-236-5

ISBN PDF 978-91-7797-237-2

Tryckeri: Universitetsservice US-AB, Stockholm 2018

Distributör: Institutionen för språkdidaktik, Stockholms universitet

# Författarens tack

Nu när denna avhandling trycks finns många att tacka. Först ett stort tack till er anonyma lärare som deltog i de samtal som denna avhandling bygger på. Utan er insats hade inte detta kunnat skrivas, tack!

Största tack till mina handledare Kristina Danielsson, Viveca Lindberg och Gustaf B Skar! Tack för att jag har fått ta del av ert stora kunnande, er omtanke och uppmuntran under dessa år. När jag umgås med er känner jag mig klokare än vad jag är, och det är härligt! Tack Nina, tack Vivve och tack Gustaf!

Stort tack till Inger Lindberg, Tammi Gustafsson Nadel och Synnøve Matre som på 50- och 90-procentseminarierna kritiskt och konstruktivt har granskat mitt avhandlingsarbete. Tack för mycket värdefull respons! Tack Inger, tack Tammi och tack Synnøve!

Stort tack till min käre vän Erik Lidbaum! Tack för att du alltid har tagit dig tid att läsa och kommentera både artiklar och kappautkast! Det finns få, om någon, med samma blick för både detaljer och helhet i text. Tack Erik!

Stort tack till mina kära vänner och tidigare ämneskolleger Anna Berge, Karin Lindqvist och Ewa Mölleryd! Utan er hade inget av det här varit möjligt! Tack Anna, tack Karin och tack Ewa!

Stort tack doktorandkolleger! Särskilt tack till Björn Kindenberg, Maria Nilsson, Enni Paul, Jeanette Toth och Jenny Uddling för att ni på olika sätt har tagit del i detta arbete. Torsdagsfika med hembakade kärleksmums till er!

Stort tack till Institutionen för språkdidaktik! Särskilt tack till Gunilla Molloy, Eva Nilson, Tore Nilsson, Sari Vuorenpää och Barbro Westlund för goda råd och läsningar under detta avhandlingsprojekts framväxt. Tack!

Stort tack till K & A Wallenbergs Stiftelse och Humanistiska fakultetens stipendienämnd! Tack vare era donationer har jag kunnat delta i ett flertal konferenser och kurser i Sverige och utomlands. Tack!

Till sist, största tack till min fina familj! Tack för att ni har haft förståelse för de perioder av isolering som krävdes för att ro detta i land. Tack Jennifer, tack Edwin, Nelly och Spencer – era hjärtan bultar i mig.

Stockholm, 4 maj 2018

*Per Blomqvist*





# Avhandlingens artiklar

## **Artikel 1**

Blomqvist, P., Lindberg, V. & Skar, G. B. (2016). Vad behöver eleverna undervisning i för att utveckla sitt skrivande? Förväntningsnormer och didaktiska beslut i svensklärares bedömningsamtal.

I: *Acta Didactica Norge, Vol. 10*, s. 1–22.

## **Artikel 2**

Blomqvist, P. (2018). Vad uppmärksammar lärare i samtal om skrivbedömning? Svensklärares normer för beslut om summativ bedömning.

I: *Nordisk Tidskrift för Allmän Didaktik, Vol. 4*, s. 34–55.

## **Artikel 3**

Blomqvist, P. (2018). Är vi överens om det? Samspel och samstämmighet i svensklärares bedömningsamtal om gymnasieelevers skrivande.

I: *Nordic Journal of Literacy Research, Vol. 4*, s. 1–21.



# Innehåll

Bakgrund .....	1
Bedömning i svensk skola .....	1
Skrivande och skrivbedömning i ämnet svenska .....	3
Syfte och frågeställningar .....	5
Avhandlingens delar .....	7
Teoretiska utgångspunkter .....	9
Begreppsdefinitioner .....	9
Skrivbedömning som tolkningar i flera led – didaktiska perspektiv .....	10
En hermeneutisk teori för pedagogisk textrespons .....	11
Lärares läsning som en didaktisk hermeneutik .....	12
Skrivbedömning som didaktisk handling .....	13
Summativa bedömningars didaktiska funktion .....	14
Skrivbedömning i och genom interaktion – dialogiska perspektiv .....	15
Dialogiska begrepp för att undersöka lärares samtal om skrivbedömning .....	16
Skrivbedömning som reflekterande praktik i en tolkningsgemenskap .....	17
Tidigare forskning .....	21
Lärares skrivbedömningar för elevers skrivutveckling .....	21
Lärares summativa skrivbedömningar: beslutsprocesser och beslut .....	24
Sammanfattning av forskningsläget .....	29
Metodiska val och empiri .....	31
Metoder för datainsamling .....	31
Samtal i fokusgrupper .....	31
Frågeformulär .....	32
Forskningsetiska överväganden .....	33
Urval av deltagare .....	34
Undersökningens genomförande .....	35
Insamlade och urval av stimulusmaterial till studien .....	35
Genomförande och utformning av samtalen i fokusgrupper .....	38
Analysmetoder .....	40
Topikanalys .....	40
Initiativ- och responsanalys .....	40
Analysprocess .....	41
Transkription av materialet .....	41

Analys av topiker .....	42
Analys av initiativ och respons .....	44
Giltighet, tillförlitlighet och generaliserbarhet .....	45
<b>Artiklarna i sammanfattning .....</b>	<b>49</b>
Artikel 1. Lärares förväntningsnormer och didaktiska beslut .....	49
Artikel 2. Lärares normer för beslut om summativ bedömning .....	50
Artikel 3. Lärares interaktion, beslutsprocesser och samstämmighet .....	51
<b>Diskussion .....</b>	<b>53</b>
Avhandlingens bidrag .....	53
Diskrepans mellan uttryck för kvalitetsuppfattningar och normer i olika led av en skrivbedömningsprocess .....	53
Variationer i samsyn och samstämmighet mellan lärare och lärargrupper .....	56
Skillnader i interaktionella mönster i lärargruppernas beslutsprocesser .....	58
Didaktiska implikationer och förslag på vidare forskning .....	60
<b>Summary .....</b>	<b>63</b>
Introduction .....	63
Aim and research questions .....	64
Theoretical framework and key concepts .....	64
Empirical material and methods .....	65
Results .....	66
Discussion .....	67
<b>Referenser .....</b>	<b>71</b>
<b>Bilagor .....</b>	<b>83</b>
Bilaga A: Förfrågan om att delta i ett forskningsprojekt .....	84
Bilaga B: Förfrågan om att bidra med elevtexter .....	85
Bilaga C: Lärargrupperna .....	86
Bilaga D: Frågeformulär .....	87
Bilaga E: Transkriptionstecken .....	88
Bilaga F: IR-kodning .....	89

# Bakgrund

- 2 Signe: en fantastisk text  
3 Sara: [jaa  
4 Sofia: [precis  
5 Siv: ett nöje att läsa den  
6 Saga: tycker jag också  
(*skratt*)  
7 Sofia: rolig  
8 Saga: ja  
9 Signe: inte ofta man skrattar liksom högt för sig själv  
(.) intressant –  
10 Sofia: [bra inledning –  
11 Signe: [fantastiskt fint flyt  
(.) variation  
(.) mycket driven

I ovanstående samtal ger fem lärare uttryck för att den elevtext som de just har läst är både välskriven och underhållande. Eftersom samtalet sker mellan svensklärare i gymnasieskolan ingår det i en specifik didaktisk kontext och uppfattningarna kan sägas återspegla förväntningar lärarna har på elevers skrivande. I det här fallet tycks förväntningarna ha överträffats.

Denna avhandling handlar om hur skrivbedömning tar form i samtal mellan lärare. Särskilt intresse riktas mot vilka uppfattningar och normer lärarna uttrycker när de gemensamt bedömer elevers skrivande, vad interaktionen kännetecknas av och hur lärarna använder skrivbedömningar för didaktiska beslut och betygssättning. Ett sådant perspektiv innebär ett fokus på de tolkande och meningsskapande processer som bedömning av komplexa handlingar innebär.

## Bedömning i svensk skola

Svensk skola har ett mål- och kunskapsrelaterat betygssystem. I korthet innebär det att elever ska bedömas i förhållande till nationella kunskapskrav. Dessa krav utgår från målen för ämnet och står även i relation till det innehåll som har behandlats i undervisningen (Skolverket, 2011).

Dagens betygssystem är det tredje sedan början av 1960-talet och betygsskalan från F till A är den fjärde under samma period. Med andra ord har svensk skola och det svenska betygssystemet på relativt kort tid genomgått betydande förändringar (Pettersson, 2010; Selghed, 2010). Samtidigt med dessa förändringar har kraven på lärares pedagogiska bedömningsarbete ökat, inte enbart i Sverige utan även internationellt (Gipps, 1999). Förutom att vara ämneskunniga pedagoger är lärares förmåga att bedöma elevers prestationer avgörande för om deras bedömningar och deras betygssättning ska kunna betraktas som rimliga i relation till den undervisning som ges (Biggs & Tang, 2007; Daugherty et al., 2012; Harlen, 2005; Wiliam, 2007).

I Sverige formuleras kraven på att bedömning, undervisning och lärande ska stämma överens i bl.a. Skollagen och Skolförordningen som att ”elever ges den ledning och stimulans som de behöver i sitt lärande för att de utifrån sina egna förutsättningar ska kunna utvecklas så långt som möjligt enligt utbildningens mål” (Skollagen, 2 kap §9, 3 kap. §3, 2010:800) och att ”eleverna genom strukturerad undervisning ges ett kontinuerligt och aktivt lärarstöd i den omfattning som behövs för att skapa förutsättningar för att eleverna når de kunskapskrav som minst ska uppnås och i övrigt utvecklas så långt som möjligt inom ramen för utbildningen” (Skolförordningen, 3 kap. §4, 2011:185). Vidare ger Läroplanen (Gy 2011) generella riktlinjer om att läraren, vid betygssättning, ska ”utnyttja all tillgänglig information om elevens kunskaper i förhållande till de nationella kunskapskrav som finns för respektive kurs” och ”allsidigt utvärdera varje elevs kunskaper” (Skolverket, 2011, s. 15). Av ämnesplanerna framgår vad undervisningen ska åstadkomma, dvs. vilka förmågor eleven ska utveckla, och för respektive kurs anges centralt innehåll och kunskapskrav som ligger till grund för betygssättning av elevernas kunskaper. Som stöd för detta bedömningsarbete finns exempelmaterial och modeller för hur lärare gemensamt utifrån olika syften kan arbeta med sambedömning av elevers prestationer och samplanering av ämnesavsnitt och kurser (Skolverket, 2014). Flera av landets kommuner har dessutom genomfört betydande satsningar på att kompetensutveckla lärare i hur de kan samarbeta i bedömningsfrågor för att öka likvärdigheten i svensk skola. Sambedömning har blivit något som lärare ägnar betydande tid åt, framför allt vid bedömning av nationella prov (Jönsson & Thornberg, 2014; SOU, 2016).

Detta till trots har såväl Skolverkets uppföljningar och Skolinspektionens granskningar (t.ex. Skolverket, 2000, 2007, 2009b; Skolinspektionen, 2010, 2013, 2015) som en rad svenska studier (t.ex. Jönsson & Svingby, 2007; Korp, 2006; Lindberg, 2002; Tengberg et al., 2017) i olika perspektiv visat

att det finns variationer i lärares bedömningar. Det handlar om variationer i samsyn av vad de grundar sina bedömningar på och variationer i samstämmighet mellan olika lärares bedömningar av nationella prov eller vid betygssättning. Riksrevisionen har i tre rapporter (2004, 2011, 2013) granskat statens insatser för ökad likvärdighet avseende betygssättning. I den senaste rapporten påtalar de behovet av att rikta uppmärksamheten mot direkta förutsättningar för ökad likvärdighet, dvs. hur undervisningen genomförs, snarare än mot indirekta förutsättningar som skolornas kvalitetsarbete. Detta motiveras utifrån forskning om att strukturerad undervisning med aktivt lärarstöd är det som har störst möjlighet att påverka elevers kunskapsresultat.

Mot bakgrund av detta framstår det angeläget att undersöka hur bedömning kan ta form på skolor i ett system som det svenska där lärarna, i internationell jämförelse, tilldelats stort mått av autonomi avseende betygssättning (Lundahl & Tveit, 2014; Mickwitz, 2011). Genom att undersöka lärares bedömningar i relation till ett specifikt skolämne under en specifik tid och på en bestämd plats kan man få en bild av hur bedömning gestaltar sig i ett mål- och kunskapsrelaterat betygssystem som det svenska. Detta kan i sin tur synliggöra några av de svårigheter och överväganden som lärare ställs inför vid bedömning *av* och *för* elevers lärande.

## Skrivande och skrivbedömning i ämnet svenska

Svenskämnet har ett betydande ansvar för elevers skrivande och skrivutveckling. Under skoltiden och framför allt i gymnasieskolan ska eleverna ges förutsättningar att tillägna sig redskap för att skriva alltmer komplicerade texter på ett sätt som fungerar i olika sammanhang och för olika mottagare (Skolverket, 2011). I ämnesplanen för svenska i gymnasieskolan uttrycks detta som att eleverna ”ska ges möjlighet att bygga upp en tillit till sin egen språkförmåga och tillägna sig de språkliga redskap som krävs för vardags- och samhällsliv” och ”utveckla sådana kunskaper om muntlig och skriftlig kommunikation som behövs i arbetslivet och för vidare studier” (Skolverket, 2017). För att ge eleverna dessa möjligheter behöver svensklärare utforma en skrivundervisning som bygger på återkommande analys av var eleverna befinner sig kunskapsmässigt i förhållande till vart eleverna ska i specifik undervisningskontext (William & Thompson, 2007). Lärare behöver med andra ord få syn på vad eleverna kan och vet och utifrån det planera och forma en utvecklande och stöttande skrivundervisning. Detta gör förmågan

att bedöma elevers skrivande till en didaktisk kärnkompetens (Evensen, 2009; Hoel, 2001).

Förutom att använda bedömningar av elevers skrivande för didaktiska beslut om skrivundervisning ska lärare även kunna bedöma elevers skrivförmåga på ett likvärdigt och rättssäkert sätt (Skolverket, 2011). Sådan resultatrapportering kan ske vid olika tidpunkter i en kurs men oftast i samband med att någon form av omdöme ges eller kursbetyg sätts. Tillsammans med de ämnesplaner och kunskapskrav som styr skolans verksamhet blir lärarens omdömen och betyg ett "message system" (Bernstein, 1971) som betonar vad som är viktigt att lära sig i en specifik undervisningskontext. Hur lärare formulerar och kommunicerar dessa bedömningar och när de äger rum kan därför få betydande påverkan på elevers fortsatta lärande, deras studiemotivation och deras förståelse för målen med studierna.

Det finns idag en stor samstämmighet bland forskare om att likvärdig utbildning förutsätter att lärare sinsemellan delar uppfattningar om vad och hur elevers prestationer ska bedömas. I arbetet för ökad likvärdighet föreslår en rad forskare (t.ex. Crisp, 2010; Sadler, 2005; Wyatt-Smith et al., 2010) att lärare sinsemellan i olika konstellationer diskuterar bedömningar av elevers texter. I sådana tolkningsgemenskaper (Berge, 1996, 2009) kan olika uppfattningar mötas och problematiseras vilket kan leda till gemensamma förhållningssätt för hur elevers skrivande ska bedömas i förhållande till styrdokumentet, något som på sikt kan leda till ökad samsyn och samstämmighet i de skrivbedömningar lärare gör. Ett sådant arbete framstår som särskilt nödvändigt med tanke på att skrivbedömningar ofta har visat sig utgå från enskilda lärares outtalade förväntningar på elevers skrivförmåga (Berge 2005; Fasting et al., 2009; Jølle, 2014; Wyatt-Smith & Klenowski, 2013). En stor del av den forskning som finns om skrivbedömning har dock inriktats på att undersöka enskilda lärares bedömningar, ofta genom så kallade *think-aloud*-studier (t.ex. Crisp, 2007, 2010; Suto & Greatorex, 2008). Dessa studier har visserligen bidragit med ökade kunskaper om lärares kognitiva beslutsprocesser när de bedömer elevers skrivande men inte hur skrivbedömning sker genom förhandlingar mellan lärare och hur en ömsesidig förståelse gradvis kan växa fram i kollegiala samtal.

I Sverige har studier om svensklärares skrivbedömningar till stor del inriktats på att undersöka variationer i enskilda lärares bedömningsbeslut (Aldrin, 2015; Borgström & Ledin, 2014; Ciolek Laerum, 2009; Östlund-Stjärnegårdh, 2002) eller reliabilitets- och validitetsproblematik avseende skrivuppgifter och lärares bedömningar av dem i undervisningen (Jönsson & Svingby, 2007; Skar, 2013). Dessutom har ett flertal studier på olika sätt



behandlat skrivbedömning i relation till lärarrespons på elevtexter (Bergman-Claeson, 2003; Kronholm-Cederberg, 2009; Norberg Brorson, 2007; Parmenius Swärd, 2008) eller belyst lärares bedömningar av textkvaliteter och betygssättning avseende elevers skrivförmåga (Ciolek Laerum, 2009; Östlund-Stjärnegårdh, 2002).

Vad som däremot har uppmärksammats i lägre grad är svensklärares samtal om skrivbedömning. Detta gör att det saknas kunskaper om hur beslutsprocesser ser ut när lärare som grupp bedömer elevers skrivande och vilka uppfattningar och normer som styr deras beslut om skrivundervisning och betygssättning. I denna studie riktas därför uppmärksamheten mot hur skrivbedömning sker i interaktion mellan lärare och hur de normer som uttrycks ligger till grund för de beslut lärare gemensamt fattar. På så sätt synliggörs vad lärare uttrycker är centralt att diskutera i en specifik undervisnings- och bedömningskontext vilket, utöver bidraget till ökad kunskap om skrivbedömning, också har implikationer för ett didaktiskt arbete med elevers skrivande i gymnasieskolan.

## Syfte och frågeställningar

Denna avhandling undersöker lärares samtal om skrivbedömning. Kunskapsobjektet är skrivbedömning som tolkande och meningsskapande kollektiva processer och det studieobjekt som undersöks är svensklärares samtal om skrivbedömning i gymnasieskolan.

Studiens syfte är att utifrån didaktiska och dialogiska perspektiv beskriva och analysera hur dessa bedömningssamtal tar form, vilka uppfattningar och normer lärarna uttrycker samt vilka beslut de fattar. Avhandlingen har alltså ett dubbelt fokus både på *hur* lärarna samtalar och *vad* de samtalar om. De övergripande forskningsfrågorna rör därför vad som utifrån dialogiska perspektiv karakteriserar lärarnas samtal om skrivbedömning, vilka normer lärarna uttrycker och vilka beslut de fattar. Dessa övergripande forskningsfrågor besvaras genom tre delstudier som presenteras i de tre vetenskapliga artiklar som ingår i avhandlingen. Genom dessa delstudier söks svar på följande mer specificerande frågor:

Vilka kvalitetsuppfattningar och förväntningsnormer uttrycker lärarna i samtal om elevers skrivande?

Vilka didaktiska beslut fattar lärarna utifrån bedömningen av elevernas skrivande?

Vilka normer för beslut om summativ bedömning uttrycker lärarna och hur ges dessa normer prioritet vid bedömning av enskilda elevers skrivförmåga?

Hur interagerar lärarna i samtal om summativ bedömning av elevers skrivande?

Hur samstämmiga ger lärarna uttryck för att vara?

I en första artikel undersöktes lärarnas uttryck för kvalitetsuppfattningar och förväntningsnormer vid bedömning av elevers texter samt hur bedömningsinformation om elevers skrivande återkom i lärarnas didaktiska beslut om skrivundervisning. Studien svarade därmed mot avhandlingens två första forskningsfrågor: ”Vilka kvalitetsuppfattningar och förväntningsnormer uttrycker lärarna i samtal om elevers skrivande?” och ”Vilka didaktiska beslut fattar lärarna utifrån bedömningen av elevernas skrivande?”.

I en andra artikel undersöktes lärarnas uttryck för normer vid summativ bedömning av elevers skrivande och vilka normer som gavs prioritet när enskilda elevers skrivande bedömdes i förhållande till kursen svenska 1. De forskningsfrågor som studien besvarade var: ”Vilka normer för beslut om summativ bedömning uttrycker lärarna och hur ges dessa normer prioritet vid bedömning av enskilda elevers skrivförmåga?”.

I en tredje artikel undersöktes slutligen lärarnas beslutsprocesser och samstämmighet i samtal om summativ bedömning av elevers skrivande. Studien svarade mot forskningsfrågorna: ”Hur interagerar lärarna i samtal om summativ bedömning av elevers skrivande?” och ”Hur samstämmiga ger lärarna uttryck för att vara?”

Empirin för de tre artiklarna är huvudsakligen densamma, tre svensk-lärargruppernas samtal om skrivbedömning, och artiklarna förhåller sig teoretiskt och metodiskt till varandra genom att empirin analyseras och diskuteras med hjälp av begrepp och modeller från didaktiska och dialogiska teorier.

Det bör redan här påpekas att de samtal som bildar empiri för avhandlingens är arrangerade, med syftet att generera data för studien (se *Metodiska val och empiri*). Samtalen i sig är visserligen autentiska genom att lärarna bedömer elevtexter som är skrivna av gymnasieelever som har gått samma kurs som bedömningarna avser men bedömningen gäller inte elever som lärarna själva undervisar. Elevtexterna är med andra ord sådana som lärarna skulle kunna bedöma i relation till sin undervisning men lärarna vet inget mer om eleverna än det som de ser i deras texter. Det är heller inte lärarna som har utformat de skrivuppgifter eller skrivprocesser vars produkter (texter) de bedömer. Följden av detta blir att studien alltså inte undersöker lärares skrivbedömningar, skrivundervisning eller betygssättning av egna elever,

som konkreta handlingar i en lärandemiljö, utan lärares samtal om detta utifrån ett antal elevtexter som de förses med. På så sätt liknar samtalen de som en expertpanel av bedömare genomför och det är genom lärarnas kommunikativa handlingar i dessa samtal – vad lärarna gör, hur de resonerar och vilka bedömningar samtalen resulterar i – som uppfattningar och normer om skrivbedömning synliggörs. Genom att lärargruppernas samtal om skrivbedömning utgår från samma elevtexter och struktureras utifrån samma frågor blev det även möjligt att jämföra bedömarvariationer inom och mellan lärargrupperna.

## Avhandlingens delar

Avhandlingen består av en kappa och tre vetenskapliga artiklar. I kappans inledande kapitel ges en bakgrund till studien med en introduktion av forskningsfältet skrivbedömning och en presentation av studiens övergripande syfte och frågeställningar. I andra kapitlet *Teoretiska utgångspunkter* presenteras de centrala begrepp och modeller som har använts för att analysera och diskutera svensklärargruppernas samtal om skrivbedömning. I tredje kapitlet *Tidigare forskning* ges en tematiserad överblick av tidigare studier på området i relation till studiens syfte och frågeställningar. I fjärde kapitlet *Metodiska val och empiri* presenteras metoder för datainsamling och analys, sammanlagd empiri, urval av deltagare och hur analysprocesserna har sett ut. Dessutom diskuteras avhandlingens giltighet, tillförlitlighet och generaliserbarhet. I femte kapitlet *Artiklarna i sammanfattning* ges en kortare presentation av de tre artiklarna och i det avslutande kapitlet *Diskussion* presenteras och diskuteras avhandlingens resultat. Dessutom diskuteras didaktiska implikationer och förslag ges på fortsatt forskning.



# Teoretiska utgångspunkter

Denna avhandling undersöker skrivbedömning som tolkande och menings-  
skapande kollektiva processer och det som studeras är svensklärares samtal  
om skrivbedömning i gymnasieskolan. De teoretiska perspektiv som läggs  
på dessa samtal är didaktiska och dialogiska.

Didaktik kan bygga på olika teoretiska grunder (James, 2012; Säljö,  
2000) men om man med didaktik avser lärares reflektiva arbete, som att pla-  
nera undervisning och bedöma elevers kunskaper i relation till styrdokument  
och det innehåll undervisningen har behandlat (Westbury, 2000), är didak-  
tiska och dialogiska teorier tätt sammankopplade. Lärares förmåga att kom-  
municera bedömningar och kontinuerligt utvärdera undervisningen i relation  
till elevernas kunskaper och den kontext lärandet ingår i blir något som kan  
bli avgörande för elevers möjligheter till fortsatt lärande (Jank & Meyer,  
1997). Dessutom är ett dialogiskt samarbete mellan kolleger en nödvändig  
del i ett gemensamt didaktiskt utvecklingsarbete (Gundem & Hopmann,  
2002; Hopmann, 1997).

För att förstå vad skrivbedömning innebär och hur samtal om detsamma  
kan förstås i didaktiska och dialogiska perspektiv struktureras detta kapitel  
utifrån följande två rubriker: ”Skrivbedömning som tolkningar i flera led –  
didaktiska perspektiv” och ”Skrivbedömning i och genom interaktion – dia-  
logiska perspektiv”. Kapitlet inleds med en presentation av centrala begrepp  
från avhandlingens syfte och frågeställningar.

## Begreppsdefinitioner

Centrala begrepp som ingår i avhandlingens syfte och forskningsfrågor är:  
kvalitetsuppfattning, förväntningsnorm, summativ bedömning och samstäm-  
mighet.

Begreppet *kvalitetsuppfattning* (Zandén, 2010) används för att benämna  
vad lärarna riktar sin uppmärksamhet mot när de diskuterar bedömning av  
elevers skrivande. Dessa kvalitetsuppfattningar är uttryck för vad lärare an-  
ser vara relevant att bedöma vid en viss tidpunkt i utbildningen och betraktas

därför som normativa. Dessa normer är dock inte synonyma med de regler som lärarna behöver förhålla sig till, dvs. styrdokumentet (jfr *formella normer*, Angelöw & Jonsson, 2000), utan kan vara *informella* och bygga på underförstådda uppfattningar eller implicita antaganden hos lärarna om vad skrivbedömningar ska inriktas mot (Evensen et al., 2016). Sådana informella normer kan vara personliga men i denna studie riktas intresset mot de normer som lärarna ger uttryck för att *dela* när de diskuterar skrivbedömning. På så sätt kan dessa normer ge en bild av hur styrdokumentens formuleringar är internaliserade i de konkreta skrivbedömningar som lärare på olika skolor gör.

De normer som lärarna ger uttryck för avseende elevtexters kvaliteter benämns *förväntningsnormer*. Detta begrepp hämtas från det norska skrivforskningsprojektet *NORM-projektet* (Berge et al., 2017; Evensen et al., 2016; Matre et al., 2016) där förväntningsnormer har utarbetats i samarbete mellan skrivforskare och erfarna norsklärare för grundskolans årskurs 5 och 8. Dessa förväntningsnormer används både vid formativ och summativ bedömning och syftar bland annat till att öka medvetenheten hos såväl elever som lärare om hur olika textkvaliteter kan benämnas och förstås.

I tillägg undersöks även de samlade överväganden som lärarna gör vid *summativ bedömning* av elevers skrivande, dvs. när elevers skrivuppgifter ska vägas samman till ett betygsomdöme för en kurs, och vad som kännetecknar interaktionen i lärargruppernas beslutsprocesser (Allal, 2013; Wyatt-Smith et al., 2013). Interaktionella mönster i lärarnas bedömningssamtal ställs även i relation till lärarnas uttryck för *samstämmighet*, dvs. de betygsomdömen som lärarna avger (Jönsson & Thornberg, 2014; Sadler, 1989), för att synliggöra hur lärarnas beslut om summativ bedömning varierar inom och mellan lärargrupperna.

## Skrivbedömning som tolkningar i flera led – didaktiska perspektiv

Gemensamt för de skrivdidaktiska modeller som har utvecklats under senare tid, exempelvis *genreskrivande* (Kuyumcu, 2004), *processkrivande* (Strömquist, 2007) och *skrivhjulet* (Berge et al., 2017), är att de bygger på både kognitivt och socialt inriktad skrivforskning som betraktar elevers skrivande och skrivutveckling som något som sker i olika steg, i processer, och där texter blir till och utvecklas i sociala och dialogiska sammanhang (Bakhtin, 1981; Hoel, 2001; Vygotsky, 1978). Återkommande betonas hur olika typer

av respons- och gensvarsarbete kan stötta elevers skrivutveckling och ge elever möjlighet att utveckla sådana språk- och textkunskaper som krävs för att kunna hantera såväl funktionella som formella krav i skrivandet av olika texter (Dysthe, 2011; Evensen, 1997; Gibbons, 2013). I ett sådant skrivdidaktiskt arbete är formativ bedömning ett bärande element. Lärarens uppgift blir att planera en skrivundervisning som möter elevernas behov att utveckla sitt skrivande genom ett lärande som sker i ett samspel mellan sociala och kognitiva handlingar.

### **En hermeneutisk teori för pedagogisk textrespons**

Ett antal forskare (Huot, 2002; Igland, 2009; Phelps, 2000) menar att det länge fanns en alltför snäv beskrivning *av* och *förståelse för* vad lärares skrivbedömningsarbete innebär. Kritiken består framför allt i att studier om skrivbedömning fram till och med 1990-talet haft ett nästan ensidigt fokus på pedagogisk textrespons i form av skriftliga kommentarer till elevtexter. Dessa studier, exempelvis Straub (2000), anses visserligen ha haft betydelse för utvecklandet av responsmetoder men de har inte lyckats beskriva skrivbedömning i en ”theoretical, pedagogical, or communicative context” (Huot, 2002, s. 111). Ett sådant ensidigt fokus på lärares responsstilar är en ”konvensjonell responsoppfatning som er altfor snever og instrumentell til å fortene karakteristikken responsteori” (Igland, 2009, s. 25) och som varken behandlar betydelsen av hur lärare läser och tolkar elevers texter eller hur lärare använder och kommunicerar bedömningar av elevers skrivande. Phelps (2000) föreslår ett vidgat teoretiskt perspektiv för att beskriva lärares didaktiska skrivbedömningsarbete, en hermeneutisk inriktad teori med fokus på de tolkningsprocesser som skrivbedömning innebär. Igland (2009) benämner en sådan teori *en eftertankens didaktik* som lyfter fram komplexiteten i lärares responshandlingar och vilka kunskaper och erfarenheter skrivbedömningsarbetet kräver.

Den modell för pedagogisk textrespons som Phelps (2000) föreslår delar in lärares skrivbedömningar i tre faser. Den första fasen behandlar läsning av elevers texter som en *didaktisk hermeneutik* (”pedagogical hermeneutics”), den andra behandlar bedömning och återkoppling på elevtexter som *didaktisk textkritisk kunskap* (”pedagogical criticism”) och den tredje didaktiska beslutsfattande om vad den fortsatta undervisningen ska inriktas på som *didaktisk handling* (”pedagogical action”). Lärares planering av skrivundervisning blir på så sätt ett resultat av de läsningar och bedömningar de gör

och pedagogisk textrespons beskrivs som ”an art of teaching” (ibid., s. 105), inte något som ligger vid sidan *av* eller *efter* undervisningen.

### **Lärares läsning som en didaktisk hermeneutik**

När Huot (2002) vill skifta fokus från *hur* lärare ger respons på elevers texter till *varför* de gör det, innebär det att närmare undersöka och diskutera vad det egentligen innebär att läsa som en lärare. Sådan läsning ska inte ses som väsensskild från annan typ av läsning men den sker i en didaktisk kontext vilket får ett antal konsekvenser. Exempelvis menar Huot att själva syftet med att lärare läser elevers texter oftast är att kunna säga något om deras skrivande. Detta innebär att läsningen har ett dubbelt fokus på att både bedöma vad som *finns* och vad som *inte* finns i texten, en läsning som Igland (2013) beskriver som en pendling mellan att bedöma vad texten *är* och vad den *kunde vara*. Dessutom poängterar Phelps (2000) att bedömningens syfte inte enbart är att just den texten som bedöms ska bli bättre utan att läraren vill förbättra och förändra elevens skrivande i stort:

changes in the writer's available options, the writer's intentions, the writer's consciousness, the writer's future choices, the writer's appreciative systems and strategic repertoire, the writer's theory of writing changes in the course of a learner's course of development (as writer, thinker, responsible user of rhetoric) and its outcome (s. 104).

Phelps (2000) lyfter fram att lärares skrivbedömningar bör ses som ”most fundamentally reading, not writing” (s. 93) och betonar därmed vikten av att även studera vad som föregår olika responshandlingar, inte enbart responshandlingen i sig. Följen av ett sådant förhållningssätt blir att fokus hamnar på de antaganden och bedömningar som lärare gör om elevers skrivförmåga utifrån läsning av något partikulärt, dvs. den konkreta elevtexten, och hur dessa bedömningar används för olika beslut (Skar, 2013). Lärares läsning kräver därför en *didaktisk textkritisk kunskap* som både är subjektivt deltagande och objektivt granskande, och som innebär att läraren ska kunna koncentrera sin bedömning till ett urval kvaliteter i elevers texter, inte alla (Igland, 2013; Phelps, 2000). Målet med sådan läsning är att lärarresponsen ska förse eleverna med en ”manageable amount of information” (Huot, s. 132) vilket kräver att lärare både har teoretiska kunskaper om text, språk och skrivutveckling men även praktisk erfarenhet av att ge respons som kan fungera i dialog med eleverna (Dysthe & Igland, 2003; Gudem & Hopmann,



2002). I lärares tolkningsprocesser av elevers texter ingår alltså även att överväga hur responsen ska nå de elever den angår.

Vidare argumenterar Huot (2002) för en större medvetenhet om de olika didaktiska och kontextuella faktorer som påverkar lärares läsning, exempelvis vad undervisningen har behandlat, om det är ett första textutkast eller en slutversion eller om texten är skriven i början eller mot slutet av en kurs. Dessa faktorer påverkar hur lärare läser elevers texter och innebär att lärares läsning "constitutes a particular kind of literacy event" (ibid., s. 119).

### **Skrivbedömning som didaktisk handling**

Didaktisk handling ("pedagogical action", Phelps, 2000) innebär att lärare använder bedömningar av elevers texter för att fatta beslut om en utvecklande och stöttande skrivundervisning (jfr Gibbons, 2013). När flera studier som rör skrivbedömning har tenderat att fokusera på lärares textrespons till enskilda elever föreslår Phelps (2000) att lärares bedömningar och textrespons i högre grad bör kunna kommuniceras med eleverna som grupp i form av undervisning. Den teori för pedagogisk textrespons som hon för fram blir på detta sätt en teori med lärandet i fokus: lärandet hos den enskilde eleven som skriver texten, hos klassen som tar del av undervisningen och hos läraren som lär av elevernas texter. Genom att identifiera var eleverna befinner sig kunskapsmässigt kan läraren överväga vilken undervisning eleverna behöver för att utveckla sitt skrivande och utifrån det fatta beslut om passande uppgifter och aktiviteter (William & Thompson, 2007).

När Sadler (1989, 1998, 2013) diskuterar den formativa bedömningens funktion menar han att det framför allt handlar om hur lärares bedömningar av kvaliteter i elevers prestationer och arbeten används för att forma och ge förutsättning för elever att utveckla sina förmågor. Han framhåller också att effektiv feedback inte ska ses som att elever enbart, genom exempelvis lärarens skriftliga textrespons, informeras om hur de har lyckats eller misslyckats med sina uppgifter utan att det krävs att läraren även fattar beslut om en undervisning som kan stötta elevers lärande framöver. Det innebär att om respons från läraren till eleverna ska bli effektiv behöver lärarens analys omsättas i aktiviteter som kan ge eleverna möjlighet att utveckla sitt lärande i handling, inte stanna vid att i ord beskriva elevernas behov av utveckling i mer allmänna ordalag för: "learners do not always learn much purely from being told, even when they are told repeatedly in the kindest possible way" (Sadler, 2013, s. 16). När istället aktiviteter utformas som i olika steg ger eleverna möjligheter att utveckla sina kunskaper och färdigheter kan under-

visningen på ett bättre sätt möta elevernas lärandebehov (Gibbons, 2013; Wiggins, 1998). Black och Wiliam (2018) beskriver detta som att:

the teacher elicits evidence from the class, and, on the basis of that evidence, makes inferences about what next steps in instruction are likely to be most effective (s. 8)

Skrivbedömning som didaktisk handling innefattar med andra ord både planering, genomförande och utvärdering av undervisning.

### **Summativa bedömningsars didaktiska funktion**

Gemensamt för flera av de modeller som beskriver lärares summativa bedömningar av elevers prestationer (se t.ex. Black & Wiliam, 2006; Brookhart, 2010; Moss et al., 2006; Sadler, 1998) är att de beskriver bedömningsprocesser där lärares tolkningar i olika steg leder fram till någon form av beslut om omdöme. Allal (2013) beskriver det som att lärare först utifrån tolkningar av styrdokumentens formuleringar utformar undervisning och examinerande uppgifter, därefter bedömer elevers enskilda prestationer i förhållande till (tolkningar av) kriterier för att slutligen väga samman elevers olika prestationer under en kurs uttryckt som ett betyg eller omdöme. Denna modell liknar i mångt och mycket Harlens *ramverk för summativ bedömning* (2012) där lärarens summativa bedömningar utgår från ett urval *elevuppgifter/aktiviteter* ("collection of evidence"), *tolkar* materialet ("interpretation of evidence"), som läraren bedömer i förhållande till kriterier av något slag ("judgement of achievement, criterion-referenced") för att *rapportera* någon form av omdöme ("report on achievement").

Harlen (2012) lyfter även fram summativa bedömningsars didaktiska funktion och menar att lärare genom att återkommande jämföra elevers kunskaper med förväntade kunskapsnivåer i en viss kurs får information som kan användas för beslut om vad undervisningen ska inriktas på. Detta betyder att lärares förmåga att summativt bedöma elevers kunskaper blir central för att kunna uppnå likvärdighet i både resultatrapportering och undervisning (Black et al., 2011; Evensen, 2010; Pedder & James, 2012; Skar, 2013).

Flera bedömningsforskare har på liknande sätt diskuterat det problematiska i att upprätthålla en tydlig åtskillnad mellan olika bedömningsars funktion i utbildningssammanhang eftersom det är svårt att med någon större precision definiera vad som ska anses vara just formativ respektive summativ bedömning (bl.a. Black, 2003; Gipps, 1994; Newton, 2007). Harlen (2005) menar i likhet med flera andra forskare (Bennett, 2011; Black & Wiliam,

2018; Brookhart, 2010; Evensen, 2010; Hoover & Abrams, 2013; Pedder & James, 2012) att all bedömning i pedagogisk verksamhet istället bör betraktas som redskap för att utveckla undervisning och ge elever bättre möjligheter att lära sig. Olika bedömningar får, beroende på hur de används, betydelse för undervisning och elevers lärande. Summativ bedömning ska därför inte likställas med betygssättning; det är inte *när* bedömningar görs som avgör om den är formativ eller summativ utan *vad* bedömningar används till (Newton, 2007; Taras, 2005, 2009). När läraren summerar vad en elevs kunskaper motsvarar i förhållande till kursens mål och kriterier får läraren information om vad eleven kan vid en viss tidpunkt i en given undervisningskontext. En sådan lägesrapport kan sedan bli underlag för olika beslut och handlingar (Harlen, 2012).

## Skrivbedömning i och genom interaktion – dialogiska perspektiv

Dialogisk teori (även: *dialogism*) utgår från sociolingvistisk och sociokulturell teoribildning som framför allt betraktar språk som kommunikativa handlingar mellan människor (Bakhtin, 1981; Vygotsky, 1978). I korthet innebär det att individers meningsskapande ses som både dynamiska och sociala processer, något som sker *i* och *genom* interaktion med andra (Linell, 2001; Rommetveit, 1992).

Linell (2001, 2009) beskriver dialogismen som ett alternativ till den monologism som han anser har genomsyrat en stor del av tidigare forskning. Till skillnad från monologismen är dialogismen intresserad av kommunikation som integrerad del av de kontexter som den ingår i genom individers gemensamma meningsskapande (jfr Bakhtin, 1981; Wertsch, 2009). I ett dialogiskt perspektiv är alltså kontextbegreppet dynamiskt; kommunikation skapas och omskapas beroende på vilken situationen är, vilka som ingår i den, var den sker och när. Varje kommunikationssituation består sålunda av yttranden som i olika grad både är svar på tidigare yttranden och ger upphov till framtida yttranden (Linell, 2001, 2009).

De samtal om skrivbedömning mellan lärare som denna avhandling undersöker betraktas som interaktiva, dynamiska och kollaborativa processer mellan professionella aktörer (Gipps, 1999) och analysen inriktas på hur de deltagande lärarna interagerar när de tillsammans diskuterar bedömning av elevers skrivande samt fattar beslut om utvecklande och stöttande skrivun-

dervisning och summativ bedömning av elevers skrivförmåga (jfr ”dialogical interacts”, Linell & Marková, 1993).

## **Dialogiska begrepp för att undersöka lärares samtal om skrivbedömning**

För att undersöka lärarnas samtal om skrivbedömning utgår analysen från följande begrepp: kommunikativt projekt, dubbel dialogism, dominans, dynamik och koherens.

Ett *kommunikativt projekt* definieras som en sekvens av kommunikativa handlingar mellan individer när de tillsammans försöker lösa ett problem eller komma överens om något (Linell, 2009). Kommunikativa projekt kan omfatta hela samtal men även vara mer begränsade *lokala kommunikativa projekt* där individer behandlar ett ämne i ett antal turer. Dessa mindre projekt har ofta till uppgift att bidra till genomförandet av mer övergripande projekt och kan därför visa hur olika kommunikativa projekt samverkar för att komma fram till en lösning på problemet eller uppgiften.

I de bedömningssamtal som utgör denna avhandlings empiri består lärarnas kommunikativa projekt av att lösa olika bedömningsuppgifter som de accepterat att göra inom ramen för samtal i fokusgrupper (se *Metodiska val och empiri*). För att göra det behöver lärarna ofta lösa olika mindre kommunikativa projekt på vägen, exempelvis kan de behöva komma överens om gemensamma förhållningsregler för hur bedömningen ska gå till. Ett exempel på det är följande utdrag från Corneaskolan när lärarna diskuterar summativ bedömning av elevers skrivande:

35 Cia: men D kan man ju prata om i betyg

36 Camilla: jaha

37 Carina: fast vi har ju inte dom andra momenten så

(.) vi använder bara dom här

(.) huvudbetygen så att säga

38 Cia: ja

39 Christel: okej men jag hade –

40 Cia: nä men då är det ju E då

I dessa sex turer behandlar fyra lärare ett mindre kommunikativt projekt och kommer överens om att de ska utgå från betygen E, C och A. När de har enats om detta fortsätter det gemensamma arbetet med att lösa det övergripande kommunikativa projektet som är att diskutera vilket betygsomdöme elevens skrivande motsvarar.

Vidare använder Linell (2001), influerad av Bakhtin, begreppet *dubbel dialogism* för att beskriva hur deltagare i kommunikativa projekt inte enbart förhåller sig dialogiskt till varandra utan även till de kontexter som omger själva samtalet. Bakhtin (1986) beskriver detta dubbla genom att beskriva hur olika typer av yttranden bör betraktas som delar av kommunikationskedjor där individers enskilda yttranden både är ett resultat av tidigare yttranden (bakåt) och vem eller vilka de vänder sig till (framåt). Med detta synsätt blir de normer och beslut som lärarna ger uttryck för både socialt producerade och reproducerade, påverkade av den didaktiska kontext de både är en del av och är med och skapar (jfr Marková et al., 2007). Kontextuella faktorer påverkar även hur samtalen genomförs och ger dem en striktare struktur än vad som kännetecknar vardaglig konversation (jfr ”institutionella diskurser”, Linell, 2009).

Centrala begrepp för att studera hur individer samarbetar interaktionellt i ett samtal är *dominans*, *dynamik* och *koherens* (Linell, 2011; Linell & Gustavsson, 1987). Med *dominans* avses de interaktionella styrkeförhållanden mellan lärarna, dvs. hur olika lärare, genom att initiera nya ämnen och svara på andras initiativ i olika grad bidrar till att diskussionerna hålls igång. Detta ger en bild av vilka lärare som styr samtalen och vad som kännetecknar det dialogiska samarbetet mellan lärarna i diskussionen.

Ett samtals dynamik beskriver sedan själva rörelsen och intensiteten i samtalet, dvs. hur lärarnas initiativ och respons styr samtalet i olika riktningar (Linell & Gustavsson, 1987). Enskilda lärares initiativ kan med olika styrka föra in nya aspekter och ämnen i samtalet och beroende på hur övriga lärare responderar synliggörs hur lärarna använder och tar hjälp av varandras yttranden i beslutsprocesserna. Det senare beskriver även ett samtals koherens, dvs. hur lärarna bemöter och relaterar till varandras yttranden, vilket säger något om hur väl det dialogiska samspelet fungerar.

## **Skrivbedömning som reflekterande praktik i en tolkningsgemenskap**

Vikten av att lärare samarbetar i bedömningsfrågor är något som återkommande lyfts fram i forskning om skrivbedömning. Lärares gemensamma diskussioner om hur elevers skrivförmåga ska bedömas kan öka deras samsyn om hur styrdokument ska tolkas, stärka likvärdigheten och samstämmigheten i deras beslut och bidra till att utveckla ett gemensamt ämnesspråk för skrivbedömning (Borgström & Ledin, 2014; Klenowski & Wyatt-Smith, 2010; Matre & Solheim, 2016; Skar & Jølle, 2017). I lärares förhandlingar

om exempelvis vilket betyg en elevs skrivande motsvarar ställs egna tolkningar mot andras vilket gör att olika förståelser kan synliggöras och bli underlag för diskussioner. En sådan reflekterande praktik kan bidra till ökad samsyn mellan lärarna om *vad* bedömningen ska inriktas på, *hur* den ska genomföras och *vad* den ska användas till (Colombini & McBride, 2012; Evensen, 2012; Jølle, 2014; Smaill, 2013).

Wenger (1998) föreslår att individers deltagande i en *praktikgemenskap* ("community of practice") kan bidra till att stärka och forma grupperns gemensamma riktlinjer för arbetet där de utifrån *gemensamma mål* ("mutual engagement") och ett ömsesidigt ansvar att lösa *samma uppgift* ("joint enterprise") utvecklar *delade uppfattningar* om hur uppgifter kan lösas ("shared repertoire"). Syftet med praktikgemenskaper är alltså framför allt att genom individers förhandlingar och omförhandlingar komma fram till gemensamma förhållningssätt att lösa problem i samband med att uppgifter ska lösas, inte att komma fram till själva svaret på uppgiften. Dessa processer är dynamiska och bygger på att deltagare delger sina tidigare erfarenheter av hur problem har lösts för att utifrån dem diskutera nya förhållningssätt. Praktikgemenskapen beskrivs därför som en dialogisk lärandeprocess mellan de deltagare som ingår i den och sker i relation till de kontexter den ingår i.

För skrivbedömningens del är sådana praktikgemenskaper, eller *tolkningsgemenskaper* ("interpretive community", Fish, 1980; Berge, 1996, 2009), nödvändiga med tanke på alla de tolkningsled som lärares skrivbedömningar består av där lärares läsningar och bedömningar av elevers skrivande har visat sig variera bland annat beroende av bedömarens bakgrund, utbildning och erfarenheter (Huot, 2002; Igland, 2013; Phelps, 2000). Därtill förståelsen av att språkligt utformade kriterier aldrig kan bli så entydiga att de förstås likadant av alla lärare utan behöver relateras till den undervisningskontext bedömningen ingår i och hur elevers prestationer förhåller sig till övriga formuleringar i styrdokumentet (Maxwell, 2010; Sadler, 1987). Jølle (2015b) beskriver denna problematik som en konflikt mellan positivistiska och hermeneutiska föreställningar om kriteriebaserad bedömning. Å ena sidan förutsätts det som ska mätas (skrivande) vara mätbart som något "entydig, sikkert og presist", å andra sidan är skrivande "komplekst, flertydig og unikt" (s. 17). Lärare behöver därför, föreslår Jølle, handla unikt genom att ha en utarbetad bred repertoar för bedömning, en repertoar som dessutom delas av lärarna som ingår i tolkningsgemenskapen. En fungerande dialog mellan lärare om hur elevers prestationer ska bedömas utifrån styrdokument, undervisning och uppgifter beskrivs som något som kan stärka lärarnas bedömarförmåga att agera i enskilda fall.

I skrivbedömningsforskning från exempelvis Australien, Kanada, Nya Zeeland, Storbritannien och USA benämns lärares diskussioner om sambedömning ”social moderation” eller ”teacher moderation meetings” (Connolly et al., 2011; Maxwell 2010; Reid, 2007; Smaill, 2013; Wyatt-Smith et al., 2010, 2017). Sådan sambedömning innebär att lärare i grupp diskuterar bedömning av elevers texter i förhållande till nationella kriterier, vilket bland annat motiveras av att lärare i exempelvis Australien och Nya Zeeland, precis som i Sverige, har tilldelats stort mått av autonomi vid betygssättning. Dessa bedömningsmöten kan stärka likvärdigheten i olika lärares skrivbedömningar. Förutom att sambedömning sker mellan lärare på enskilda skolor möts även lärare från olika skolor, så att deras olika erfarenheter av att undervisa och bedöma olika elevgrupper får mötas (Klenowski & Wyatt-Smith, 2010).

Utvärderingar av lärares gemensamma diskussioner om skrivbedömning som ”social moderation” visar dock att denna typ av sambedömning inte per automatik leder till ökad samsyn och samstämmighet. Exempelvis visar Maxwell (2010) att bedömningsamtal som i alltför hög grad inriktas på att gruppen ska komma fram till ett gemensamt beslut kan hämma möjligheten att problematisera och fördjupa resonemang om på vilka grunder skrivbedömningar görs. Detta kan leda till att lärarna inte bär med sig en fördjupad kunskap om bedömning när de bedömer i ett annat sammanhang. När Smaill (2013) utvärderar bedömningsens roll utifrån en nya zeeländsk kontext menar hon att ”social moderation” visserligen *kan* leda till att lärare utvecklar sin bedömarförmåga men att det också kan innebära ”missed learning opportunities and unintended learning” (s. 251). Hon föreslår därför att sambedömning bör ha en högre målsättning än att lärare enbart ska dela uppfattningar om hur kriterier ska användas; lärare ska även använda dessa diskussioner ”to strengthen their judgement making and evidence collection” (ibid.). Förutom att lärare tillsammans bedömer och kommer överens om vilka kvalitetskriterier (”quality benchmarks”, Maxwell, 2002a) elevers texter ska motsvara för ett visst betyg, behöver lärare även utveckla kunskaper om hur skrivuppgifter ska utformas i relation till styrdokumentens mål och vilka förutsättningar elever ges för att utveckla sitt skrivande under en kurs.

I likhet med skandinaviska forskare som Berge (2009) och Skar (2013) lyfter Smaill (2013) därmed fram validitetsaspekten i lärares skrivbedömningar och betonar vikten av att lärare tillsammans diskuterar vilka uppgifter elever kan ges i en kurs och hur elevernas skrivprocesser kan utformas; i ett likvärdighetsperspektiv behöver lärare även kunna göra tolkningar av texternas betydelse som bedömningsinformation i relation till given undervis-

ningskontext och utifrån det utveckla skrivundervisningen (jfr Black et al., 2010). Berge (2009) föreslår att ett sådant gemensamt bedömningsarbete mellan lärare bör kännetecknas av ett samspel på två nivåer. Först en *normativ* nivå där lärare i samtal delger sina uppfattningar om kvaliteter i elevers texter vilket kan leda till ökad samstämmighet. Därefter en *praktisk* nivå som innebär att lärare även diskuterar skrivbedömningar i relation till elevers lärande och den skrivundervisning som ges. Ett sådant tillvägagångssätt kan på sikt stärka likvärdigheten i lärares skrivbedömningar genom att pålitligheten i de resultatbedömningar som görs stärks samtidigt som skrivundervisningen utvecklas (jfr Skar, 2013). Dessa tankar ligger till grund för *NORM-projektets* upplägg och genomförande (Berge et al. 2017; Evensen et al., 2016) där ansvaret för elevers skrivutveckling är gemensamt för alla lärare som undervisar i skrivande på skolan och där bedömning blir en kollektiv angelägenhet som tränas explicit genom diskussioner utifrån läsning av elevers texter (jfr Sadler, 1998). I dessa diskussioner kan lärares olika uppfattningar om kvaliteter i elevers skrivande synliggöras och leda till ett meningsutbyte som kan utveckla ett gemensamt ämnesspråk om skrivande, bedömning och lärande (Evensen, 2009; Jølle, 2014; Matre & Solheim, 2016). Bedömning betraktas alltså som en i grunden normativ aktivitet där lärares möjligheter till att gemensamt problematisera skrivbedömning blir en förutsättning för likvärdighet.



# Tidigare forskning

Följande genomgång av tidigare forskning är tematiserad utifrån avhandlingens syfte och forskningsfrågor med fokus på lärares skrivbedömningar som ett didaktiskt och dialogiskt arbete. De två rubriker som strukturerar genomgången benämns ”Lärares skrivbedömningar för elevers skrivutveckling” som avser studier om lärares textrespons och formativa bedömningar och ”Lärares summativa skrivbedömningar: beslutsprocesser och beslut” som avser studier om lärares bedömningar av elevers skrivförmåga.

Litteraturgenomgången är ett försök att ringa in aktuell och för denna avhandling relevant forskning om skrivbedömning. Genomgången utgår framför allt från svensk och övrig nordisk skrivbedömningsforskning från senare år men kompletteras även med några internationella utblickar. Litteraturgenomgången avslutas med en sammanfattning av forskningsläget och denna avhandlings bidrag till forskningsfältet.

## Lärares skrivbedömningar för elevers skrivutveckling

En betydande del av den forskning som finns om lärares skrivbedömningar har formativa ansatser och undersöker hur lärares bedömningar kan ge elever möjlighet att utveckla sitt skrivande. Flera av dessa studier behandlar lärares skriftliga kommentarer till elevtexter, både på textutkast och färdiga produkter (texter), och den påverkan lärares respons har på elevers skrivande, dvs. hur responsen förstås, uppfattas och används av eleverna.

När det gäller vilken påverkan lärares skriftliga kommentarer får på elevers skrivande pekar forskningsresultat i delvis olika riktningar. Norberg Brorson (2007) som har följt högstadielevens skrivande under tre terminer visar att eleverna är osäkra på vad lärarnas textrespons egentligen innebär och att eleverna har svårt att ta till sig den respons som ges för att förbättra sina texter. Bueie (2015) däremot, som även hon undersöker högstadielevens förståelse av lärares skriftliga kommentarer men som respons på skrivprov, visar att eleverna i huvudsak förstår lärarnas kommentarer men att kommentarerna får liten påverkan på hur eleverna förbättrar sina texter när

de inte följs av aktiviteter i den fortsatta skrivundervisningen. Skillnader i hur elever förstår och använder lärares textrespons kan givetvis bero på hur responsen formuleras vilket är något som kan variera beroende på situation och syfte. Det är exempelvis stor skillnad om responsen mest består av lärares rättningar eller om responsen är mer utförlig återkoppling på flera olika nivåer. Men vad som ändå framstår som avgörande för om eleverna får användning av lärarnas textrespons för att utveckla sina texter är att responsen är en del av en pågående skrivprocess.

När Igland (2008) undersöker hur lärares kommentarer på elevers textutkast kan bli ett formativt redskap i skrivundervisningen kommer hon bland annat fram till att sannolikheten är låg att eleverna har någon nytta av dem alls när bedömningarna inte diskuteras tillsammans med eleverna i relation till den undervisning som ges. Den respons som lärarna ger genom skriftliga kommentarer framstår som abstrakt information för eleverna vilket framför allt hämmar de skrivovana elevernas skrivutveckling. I bästa fall rättar eleverna till stavningen och strukturen i sina texter, men det får dem inte att utveckla texterna på en mer övergripande nivå; lärarnas kommentarer får därför ingen betydande påverkan på elevernas skrivutveckling i stort. Även när Eriksen (2017) undersöker formativa aspekter av lärares skriftliga kommentarer till elevtexter visar resultatet att de gymnasieelever som ingår i studien visserligen uppfattar att de kan ha nytta av lärarresponsen men att de inte upplever att de har det. Orsaken till det visar sig bland annat vara att eleverna upplever det svårt att göra genomgripande förändringar av en text när lång tid har gått mellan skrivtillfället och lärarnas bedömning.

Liknande resultat presenterar Kronholm-Cederberg (2009) som visar att när lärarnas kommentarer enbart ges på slutprodukter utan möjlighet till bearbetning för eleverna får de liten påverkan på elevernas möjligheter att utveckla sitt skrivande. Parmenius Swärd (2008) visar även att eleverna kan uppfatta att skolskrivandet är på låtsas och att de råd och tips som de får av läraren inte hjälper dem vidare i skrivandet; det skrivande som framträder beskrivs som ”produktinriktat och betygsorienterat och har inte som mål att eleverna ska utveckla sitt individuella skrivande i ett processperspektiv” (s. 229). Att lärarnas kommentarer till elevernas texter framför allt inriktas på att belysa brister på lokal textnivå bidrar inte heller till att elevernas motivation och lust att utveckla sitt skrivande ökar.

Kronholm-Cederberg (2009) visar även hur elever kan reagera på olika sätt när de tar del av lärares textkommentarer och beskriver deras reaktioner i termer av acceptans, ambivalens och protest vilket även det, på olika sätt, påverkar deras möjligheter att med hjälp av lärarens textkommentarer ut-

veckla sina texter och sitt skrivande. Detta är något som även Bueie (2014) belyser när hon beskriver hur högstadieelever kan ta till sig lärarens textkommentarer på olika sätt. En del elever säger sig förstå och acceptera lärarens kommentarer medan andra elever säger sig antingen förstå men gör motstånd eller inte förstå lärarens kommentarer. Störst problem, menar hon, blir det om eleverna inte förstår men samtidigt accepterar lärarens bedömning. Det leder inte till någon dialog om vad bedömningarna grundar sig på, något som kan ske med elever som åtminstone visar motstånd.

Skar (2013) vidgar innebörden av skrivbedömningens roll för elevers skrivutveckling och undersöker skrivbedömning i relation till den skrivundervisning som bedrivs i olika svenskklassrum i gymnasieskolan: vilka skrivuppgifter lärarna utformar, hur skrivande sker och vilka möjligheter eleverna ges att utifrån lärarnas skrivuppgifter och skrivbedömningar lära sig vad de förväntas kunna. I likhet med andra studiers resultat (Norberg Brorson, 2007; Parmenius Swärd, 2008) visar studien att eleverna inte bedöms i relation till den skrivundervisning som de har tagit del av och att eleverna i övrigt ofta lämnas ensamma i sitt skrivande.

Internationella studier som undersöker hur lärares textrespons förstås, uppfattas och används av elever för att utveckla sina texter och sitt skrivande redovisar liknande resultat. Exempelvis visar Calhoon-Dillahunt & Forrest (2013) i en brittisk kontext att elever visserligen uppskattar lärares konkreta förbättringsförslag men att de ändå har svårt att använda lärarnas kommentarer när de ska förbättra sina texter vilket även kopplas samman med att eleverna lämnas ensamma i den fortsatta skrivprocessen. Richardson (2000) undersöker hur amerikanska elever, från mellanstadiet till gymnasiet, använder lärares textrespons och visar att eleverna framför allt korrigerar sina texter, dvs. rättar felaktigheter i dem, men utan att utveckla texterna i stort. Detta beskrivs bland annat som en konsekvens av att eleverna försöker förstå vad lärarna vill ha istället för att våga sig på att ta egna initiativ.

I diskussioner om hur lärares skrivbedömningar i högre grad kan bidra till att utveckla elevers skrivande menar forskare, bland andra Gibbs och Simpson (2004), att elever behöver vara införstådda med skrivuppgiftens syfte och vilka krav som ställs och att elever behöver få lära sig hur olika textkvaliteter samspelar när texter skrivs. Higgins m.fl. (2001) betonar även att lärares skrivbedömningar ska ske i dialog med eleverna under skrivandets gång så att eleverna ges möjlighet att förstå vad som förväntas av dem under tiden de skriver, istället för att efteråt informeras om hur de skulle ha skrivit (s. 274). Detta är något som återkommer i studier av hur elevers skrivande kan utvecklas genom skrivportföljer (t.ex. Freccero & Pousette, 2005; Krogh,

2007). Det betonas att när lärares textrespons ingår i en processororienterad skrivundervisning och omges av diskussioner mellan elever och lärare om hur texter skrivs kan eleverna ges ökade möjligheter att bygga upp ett metaspråk om texter och skrivande, något som kan hjälpa dem att aktivt delta i sin läroprocess och därigenom utveckla sitt skrivande (jfr Dysthe & Hertzberg, 2009).

Forskning inom det norska skrivforskningsprojektet *NORM-projektet* har även genom att studera effekter av att bedriva skrivundervisning med utgångspunkt i förväntningsnormer visat att elevers skrivande kan förbättras avsevärt om lärarnas skrivbedömningar är en integrerad del av skrivundervisningen (Berge et al., 2017). Förväntningsnormerna används för att både bedöma elevernas skrivande och forma den skrivundervisning som ges i dialog med eleverna, vilket har visat sig öka elevernas möjligheter att utveckla sitt skrivande. I likhet med ett flertal studier (Dysthe, 2011; Hattie & Timperley, 2007; Marshall, 2004; Parr & Timperley, 2010; Sadler 2009, 2013) betonas att när lärares formativa bedömningar kommuniceras med hjälp av kriterier av något slag behöver bedömningarna diskuteras tillsammans med eleverna så att det konkreta skrivandet relateras till uppgiften och målen för utbildningen.

## Lärares summativa skrivbedömningar: beslutsprocesser och beslut

Forskning om svensklärares summativa bedömningar av elevers texter har visat betydande variationer i de beslut olika lärare fattar. Dessa variationer yttrar sig både som bristande samsyn dvs. *vad* de grundar sina bedömningar på och bristande samstämmighet dvs. *vilket betyg* de ger på samma text (jfr Jönsson & Thornberg, 2014). Exempelvis lät Aldrin (2015) drygt hundra svensklärare på högstadiet enskilt betygsätta och motivera sina bedömningar av samma elevtext i årskurs 8. Resultatet visar att både samsynen och samstämmigheten mellan lärarna är låg. Lärarna tar i olika grad hänsyn till olika textkvaliteter när de motiverar sin bedömning; en del lärare uppfattar att texten håller sig till genren, har tydliga tankegångar och en fungerande meningsbyggnad medan andra lärare uppfattar att genreanpassningen inte fungerar, att textens struktur brister och att meningsbyggnaden är problematisk. Vidare varierar lärarnas betygssättning med upp till fyra olika betygssteg, från E till B, där ungefär lika stor andel av lärarna anger betygen E, D och C (30, 34 respektive 32% av lärarna) och en mindre andel betyget B (4%).

Liknande resultat om bristande samsyn och samstämmighet i svensklärares skrivbedömningar redovisas i såväl tidigare som senare studier. Exempelvis visar Östlund-Stjärnegårdh (1999) att lärares betygssättning av det nationella provets skrivdel för gymnasieskolan, med dåvarande betygsskala IG-MVG, i genomsnitt skiljer med mer än ett betygssteg och att externa bedömares betygssättning är mindre generös än betygssättande lärares (2002). Ciolek Laerum (2009) visar att lärares betygssättning av det nationella skrivprovet i årskurs 9, med nuvarande betygsskala F-A, skiljer med upp till tre betygssteg.

Dessutom fäster lärare olika stor vikt vid vilka kvaliteter i elevernas texter som har betydelse vid summativ bedömning. Det handlar om att de främst beaktar texters helhet och innehåll (jfr Östlund-Stjärnegårdh, 1999, 2002), texters yta, stavning och den lokala textnivån (jfr Kronholm-Cederberg, 2009), texters språk och form framför texttyp och funktion (Hedeboe, 2007) eller texters sätt att problematisera resonemang och göra kopplingar till andra källor (Stolt, 2016). Även om studier kan redovisa att lärare och andra bedömare uppvisar viss överensstämmelse i vad de grundar sina beslut om summativ bedömning på, finns betydande variationer mellan vad enskilda lärare utgår från och vilka beslut de fattar baserat på det. Tentativa förklaringar till bristande samsyn och samstämmighet i svensklärares skrivbedömningar utgår bland annat från lärares olika erfarenheter av att bedöma och bedömares skilda roller och olika tillgång till bedömningsinformation (t.ex. Aldrin, 2015; Lundgren, 2013; Östlund-Stjärnegårdh, 2009).

För att öka lärares samsyn och samstämmighet vid summativ bedömning av elevers skrivande förespråkar en rad svenska forskare (Borgström & Ledin, 2014; Ciolek Laerum, 2009; Tengberg et al., 2017; Östlund-Stjärnegårdh, 2002) att lärare oftare möts för att diskutera kvaliteter i elevers texter och de föreslår sambedömning mellan lärare som en framkomlig väg i det arbetet. Detta är förslag som har presenterats långt tidigare i svensk skrivbedömningsforskning, t.ex. att bedömningars reliabilitet kan öka om två eller tre lärare bedömer samma elevtext (Björnsson, 1960; Larsson, 1984), men är något som senare studier påvisat är mycket svårt att uppnå. Exempelvis visar Borgström och Ledin (2014) att en tolkningsgemenskap bestående av erfarna och ämneskunniga lärare inte garanterar ökad samstämmighet i bedömningar av skrivprov även om variationerna minskar om fler än en lärare bedömer samma elevtext. När Tengberg m.fl. (2017) undersöker hur svensklärare kan öva på att utveckla sin förmåga att bedöma elevers skrivande summativt konstaterar de att konsistensen i lärarnas bedömningar, dvs. deras sätt att rangordna texter på ett likartat sätt, kan öka relativt snabbt

medan det är betydligt svårare att närma sig ökad samstämmighet i betygsättning. Vilka orsakerna till detta är behöver dock enligt forskarna undersökas mer ingående genom att exempelvis analysera de beslutsprocesser som lärare deltar i när de i grupp diskuterar bedömning av elevtexter.

Berge (1996, 2002) har bl.a. synliggjort och problematiserat vad han ser som brister i de tolkningsgemenskaper som utmärker norsklärares skrivbedömningspraktiker. Ett antal norska studier har sedan undersökt hur lärares gemensamma bedömningsarbete kan ta form *i* och *genom* samtal. Exempelvis undersökte Jølle (2014, 2015a) hur norska högstadielärares bedömar-kompetens kan utvecklas över tid genom deltagande i professionella träningsprogram och hur lärare använder olika strategier vid parbedömning av elevers texter. Materialet bestod av bedömarialoger mellan lärare i en nationell bedömarpanel och undersöktes som interaktiva situationer avseende lärares samarbete vid summativ bedömning av elevers texter. I en första studie (Jølle, 2014) undersöktes dels hur lärarna responderar på varandras yttranden, om de accepterar eller avslår eller något däremellan, dels vilka referenser lärarna använder när de bedömer texters kvaliteter. I en andra studie (Jølle, 2015a) undersöktes sedan vilka strategier lärarna använder när bedömningen styrs av uppgiftsmatriser. Resultat från dessa båda studier visar dock att den bedömarträning lärarna tog del av i bedömarpanelerna enbart i mindre utsträckning utvecklade deras sätt att använda bedömningsredskapen och att de konsensusbeslut som fattades av lärarna inte byggde på fördjupade diskussioner. Dessutom använde sig lärarna i betydande omfattning av både formativa och privata strategier för beslut om summativ bedömning av elevers skrivande vilket beskrivs som ”a major contributor of making large-scale writing assessments quite blurry” (Jølle, 2015a, s. 466).

Även Matre och Solheim (2015, 2016) har i ett antal studier belyst norska lärares bedömningsprocesser vid summativ bedömning av elevers skrivande. Bland annat undersökte de lärares beslutsprocesser vid bedömning av elevtexter i årskurs 4 och 7 avseende hur lärarna argumenterade, interagerade och använde sig av bedömningsverktyg. Ett övergripande syfte var att se hur kvaliteten i lärarnas skrivbedömningar kunde öka genom deltagande i kollegiala samtal om sambedömning och om dessa samtal öppnade för möjligheter att ifrågasätta och argumentera för olika läsningar och bedömningar. Resultat från dessa studier visar att lärarna använde bedömningsverktygen på olika sätt i beslutsprocesserna. En del lärare framstod som mycket instrumentella genom att de checkade av ett bedömningskriterium i taget medan andra var mer flexibla och utgick både från kriterierna och från sina kunskaper *om* och erfarenheter *av* att bedöma elevtexter. Detta gjorde att de

senares bedömningar i högre grad bestod av tolkningar av kvaliteter i elevers skrivande i stort. Genom bedömarträning tenderade dock lärarna att tillägna sig en mer flexibel förståelse för hur bedömningskriterierna (förväntningsnormerna) kunde användas och utvecklade ett gemensamt ämnesspråk om skrivbedömning. Detta förutsatte dock ett aktivt deltagande från de olika lärarna i bedömningssamtalen; att lärare gavs dialogiskt utrymme att tillsammans diskutera olika aspekter av vad bedömningar kan grunda sig på ("dialogic spaces", Matre & Solheim, 2016).

Skar och Jølle (2017) undersökte om lärare genom bedömarträning kan minska variationer mellan olika lärares bedömningar av elevers skrivande. I studien deltog omkring åttio lärare som undervisade i årskurserna 5 och 8, och som under minst två år utbildades i skrivbedömning för att i förlängningen utgöra en nationell bedömarpanel i de nationella skrivproven. Dessa lärare deltog i återkommande seminarier i form av två till tre dagars utbildning upp till fyra gånger årligen, där de diskuterade bedömningar av elevtexter som del av en tolkningsgemenskap. Resultat visar dock att bedömarträningen inte ledde till att lärarnas skrivbedömningar förbättrades i tillräckligt hög grad för att anses vara pålitliga i bedömningar av nationella skrivprov (jfr Skar, Thygesen & Evensen, 2017). Visserligen gav lärarna uttryck för att vara överens på en holistisk nivå men när de sedan ska motivera sina bedömningar analytiskt, vad som avgjorde bedömningen, kan några peka på kvaliteter i textstruktur, andra på kommunikativa aspekter eller ämnesinnehållet, eller enkelt konstatera att texten har "Something that strikes you" (Skar & Jølle, 2017, s. 12). Utöver det hade lärarna olika uppfattningar om vilka aspekter av skrivandet som eleverna behöver utveckla för att höja sin skrivförmåga.

Flera norska studier har på detta sätt medverkat till att skrivbedömningsforskning gått från att i huvudsak intressera sig för enskilda lärares bedömningar till att fokusera på lärares gemensamma beslut och beslutsprocesser, från att i huvudsak vara kognitivt inriktad till att anta sociokulturella och dialogiska perspektiv. Jølle (2015a) identifierade detta som en internationell trend och beskriver två traditioner inom forskningsfältet skrivbedömning: en teknisk-positivistisk tradition som inriktas på att undersöka vilka textuella aspekter lärare grundar sina bedömningar på och som menar att tydligare kriterier kommer att leda till ökad likvärdighet i lärares bedömningar och en hermeneutiskt orienterad tradition som inriktas på att undersöka meningskapande processer i lärares bedömningar och som lyfter fram att flera aspekter utöver kriterier påverkar lärares bedömningar och därför måste tas hänsyn till (jfr Berge, 2002; Cooksey et al. 2007). Den senare traditionen

poängterar vikten av att undersöka lärares bedömningshandlingar som meningsskapande processer i dialog för att därigenom få kunskaper som säger något om den komplexitet bedömning av skrivande verkligen innebär. Sådana studier kan synliggöra vad som ligger bakom lärares beslut om skrivbedömning och utöka lärares metaspråk om skrivbedömning, kunskaper som beskrivs som centrala för lärares didaktiska professionsutveckling (Evensen, 2012; Matre & Solheim 2015).

När Wyatt-Smith m.fl. (2013) undersökte australiska lärares gemensamma beslutsprocesser vid summativ bedömning av elevers skrivande, identifierade de tre typer av bedömningskriterier som skrivlärarna använde sig av: *explicita kriterier* dvs. föreskrivna kriterier från styrdokument, *latenta kriterier* dvs. ospecificerade och uttalade kriterier som lärare aktiverar i specifik bedömningsituation och *metakriterier* dvs. regler för hur explicita och latenta kriterier ska användas eller inte användas vid summativa bedömningar. Dessa kriterietyper beskrivs som att de både är beroende av varandra ("inter-dependent") och står i relation till varandra ("interrelated") och motiveras av att alla behövs för att fånga upp det komplexa i elevers skrivande där allt inte går att förutse på förhand med enbart explicita kriterier. Kriterierna är enbart en del av beslutsprocesserna där även sociala och kognitiva processer får betydelse för hur kriterier används för att beskriva texters värde. En del lärare refererar i stort sett enbart till kriterierna medan andra lärare knappt nämner dem alls. För att hitta förklaringar till vad som påverkar lärares bedömningar bortom kriterierna undersöks lärares skrivbedömningar som social praktik (Adie et al., 2012). Fokus för dessa studier är att synliggöra hur en ömsesidig förståelse gradvis kan växa fram genom att lärare kontinuerligt deltar i kollegiala samtal om bedömning av elevers skrivande.

Även i amerikanska, brittiska och nyzeeländska studier av lärares samarbete vid summativa bedömningar av elevers skrivande (t.ex. Colombini och McBride, 2012; Reid, 2007; Smaill, 2013) finns flera studier som undersöker lärares beslutsfattande processer och hur lärargrupper gemensamt kan utveckla bedömarkompetens med hjälp av professionella samtal. Vad flera av dessa studier betonar är att svårigheter, det komplexa och dynamiska i lärares skrivbedömningar, är något som måste bejakas för att kunna bidra till att enskilda lärare på sikt utvecklar en ökad överensstämmelse i sina skrivbedömningar. Med ett hermeneutiskt utforskande perspektiv på lärares diskussioner om skrivbedömning blir det ett istället ökat fokus på de skillnader och spänningar som uppstår i lärarnas olika tolkningar av kvaliteter i elevernas



skrivande och hur dessa ska bedömas i relation till skrivuppgifter och kurs eller årskurs.

## Sammanfattning av forskningsläget

Ett flertal studier om skrivbedömning har belyst lärares skriftliga respons till elevers texter och vilka konsekvenser dessa bedömningar kan ha för elevers fortsatta skrivutveckling (Bueie, 2017; Eriksen, 2017; Igländ, 2008; Kronholm-Cederberg, 2009; Norberg Brorsson, 2007; Parmenius Swärd, 2008). Ett huvudsakligt fokus har därmed varit att undersöka hur lärares skrivbedömningar som skriftliga kommentarer kan forma elevers lärande och fortsatta skrivutveckling dvs. hur elever med hjälp av lärares bedömningar kan komma vidare i läroprocessen. Vad få studier har undersökt är däremot hur lärares bedömningar av elevers skrivande används för att forma skrivundervisningen (se dock Berge et al., 2017; Berge & Skar, 2015; Evensen et al., 2016; Skar, 2013). Detta gör att vi vet relativt lite om hur lärare, utifrån bedömningar av elevers skrivande, fattar beslut om skrivundervisning som kan möta elevers lärandebehov i en viss didaktisk kontext (jfr Harlen, 2005; Wiliam & Thompson, 2007; Wiliam, 2011).

Vidare har svensk och övrig nordisk forskning om lärares summativa bedömningar av elevers skrivande undersökt bedömarvariationer och visat en genomgående låg bedömaröverensstämmelse (se t.ex. Aldrin, 2015; Borgström & Ledin, 2014; Ciolek Laerum, 2009; Fasting et al., 2009; Östlund-Stjärnegårdh, 1999). Dessa studier jämför enskilda lärares bedömningar av skrivuppgifter, ofta betygssättning av skrivprov, med andras och vad olika lärare grundar sina summativa bedömningar på. Däremot saknas forskning om hur lärare i grupp tillsammans bedömer elevers skrivande i relation till ämnens mål och kursers kunskapskrav där bedömningen utgår från mer än en text. Detta gör att kunskaperna är begränsade om hur lärare gemensamt formulerar och motiverar summativa bedömningar av elevers skrivande utifrån bedömningsinformation baserad på mer än enskilda skrivuppgifter och i förhållande till annat än enbart specifika uppgiftsmatriser. Samtidigt är det just lärares förmåga att formulera och kommunicera bedömningar av elevers skrivande under kursers gång som framhålls som något som kan bli avgörande för elevers likvärdiga möjligheter att utveckla sitt skrivande (Berge & Skar, 2015; Evensen, 2012; Parr & Timperley, 2010). När lärare tillsammans diskuterar hur styrdokument ska tolkas i relation till bedömningar av elevers prestationer förväntas de även utveckla en gemensam förståelse för vad

skrivbedömningar ska gälla vilket kan stödja arbetet för likvärdig och rätts-säker bedömning (Behizadeh & Engelhardt, 2011; Klenowski & Wyatt-Smith, 2010).

Slutligen har undersökningar av lärares gemensamma bedömningsprocesser identifierat lärares svårigheter att fatta konsensusbeslut om skrivbedömning utifrån fördjupade diskussioner och genom att använda bedömningsredskap på ett flexibelt sätt (se bl.a. Jølle, 2015a; Matre & Solheim, 2016). Dessa studier har bidragit med betydande kunskaper om hur sambedömning mellan lärare sker i de beslutsprocesser som lärare deltar i, men eftersom studierna är få och inriktade på norska lärares skrivbedömningar på mellan- och högstadiet vet vi ändå lite om hur sambedömning tar form mellan lärare i den svenska gymnasieskolan. För att kunna säga något om hur skrivbedömning sker i ett specifikt ämne i relation till en viss undervisningskontext behövs därför studier som undersöker hur svensklärare resonerar och kommer överens när de tillsammans bedömer elevers skrivande (Tengberg et al., 2017).

Sammanfattningsvis saknas alltså svensk forskning om skrivbedömning som undersöker vad som kännetecknar svensklärares beslutsprocesser när de tillsammans diskuterar skrivbedömning. Detta gör att det finns begränsade kunskaper om hur lärare interagerar i sådana samtal, vilka normer lärare ger uttryck för och vilka beslut de fattar vid olika typer av skrivbedömningar – *av* och *för* elevers lärande.

# Metodiska val och empiri

Detta kapitel inleds med en presentation av avhandlingens metodiska val för datainsamling och sammanlagda empiri. Därefter redogörs för urval av deltagare och textmaterial samt undersökningens genomförande, analysmetoder och analysprocesser. Slutligen diskuteras studiens tillförlitlighet, giltighet och generaliserbarhet i relation till studiens metoder och empiri.

## Metoder för datainsamling

### **Samtal i fokusgrupper**

För att få tillgång till en empiri som kunde ge svar på vad som karaktäriserar lärarnas skrivbedömningar som samtal och vilka normer och beslut lärarna gav uttryck för användes samtal i fokusgrupper.

Fokusgrupper är en metod för datainsamling som förespråkas för att studera hur människor i interaktion ”samfällt skapar mening i en företeelse” (Bryman, 2011, s. 449) och innebär att data samlas in genom gruppinteraktion om ett ämne som forskaren har bestämt (Morgan, 1996). Dessa samtal leds av en intervjuare vars uppgift är att förse deltagarna med samtalsämnen men som i övrigt inte aktivt deltar i samtalen. Fördelen med detta sätt är att forskaren i högre utsträckning kan komma åt föreställningar om vad deltagarna anser angeläget att samtala om (Wibeck, 2010). Tre fokusgrupper brukar anges som ett minimum vid samtal i fokusgrupper som datainsamlingsmetod och det optimala är att varje grupp består av minst fyra deltagare och helst inte fler än tio. På så sätt kan samtalen ge önskvärd bredd samtidigt som alla deltagare ges möjlighet att komma till tals (Krueger & Casey, 2014; Wibeck, 2010).

I denna studie fick tre svensklärargrupper med fem till sex lärare per grupp diskutera bedömning av elevers skrivande utifrån enskild läsning av autentiska elevtexter skrivna av för dem okända elever. De fick även tillsammans diskutera vilken skrivundervisning eleverna skulle behöva för att utveckla sitt skrivande och summativt bedöma elevers skrivförmåga i förhål-

lande till gymnasiekursen svenska 1. En närmare beskrivning av hur samtalen utformades följer senare i detta kapitel, under rubriken ”Genomförande och utformning av samtalen i fokusgrupper”.

Empirin för avhandlingens delstudier består huvudsakligen av ljud- och filmupptagningar från tre svensklärargrupperns samtal i fokusgrupper. Därutöver består empirin av ett frågeformulär som lärarna besvarade i anslutning till samtalen. I *Tabell 1* redovisas sammanlagd empiri från samtalen och hur lång tid varje lärargrupp ägnade åt respektive samtalsfråga. Tidsangivelserna är avrundade till hela minuter.

*Tabell 1* Översikt av empiri från lärargruppernas samtal i fokusgrupper

	Corneaskolan	Irisskolan	Scleraskolan	Sammanlagt
<i>Samtal om textbedömning</i>	39 min.	48 min.	44 min.	131 min.
<i>Samtal om undervisning</i>	12 min.	11 min.	12 min.	35 min.
<i>Samtal om betygsomdöme</i>	12 min.	13 min.	14 min.	39 min.
<i>Sammanlagt</i>	63 min.	72 min.	70 min.	205 min.

Notera: min.=minuter

I avhandlingens första artikel användes empiri från lärarnas *samtal om textbedömning* och lärarnas *samtal om undervisning* (ca 165 min.), i den andra artikeln användes empiri från lärarnas *samtal om textbedömning* och lärarnas *samtal om betygsomdöme* (ca 170 min.) och i den tredje artikeln användes lärarnas *samtal om betygsomdöme* (ca 40 min).

## Frågeformulär

I anslutning till lärargruppernas samtal i fokusgrupper fick lärarna enskilt besvara ett antal frågor i ett frågeformulär. Ett antal av dessa frågor syftade till att få tillgång till biografiska data och information om gymnasieskolorna. Dessa bakgrundsdata presenteras översiktligt under ”Urval av deltagare” nedan, och mer i detalj i *Bilaga C*. Lärarna fick även utvärdera de samtal i fokusgrupper som de just hade deltagit i och ombads också att enskilt bedöma elevernas skrivande i förhållande till kursen svenska 1 (*Bilaga D*).

Lärarnas utvärderingar av samtalen i fokusgrupper och deras bedömningar av elevernas skrivande används i studiens tredje artikel. I den ställs mått på lärarnas interaktionella samarbete och deras uttryck för samstämmighet i

relation till enskilda lärares uppfattningar om hur samtalen hade genomförts och vilket betygsomdöme de ansåg att elevernas skrivande motsvarade. På detta sätt kunde bland annat grad av samstämmighet inom lärargrupperna synliggöras.

Vidare innehöll frågeformuläret ett antal frågor om hur lärarna uppfattade hur arbetet med skrivbedömning organiserades på skolan, exempelvis avseende hur ofta och när lärarna ansåg att de träffades för att sambedöma elevers skrivande och samplanera skrivundervisning. Även dessa svar används i avhandlingens tredje artikel för att ge en bild av vilka möjligheter lärarna har getts att tillsammans öva att diskutera skrivbedömning på skolorna. I *Bilaga D* framgår vilka frågor lärarna fick besvara för att enskilt utvärdera samtalen i fokusgrupper och bedöma elevernas skrivande.

### **Forskningsetiska överväganden**

Avhandlingen har tagit hänsyn till Vetenskapsrådets krav (2017) på *information, samtycke, nyttjande* och *anonymitet*. De lärare som deltog i studien har muntligen och skriftligen informerats om att deltagande i studien är frivillig. De har även informerats om vad studien syftar till och vad den ska användas till. Förutom att informera om detta i det brev som skickades till ämnesansvarig med en förfrågan om att delta i ett forskningsprojekt (*Bilaga A*) var detta något som muntligen förtydligades i samband med att lärargrupperna samlades för samtal.

För att även möta kravet på anonymitet har lärarnas och skolornas namn anonymiserats. Detta gäller även de elevtexter som användes som fokusmaterial i samtalen. Däremot redovisas i artiklarna följande uppgifter om gymnasieskolorna: hur gymnasieskolorna drivs (kommunal regi eller friskola), geografiskt läge (innerstad eller närförort), vilka nationella program som ges och ungefär hur många elever skolorna har. Dessutom följande uppgifter om lärarna: ålder, kön, yrkeserfarenhet i antal år, tjänster (undervisningsämnen och eventuell karriärtjänst) samt anställningsår som gymnasielärare på aktuell gymnasieskola. Anledningen till att dessa uppgifter redovisas är att delar av studiens resultat diskuteras i relation till lärarnas bakgrund och de gymnasieskolor som de är anställda vid. Eftersom studien gjorts i en storstadsregion i Sverige innebär dock inte dessa uppgifter att det röjer vare sig vilka gymnasieskolorna är eller lärarnas identiteter.

## Urval av deltagare

Tre svensklärargrupper från tre gymnasieskolor ingår i studien. Lärargrupperna består av fem till sex lärare per skola varav samtliga lärare undervisade i svenska innevarande läsår (2014–2015) och hade erfarenhet av att undervisa i kursen svenska 1. Majoriteten av lärarna, femton av sjutton, kan sägas ha betydande erfarenhet av att undervisa i svenska i gymnasieskolan; sex av dem har undervisat i mer än fem år, nio i mer än femton år. Två av lärarna har mer begränsad erfarenhet och undervisade sitt första år.

I *Tabell 2* presenteras antal deltagande lärare från respektive skola och antal yrkesår som gymnasielärare (Blomqvist et al., 2016).

*Tabell 2* Lärare och yrkesår

	N	Yrkesår		
		Min	Max	Medel
Corneaskolan	6	13	31	19
Irisskolan	5	1	16	7
Scleraskolan	6	9	23	16

Notera: N=antal deltagare.

Som *Tabell 2* visar skiljde sig det genomsnittliga antalet yrkesår som gymnasielärare åt mellan lärargrupperna. På Corneaskolan var det genomsnittliga antalet yrkesår för lärarna 19 år med ett intervall på 13 till 31 års yrkeserfarenhet medan det genomsnittliga antalet yrkesår för lärarna på Irisskolan var 7 år med ett intervall på 1 till 16 år. På Scleraskolan var lärarnas genomsnittliga antal yrkesår 16 år med ett intervall på 9 till 23 år.

De tre gymnasieskolorna skiljde sig även åt avseende huvudman, geografiskt läge och profil: en av skolorna drivs i kommunal regi, två är fristående skolor; två av skolorna ligger i närförort, en i innerstaden; alla skolor erbjuder högskoleförberedande program, en av dem även yrkesförberedande program. Sammantaget motiveras urvalet gymnasieskolor och de svensklärare som är verksamma vid dem av att de representerar viss variation som kan säga något om hur lärare samtalar om skrivbedömning på dessa skolor (Cohen et al., 2011).

Hur gjordes då detta urval av lärare och skolor? I början av augusti 2014 kontaktades fem svensklärare på fem gymnasieskolor i en storstadsregion i Sverige med en ”Förfrågan om att delta i ett forskningsprojekt” (*Bilaga A*). Att denna förfrågan ställdes direkt till lärarna och inte till rektor eller annan skolledare motiveras av vikten av att hitta lärare och lärargrupper som var

positivt inställda till att delta i projektet. Majoriteten av de fem lärarna som kontaktades var ämnesansvariga i svenska på skolan och deras kontaktuppgifter fanns att finna på skolornas hemsidor.

Av dessa fem lärare återkom fyra redan inom två veckor med positiva besked. De meddelade att en grupp svensklärare på deras skola ville delta i projektet. Den femte läraren meddelade däremot att det inte fanns intresse av att delta i projektet. När tider för besök skulle bokas, avsikten var att besöka skolorna redan samma höst, visade det sig att en av de positivt inställda lärargrupperna inte kunde hitta tider för samtal under höstterminen eftersom det krockade med annat arbete på skolan. Detta gjorde att antalet deltagande svensklärargrupper i studien blev tre.

## Undersökningens genomförande

### **Insamlande och urval av stimulusmaterial till studien**

För att kunna genomföra studiens samtal i fokusgrupper behövdes ett antal elevtexter som kunde fungera som underlag för diskussionerna, ett så kallat *stimulusmaterial* (Wibeck, 2010). Under våren 2014 togs därför kontakt med några tidigare ämneskolleger i svenska som innevarande år undervisade i kursen svenska 1. En förfrågan om att bidra med elevtexter till studien skickades per e-brev till dem (*Bilaga B*).

I och med att svensklärarna på skolan tillsammans planerade kursens olika ämnesavsnitt och examinationer kunde min förfrågan vara specifik avseende vilka texter som behövdes. Förslaget var att de skulle bidra med hela elevers skrivportföljer innehållande en utredande text, en litterär essä, en argumenterande text och "något mer", t.ex. skönlitterära texter, metatexter, reflektioner, gensvar, processtexter eller annat. Min förfrågan ledde till att det någon månad senare, i juni 2014, fanns ett material bestående av sju elevers hela skrivportföljer från kursen svenska 1 att tillgå.

Att använda sju elevers hela skrivportföljer från kursen svenska 1 var utifrån det planerade upplägget med studien, samtal i fokusgruppsessioner à 120 minuter, inte möjligt. Inom denna tidsrymd skulle lärarna dessutom hinna läsa elevtexterna enskilt. Därför behövde ett texturval göras vilket gjordes i två steg: först utifrån frågan "Vilka elevers texter skulle ingå i studien?" och därefter "Vilka av dessa elevers texter och vilka texttyper skulle ingå i studien?".

### *Vilka elevers texter skulle ingå i studien?*

För att göra ett urval av vilka elevers texter som skulle ingå i studien genomfördes kortare intervjuer med de tre lärare som hade lämnat in skrivportföljer. Förhoppningen var att lärarna i dessa intervjuer kunde ge en tydligare bild av elevernas skrivande och därigenom underlätta urvalsprocessen.

Intentionen var att stimulusmaterialet skulle bestå av elevtexter av varierande kvalitet som visade på olika aspekter av skrivförmåga och att lärarna därför skulle utmanas att diskutera bedömning ur flera olika perspektiv. Tidigare studier på området (Adie et al., 2012; Falk & Ort, 1998) har visat att just en betydande bredd av elevers prestationer har större chans att leda till att lärares olika uppfattningar kommer till uttryck vid sambedömning. Eleverna fick därför gärna ha presterat lite ojämnt under kursen genom att visa upp både brister och fördelar i sina texter. Samtidigt eftersträvade jag viss men inte alltför tydlig spridning av olika elevers skrivförmåga. Under intervjuerna med de tre lärarna ställdes följande frågor:

Vad uppfattar du utmärkte den här elevens skrivande?

Varför är den här elevens texter intressanta att diskutera?

Vilka mål och kunskapskrav bedömer du att elevens skrivande motsvarar?

Utifrån lärarnas svar föll valet på tre elevers skrivande som enligt läraren uppvisade både bredd och variation i sitt skrivande.

Här följer utdrag ur lärarnas beskrivningar av de tre eleverna vars texter ingår i studien. I citaten har pronomen som ”han” och ”hon” samt och ”hans” och ”hennes” bytts ut mot ”eleven” och ”elevens”.

Den första eleven beskrevs av sin lärare som ”mycket ojämn i sitt skrivande”:

*Språkligt sett finns det många brister. Det gör texterna svårbedömda. Eleven har fattat uppgifterna, har förstått idéerna, men får inte riktigt ihop det. Den bästa texten är essän som är betydligt bättre än den utredande texten.*

Denna elev fick E i kursbetyg och bedömdes ha en skrivkompetens motsvarande just E.

Den andra eleven beskrevs av sin lärare som ”en stark skribent” och kan på många sätt sägas vara en direkt motsats till den tidigare eleven. Läraren uttryckte det som att:



*Eleven är inte beroende av min undervisning. Texterna blir bra ändå. Eleven har en självsäkerhet och är heller inte mottaglig i klassrummet för undervisning, är självständig, har en jämn hög nivå, är intressant ur bedömningssynvinkel för att eleven är hemlig med vad den tänker i klassrummet men från en text till en annan så kan eleven ändå ha tagit till sig undervisningen. God läsförmåga, förstår uppgift och vet vad som ska göras. Eleven formar en egen idé vad den vill skriva, självständig. God textstruktur.*

Denna elev fick B i kursbetyg och bedömdes ha en skrivkompetens motsvarande A-C.

Den tredje eleven beskrevs av sin lärare som en elev som ”växte mycket under kursens gång” men som ofta lämnade in sina texter ”i sista sekund”. Detta gjorde att eleven beskrevs som ”svårundervisad”:

*Eleven arbetade väldigt ryckvis kan man säga, det hände ingenting, ingenting, ingenting... dålig närvaro, frånvarande, men sen när eleven väl satt i gång med att skriva då hände saker. När deadline närmade sig, när det var vid slutet av en skrivprocess skrev eleven dygnet runt och skickade flera utkast på kort tid. Eleven fick då lärarrespons av mig. Detta gör att jag kunde följa elevens skrivprocess och behövde inte misstänka fusk. Däremot blev elevens skrivande lidande av att allt skrevs på slutet. Strukturen i texterna hade t.ex. varit mycket bättre om eleven hade skrivit under hela processen. Det är jag övertygad om.*

Denna elev fick A i kursbetyg och bedömdes av sin lärare ha en skrivkompetens motsvarande just A.

### ***Hur många och vilka elevtexter skulle ingå i studien?***

Centralt för studiens utformning var att lärarnas bedömningar av elevers skrivande skulle avse mer än bedömning av enskilda elevtexter. Anledningen till detta var att lärarnas skrivbedömningar även skulle motsvara de bedömningar lärare gör i förhållande till en specifik kurs. Elevtexterna i stimulusmaterialet behövde dessutom vara jämförbara dvs. bygga på samma skrivuppgift och skrivas under liknande förutsättningar. Detta för att lärargrupperna skulle kunna bedöma texterna och elevernas skrivande på ett likvärdigt sätt.

Utifrån studien begränsade tidsomfång var det inte möjligt att använda mer än två texter av varje elev. Vilka texter det skulle bli avgjordes av att lärarna som bidragit med textmaterialet genomgående hade lyft fram två texter i elevernas skrivportföljer som hade haft extra betydelse för deras bedömningar av elevernas skrivande i kursen. Dessa två texter var den utredande texten och den litterära essän, texter som även kunde motiveras utifrån

såväl ämnesplanens mål och centrala innehåll som kursens kunskapskrav avseende skrivande i svenska 1:

#### Betyget E

Eleven kan skriva argumenterande text och andra typer av texter, som är sammanhängande och begripliga samt till viss del anpassade till syfte, mottagare och kommunikationssituation. Eleven kan i huvudsak följa skriftspråkets normer för språkriktighet.

(Skolverket, 2017, s. 9)

Eftersom argumenterande text är den enda texttyp som specifikt anges i både centralt innehåll och kunskapskrav för svenska 1 ingår studiens textmaterial, den utredande texten och den litterära essän i vad som anges som ”andra typer av texter” (ibid.).

Sammanlagt utgjorde textmaterialet sex texter skrivna av tre elever: en utredande text och en litterär essä från varje elev. Texterna avidentifierades och försågs med enhetligt teckensnitt, textstorlek och radavstånd. Inför användning märktes även texterna som ”Text 1”, ”Text 2” o.s.v. till ”Text 6” för att underlätta för lärarna att referera till dem under samtals gång.

### **Genomförande och utformning av samtalen i fokusgrupper**

Under höstterminen 2014 besöktes de tre lärargrupperna på respektive gymnasieskola där de undervisade. Samtalen genomfördes under lärarnas ordinarie mötestid och inleddes med en kortare presentationsrunda där lärarna introducerades för studiens upplägg och genomförande: Lärarna fick veta att de enskilt skulle få läsa ett antal elevtexter (stimulusmaterialet) för att sedan diskutera dem i grupp utifrån specifika frågor som de skulle förses med allt eftersom. Elevernas texter introducerades sedan i två omgångar med början i de tre utredande texterna, vilket motsvarade den ordning i vilken eleverna hade skrivit texterna under kursen. Innan lärarna fick läsa texterna enskilt fick de även information om hur uppgiftsinstruktionerna för skrivuppgifterna hade formulerats avseende ämne för texten och formella krav samt hur skrivprocessen hade sett ut. Detta för att samtliga lärare skulle känna till skrivsituationens kontext.

Lärarna fick sedan tid att enskilt läsa elevtexterna: de tre utredande texter om språklig variation och tre litterära essäer om existentiella frågor. Om de ville fick de gärna göra markeringar eller kommentarer direkt i elevtexterna för att enklare kunna återvända till dem under samtals gång. De skulle

meddela när de ansåg att de var färdiga med läsningen vilket de gjorde efter ca 20 minuter. Samtalen i fokusgrupp ramades sedan in av följande fyra frågor:

Vad kan ni säga om den här elevens skrivande?

*Frågan ställdes till varje enskild elevtext dvs. sex gånger under varje lärargrups samtal*

Hur skulle ni rangordna de här texterna?

*Frågan ställdes efter samtal om tre elevtexter av samma texttyp dvs. två gånger under varje lärargrups samtal*

Vad behöver eleverna undervisning i för att utveckla sitt skrivande?

*Frågan ställdes efter samtal om tre elevtexter av samma texttyp dvs. två gånger under varje lärargrups samtal*

Vilket betygsomdöme motsvarar den här elevens skrivande?

*Frågan ställdes efter samtal om alla tre elevers två texter dvs. en gång under varje lärargrups samtal*

Den första frågan ”Vad kan ni säga om den här elevens skrivande?” innebär att lärarna diskuterade kvaliteter i enskilda elevtexter medan den andra frågan ”Hur skulle ni rangordna de här texterna?” innebär att de bedömde elevtexterna i förhållande till varandra. Den tredje frågan ”Vad behöver eleverna undervisning i för att utveckla sitt skrivande?” är didaktiskt inriktad och innebär att lärarna diskuterade lämplig undervisning för att eleverna ska ges möjlighet att förbättra sitt skrivande och den fjärde frågan ”Vilket betygsomdöme motsvarar den här elevens skrivande” innebär att elevernas skrivförmåga ska bedömas i relation till kriterier.

Lärarna hade i dessa samtal inte tillgång till något av de styrdokument som reglerar bedömningens inriktning och genomförande i gymnasieskolan. Fördelen med ett sådant tillvägagångssätt är att samtalen inte struktureras utifrån bedömningsredskapen utan ger plats åt lärarna att uttrycka vad de anser vara relevant att diskutera och bedöma i den givna kontexten (jfr Matre & Solheim, 2016; Smaill, 2013). Detta ökar möjligheterna att få syn på vilka bedömningsnormer som kan tänkas vara internaliserade i lärarnas bedömningar. Samtidigt, med tillgång till styrdokumentet, skulle lärarna haft möjlighet att i samtalen hänvisa till dem mer explicit.

I diskussionen om vilket betygsomdöme elevernas skrivande motsvarar avstod en lärare på Scleraskolan från att delta. Läraren motiverade det med att osäkerheten om kunskapskraven för svenska 1 var för stor.

## Analysmetoder

De analysmetoder som har använts för att analysera lärarnas samtal i fokusgrupper är topikanalys och initiativ-respons-analys (Linell, 2011; Linell & Gustavsson, 1987).

### Topikanalys

Topikanalys bygger på dialogisk teori och inriktas på att belysa vad grupper uttrycker som relevant och viktigt att samtala *om* när de tillsammans löser en uppgift (*kommunikativt projekt*, Linell, 2001; 2009). Topiker definieras som "units of meaning rather than of expressions" (Linell, 2001, s. 233) och kan därför återspegla vad individer i en grupp dialogiskt uttrycker för uppfattningar om ett ämne. I denna studie innebär det att när en lärargrupp exempelvis diskuterar en elevtexts användning av källor identifieras inte topiken enbart utifrån själva ämnet, dvs. textens källanvändning, utan det lärarna gemensamt uttrycker *om* källanvändningen, som att källanvändningen fungerar väl i sitt sammanhang eller att källanvändningen behöver förtydligas. Eftersom det krävs samarbete mellan minst två deltagare i ett samtal för att en topik ska anses dialogiskt etablerad i ett samtal (jfr *minimal kommunikativ interaktion*, Linell, 2001) kan analystekniken ge en bild av vilka uppfattningar och normer som lärarna i samtal om skrivbedömning delar, inte enskilda lärares uppfattningar och normer.

I avhandlingens första artikel används topikanalys för att beskriva och analysera lärarnas uttryck för förväntningsnormer vid bedömning av enskilda elevtexter och ställa dem i relation till deras uttryck för didaktiska beslut. I studiens andra artikel används topikanalys för att undersöka lärarnas beslut vid summativ bedömning och vilka normer som ges prioritet vid betygssättning av enskilda elevers skrivande.

### Initiativ- och responsanalys

Även initiativ-respons-analysen bygger på dialogisk teori men inriktas på hur lärarna interagerar genom att relatera till varandras yttranden (Linell, 2011; Linell & Gustavsson, 1987).

Analysen innebär att yttranden poängsätts utifrån grad av initiativ och respons. Principen för denna IR-kodning är att turer som för in ett nytt ämne i samtalet och explicit begär respons av övriga deltagare, ett så kallat *fritt, begärande initiativ* (Linell, 2011), får högst poäng (värde 6) medan turer som enbart responderar genom att i korthet uppmärksamma en tur, en så

kallad *minimal adekvat respons* (ibid.) får lägst poäng (värde 1). Däremellan finns olika grader av interaktionell styrka i enskilda turer avseende kombinationen av initiativ och respons (värde 2–5). På så sätt ger IR-kodning kvantitativa mått på interaktion i ett samtal. Exempelvis synliggörs lärarnas individuella IR-index, dvs. summan av enskild lärares poängsatta turer delat med antal yttranden från samme lärare, olika lärares interaktionella bidrag och hur den interaktionella dominansen fördelar sig mellan lärarna i diskussionen (IR-differens). Vidare belyser ett IR-index på gruppnivå den genomsnittliga interaktionella dynamiken i samtalet där ett högt IR-index innebär att samtalet byggs av flera högt poängsatta repliker dvs. initiativ som initierar nya ämnen och explicit begär respons från övriga deltagare (Linell & Gustavsson, 1987). I *Bilaga F* finns en förteckning över hur yttranden har kodats och poängsatts.

I avhandlingens tredje artikel används IR-kodning för att med hjälp av interaktionella mått belysa dominans, dynamik och koherens i lärargruppernas bedömningssamtal. Dessa mått ställs sedan i relation till lärarnas uttryck för samstämmighet i samtalen.

## Analysprocess

### **Transkription av materialet**

Det analytiska arbetet inleddes med att det empiriska materialet från lärargruppernas bedömningssamtal transkriberades i sin helhet. I avhandlingens två första artiklar inriktades analysen på den verbala kommunikationen i bedömningssamtalen och dessa transkriptioner är därför en viss förenkling av de konventioner som annars kännetecknar samtalsanalytisk forskning, avseende t.ex. angivelser av pauslängd, taltempo och icke-verbala handlingar (jfr Anward & Nordberg, 2005). Filmupptagningarna från samtalen användes här enbart för att med högre grad av säkerhet urskilja vilken deltagare som yttrade vad i de ibland intensiva samtalen med snabba replikskiften.

Inför avhandlingens tredje artikel förfinades dessa transkriptioner genom att ljud- och videoupptagningar från de tre lärargruppernas bedömningssamtal kompletterades med mer noggranna angivelser av t.ex. pausers längd, lärares intonation och gester. Detta för att möjliggöra en analys av initiativ och respons i samtalen. I *Bilaga E* finns en förteckning över de transkriptionstecken som har använts.

## Analys av topiker

Betydande delar av analysarbetet har inneburit en kollektiv process där materialet i olika omgångar har lästs, diskuterats och analyserats tillsammans med forskarkolleger (jfr *validering med forskarsamhället*, Kvale & Brinkmann, 2009). Till att börja med skedde detta genom att delar av transkriptionerna från bedömningsamtalen lästes och diskuterades utifrån frågan ”Vad lägger ni märke till i dessa samtal?”. Dessa samtal gav upphov till några första tentativa resultat om vad som kännetecknade lärarnas samtal om skrivbedömning och vilka uppfattningar, normer och beslut lärarna gav uttryck för i olika faser av bedömningsprocessen.

Topikanalysen innebar sedan en mer grundlig bearbetning av empirin. Utifrån transkriptionerna från samtalen identifierades samtliga topiker i de tre lärargruppernas samtal och definierades som *kvalitetsuppfattningar* om elevers texter, *didaktiska beslut* om skrivundervisningens inriktning eller *summativa bedömningar* av elevers skrivförmåga i relation till svenska 1. För att illustrera hur topikanalysen har gått till följer här ett utdrag ur Corneaskolans samtal om summativ bedömning av elevers skrivande utifrån frågan ”Vilket betygsomdöme motsvarar elevens skrivande?” (Blomqvist, 2018a).

125 Cia: den här tycker jag vi sätter ett E på

126 Carina: nä

(.) lite högre tror jag nog

127 Cia: ett plus? ett C?

128 Carina: nä inte ur dispositionssynpunkt så e det ju inte det

(.) men ett E plus tycker jag utan tvekan

(.) vad det nu är för skillnad på det?

(.) men inte riktigt C-mässig än

129 Cia: nä det kan jag hålla med om

(.) nä det tycker inte jag heller

(.) E då

Cia föreslår inledningsvis betygsomdömet E (tur 125), vilket Carina invänder mot ”nä, lite högre tror jag nog” (tur 126). Cia föreslår då ”ett plus? ett C?” (tur 127) men Carina menar att betyget C inte är möjligt ”ur dispositionssynpunkt” (tur 128) men ett ”E plus”. De båda uttrycker enighet om att elevens skrivförmåga motsvarar ett E för att den är ”inte riktigt C-mässig än” (Carina, tur 128), vilket Cia håller med om (tur 129). Den topik som här etableras i fem turer mellan två lärare är en kvalitetsuppfattning som uttrycker att vissa textkvaliteter i elevens skrivande brister för ett högre betyg än E

vilket implicerar en bedömningsnorm om att samtliga kvaliteter i elevers texter ska motsvara samma kunskapskrav för ett visst betygsomdöme.

Genom att på detta sätt identifiera topiker i lärargruppernas bedömnings-samtal framträdde tematiska mönster i vad lärarna riktade bedömningarna mot (jfr Marková et al., 2007). För att synliggöra mönster av topiker som uttryckte liknande kvalitetsuppfattningar sammanfördes dessa i kategorier. Denna del av analysen innebar både en strukturering och reducering av materialet (Kvale & Brinkmann, 2009) men var nödvändigt för att kunna synliggöra vad lärarna uppmärksammade och uttryckte för kvalitetsuppfattningar och normer i olika faser av bedömningsprocessen. Utöver det kunde antalet topiker som uttryckte kvalitetsuppfattningar inom vardera kategorin synliggöra vad lärarna framför allt grundade bedömningarna på.

En svårighet i analysarbetet bestod i att lärarna sällan diskuterade ett ämne i taget utan uppehöll sig vid flera aspekter av bedömning samtidigt. Exempelvis kunde en lärare tidigt i ett samtal introducera ett ämne som först blev responderat på långt senare. Detta är till viss del en följd av studiens upplägg med samtal i fokusgrupper där lärarna efter att ha fått en bedömningsuppgift lämnades ensamma att lösa den (Wibeck, 2010). Samtidigt är det ett resultat av den dynamik som kännetecknar de flesta diskussioner där uppfattningar och åsikter växer fram i ett samtal allt eftersom (Palmér, 2008). För analysens resultat har detta inte haft någon betydelse; så länge ämnet har behandlats av minst två lärare i minst tre turer under samtalets gång har det identifierats som topik. Initiativ som inte blev responderade ingår däremot inte i studiens resultat.

Topikanalysen låg sedan till grund för att ställa lärarnas uppfattningar och normer i olika faser av en skrivbedömningsprocess i relation till varandra (jfr Phelps, 2000; Harlen, 2012). Först ställdes lärarnas uttryck för förväntningsnormer i relation till deras didaktiska beslut om skrivundervisning (artikel 1), därefter ställdes förväntningsnormer i relation till lärarnas beslut om summativ bedömning av elevers skrivande med fokus på vilka normer som gavs prioritet vid bedömning av enskilda elevers skrivande (artikel 2). På så sätt har analysen av topiker synliggjort i vilken utsträckning lärarna använde egen formulerad bedömningsinformation om enskilda elevtexters kvaliteter för didaktiska beslut om skrivundervisning och summativ bedömning av skrivförmåga.

## Analys av initiativ och respons

Initiativ-respons-analysen innebar att transkriptionerna från de tre lärargruppernas samtal om summativ bedömning av elevernas skrivande IR-kodades (Linell, 2011; Linell & Gustavsson, 1987). För att illustrera hur IR-kodningen gick till och vad den synliggjorde i lärargruppernas samtal följer här ett utdrag från Scleraskolans samtal om summativ bedömning av elevers skrivande (Blomqvist, 2018b):

- 51 > Saga: skulle vi tänka annorlunda om dom kom i annan ordning?  
(.) att först skrev eleven att kärleken är dum –
- 52 <^ Signe: ja för då skulle jag känna att oj vad den här eleven är på väg språkligt asså  
(.) det e ju –  
(.) jaa
- 53 <=> Saga: skulle man tänka mer mot  
(.) att titta här har hon ju faktiskt börjat nå  
(.) eller han
- b* (mmm x 3)  
(.) då har den ju tagit till sig om min kommentar om det där

Saga inleder med att initiera ett nytt ämne för diskussionen och explicit begära respons från de övriga lärarna genom frågan: ”skulle vi tänka annorlunda om dom kom i annan ordning?” (tur 51). Signe responderar genom att acceptera och till viss del även utveckla Sagas initiativ men utan att själv explicit begära respons ”ja för då skulle jag känna att oj vad den här eleven är på väg språkligt asså (.) det e ju – (.) jaa” (tur 52) varpå Saga knyter an till egen tur och följer upp den med ”skulle man tänka mer mot (.) att titta här har hon ju faktiskt börjat nå /---/” (tur 53) vilket backas upp av flera lärare (*b*). Saga behåller på detta sätt initiativet i samtalet och ger därigenom uttryck för att interaktionellt styra samtalet.

Genom IR-kodning där lärarnas enskilda turer poängsattes från 1 till 6 efter grad av initiativ och respons (*Bilaga F*) kunde såväl enskilda lärares bidrag i samtalen som interaktionella mönster i lärargruppernas samtal beskrivas med mått på dominans, dynamik och koherens. Lärarnas dialogiska samarbete och deras uttryck för samstämmighet ställdes även i relation till enskilda lärares uppfattningar om hur samtalen genomfördes och deras betygsättning av elevernas skrivande.

Även denna del av analysarbetet har i omgångar diskuterats och problematiserats tillsammans med forskarkolleger och inneburit flera överväganden och ställningstaganden. Enklast att koda var lärares uttryck för acceptans



(<, 2 poäng) medan det ibland var svårt att avgöra den interaktionella styrkan i enskilda turer avseende om turen explicit begärde respons från övriga lärare eller inte. Exempelvis kan ett yttrande som ”men källförteckningen” som anknyter till huvudinnehållet för diskussionen i tidigare tur kodas som aningen en tur som ”genom en fråga eller begäran vill ha respons” (<>, 4 poäng) eller en tur som inte begär respons (<^, 3 poäng). Här har prosodin fått avgöra där ”men källförteckningen?” kodas som en fråga, dvs. ett begärande initiativ, medan ”men källförteckningen –” kodas som ett konstaterande, dvs. ett icke-begärande initiativ. Att avgöra detta har inneburit flera omlýssningar av materialet vilket i sin tur har gjort transkriptionerna mer och mer detaljerade.

För att testa reliabiliteten i IR-kodningen medkodade en expert på området drygt 10% av de sammanlagt 431 replikerna från lärargruppernas samtal om summativ bedömning (Blomqvist, 2018b). Reliabiliteten i kodningen uppgick till 80 %. Linell och Gustavsson (1987) menar att en överensstämmelse på 75 % vid denna typ av IR-kodning av dialoger får anses vara godkänd. Framför allt eftersom kategorierna i detta IR-system är flera och skillnaderna mellan dem små (*Bilaga F*). Dessutom är variationer i IR-kodningen ett resultat av ”den inneboende otydlighet, vaghet och flertydighet som vidlåder replikernas interaktionella status” (ibid., s. 205). Det innebär att till nästan varje replik är två eller fler alternativa kodningar möjliga. Att analyser kan leda till delvis olika alternativ innebär däremot inte att resultaten behöver påverkas på ett genomgripande sätt.

## Giltighet, tillförlitlighet och generaliserbarhet

Frågan om en studies giltighet har att göra med om studien verkligen undersöker det den säger sig undersöka, eller mäter det den säger sig mäta (*intern validitet*, Bryman, 2011). Till stor del avgörs denna giltighet av om metoder för datainsamling och analys står i relation till studiens syfte och frågeställningar.

I studien används fokusgrupper för att undersöka lärares samtal om skrivbedömning. En fördel med fokusgruppsamtal är att de kan synliggöra hur individers olika uppfattningar tas emot av resten av gruppen (Wibeck, 2010) vilket ökar möjligheten att få syn på uppfattningar som lärarna på en skola ger uttryck för att dela, något som i sin tur kan ses som uttryck för rådande normer vid dessa gymnasieskolor. Samtidigt kan fokusgrupper beroende på gruppens sammansättning ha en negativ inverkan på individers

deltagande och vara begränsande (Marková et al., 2007). I mer homogena grupper där deltagarna känner varandra väl finns även en risk att yttranden lämnas utsagda eftersom de anses vara underförstådda och för triviala att behandla (Myers, 2004).

På ett sätt kan svensklärarna i denna studie sägas vara en homogen grupp eftersom de undervisar i samma ämne vid samma gymnasieskola men det är enbart vid en av skolorna (Irisskolan) som lärarna träffas mer regelbundet för att tillsammans bedöma elevers skrivande och planera undervisningen. Detta faktum, tillsammans med den systematiska inramningen av studiens fokus-samtal, kan förhoppningsvis bidra till att lärarna inte avfärdar uppfattningar som självklara och triviala utan uttrycker sådant som de anser vara viktigt och nödvändigt att få sagt. En fördel med en mer homogen grupp, som i det här fallen innebär en grupp av svenskämneskolleger vid samma gymnasieskola, är att samtalen kan ge inblick i redan befintliga samtalsmönster på dessa skolor och att deltagarna har gemensamma erfarenheter som de kan förtydliga och förankra sina resonemang med hjälp av (Kitzinger, 1994; Wibeck, 2010).

Trots att samtalen i fokusgrupp ägde rum på respektive gymnasieskola kan de inte ses som opåverkade från yttre faktorer. En forskares närvaro och det faktum att samtal spelas in har alltid viss inverkan på de diskussioner som förs (Speer och Hutchby, 2003). För att minska denna påverkan genomfördes fokussamtalen som semistrukturerade intervjuer där lärarna fritt behandlade de frågor som strukturerade samtalen (Bryman, 2011). Lärarna satt även vid ett eget bord under diskussionerna och de avslutade själva samtalen när de uttryckte att de hade fullgjort uppgiften. Den tekniska apparaturen för ljudinspelning var placerad på bordet där lärarna satt och videokameran några meter bort. Lärarna gav under samtalen varken uttryck för att prata *om* eller *till* inspelningsapparaturen utan vände sig mot varandra.

De analysmetoder som används i studien inriktas på att belysa *vad* lärarna diskuterar och *hur* de gör det. Med topikanalys synliggörs lärarnas uppfattningar och normer och initiativ-respons-analys belyser själva interaktionen i samtalen. En begränsning med topikanalys är att analysen enbart tar fasta på yttranden som blir bemötta verbalt av andra samtalsdeltagare. Det finns alltså en risk att uppfattningar och normer som lärarna egentligen delar inte kommer med i analysen på grund av att lärare väljer att inte uttrycka detta verbalt (jfr tidigare diskussion om homogena grupper). I de fallen skapas ingen topik och dessa icke-topikaliserade initiativ ingår inte i analysen (Linell, 2001). Vidare är det en begränsning med initiativ-respons-analysen att den enbart tar hänsyn till yttrandets grad av just initiativ och respons, inte till

exempelvis hur långa yttrandena är. En talmängdsanalys hade gett fler mått på interaktionen i lärargruppernas samtal samtidigt som även den kan ge en missvisande bild av vilka som dominerar ett samtal eftersom den som pratar mest inte behöver ha störst påverkan på gruppens samtal och de beslut gruppen fattar. Samtidigt kan även enskilda lärares IR-index bli missvisande eftersom det anger ett medelvärde av interaktionell styrka i de yttranden som avges. Det innebär att en lärare som bidrar med ett mycket begränsat antal repliker ändå kan få ett högt IR-index förutsatt att replikerna poängsätts högt. Detta var dock inte fallet i denna studies samtal. En annan begränsning i studiens analys är att den IR-kodning som tillämpas (Linell, 2011) är utformad för analys av dyader, dvs. analys av transkriptioner mellan två personer. I sådana samtal riktas alltid ett yttrande till enbart en individ. Detta medför bland annat att transkriptionerna av lärarnas samtal kan verka mer fragmentariska än vad samtalen var. Exempelvis kan en lärare med ett initiativ rikta sig till flera lärare samtidigt, vilket kan leda till tre liknande responser på ett och samma yttrande, t.ex. tre balanserande turer (<^).

Frågan om en studies tillförlitlighet har till stor del att göra med att konsistensen är hög och att studien innehåller utförliga redogörelser av alla de val som ingår i forskningsprocessen. Syftet med att studien är transparent är att läsare och andra forskare ska kunna uppskatta rimligheten i de resultat som presenteras. Dessutom ska det framgå hur forskaren själv kan ha påverkat undersökningens resultat och hur en sådan påverkan har minimerats. Dessa krav på *intern reliabilitet* (Bryman, 2011) har inneburit att arbetet omgetts av en kontinuerlig och kritisk granskning av forskarkolleger. Studiens syften, forskningsfrågor, teorier, metoder och resultat har vid flera tillfällen presenterats och diskuterats i olika typer av seminarier och vid konferenspresentationer. Detta innebär att flera forskarkolleger från olika högskolor och universitet har haft insyn i forskningsprocessen under längre tid och har haft möjlighet att ge synpunkter på den. Exempelvis har analyserna av topiker och kategorisering av dem till stora delar gjorts tillsammans med forskarkolleger, liksom analysen av initiativ och respons som även har medkodats av en expert på området. Detta har gjort att studiens material, transkriptionerna från ljudupptagningarna av de tre lärargruppernas bedömningssamtal, vid återkommande tillfällen diskuterats och analyserats i dialog med forskarkolleger. Utifrån dessa diskussioner och granskningar har de enskilda artiklarna försetts med redogörelser av hur analysprocesserna har gått till och vilka svårigheter de inneburit.

Ett annat giltighetskriterium har med en studies överförbarhet att göra, dvs. om det går att jämföra resultaten i en studie med annan miljö (*extern*

*validitet*, Bryman, 2011). Eftersom studien bygger på ett begränsat material som även är begränsat till sin kontext finns ingen möjlighet att dra generella slutsatser om svensklärares skrivbedömning. Däremot kan studien bidra till viss överföring utifall de som läser studien känner igen sig i de exempel som ges. Då kan resultatet få relevans och giltighet genom att det exempelvis används som diskussionsunderlag för att diskutera eller utvärdera svensklärares skrivbedömningar över lag. Detta gäller även de metoder som har använts för att de skrivbedömningar lärare gör i didaktiska och dialogiska perspektiv.

# Artiklarna i sammanfattning

Denna avhandling bygger på tre vetenskapliga artiklar. Teoretiskt och metodiskt hör dessa ihop genom att den huvudsakliga empirin, de tre lärargruppernas samtal i fokusgrupper om skrivbedömning, analyseras och diskuteras med hjälp av begrepp och modeller från didaktiska och dialogiska teorier.

Här följer kortare sammanfattningar av artiklarna. Kapitlet är framför allt till för att underlätta läsningen av avhandlingen som helhet genom att redovisa vad de olika artiklarna behandlar, vilka forskningsfrågor de svarar mot, vilket material artiklarna bygger på, vilka metoder som har använts och vilka resultaten är. Artiklarna återfinns i sin helhet i slutet av avhandlingen.

## Artikel 1. Lärares förväntningsnormer och didaktiska beslut

Studien undersökte svensklärares uttryck för kvalitetsuppfattningar och förväntningsnormer vid bedömning av elevers texter samt hur bedömningsinformation om elevernas skrivande användes i lärarnas didaktiska beslut om skrivundervisning. Studien svarade mot följande två forskningsfrågor: ”Vilka kvalitetsuppfattningar och förväntningsnormer uttrycker lärarna i samtal om elevers skrivande?” och ”Vilka didaktiska beslut fattar lärarna utifrån bedömningen av elevernas skrivande?”.

Empirin byggde på lärarnas diskussioner i fokusgrupper utifrån frågorna: ”Vad kan ni säga om den här elevens skrivande?” och ”Vilken undervisning skulle de här eleverna behöva för att utveckla sitt skrivande?”. Samtalen analyserades med topikanalys (Linell, 2001) vilket gjorde det möjligt att belysa lärarnas uttryck för delade uppfattningar om elevernas texter och deras beslut om skrivundervisning. Dessa kvalitetsuppfattningar och beslut implicerade normer om hur elever förväntas kunna skriva och hur de ska lära sig det i relation till gymnasiekursen svenska 1.

Det teoretiska perspektiv som lades på dessa samtal var didaktiskt och inriktades på att undersöka hur lärare formulerade bedömningsinformation om elevers skrivande och hur de sedan använde den för beslut om skrivunder-

visningens innehåll. För att synliggöra relationen mellan lärarnas uttryck för förväntningsnormer och deras didaktiska beslut användes Phelps (2000) modell för pedagogisk textrespons där lärares läsning av elevers texter ligger till grund för bedömning, respons och didaktisk handling.

Resultatet visade att lärarnas bedömningar av elevernas skrivande behandlade fler och delvis andra aspekter av skrivförmåga än vad deras didaktiska beslut sedan omfattade. De kvalitetsuppfattningar och förväntningsnormer som lärarna uttryckte i sina bedömningar av elevernas texter inriktades huvudsakligen på texters kommunikativa kvaliteter, den stilistiska utformningen och textupbyggnaden, följt av ämnesinnehåll och källanvändning. De didaktiska beslut som fattades utifrån bedömningen av elevernas texter inriktas däremot i stort sett enbart på att eleverna behövde skrivundervisning i att disponera text och att använda källor. Detta ledde till slutsatsen att överensstämmelsen var låg mellan lärarnas beslut om skrivundervisningens inriktning, dvs. vad eleverna behöver undervisning i för att utveckla sitt skrivande, och deras förväntningar på vad eleverna ska kunna. De didaktiska besluten omfattade därmed i begränsad utsträckning den bedömningsinformation som lärarna själva hade formulerat.

## Artikel 2. Lärares normer för beslut om summativ bedömning

Studien undersökte svensklärarnas uttryck för normer vid summativ bedömning av elevers skrivande och vilka normer som gavs prioritet när enskilda elevers skrivande bedömdes i förhållande till kursen svenska 1. Den forskningsfråga som studien besvarade är: "Vilka normer för beslut om summativ bedömning uttrycker lärarna och hur ges dessa normer prioritet vid bedömning av enskilda elevers skrivförmåga?".

Empirin byggde på var lärarnas diskussioner i fokusgrupper utifrån frågorna: "Vad kan ni säga om den här elevens skrivande?", "Hur skulle ni rangordna de här elevernas skrivande?" och "Vilket betygsomdöme motsvarar den här elevens skrivande?". Även dessa samtal analyserades med topikanalys (Linell, 2001), vilket gjorde det möjligt att belysa lärarnas uttryck för kvalitetsuppfattningar om elevers skrivande vilka implicerar normer för beslut om summativ bedömning av elevers skrivande i relation till gymnasiekursen svenska 1.

För att beskriva och analysera lärarnas normer i olika faser av en bedömningsprocess användes Harlens (2012) *ramverk för summativ bedömning*.

Faserna beskrivs som att lärare från ett urval elevuppgifter/-aktiviteter ("collection of evidence"), tolkar materialet ("interpretation of evidence"), bedömer elevers prestationer i förhållande till kriterier av något slag ("judgement of achievement, criterion-referenced") och rapporterar någon form av omdöme ("report on achievement"). Även här användes topikanalys (Linell, 2001) för att beskriva och analysera vilka uppfattningar och normer lärarna gav uttryck för i dessa samtal men nu avseende deras summativa bedömningar.

Resultatet visade att lärarna framför allt uppmärksammade texters kvaliteter i relation till kunskapskravens betygssteg och uttryckte att samtliga textkvaliteter ska motsvara samma betygssteg för ett visst betygsomdöme. Detta innebär att lärarnas beslut om summativ bedömning i hög grad bestämdes av brister i enskilda textkvaliteter, främst avseende texters språkliga stil och struktur. Samtidigt uttryckte lärarna betydande hänsyn åt aspekter som elevers ålder och skrivuppgifters svårighetsgrad, elevers skrivutveckling samt det nationella kursprovet. Dessa normer var oftast kompensatoriska genom att de användes för att väga upp brister i texternas kvaliteter, vilket visade sig få avgörande betydelse för lärarnas beslut om summativ bedömning. Lärargrupperna uppvisade sinsemellan en betydande variation i samsyn avseende vad de grundade beslut om summativ bedömning av elevers skrivande på och textkvaliteter jämfördes snarare med varandra än uttrycktes som grader av måluppfyllelse.

### Artikel 3. Lärares interaktion, beslutsprocesser och samstämmighet

Studien undersökte svensklärarnas beslutsprocesser och samstämmighet i samtal om summativ bedömning av elevers skrivande. Studien svarade mot de två forskningsfrågorna: "Hur interagerar lärarna i samtal om summativ bedömning av elevers skrivande?" och "Hur samstämmiga ger lärarna uttryck för att vara?"

Empirin byggde på lärarnas diskussioner i fokusgrupper utifrån frågan "Vilket betygsomdöme motsvarar den här elevens skrivande?" samt ett frågeformulär där enskilda lärare i anslutning till bedömningssamtalet dels bedömde elevernas texter utifrån frågan: "Vilket betygsomdöme anser *du* att de tre elevernas skrivande bäst motsvarar?" (ange *ett* betygsomdöme från A till F), dels utvärderade samtalen utifrån frågorna: "I vilken utsträckning anser du att du fick komma till tals under samtalet?" och "I vilken utsträck-

ning fick samtalet dig att se nya perspektiv eller kvaliteter i elevernas skrivande?” genom att på en sexgradig skala ange en siffra från *till stor del* (6) till *inte alls* (1).

Det teoretiska perspektiv som lades på dessa samtal var dialogiskt (Linell, 2011) och det som undersöktes var interaktionella mönster avseende dominans, dynamik och koherens. Kvantitativa mått på vad som kännetecknade lärargruppernas interaktion ställdes i relation till deras uttryck för samstämmighet inom och mellan lärargrupperna.

Resultatet visade att lärarna uttryckte hög grad av samstämmighet i bedömningssamtalen. Däremot överensstämmer de bedömningar som lärarna gjorde enskilt i anslutning till samtalen enbart till viss del med gruppens beslut. Lägst grad av samstämmighet uppvisade den lärargrupp som i samtalen framstod som mest samstämmiga. Beslutsprocesserna i dessa samtal kännetecknades av låg intensitet där få alternativ till bedömningar prövades och där besluten summerades genom tyst bifall från en majoritet av gruppens lärare. I lärargrupper där intensiteten var högre genom att flera lärare styrde samtalet i olika riktningar och där lärarna tillsammans prövade en mängd olika förslag på bedömning överensstämde även enskilda lärares bedömningar i högre grad med gruppens beslut. Mellan lärargrupperna var bedömningsöverensstämmelsen däremot låg.



# Diskussion

I inledningskapitlet formulerades avhandlingens övergripande syfte: att utifrån didaktiska och dialogiska perspektiv beskriva och analysera hur bedömningsamtal mellan svensklärare tar form, vilka normer lärarna ger uttryck för och vilka beslut de fattar. Avhandlingen har alltså ett dubbelt fokus både på *hur* lärarna samtalar om skrivbedömning och *vad* de samtalar om. Detta dubbla fokus har gjort det möjligt att synliggöra vad som kännetecknar interaktionen i lärarnas samtal och vad lärarna uppmärksammar i olika led av en bedömningsprocess.

Detta kapitel inleds med en presentation och diskussion av vad resultaten från de tre delstudierna ger svar på och hur de kan förstås i relation till de teorier och den tidigare forskning som har presenterats. Därefter diskuteras didaktiska implikationer för svensklärares arbete med skrivbedömning. Slutligen ges förslag på vidare forskning.

## Avhandlingens bidrag

Denna avhandling har i tre vetenskapliga artiklar synliggjort hur skrivbedömning tar form i samtal mellan svensklärare på tre gymnasieskolor. Utifrån de övergripande forskningsfrågorna är avhandlingens viktigaste bidrag att den (1) synliggör en diskrepans mellan lärarnas uttryck för kvalitetsuppfattningar och normer i olika led av en skrivbedömningsprocess, (2) belyser variationer i samsyn och samstämmighet mellan lärare och lärargrupper samt (3) visar skillnader i interaktionella mönster i lärargruppernas beslutsprocesser. I det följande sammanfattas och diskuteras dessa resultat mer ingående.

### **Diskrepans mellan uttryck för kvalitetsuppfattningar och normer i olika led av en skrivbedömningsprocess**

De kvalitetsuppfattningar och förväntningsnormer som svensklärarna uttryckte i samtal om elevernas texter inriktades framför allt på kommunikativa kvaliteter, den stilistiska utformningen och strukturen, men även på äm-

nesinnehåll och källanvändning. Det innebär att lärarna uppmärksammade ett flertal textkvaliteter på olika nivåer och gav uttryck för en inkännande och kritisk läsning av elevernas texter vilket tyder på en betydande didaktisk textkritisk kunskap (Hoel, 2001; Huot, 2002; Phelps, 2000; Weigle, 2002).

Däremot, när lärarna i de efterföljande diskussionerna om vilken undervisning eleverna behöver för att utveckla sitt skrivande och vilket betygsomdöme elevernas skrivande motsvarade, behandlade de enbart ett fåtal av dessa textkvaliteter. Lärarnas didaktiska beslut om en stöttande och utvecklande skrivundervisning inriktades i stort sett enbart på att eleverna behövde undervisning i att strukturera text och att använda källor medan besluten om summativ bedömning i huvudsak bestämdes av texternas stilistiska utformning och struktur. Detta gjorde att det uppstod en diskrepans mellan vilka textkvaliteter lärarna uppmärksammade i olika led av bedömningsprocessen.

I likhet med tidigare svensk forskning om lärares textrespons (Ciolek Laerum, 2009; Hedeboe, 2007; Östlund-Stjärnegårdh, 2002) tar lärarna hänsyn till flera olika textkvaliteter när de bedömer elevers skrivande. Detta tyder på en väl utarbetad praktik hos lärarna att bedöma texters olika kvaliteter vilket kan förklaras med att lärare av tradition ägnar betydande tid åt just detta: att formulera och motivera skrivbedömningar av elevers texter, ofta som individuell skriftlig respons (Bergman Claeson, 2003; Norberg Brorson, 2007). Det kan också förklaras av att svensklärare kontinuerligt förses med bedömningsverktyg för denna typ av textbedömning genom de bedömningsanvisningar och elevexempel som återfinns i materialet för de nationella skrivproven i svenska. I det materialet får lärarna bland annat ta del av instruktioner för hur elevtexter ska bedömas på både holistisk och analytisk nivå där de mer specifika bedömningskriterierna ska underbygga en helhetsbedömning av elevernas skrivande.

Lärarna tycks däremot ha en mer begränsad vana att utifrån bedömningar av elevers texter diskutera och planera för en utvecklande och stöttande skrivundervisning. Deras didaktiska beslut omfattade i mycket begränsad utsträckning de textkvaliteter som de hade ägnat störst uppmärksamhet åt i tidigare samtal om bedömning av elevernas texter. Framför allt ägnades liten uppmärksamhet åt att diskutera och fatta beslut om skrivundervisning som kunde ge eleverna möjlighet att utveckla kommunikativa kvaliteter i sina texter, exempelvis föra logiska välbalanserade resonemang och anpassa stil efter kommunikationssituation, något som de betonade i samtalen om elevers texter. Naturligtvis går det inte att hävda att dessa didaktiska beslut om skrivundervisning, omvandlade till undervisningspraktik, *inte* skulle kunna leda till att eleverna ges möjligheter att utveckla sitt skrivande. Vad som

framstår som problematiskt är dock att de didaktiska besluten inriktades på ett fåtal aspekter och att lärarna snarare tycktes redogöra för en redan befintlig skrivundervisningspraktik och vilken nytta eleverna kunde ha av den, istället för att diskutera en skrivundervisning som kunde möta just dessa elevers lärandebehov. De didaktiska besluten föreföll motsvara en undervisning som redan *är*, inte en som *kan bli*.

På liknande sätt som tidigare studier har belyst problematiken med att lärarens textrespons inte följs av aktiviteter i undervisningen (Bueie, 2015; Igland, 2008; Eriksen, 2017) finns en risk att lärarnas didaktiska beslut inte kommer att leda till aktiviteter som får någon större påverkan på elevernas skrivutveckling eftersom de inte formas utifrån lärarnas analys av elevernas lärandebehov. Lärarnas didaktiska beslut inriktades på aktiviteter som eventuellt rättar till fel i elevernas texter (Richardson, 2000) men de innebär inte att eleverna får tillgång till de kunskaper som krävs för högre betyg, vilket är problematiskt i ett likvärdighetsperspektiv (Calhoon-Dillahunt & Forrest, 2013; Parmenius Swärd, 2008; Skar, 2013).

Vad kan denna diskrepans mellan lärarnas textbedömningar och deras didaktiska beslut bero på? En förklaring kan vara att de generella principer för formativ bedömning som har haft stort genomslag i svensk skola under 2000-talets två första decennier framför allt har betonat vikten av lärares återkoppling till elever och hur elever genom lärares respons kan forma sitt lärande (t.ex. Black och Wiliam, 1998; 2009). I och med det har ansvaret för att eleverna utvecklar nya kunskaper och förmågor till stor del lämnats över till eleverna och lärarnas roll har blivit att förse eleverna med användbar information, dvs. bedömningar, för att göra det. Även om forskning om formativ bedömning sedan 1960-talet och framåt (Scriven, 1967; Sadler, 1989) har betonat att bedömningsinformation om elevers kunskaper också ska användas till att forma den fortsatta undervisningen verkar den aspekten inte ha fått lika stort genomslag i diskussioner om formativ bedömning, åtminstone inte i de samtal som här undersökts.

När Phelps (2000) diskuterar hur en teoretisering av skrivbedömning kan bidra till att utveckla en mer reflekterad didaktisk skrivbedömningspraktik, lyfter hon fram att en sådan teori kan användas för att utvärdera vilka konsekvenser skrivbedömningar får och balansera dem gentemot andra möjligheter. En sådan teori är alltså inte menad att ge ett färdigt svar på hur skrivundervisning ska utformas utan fungera som ramverk för ett didaktiskt utvecklingsarbete där lärare som del av en praktikgemenskap prövar och utvärderar vad skrivbedömningar och skrivundervisning i en viss undervisningskontext kan ge. Phelps frågar hur vi annars kan "appraise the value of response to

student writing without examining empirically the learning that takes place?” (ibid., s. 96). Jølle (2015b) diskuterar på ett liknande sätt vikten av att även diskutera vilka följder olika bedömningshandlingar får.

Även lärarnas summativa bedömningar av elevernas skrivande omfattade i mycket begränsad utsträckning de textkvaliteter som lärarna ägnat störst uppmärksamhet åt i samtalen om bedömning av elevernas texter. Framför allt innebar det att texternas övergripande kommunikativa kvaliteter övergavs till förmån för den stilistiska utformningen och texternas struktur. En förklaring till detta snävare fokus i samtalen om summativ bedömning är den icke-kompensatoriska bedömningsnorm som alla tre lärargrupper återkommande gav uttryck för och som innebär att samtliga kvaliteter i elevernas texter ska motsvara samma betygssteg för ett visst betygsomdöme. Denna bedömningsnorm kan i sin tur förklaras av de allmänna regler som styr bedömning och betygssättning i gymnasieskolan (Gymnasieförordningen, 8 kap. 2 §) och ledde till att lärarnas beslut om summativ bedömning fattades utifrån den eller de brister i elevernas texter som motiverade ett lägre betyg. Flera forskare (bl.a. Goodwin, 2016; Marshall, 2004; Sadler, 2009; Wyatt Smith et al., 2013) varnar för att detta innebär en förenkling av vad skrivbedömning är, vilket kan få negativa konsekvenser för hur bedömningar kommuniceras och används i relation till undervisningen och de mål som ska behandlas.

En ytterligare konsekvens av lärarnas uttryck för en icke-kompensatorisk bedömningsnorm blev att textkvaliteter snarare vägdes mot varandra än bedömdes i förhållande till grader av måluppfyllelse. Motiveringarna för besluten om summativ bedömning anknöt varken till ämnesplanens mål, centralt innehåll eller kursens kunskapskrav, utan uttrycktes i generella termer som att texterna hade ”ett ojämnt språk” och ”vissa brister i språkhanteringen” eller att elevens skrivande ”brister för ett högre betyg” och ”är för ostrukturerat för ett högre betyg”. Mer explicita än så var inte lärarnas motiveringar utan besluten byggde istället på en till synes underförstådd och outtalad gemensam förståelse av vilket betyg elevers skrivande motsvarade (Sadler, 1989; Wyatt-Smith et al., 2013). Detta visade sig också få konsekvenser för samstämmighet i lärarnas bedömningar.

## **Variationer i samsyn och samstämmighet mellan lärare och lärargrupper**

Lärarnas samtal om summativ bedömning av elevernas skrivande uppmärksammade framför allt texternas kvaliteter i relation till kunskapskravens

betygssteg. Men lärarna visade även betydande hänsyn till skrivuppgifternas svårighetsgrad, förväntningar på elevernas skrivande i årskurs 1, elevernas skrivutveckling i kursen och bedömningsanvisningarna från det nationella kursprovet i svenska. Därmed bekräftas bilden av att lärares summativa bedömningar av elevers skrivande innebär att ”much more is in play than a set of stated expectations” (Wyatt-Smith et al., 2010, s. 72) och att en sammanvägning av elevers skrivförmåga till ett betygsomdöme kräver tolkningar på flera olika nivåer (Allal, 2013; Black & Wiliam, 2006; Brookhart, 2010; Harlen, 2012; Moss et al., 2006; Sadler, 1998).

I styrdokumenterna för gymnasieskolan betonas att lärare vid betygssättning ska ta hänsyn till flera olika aspekter. Exempelvis ska lärare göra ”en helhetsbedömning av de kunskaper eleven har visat vid avslutad kurs och jämföra dessa med kunskapskraven” (Skolverket, 2012) och ”utnyttja all tillgänglig information om elevens kunskaper i förhållande till de nationella kunskapskraven och göra en allsidig bedömning av dessa kunskaper” (SKOLFS 2011:144, avsnitt 2.5). Lärare ska även visa bedömningshänsyn till elevernas progression i en kurs genom att ”läraren vid betygssättningen av kursen inte lägger ihop resultaten på alla prov och uppgifter för att få fram en medelprestation” (Skolverket, 2012, s. 23) och till resultat på det nationella kursprovet för att ”bedömningsgrunderna ska bli så enhetliga som möjligt över landet” (Gymnasieförordningen, 8 kap. 3§). Att väga samman en elevs prestationer till ett betyg med hänsyn till alla dessa aspekter är allt annat än lätt, särskilt när det handlar om bedömning av en så komplex handling som skrivande (Moss, 1994; Sadler, 1985).

De *variationer i samsyn* som lärarna gav uttryck för handlade främst om att de olika lärargrupperna baserade sina beslut om summativ bedömning på delvis olika grunder. Detta yttrade sig dock inte som att lärarna hade diametralt olika uppfattningar om vad bedömningen skulle avse utan snarare om vilka bedömningsnormer som gavs prioritet vid bedömning av enskilda elevers skrivande. Exempelvis baserade en av lärargrupperna sitt beslut om summativ bedömning av en elevs skrivande enbart på textkvaliteter för att senare vid bedömning av en annan elevs skrivande ta betydande hänsyn till bedömningsgrunder som elevens progression under kursen och det nationella skrivprovet. Eftersom hänsynstaganden till annat än texters kvaliteter i regel var kompensatoriska ledde detta till att olika elevers skrivande bedömdes på skilda grunder.

De *variationer i samstämmighet* som lärarna gav uttryck för kan ses som en konsekvens av att lärarna prioriterade olika bedömningsgrunder vid bedömning av enskilda elevers skrivande, men även av att de inte explicit ut-

tryckte nivåer av elevers skrivförmåga i relation till styrdokumentet. Istället blev aspekter som bedömningsuppgifters relevans, förväntningar på elevernas skrivande utifrån ålder och årskurs samt elevers progression i kursen styrande. Detta innebar i sin tur att lärarnas skrivbedömningar blev avhängiga deras erfarenheter av att undervisa och bedöma specifika elevgruppers skrivande vilket även tidigare studier har visat är en betydande faktor vid lärares betygssättning av elevers skrivande (Barkaoui, 2011; Crusan et al., 2016; Suto & Greatorex, 2008; Wiseman, 2012).

Lärargruppernas summativa bedömningar av de tre elevernas skrivande skiljde sig åt med upp till fyra betygssteg. I ett fall var lärargruppen överens, i ett annat fall skiljde ett betygssteg. Det fanns däremot inget mönster som visade att en lärargrupp genomgående satte högre eller lägre betyg än övriga.

### **Skillnader i interaktionella mönster i lärargruppernas beslutsprocesser**

När Wenger (1998) beskriver hur deltagare i en praktikgemenskap tillsammans kan utveckla delade uppfattningar om hur uppgifter kan lösas lyfter han fram att de inte enbart behöver ha ett gemensamt fokus på vilken uppgiften är; samtalen ska även genomsyras av förhandlingar och omförhandlingar där alla individer deltar (s. 165). Genom att deltagarna delger och problematiserar sina uppfattningar i relation till tidigare erfarenheter av hur problem har lösts bidrar det till dynamik i samtalet. Denna interaktionella dynamik beskrivs som nödvändig för att gruppen ska utveckla nya gemensamma förhållningssätt.

Ett kännetecknande drag för lärargruppernas beslutsprocesser i samtalen var att de uttryckte sig mycket tentativt och vände sig till varandra med frågor och förslag. Matre och Solheim (2016) beskriver detta som ett kännetecknande drag för lärare som är medvetna om de tolkningshandlingar betygssättning innebär; även om lärarna besitter goda kunskaper i och har omfattande erfarenheter av skrivbedömning generellt är det ingen enkel eller självklar sak att bedöma enskilda elevers skrivande. Lärarna använde på detta sätt varandra som resurser för att lösa bedömningsuppgifterna; återkommande bekräftade och utvecklade de varandras resonemang genom så kallade balanserande turer (Linell & Gustavsson, 1987) vilket gjorde att samtalen överlag kan beskrivas som ett väl fungerande reflekterande samspel mellan lärarna.

Vidare var lärargruppernas samtal mycket konsensusinriktade. Trots att de inte uppmanades att nödvändigtvis bli överens om ett betygsomdöme var det ändå något som alla tre lärargrupper uttryckligen eftersträvade att bli. Ingen enskild lärare eller en mindre grupp av lärare drev en egen linje gentemot övriga. Istället sökte lärarna efter gemensamma ståndpunkter att motivera beslut om summativ bedömning med. Varje lärargrupp kom även fram till ett gemensamt beslut om vilket betygsomdöme de tre elevernas skrivande motsvarade genom att avsluta varje betygssamtal med något som kan liknas vid konsensusbeslut där de öppet sammanfattade gruppens beslut utan att någon lärare uttryckte någon invändning. Även om flera studier (bl.a. Jølle, 2014; Wyatt-Smith & Klenowski, 2013) har beskrivit uttryck för lärares vilja till konsensusbeslut riskerar det att leda till att bedömningsamtalen blir slätstrukna och att de beslut som fattas inte föregås av diskussioner där olika argument problematiserats (Colombini & McBride, 2012; Maxwell, 2010). När konsensus ses som målet kan risken även vara att lärare hämmas att lyfta komplexa problem och att samtalen därigenom blir mindre utforskande än vad de kunde vara (Mercer, 2000).

Samtidigt som det finns flera gemensamma kännetecken i lärargruppernas interaktion visade analysen av dominans, dynamik och koherens betydande skillnader i hur det interaktionella samarbetet fungerade i de tre lärargrupperna. När det gäller dominans delade majoriteten av lärarna på Corneaskolan och Scleraskolan talutrymmet och flera lärare bidrog med ett antal styrande initiativ. Detta gjorde att de interaktionella styrkeförhållandena var mer jämnt fördelade i dessa lärargrupper än i lärargruppen på den tredje skolan, Irisskolan, där två av fem lärare dominerade talutrymmet och bidrog med majoriteten av gruppens styrande initiativ. När det gäller dynamik och koherens, dvs. mått på hur det interaktionella samarbetet fungerade genom att lärarna utvecklade varandras resonemang och bidrog med olika infallsvinklar till diskussionen, var de måtten genomgående högre i lärargrupperna på både Corneaskolan och Scleraskolan än i lärargruppen på Irisskolan. Detta visade sig bland annat i hur lärargrupperna i olika utsträckning tog ställning till och problematiserade olika alternativ innan beslut om summativ bedömning fattades.

När lärarna i anslutning till samtalen enskilt betygssatte elevernas skrivande var samstämmigheten betydligt högre på både Corneaskolan och Scleraskolan än vad den var på Irisskolan. På de två förstnämnda skolorna var lärarna, inom gruppen, helt eniga i sin betygssättning av en elevs skrivande medan lärarnas betyg på de två andra elevernas skrivande skiljde med ett eller två betygsteg genom att en eller två lärare i gruppen satte ett annat

betyg än de övriga lärarna. På Irissskolan var lärarna inte överens om något betyg. Avseende betyg på en av elevernas skrivande skiljde det med ett betygssteg genom att en lärare satte ett D, övriga E. Betygen på de två andra elevernas skrivande skiljde sig däremot desto mer åt. I ena fallet var betygsdifferensen mellan de fem lärarna tre betygssteg (D till B) och i det andra fallet fyra betygssteg (D till A). I båda fallen var det inte mer än två av lärarna som var eniga i sin betygssättning.

För att hitta förklaringar till lärargruppernas olika interaktionella mönster kan man se till lärargruppernas sammansättning. Tendensen var att lärare med lång yrkeserfarenhet, lärare med ämnesansvar och förstelärare dominerade samtalen. Det var de som visade självförtroende att testa olika alternativ, de var prestigelösa och tydliga med att det är tolkningar bedömning innebär vilket öppnade för större dialogiskt utrymme att i gruppen problematisera skrivbedömningar (Matre & Solheim, 2016). Detta blev särskilt tydligt på Irissskolan där skillnaden mellan de fem lärarnas yrkeserfarenhet var som störst; två av lärarna hade varit svensklärare i över tio år vilket är betydligt längre än de övriga lärarna varav två var nyexaminerade. På Corneaskolan och Scleraskolan hade flera av lärarna betydligt längre yrkeserfarenhet som gymnasielärare. Detta visar bland annat på behovet av återkommande samtal om skrivbedömning mellan lärare där även mer oerfarna kolleger är med och aktivt formar gemensamma riktlinjer för gruppens bedömningsarbete (Baird et al., 2004; Little et al., 2003; Wenger, 1998).

## Didaktiska implikationer och förslag på vidare forskning

I den forskningsöversikt om formativ bedömning som Black och Wiliam (1998) presenterade för två decennier sedan konstaterade de att det saknas en entydig definition av vad formativ bedömning är. Den definition de angav löd ”all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” (s. 7–8). Deras betoning av att bedömningsinformation ska användas för att forma undervisning och lärandeaktiviteter förefaller lika aktuell nu som då.

Framför allt behöver lärare utveckla kunskaper om hur bedömningar av elevers skrivande kan användas för didaktiska beslut om skrivundervisning. I ett sådant didaktiskt arbete blir lärares utmaning att utforma undervisningsaktiviteter där eleverna tillsammans ges möjlighet att utveckla sitt skrivande



i relation till målen för kursen (Marshall, 2004; Phelps, 2000). Skrivundervisningen utvärderas sedan i förhållande till dessa undervisningsaktiviteter vilket för läraren blir underlag för fortsatt didaktisk planering och för eleverna fortsatt lärande (Black & Wiliam, 2018; Wiliam & Thompson, 2007).

För att lärare även ska utveckla en gemensam förståelse för de komplexa tolkningsprocesser summativ bedömning av elevers skrivande innebär behöver de aktivt delta i samtal där de utifrån konkreta elevexempel återkommande diskuterar och problematiserar bedömningar av elevers skrivande i relation till uppgift och specifik undervisningskontext (Berge, 2009). Dessa diskussioner bör inte alltid ha som mål att nå konsensusbeslut utan fokusera på de meningsskiljaktigheter som finns. På så sätt kan perspektiven och förståelsen vidgas och bidra till att utveckla lärares gemensamma metaspråk för att tala om texter, skrivande och bedömning (Colombini & McBride, 2012; Jølle, 2014; Matre & Solheim, 2016). Vidare, för att lärares samsyn och samstämmighet ska kunna öka mellan olika skolor på ett likvärdigt sätt, bör dessa diskussioner ske i dialog med såväl kolleger från andra skolor som övriga experter på området (Timperley et al., 2007). Detta innebär ett nära samarbete mellan skolans samtliga aktörer och en förståelse för att variationer i olika lärares bedömningar inte beror på enskilda lärares bristande bedömarförmåga utan på brister i tolkningsgemenskapen (Jølle, 2015a; Wyatt-Smith et al., 2017).

Forskningsöversikten av Black och Wiliam (1998) avslutas med en rad förslag på vad framtida studier kan inriktas på. Eftersom även dessa förslag fortfarande är aktuella tar jag utgångspunkt i ett par av dem och föreslår att framtida studier om skrivbedömning kan undersöka:

Vilka tolkningsramar använder sig lärare av när de bedömer elevers skrivande?

Vad karakteriserar det skrivande som sker i olika klassrum i relation till lärares skrivbedömningar?

Hur fördelas ansvaret mellan lärare och elever i dessa skrivprocesser?

Vilka uppfattningar och föreställningar har lärare om elevers skrivande och om sin egen roll som bedömare?

Sådana studier skulle kunna bidra med ytterligare kunskaper om skrivbedömning som tolkande och meningsskapande processer mellan lärare och bli ett stöd för didaktiskt utvecklingsarbete i skolan.



# Summary

**Title:** Teachers' conversations about writing assessment. Norms, pedagogical decisions and grading

## Introduction

Teachers' ability to communicate assessments and continuously evaluate teaching in relation to students' knowledge and the context in which learning takes place can be crucial for students' opportunities for further learning (Jank & Meyer, 1997). In addition, dialogues between colleagues become an essential part of the pedagogical work for subject development (Bolt, 2011; Gundem & Hopmann, 2002).

As part of the work of increasing the equivalence in teachers' writing assessments, a number of studies (Crisp, 2010; Sadler, 2005; Wyatt-Smith et al., 2010) suggest that teachers discuss assessments of student texts. In such interpretative communities (Fish, 1980; Berge, 2009) teachers' different perceptions and norms can be made visible and problematized, which can lead to common approaches for how students' writing can be judged in relation to the teaching and steering documents. Such work appears to be particularly necessary since several studies have shown that teachers, when assessing students' writing, base their assessment on individual expectations of student writing skills (Berge 2005; Fasting et al., 2009; Jølle, 2014; McBride, 2012; Wyatt-Smith & Klenowski, 2013).

Despite this, a great deal of the research on writing assessment has focused on examining individual teacher's assessments, often through so-called think-aloud protocols (TAP). These studies have helped to visualize teachers' cognitive decision-making processes when assessing students' writing, but not how writing assessments take place through negotiations between teachers and how mutual understanding can progressively emerge in collegial conversations.

## Aim and research questions

This dissertation investigates Swedish teachers' conversations about writing assessment in upper secondary school. The overall aim of the study is to describe and analyze teachers' assessment conversations from both dialogical and pedagogical perspectives regarding how the teachers talk and what they pay attention to when assessing students' writing. In relation to this aim the following three research questions were posed:

What characterizes the teachers' conversation about writing assessment?

What assessment norms do the teachers express?

What formative and summative decisions do they make?

These research questions are answered in three studies presented in the three scholarly articles included in the thesis. In these three studies the following more specific questions are answered: 1) What quality standards and assessment norms do the teachers express in their conversations about writing assessment of students' writing?, 2) What pedagogical decisions do the teachers make from the assessment information about students' writing?, 3) What assessment norms do the teachers express when grading students' writing?, 4) How do the teachers interact in conversation about summative assessment of students' writing? 5) How consistent are the teachers' writing assessment within and between the groups?

## Theoretical framework and key concepts

The study combines dialogical and pedagogical perspectives. From a dialogical perspective, the assessment conversations are considered as social acts where teachers in interaction collectively create meaning. The assessment conversations are treated as community projects (Linell, 2009) where the teachers in interaction help each other to solve different assessment tasks. These communicative projects are treated as not only related to each other but also to the contexts surrounding the conversation itself. Bakhtin (1986) discusses this by describing how different types of opinions should be regarded as parts of communication chains where individual opinions are both the result of previous opinions (backwards) and to whom the utterance is addressed (forward). With this approach, the norms and decisions that the teachers express are seen as both socially produced and reproduced, influ-

enced by the educational context they are a part of and a contribution to (Marková et al., 2007).

Furthermore, the assessment conversations are considered as pedagogical in the sense that they form a part of teachers' reflective work to discuss assessment of students' work together with colleagues. The pedagogical perspective also refers to the decision-making processes such as how to use assessment information for pedagogical decisions about writing instructions ("pedagogical action", Phelps, 2000) and for grading students' writing (Harlen, 2005). The concept of quality perception (Zandén, 2010) is used to describe what teachers' pay attention to when discussing assessment of students' writing. These quality perceptions express what teachers consider relevant to assess in students' writing in a specific course and are therefore considered normative, called norms of expectations (Berge et al., 2017; Evensen et al., 2016; Matre & Solheim, 2016). Such norms may be personal, but in this study the interest focuses on the shared norms that teachers express when discussing writing assessment; norms that can say something about how subject achievement goals and criteria are internalized in teachers' concrete discussions about writing assessment. In addition, this study examines the overall considerations that the teachers make when grading students' writing, i.e. when they weigh different text qualities and other aspects together into a grade for a course (Allal, 2013; Wyatt-Smith et al., 2013).

## Empirical material and methods

Three Swedish teacher groups from three upper secondary schools are included in this study. The teacher groups consist of five to six Swedish teachers per school, comprising a total of 17 teachers, where the majority, fifteen out of seventeen, had considerable experience in teaching Swedish in upper secondary school; six of them had taught for more than five years, nine for more than fifteen years. Two of the teachers had more limited experience and taught their first year.

Focus group discussions (Bryman, 2011) is the main method used to get empirical data that could answer the research questions. In these focus group discussions, the teachers were asked to discuss the assessment of students' writing based on individual readings of authentic student texts. They were also asked to discuss how the students could develop their writing, i.e., consider pedagogical decisions for writing instructions, and summatively assess the students' writing in relation to the course Swedish 1. The other method

used to obtain empirical data was a questionnaire that the teachers answered individually after the focus group discussions. The teachers were for instance asked to evaluate the focus group discussions and to individually assess the students' writing. These answers were compared with the teacher groups' decisions on summative assessments and interactional patterns in the conversations in order to investigate the degree of consistency within each teacher group.

Methods used to analyze the assessment conversations were topic analysis and initiative-response-analysis (Linell, 2011). The topic analysis determined how the teachers interpreted each other's opinions when solving different communicative projects. Topics are created when at least two of the teachers express a common understanding about a subject and are therefore considered to reflect perceptions, opinions and attitudes which the participants share. The topic analysis revealed what the teachers paid attention to in different phases of the assessment process when discussing students' texts, writing instructions and grades. The initiative-response-analysis (Linell, 2011; Linell & Gustavsson, 1987) revealed interactional patterns of dominance, dynamics and coherence in the assessment conversations by coding each utterance based on the level of initiative and response (*Bilaga F*). Dominance refers to the interactional strengths between the teachers, i.e. how different teachers, by initiating new topics and responding to other people's initiatives, contribute to the ongoing discussions. This gives a picture of the interactional relations in the group and how different teachers control the conversation. Dynamics relates to the intensity of the conversation and to what degree the teachers' initiatives and responses steer the conversation in different directions. Conversational coherence refers to how the teachers respond to each other and relate to each other's opinions, which says something about how well the dialogue interaction works.

## Results

This dissertation reports and discusses how writing assessment takes place in conversations between Swedish teachers at three upper secondary schools.

The first study shows that teachers' assessments of students' writing focus on much more, and partly other, criteria than their pedagogical decisions do. The quality standards that teachers express in the discussions about student texts focus mostly on communicative quality, language style and text structure but also on content and the use of sources. The pedagogical decisions,

on the other hand, almost exclusively focus on text structure and the use of sources. This means that there is a tension between teachers' perceptions of qualities in students' writing and their pedagogical decisions.

The second study shows that teachers' summative assessments of students' writing focus on the text qualities in relation to the grading scale and express that all text qualities must reach all criteria for a certain grade. This means that shortcomings in the students' texts are crucial for teachers' summative assessments. Meanwhile, the teachers also express that students' age and writing instructions as well as students' writing development and the national test must be considered. These assessment norms can be compensatory and have a substantial impact on these teachers' decisions on summative assessments. The teacher groups show considerable variation in the basis for their decisions regarding summative assessment of students' writing.

Finally, the results from the third study reveals a high degree of consistency within each teacher group's discussion. However, the assessments that individual teachers make after the discussions comply only to some extent with the group's decisions. The least degree of consistency was shown by the teacher group who seemed the most consistent in the discussions. Decision-making processes in these discussions are characterized by low intensity, where few assessment alternatives are being considered. In teacher group discussions characterized by higher intensity, the teachers make a variety of suggestions for assessment. The individual teachers' assessments in those groups are more consistent with the group's decisions. Meanwhile, the overall assessment consistency between the teachers was found to be low.

## Discussion

The three main contributions of this dissertation are that it highlights (1) a discrepancy between the quality perceptions and the norms the teachers express in different phases when assessing students' writing assessment, (2) considerable variations in the teachers' expressions of norms and decisions when grading students' writing and (3) differences in interactional patterns in the teacher groups' decision-making processes.

In accordance with the results of previous Swedish studies on teachers' writing assessment (Ciolek Laerum, 2009; Hedeboe, 2007; Östlund-Stjärnegårdh, 2002), the teachers included in the present dissertation pay attention to several different text qualities when assessing students' writing.

This can be explained by the fact that teachers traditionally devote considerable time to formulating writing assessments of student texts, often as individual written response (Bergman Claeson, 2003; Norberg Brorson, 2007). This can also be explained by the fact that Swedish teachers are continuously provided with assessment tools for this kind of assessment, for instance through the instructions for the national tests in Swedish. In this material, teachers get information on how student texts are to be assessed at both holistic and analytical levels, where the more specific assessment criteria will support an overall assessment of the students' writing.

However, the teachers seem to have less experience of planning writing instructions based on assessments of student writing. To a very limited extent the pedagogical decisions of the teachers included the text qualities they paid most attention to in previous conversations about the assessment of students' texts. In particular, little attention was devoted to discussing writing instruction that could give students the opportunity to develop communicative qualities in their texts, for example, to support logical well-balanced reasoning and to adapt style to the communication situation. Of course, it cannot be argued that these pedagogical decisions, transformed into teaching practice, would not lead to opportunities for the students to develop their writing. However, it appears to be problematic when the pedagogical decisions focus on very few aspects and when the teachers seem to account for an already existing writing instructional practice, instead of discussing writing instructions that could meet the needs of these students. The pedagogical decisions seem to correspond to a teaching that already is, not one that can be.

Furthermore, the teachers' summative assessments of the students' writing included to a very limited extent the text qualities that the teachers paid most attention when assessing student texts. In particular, the overall communicative qualities of the texts were abandoned in favor of the stylistic design and the structure of the texts. An explanation for this narrower focus when discussing summative assessment of students' writing is the non-compensatory assessment standard that all three groups of teachers repeatedly express. This means that all qualities in student texts should correspond to the same standards and resulted in a strong focus on finding what shortcomings could motivate a lower grade. Although the teachers' expression of a non-compensatory assessment standard can be explained by the general rules governing assessment and grading in upper secondary school (Gymnasieförordningen, 8 kap. 2 §), it appears problematic when the teachers mainly address shortcomings in the students' texts instead of assessing their writing



overall. Several researchers (Goodwin, 2016; Sadler, 2009; Wyatt Smith et al., 2013) warn of such development and argue that it can lead to a simplified view of what writing assessment is and can adversely affect how assessments are communicated and used in relation to the teaching and the subject goals.

The variations in the teachers' expressions of norms and decisions when grading students' writing refer in particular to the fact that the different groups of teachers base their decisions on a number of different grounds. However, this did not indicate that the teachers had diametrically different perceptions of what the assessment should relate to, but rather about which assessment standards were given priority in assessing an individual student's writing. For example, one of the teacher groups base their summative decision concerning one particular student's writing entirely on textual qualities, but when assessing another student's writing, they laid particular stress on the student's progression during the course and the national writing test. Since considerations about other than textual qualities were generally compensatory, this resulted in different students' writing being assessed on different grounds. Similar patterns were found in all teacher groups' conversations about summative assessment.

These variations can be seen as a consequence of the teachers prioritizing different assessment bases in assessing the writing of individual students, but also that they did not explicitly express levels of student writing skills in relation to the steering documents. Instead, aspects such as relevance of assessment tasks, expectations of students' writing based on age and grade and students' progression through the course became guiding principles. In turn, this meant that the teachers' writing assessments were dependent on their experience of teaching and assessing the writing of specific student groups. Previous studies have shown that this can be an important factor when teachers grade students' writing (Barkaoui, 2011; Crusan et al., 2016; Suto & Greator, 2008; Wiseman, 2012). The teachers' summative assessments of the three students' writing differed by up to four grades. In one case, the teacher groups agreed on the same grade. There was, however, no pattern that showed that one of the teacher groups consistently set higher or lower grades than the others.

Variations in interactional patterns between the teacher groups can be explained by the composition of the teacher groups. The tendency is that in groups with a majority of teachers with long professional experience there is more dialogical space to problematize the assessments. Matre and Solheim (2016) describe this as a characteristic feature for teachers who are aware

that assessment means interpretation; although the teachers possess good knowledge of and have extensive experience of writing assessment in general, it is not easy or obvious to assess individual students' writing. This approach to assessment is missing at the school (Irisskolan) where several teachers were newly graduated.

At the same time, there are several similarities in how the teachers discuss assessment. A characteristic feature of the decision-making processes is that the teachers express themselves very tentatively and turn to each other with questions and suggestions. In this way, the teachers use each other as resources to solve the assessment tasks; repeatedly confirming and developing each other's reasoning through so-called balancing utterances (Linell & Gustavsson, 1987), can be seen as an expression of a well-functioning reflective interaction between the teachers.

Furthermore, the teachers are very consensus-oriented. Each teacher group came to a joint decision on what grade the three students' writing corresponded to by summarizing the group's decisions without any teacher's objection. Although several studies (Jølle, 2014; Wyatt-Smith & Klenowski, 2013) have described teachers' willingness to consensus decisions, there is a risk that decisions are taken without arguments being discussed sufficiently (Colombini & McBride, 2012; Maxwell, 2010). When consensus is seen as the goal, the risk may also be that teachers are prevented from raising complex problems, thereby making conversation less exploring than they could be (Mercer, 2000).

# Referenser

- Adie, L. E., Klenowski, V. & Wyatt-Smith, C. (2012). Towards an understanding of teacher judgement in the context of social moderation. I: *Educational Review*, 64, s. 223–240.
- Aldrin, E. (2015). Textbedömning i svenskämnet – attityder, erfarenheter och variation. I: *Forskning om undervisning och lärande*, 14, s. 62–83.
- Allal, L. (2013). Teachers' professional judgement in assessment: A cognitive act and a socially situated practice. I: *Assessment in education: Principles, Policy & Practice*, 20(1), s. 20–34.
- Angelöw, B. & Jonsson, T. (2000). *Introduktion till socialpsykologi*. Lund: Studentlitteratur.
- Anward, J. & Nordberg, B. (2005). *Samtal och grammatik: Studier i svenskt samtalspråk*. Lund: Studentlitteratur.
- Atkinson, J. Maxwell & Heritage, J. (1984). Preference organization. I: J. Maxwell Atkinson & J. Heritage (red.), *Structures of social action*, s. 53–56. Cambridge: Studies in conversation analysis.
- Baird, J.-A., Greatorex, J. & Bell, J. F. (2004). What makes marking reliable? Experiments with UK examinations. I: *Assessment in Education: Principles, Policy & Practice*, 11(3), s. 331–348.
- Bakhtin, M. (1981). *The dialogic imagination: Four Essays*. Austin/London: The University of Texas Press.
- Bakhtin, M. (1986). The Problem of Speech Genres. I: M. Holquist & C. Emerson (red), *Speech genres and other late essays* (s. 60–102). Austin: University of Texas Press.
- Barkaoui, K. (2011). Think-aloud protocols in research on essay rating: An empirical study of their veridicality and reactivity. I: *Language Testing*, 28, s. 51-75.
- Behizadeh, N. & Engelhardt, G. (2011). Historical View of the Influences of Measurement and Writing Theories on the Practice of Writing Assessment in the United States. I: *Assessing Writing*, 16 (3), s. 189–211.
- Bennett, R. A. (2011). Formative assessment: a critical review. I: *Assessment in Education: Principles, Policy and Practice*, 18(1), s. 5–25.
- Berge, K. L. (1996). *Norsksensorenes tekstnormer og doxa: en kultursemiotisk og sosioteknologisk analyse*. Doktorsavhandling, Norges teknisk-naturvitenskapelige universitet.

- Berge, K. L. (2002). Hidden Norms in Assessment of Students' Exam Essays in Norwegian Upper Secondary Schools. I: *Written Communication*, 19 (4), s. 458–492.
- Berge, K. L. (2005). Tekstskulturer og tekstkvaliteter. I: K. L. Berge et al. (2005) *Ungdommens skrivekompetanse*, Bind II. Oslo: Universitetsforlaget.
- Berge, K. L. (2009). Er tolkningsfelleskap mulig å oppnå i skriveprøver? I: H. Otnes, F. Hertzberg, L. S. Evensen & O. K. Haugaløkken (red.), *Tekstvurdering som didaktisk utfordring*. Oslo: Universitetsforlaget.
- Berge, K. L. & Skar, G. B. (2015). Ble elevene bedre skrivere? Intervensjonseffekter på elevers skriveferdigheter og skriveutvikling. I: *Rapport 2 fra Normprosjektet*. Trondheim: Høgskolen i Sør-Trøndelag, Avdeling for lærer-og tolkutdanning.
- Berge, K. L., Skar, G. B., Matre, S., Solheim, R., Evensen, L. & Thygesen, R. (2017). Introducing teachers to new semiotic tools for writing instruction and writing assessment: consequences for students' writing proficiency. I: *Assessment in Education: Principles, Policy & Practice* 24 (3), s. 1–20.
- Bergman-Claeson, G. (2003). *Tre lärare – tre världar: lärarkommentarer till elevtexter i tre gymnasieklasser*. Uppsala: Uppsala universitet.
- Bernstein, B. 1971. On the classification and framing of educational knowledge. I: M. F. D. Young (red). *Knowledge and Control: New directions for the sociology of education*. London: Collier MacMillan, s. 47–69.
- Biggs, J. & Tang, C. (2007). *Teaching for quality learning at university: what the student does*. Maidenhead: New York, NY. McGraw-Hill/Society for Research into Higher Education: Open University Press.
- Björnsson, C. H. (1960). *Uppsatsbedömning och uppsatsskrivning*. Stockholm: Almqvist & Wiksell.
- Black, P. (2003). The Nature and Value of Formative Assessment for Learning. I: *Improving Schools, Vol. 6* (3), s. 7–22.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. I: *Assessment in Education: Principles, Policy and Practice*, 5(1), 7–73.
- Black, P. & Wiliam, D. (2006). Assessment for learning in the classroom. I: J. Gardner (red.), *Assessment and learning* (s. 9–26). London, Thousand Oaks & New Delhi: SAGE Publications.
- Black, P. & Wiliam, D. (2009). *Developing the theory of formative assessment. Educational Assessment Evaluation and Accountability. Vol 21* (1), s. 5–31.
- Black, P., Harrison, C., Hodgen, J. Marshall, B. & Serret, N. (2010). Validity in teachers' summative assessments. I: *Assessment in Education Principles, Policy & Practice*, 17, s. 215–232.

- Black, P., Harrison, C., Hodgen, J. Marshall, B. & Serret, N. (2011). Can teachers' summative assessments produce dependable results and also enhance classroom learning? I: *Assessment in Education Principles, Policy & Practice*, 18 (4), s. 451–469.
- Black, P. & Wiliam, D. (2018). Classroom assessment and pedagogy. I: *Assessment in Education Principles Policy and Practice*, s. 1–25.
- Blomqvist, P. (2018a). Vad uppmärksammar lärare i samtal om skrivbedömning? Svensklärares normer för beslut om summativ bedömning. I: *Nordisk Tidskrift för Allmän Didaktik*, Vol. 4, s. 34–55.
- Blomqvist, P. (2018b). Är vi överens om det? Samspel och samstämmighet i svensklärares bedömningsamtal om gymnasieelevers skrivande. I: *Nordic Journal of Literacy Research*, Vol. 4, s. 1–21.
- Blomqvist, P., Lindberg, V. & Skar, G. B. (2016). Vad behöver eleverna undervisning i för att utveckla sitt skrivande? Förväntningsnormer och didaktiska beslut i svensklärares bedömningsamtal. I: *Acta Didactica Norge*, Vol. 10 (1), s. 1–22.
- Borgström, E., & Ledin, P. (2014). Bedömarvariation. Balansen mellan teknisk och hermeneutisk rationalitet vid bedömning av skrivprov. I: *Språk och stil*, 24, s. 133–165.
- Brookhart, S. M. (2010). *Formative assessment strategies for every classroom*. Alexandria, Va.: ASCD.
- Bryman, A. (2011). *Samhällsvetenskapliga metoder*. Malmö: Liber.
- Bueie, A. (2014). "Jeg synes det er ett fett om det er 'klarte' eller 'hadde'." Om elevers forståelse og reaksjon på lærers tekstrespons i norskfaget. I A. J. Aasen, Skaftun, A., og Uppstad, P. H. (red.), *Skriv! Les!* (s. 109–132). Bergen: Fagbokforlaget.
- Bueie, A. (2015). Summativ vurdering i formativ drakt – elevperspektiv på tilbakemelding fra heldagsprøver i norsk. I: *Acta Didactica Norge*, 9 (1), s. 1–21.
- Bueie, A. (2017). "Sliske kommentarer er de som hjelper". *Elevers forståelse og bruk av lærerkommentarer i norskfaget*. Doktorsavhandling. Oslo: Universitetet i Oslo.
- Calhoon-Dillahunt, C. & Forrest, D. (2013). Conversing in Marginal Spaces: Developmental Writers' Responses to Teacher Comments. I: *Teaching English in the Two Year College*, 40(3), s. 230–247.
- Ciolek Laerum, B. (2009). *Elever skriver och lärare bedömer: en studie av elevtexter i åk 9*. Uppsala: Universitetstryckeriet.
- Cohen, L., Manion, L. & Morrison, K. (2011). *Research methods in education*. Milton Park, Abingdon, Oxon: Routledge.
- Colombini, C. B. & McBride, M. (2012). "Storming and norming": Exploring the value of group development models in addressing conflict in communal writing assessment. I: *Assessing Writing*, 17 (4), s. 191–207.
- Connolly, S., Klenowski, V. & Wyatt-Smith, C. 2011. Moderation and consistency of teacher judgement: Teachers' views. I: *British Educational Research Journal*, s. 1–22.

- Cooksey, R. W., Freebody, P. & Wyatt-Smith, C. (2007). Assessment as Judgement-in-Context: Analysing how teachers evaluate students' writing. I: *Educational Research and Evaluation* 13 (5), s. 401–434.
- Crisp, V. (2007). Do assessors pay attention to appropriate features of student work when making assessment judgements? A paper presented at the International Association for Educational. I: *Cambridge Assessment. Cambridge: University of Cambridge. Assessment Annual Conference, Baku, Azerbaijan, 2007.*
- Crisp, V. (2010). Judging the grade: Exploring the judgement processes involved in examination grading decisions. I: *Evaluation and Research in Education*, 23(1), s. 19–35.
- Crisp, V. (2013). Criteria, comparison and past experiences: how do teachers make judgements when marking coursework? I: *Assessment in Education: Principles, Policy & Practice*, 20, s. 127–144.
- Crusan, D., Plakans, L. & Gebril, A. (2016). Writing assessment literacy: surveying second language teachers' knowledge, beliefs and practices. I: *Assessing Writing*, 28, s. 43–56.
- Daugherty, R., Black, P., Ecclestone, K., James, M. & Newton, P. E. (2012). Alternative perspectives on learning outcomes: Challenges for assessment. I: J. Gardner, (red.) (2012), *Assessment and learning* (s. 72–86). London: Sage Publications.
- Dysthe, O. (2011). 'What is the Purpose of Feedback when Revision is not Expected?' A Case Study of Feedback Quality and Study Design in a First Year Master's Programme. I: *Journal of Academic Writing*, 1(1), 135–142.
- Dysthe, O., & Hertzberg, F. (2009). Den nyttige tekstresponen: – hva sier nyere forskning? I: O.K. Haugaløkken, L.S. Evensen, F. Hertzberg, & H. Otnes (red.), *Tekstvurdering som didaktisk utfordring*. Oslo: Universitetsforlaget, s. 35–43.
- Dysthe, O. & Iglund, M. -A. (2003), Mikhail Bakhtin och sociokulturell teori. I: O. Dysthe (red.), *Dialog, samspel och lärande*. Lund: Studentlitteratur.
- Eriksen, H. (2017). Elevers oppfatning av lærers tilbakemeldingspraksis: om sammenheng mellom graden av VfL-praksis og elevenes opplevde nytte av lærers skriftlige tilbakemeldinger i skriftlig norsk. I: *Nordic Journal of Literacy Research* 3, s. 1–20.
- Evensen, L. S. (1997). Å skrive seg stor: Utvikling av koherens og sosial identitet i tidlig skrivning. I: L. S. Evensen & T. L. Hoel (red.), *Skriveteorier og skolepraksis*. Oslo: LNU/Cappelen Akademisk Forlag.
- Evensen, L. S. (2009). Underveis mot et tolkningsfelleskap: lærerstemmer om elevtekster. I: S. Matre., D. Sjøhelle, R. Solheim (red.), *Teorier om tekst i møte med skolens lese- og skrivepraksiser*. Oslo: Universitetsforlaget.

- Evensen, L. S. (2010). En gyldig vurdering av elevers skrivekompetanse? I: J. Smidt (red.), *Rammer for skriving: om skriveutvikling i skole og yrkesliv*. Trondheim: Tapir akademisk forlag.
- Evensen, L. S. (2012). Vurdering av skrivekompetanse: En kompleks utfordring. I: H. Otnes (red.), *Tekstvurdering som didaktisk utfordring*. Oslo: Universitetsforlaget.
- Evensen, L. S., Berge, K. L., Thygesen, R., Matre, S. & Solheim, R. (2016). Standards as a tool for teaching and assessing cross-curricular writing. I: *The Curriculum Journal*, 27(2), s. 229–245.
- Falk, B. & Ort, S. (1998). Sitting down to score: Teacher learning through assessment. I: *Phi Delta Kappa*, 80 (1), s. 59–64.
- Fasting, R. B., Thygesen, R., Berge, K. L., Evensen, L. S., & Vagle, W. (2009). National Assessment of Writing Proficiency Among Norwegian Students in Compulsory Schools. I: *Scandinavian Journal of Educational Research*, 53 (6), s. 617–637.
- Fish, S. (1980). *Is there a text in this class? The authority of interpretive communities*. Cambridge, MA: Harvard University Press.
- Forsberg, E. & Lindberg, V. (2010). *Svensk forskning om bedömning – en kartläggning*. Vetenskapsrådet: Vetenskapsrådets rapportserie 2:2010.
- Freccero, U. & Pousette, A. (2005). Bedömning för utveckling i ämnet svenska. I: *Bedömning av kvalitativ kunskap: Konkreta exempel från gymnasieskolan*, 24, s. 15–30. U. Freccero, T. Hortlund & A. Pousette (red.). Solna: Gymnasietidningen Fortbildningsförlaget.
- Gardner, J. (2012). *Assessment and learning*. London: Sage Publications.
- Gibbons, P. (2013). *Stärk språket, stärk lärandet: språk- och kunskapsutvecklande arbetssätt för och med andraspråkslever i klassrummet*. Stockholm: Hallgren & Fallgren.
- Gibbs, G. & Simpson, C. (2004). Does your assessment support your students' learning? I: *Journal of Teaching and Learning in Higher Education*, 1(1), s. 3–31.
- Gipps, C. (1994). *Beyond testing. Towards a theory of educational assessment*. London & New York: The Falmer Press.
- Gipps, C. (1999). Socio-Cultural Aspects of Assessment. I: *Review of Research in Education*, 24, s. 355–392.
- Goodwin, S. (2016). A many-facet Rasch analysis comparing essay rater behavior on an academic English reading/writing test used for two purposes. I: *Assessing Writing*, 30, s. 21–31.
- Gundem, B. B. & Hopmann S. (2002). *Didaktik and/or Curriculum: An International Dialogue: An International Dialogue*. New York: American University Studies.
- Harlen, W. (2005). Teachers' Summative Practices and Assessment for Learning: Tensions and Synergies. I: *The Curriculum Journal*, 16(2), s. 207–223.

- Harlen, W. (2012). On the relationship between assessment for formative and summative purposes. I: J. Gardner (red.), *Assessment and Learning*. London, California, New Delhi: Sage Publications.
- Hattie, J. & Timperley, H. (2007). The power of feedback. I: *Review of Educational Research* 77 (1), s. 81–112.
- Hedeboe, B. (2007). On the 'internal dialogue' between an examination task and pre-university students' responses. I: A. MacCabe et al. (red.) *Advances in language and education*. London: Continuum.
- Higgins, R., Hartley, P. & Skelton, A. (2001). Getting the message across: The problem of communicating assessment feedback. I: *Teaching in Higher Education*, 6, s. 269–274.
- Hirsh, Å. & Lindberg, V. (2015). *Formativ bedömning på 2000-talet – en översikt av svensk och internationell forskning*. Stockholm: Vetenskapsrådet.
- Hoel, T. L. (2001). *Skriva och samtala: lärande genom responsgrupper*. Lund: Studentlitteratur.
- Hoover, N. R & Abrams, L. M. (2013). Teachers' instructional use of summative student assessment data. I: *Applied Measurement in Education*, 26, s. 219–231.
- Hopmann, S. (1997). Wolfgang Klafki och den tyska didaktiken. I: M. Uljens (red.), *Didaktik: teori, reflektion och praktik*. Lund: Studentlitteratur.
- Huot, B. (2002). *(Re)articulating Writing Assessment for Teaching and Learning*. Logan, Utah: Utah State University Press.
- Igland, M. -A. (2008). *Mens teksten blir til*. Doktorgradsavhandling. Oslo: Universitetet i Oslo.
- Igland, M. -A. (2009). Ein ettertankens didaktikk: om forståinga av pedagogisk tekstrespons. I: O. K. Haugaløkken, L. S. Evensen, F. Hertzberg & H. Otnes (red.) *Tekstvurdering som didaktisk utfordring* (s. 24–35).
- Igland, M. -A. (2013). Kunsten å lese elevtekster. I: Skar, G. & M. Tengberg (red.), *Läsning!* Stockholm: Svenskläraryörelingen.
- James, M. (2012). Assessment in Harmony with our Understanding of Learning: Problems and possibilities. I: Gardner, J. (red.) (2012). *Assessment and learning*, s. 33–48. London: Sage Publications.
- Jank, W. & Meyer, H. (1997). Didaktikens centrala frågor. I: M. Uljens (red.), *Didaktik: teori, reflektion och praktik*. Lund: Studentlitteratur.
- Jølle, L. (2014). Pair assessment of pupil writing: A dialogic approach for studying the development of rater competence. I: *Assessing Writing*, 20, s. 37–52.
- Jølle, L. (2015a). Rater strategies for reaching agreement on pupil text quality. I: *Assessment in Education: Principles, Policy & Practice*, 22 (4), s. 458–474.



- Jølle, L. (2015b). *Vurderingsdialogen. En undersøkelse av tekstvurderings praksis ved nasjonal læringsstøttende prøve i skriving*. Doktoravhandling. Trondheim: Norges teknisk–naturvitenskapelige universitet.
- Jönsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. I: *Educational Research Review*, 2 (2), s. 130–144.
- Jönsson, A. & Thornberg, P. (2014). Samsyn eller samstämmighet? En diskussion om sambedömning som redskap för likvärdig bedömning i skolan. I: *Pedagogisk forskning i Sverige*, 19, nr 4–5, 2014, s. 386–402.
- Kitzinger, J. (1994): The methodology of Focus Groups: The importance of interaction between research participants. I: *Sociology of Health and Illness*, 16 (1), s. 103–121.
- Klenowski, V. & Wyatt-Smith, C. (2010). Standards, teacher judgement and moderation in contexts of national curriculum and assessment reform. I: *Assessment Matters*, 2, s. 107–131.
- Korp, H. (2006). *Lika chanser på gymnasiet? En studie om betyg, nationella prov och social reproduktion*. Doktorsavhandling. Malmö: Malmö högskola.
- Krogh, E. (2007). En ekstra chance. Portfolioevaluering i dansk. I: *Gymnasiepædagogik 60*. Odense: IFPR. Syddansk Universitet.
- Kronholm-Cederberg, A. (2009). *Skolans responskultur som skriftpraktik: gymnasisisters berättelser om lärarens skriftliga respons på uppsatsen*. Doktorsavhandling. Åbo: Åbo Akademi.
- Krueger, R. A., & Casey, M. A. (2014). *Focus Groups*. Thousand Oaks: Sage.
- Kuyumcu, E. (2004). Genrer i skolans språkutvecklande arbete. I: K. Hyltenstam, & I. Lindberg (red.), *Svenska som andraspråk: i forskning, undervisning och samhälle*. Lund: Studentlitteratur.
- Kvale, S. & Brinkmann, S. (2009). *Den kvalitativa forskningsintervjun*. Stockholm: Studentlitteratur.
- Larsson, K. (1984). *Skrivförmåga. Studier i svenskt elevspråk*. Lund: Liber.
- Lindberg, V. (2002). Införandet av godkändgränsen – konsekvenser för lärare och elever. I: H. Andersson (red.), *Att bedöma eller döma. Tio artiklar om bedömning och betygssättning*. Stockholm: Skolverket.
- Lindberg, V. (2011). Betyg och bedömning i svensk didaktisk forskning 1990–2009. I: L. Lindström, V. Lindberg, & A. Pettersson (red.), *Pedagogisk bedömning: att dokumentera, bedöma och utveckla kunskap*. Stockholm: Stockholms universitets förlag.
- Linell, P. (2001). *Approaching Dialogue: Talk, Interaction and Contexts in Dialogical Perspectives*. Amsterdam: John Benjamins Publishing Company.
- Linell, P. (2009). *Rethinking Language, Mind, and World Dialogically: Interactional and Contextual Theories of Human Sense-Making*. Charlotte, NC: Information Age Publishing.

- Linell, P. (2011). *Samtalskulturer. Kommunikativa verksamhetstyper i samhället, volym 1–2*. Linköping: Linköpings universitet.
- Linell, P. & Gustavsson, L. (1987). *Initiativ och respons. Om dialogens dynamik, dominans och koherens*. Linköping: Universitetet i Linköping.
- Linell, P. & Marková, I. (1993). Acts in discourse: from monological speech acts to dialogical inter-acts. I: *Journal for the Theory of Social Behaviour*, 23, s. 173–195.
- Little, J., Gearhart, M., Curry, M., & Kafka, J. (2003). Looking at student work for teacher learning, teacher community, and school reform. I: *Phi Delta Kappan*, s. 185–192.
- Lundahl, C. & Tveit, S. (2014). Att legitimera nationella prov i Sverige och i Norge – en fråga om profession och tradition. I: *Pedagogisk forskning i Sverige*, 19, nr 4–5, s. 297–323.
- Lundgren, P. (2013). *Om bedömning av skrivdelen i nationella proven i svenska i årskurs 9 och Sv1 vårterminen 2012*. Stockholm: Utbildningsförvaltningen, Stockholms stad.
- Marková, I., Linell, P., Grossen, M. & Salazar Orvig, A. (2007). *Dialogue in focus groups: Exploring socially shared knowledge*. London: Equinox.
- Marshall, B. (2004). Goals or horizons – The conundrum of progression in English: or a possible way of understanding formative assessment in English. I: *The Curriculum Journal*, 15(2), s. 101–113.
- Matre, S., Berge, K. L., Evensen, L. S., Fasting, R. B., Solheim, R., & Thygesen, R. (2016). *Developing National Standards for the Teaching and Assessment of Writing. Rapport frå forprosjekt Utdanning 2020*. Trondheim: Nasjonalt senter for skriveopplæring og skriveforskning.
- Matre, S. & Solheim, R. (2015). Writing education and assessment in Norway: Towards shared understanding, shared language and shared responsibility. I: *Educational Studies in Language and Literature vol. 15*, s. 1–33.
- Matre, S., & Solheim, R. (2016). Opening dialogical spaces: Teachers' meta-talk on writing. I: *International Journal of Educational Research* 80, s. 188–203.
- Maxwell, G. (2002a). Moderation of teacher judgements in student assessment. Brisbane: Queensland School Curriculum Council.
- Maxwell, G. (2010). Moderation of student work by teachers. I: *International encyclopedia of education*, s. 457–463. Oxford: Elsevier.
- Mercer, N. (2000). *Words and minds: How we use language to think together*. London: Routledge.
- Mickwitz, L (2011). *Rätt betyg för vem? Betygsättning som institutionaliserad praktik*. Licentiatavhandling. Stockholm: Stockholms universitet.
- Morgan, D. (1996). Focus groups. I: *Annual Review of Sociology* 22, s. 129–152.
- Moss, P. A. (1994). Can there be validity without reliability? I: *Educational Research*, 23(2), s. 6–12.

- Moss, P. A., Girard, B. J., & Haniford, L. C. (2006). Validity in Educational Assessment. I: *Review of Research in Education*, 30 (1), s. 109–162.
- Myers, G. (2004). *Matters of Opinion. Talking About Public Issues*. Cambridge: Cambridge University Press.
- Newton, P. E. (2007). Clarifying the purposes of educational assessment. I: *Assessment in Education*, 14(2), s. 149–170.
- Norberg Brorson, B. (2007). *Man liksom bara skriver*. Doktorsavhandling. Örebro: Örebro universitet.
- Palmér, A. (2008). *Samspel och solostämmor: om muntlig kommunikation i gymnasieskolan*. Uppsala: Doktorsavhandling, Uppsala universitet.
- Parmenius Swärd, S. (2008). *Skrivande som handling och möte – gymnasieelever om skrivuppgifter, tidsvillkor och bedömning i svenskämnet*. Doktorsavhandling. Malmö: Malmö högskola.
- Parr, J., & Timperley, H. (2010). Feedback to writing, assessment for teaching and learning and student progress. I: *Assessing Writing*, 15(2), s. 68–85.
- Pedder, D. & James, M. (2012). Professional learning as a condition for assessment for learning. I: Gardner, J. (red.) (2012). *Assessment and learning*, s. 33–48. London: Sage Publications.
- Pettersson, A. (2010). Bedömning av kunskap för lärande och undervisning. I: S. Eklund (red.), *Bedömning för lärande – en grund för ökat kunskapslärande. Forskning om undervisning och lärande 3, 2010*.
- Phelps, L. W. (2000). Cyrano's nose: Variations on the theme of response. I: *Assessing Writing*, 7 (1), s. 91–110.
- Reid, L. (2007). Teachers talking about writing assessment: Valuable professional learning? I: *Improving Schools*, 10, s. 132–149.
- Richardson, S. (2000). Students' Conditioned Response to Teachers' Response. I: *Assessing Writing*, 7(2), s.117–141.
- Riksrevisionen. (2004). *Betyg med lika värde? En granskning av statens insatser*. RiR 2004:11. Stockholm: Riksrevisionen.
- Riksrevisionen. (2011). *Lika betyg, lika kunskap? En uppföljning av statens styrning mot en likvärdig betygssättning i grundskolan*. RiR 2011:23. Stockholm: Riksrevisionen.
- Riksrevisionen. (2013). *Statens tillsyn över skolan – bidrar den till förbättrade kunskapsresultat?* RiR 2013:16. Stockholm: Riksrevisionen.
- Rommetveit, R. (1992). Outlines of a dialogically based social-cognitive approach to human cognition and communication. I: A. H. Wold (Ed.), *The Dialogical Alternative: Towards a Theory of Language and Mind* (s. 19–44). Oslo: Scandinavian University Press.
- Sadler, D. R. (1985). The origins and functions of evaluative criteria. I: *Educational Theory*, 35(3), s. 285–297.
- Sadler, D. R. (1987) Specifying and promulgating achievement standards. I: *Oxford Review of Education*, 13, s. 191–209.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. I: *Instructional Science*, vol. 18 (2), s. 119–144.

- Sadler, D. R. (1998). Formative Assessment: revisiting the territory. I: *Assessment in education: Principles, Policy & Practice*, 5 (1), s. 77–84.
- Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. I: *Assessment and Evaluation in Higher Education*, 30, s. 175–194.
- Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading in higher education. I: *Assessment and Evaluation in Higher Education*, 34, s. 159–179.
- Sadler, D. R. (2013). Backwards Assessment Explanations: Implications for Teaching and Assessment Practice. I: D. Lebler (red.), *Assessment in Music Education: from Policy to Practice* (s. 9–19).
- Scriven, M. (1967). The methodology of evaluation. I: R. W. Tyler, R. M. Gagne, and M. Scriven (red.), *Perspectives of curriculum evaluation, Volume I* (s. 39–83). Chicago, IL: Rand McNally.
- Selghed, B. (2010). Ett omöjligt uppdrag? Om lärares bedömning och betygssättning. I: *Dilemman i skolan: aktuella utmaningar och professionella omställningar*, s. 82–95. Kristianstad: Kristianstad University Press.
- Skar, G. (2013). *Skrivbedömning och validitet. Fallstudier av skrivbedömning i svenskundervisningen på gymnasiet*. Doktorsavhandling. Stockholm: Stockholms universitet.
- Skar, G. B., & Jølle, L. (2017). Teachers as raters: Investigation of a long term writing assessment program. I: *L1 Educational Studies in Language and Literature*, 17, s. 1–30.
- Skar, G. B., Thygesen, R., & Evensen, L. S. (2017). Assessment for Learning and Standards: A Norwegian Strategy and Its Challenges. I: S. Blömeke & J.-E. Gustafsson (red.), *Standard Setting in Education, Methodology of Educational Measurement and Assessment* (s. 225–241). New York: Springer International Publishing AG.
- Skolinspektionen. (2010). *Kontrollrättning av nationella prov i grundskolan och gymnasieskolan, 2010*. Stockholm: Skolinspektionen.
- Skolinspektionen. (2013). *Tillsyn av bedömning och betygssättning*. Stockholm: Skolinspektionen.
- Skolinspektionen (2015). Ombedömning av nationella prov 2014: ”Processerna spelar roll”. Stockholm: Skolinspektionen.
- Skolverket. (2000). Nationella kvalitetsgranskningar 2000. I: *Skolverkets rapport nr 190*. Stockholm: Skolverket.
- Skolverket (2007). *Bli ännu bättre på att sätta likvärdiga betyg*. Stockholm: Skolverket.
- Skolverket (2009b). *Bedömaröverensstämmelse vid bedömning av nationella prov*. DNR 2008:286. Stockholm: Skolverket
- Skolverket (2011). *Gymnasieskola 2011*. Stockholm: Skolverket.
- Skolverket. (2011). *Ämnesplanen i svenska*. Stockholm: Skolverket.
- Skolverket (2014). *Sambedömning i skolan – exempel och forskning*. Stockholm: Skolverket.
- Skolverket. (2017). *Ämnesplanen i svenska (rev.)*. Stockholm: Skolverket.

- Smaill, E. (2013). Moderating New Zealand's National Standards: teacher learning and assessment outcomes. I: *Assessment in education: Principles, Policy & Practice*, 20 (3), s. 250–265.
- SOU 2016:25 (2016). *Likvärdigt, rättssäkert och effektivt – ett nytt nationellt system för kunskapsutveckling*. Stockholm: Statens offentliga utredning.
- Speer, S. A., & Hutchby, I. (2003). From Ethics to Analytics: Aspects of Participants' Orientations to the Presence and Relevance of Recording Devices. I: *Sociology*, 37 (2), s. 315–337.
- Stolt, S. (2016). *Vägen till vitsordet. En studie av värderande yttranden i bedömarkommentarer på studentexamensuppsatser i modersmål och litteratur*. Doktorsavhandling. Helsingfors: Helsingfors universitet.
- Straub, R. (2000). The student, the text, and the classroom context: A case study of teacher response. I: *Assessing Writing* (7)1, s. 23–55.
- Strömqvist, S. (2007). *Skrivprocessen: teori och tillämpning*. Stockholm: Studentlitteratur.
- Suto, I. & Greatorex, J. (2008). What goes through an examiner's mind? Using verbal protocols to gain insights into the GCSE marking process. I: *British Educational Journal*, 34, s. 213–233.
- Säljö, R. (2000). *Lärande i praktiken: ett sociokulturellt perspektiv*. Stockholm: Prisma.
- Taras, M. (2005). Assessment – summative and formative – some theoretical reflections. I: *British Journal of Educational Studies*, 53(4), s. 466–478.
- Taras, M. (2009). Summative assessment: the missing link for formative assessment. I: *Journal for Further and Higher Education*, 33 (1), s. 57–69.
- Tengberg, M., Skar, G. B., Borgström, E., Sandlund, E., Sundqvist, P., Lötmarker, L. Axelsson, K. & Walkert, M. (2017). *Likvärdig bedömning av elevers språkförmågor. Preliminära resultat från ett ämnesdidaktiskt forskningsprojekt*. Karlstad: Karlstad universitet.
- Utbildningsdepartementet (2010). *Gymnasieförordningen. Svensk författningssamling 2010:2039*. Stockholm: Utbildningsdepartementet.
- Utbildningsdepartementet (2010). *Skollagen. Svensk författningssamling 2010:800*. Stockholm: Utbildningsdepartementet.
- Vetenskapsrådet. (2017). *God forskningssed*. Stockholm: Vetenskapsrådet.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.
- Wenger, E. (1998). *Communities of practice. Learning, meaning and identity*. Cambridge: Cambridge University Press.
- Westbury, I. (2000). Teaching as reflective practice: What might Didaktik teach curriculum? I: I. Westbury & W. Hopmann (red.), *Teaching as a reflective practice: The German Didaktik tradition* (s. 15–40). Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Wertsch, J. V. (2009). *Voices of the Mind. A Sociocultural Approach to Mediated Action*. Cambridge, Mass.: Harvard.

- Wibeck, V. (2010). *Fokusgrupper. Om fokuserade gruppintervjuer som undersökningsmetod*. Lund: Studentlitteratur.
- Wiggins, G. (1998). *Educative Assessment*. San Francisco: Jossey-Bass.
- Wiliam, D. (2007). Content then process: Teacher learning communities in the service of formative assessment. I: D. Reeves (red.), *Ahead of the curve: The power of assessment to transform teaching and learning*, s. 182–204. Bloomington, IN: Solution Tree.
- Wiliam, D. (2011). What is assessment for learning? I: *Studies in Educational Evaluation*, 37(1), s. 2–14.
- Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? I: C. A. Dwyer (red.), *The future of assessment: Shaping teaching and learning*, s. 53–82. Mahwah, NJ: Erlbaum.
- Wiseman, C. S. (2012). Rater effects: Ego engagement in rater decision-making. I: *Assessing Writing*, 17, s. 150–173.
- Wyatt-Smith, C., Alexander, C., Fishburn, D. & McMahon, P. (2017). Standards of practice to standards of evidence: developing assessment capable teachers. I: *Assessment in Education: Principles, Policy & Practice*, s. 24, 250–270.
- Wyatt-Smith, C., Klenowski, V., & Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: A study of standards in moderation. I: *Assessment in Education: Principles, Policy & Practice*, 17, s. 59–75.
- Wyatt-Smith, C., & Klenowski, V. (2013). Explicit, latent and meta-criteria: Types of criteria in play in professional judgement practice. I: *Assessment in Education: Principles Policy & Practice*, 20(1), s. 35–52.
- Vygotsky, L. (1978). *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Zandén, O. (2010). *Samtal om samspel. Kvalitetsuppfattningar i musiklärares dialoger om ensemblespel på gymnasiet*. Doktorsavhandling. Göteborg: Göteborgs universitet.
- Östlund-Stjärnegårdh, E. (1999). Principen och praktiken. En enkätundersökning av lärares syn på bedömning av gymnasieelevers texter. I: *Svenska i utveckling 12. FUMS Rapport 195*. Uppsala: Uppsala universitet.
- Östlund-Stjärnegårdh, E. (2002). *Godkänd i svenska? Bedömning och analys av gymnasieelevers texter*. Doktorsavhandling. Uppsala: Uppsala universitet.

# Bilagor

## Bilaga A: Förfrågan om att delta i ett forskningsprojekt

Hej!

Jag heter Per Blomqvist och är doktorand vid Institutionen för språkdidaktik, Stockholms universitet. Mitt avhandlingsarbete är inriktat på skrivundervisning i ämnet svenska på gymnasiet och jag skulle nu behöva din hjälp.

Finns det någon möjlighet att få träffa dig och några av dina kolleger i svenska (4–6 personer) under hösten för ett samtal om gymnasieelevers skrivande? Vi kommer i så fall att ha ett gruppsamtal utifrån några autentiska elevtexter från en annan gymnasieskola (kurs: svenska 1) och jag är intresserad av era tankar och åsikter om dessa texter. Jag beräknar samtalets längd till ca 60 minuter och ytterligare ca 60 minuter för en kortare genomgång och förberedelser. Sammanlagt ca två timmar alltså.

Jag är väl medveten om att ni har mycket att göra och att ni måste göra prioriteringar i ert arbete. Men kanske finns det någon sammanhållen mötestid eller del av studiedag under höstterminen som skulle kunna användas? För forskningens del skulle det vara ovärderligt att få ta del av ert samtal och jag tror även att ert gemensamma arbete i lärargruppen kan stärkas av det.

Vårt samtal kommer att spelas in och utgöra en del av mitt forskningsmaterial. Era identiteter och er skola kommer däremot inte att kunna spåras i min avhandling.

Med vänliga hälsningar

Per Blomqvist

Doktorand

Institutionen för språkdidaktik, Stockholms universitet

per.blomqvist@isd.su.se



## Bilaga B: Förfrågan om att bidra med elevtexter

Hej på er!

Jag skulle behöva er hjälp!

I höst ska jag genomföra en undersökning som innebär att några gymnasielärare i svenska ska diskutera elevers skrivande på gymnasiet utifrån autentiska elevexempel. För att kunna göra det behövs några elevportföljer från svenska 1 innehållande: 1) utredande text, 2) litterär essä, 3) argumenterande text/manus, 4) "något mer", t.ex. skönlitterära texter, metatexter, reflektioner, gensvar, processtexter eller annat.

Om ni skulle kunna bidra med det materialet vore det fantastiskt! Välj i så fall ut någon eller några elevportföljer från era klasser i svenska 1 med texter som ni själva skulle tycka vara intressanta att diskutera utifrån kvalitetsaspekter och bedömningskriterier.

Digitalt eller papper eller lite av varje – det spelar ingen roll! Texterna kommer att avidentifieras före användning.

Allt gott!

Per

## Bilaga C: Lärargrupperna

### Corneaskolan

	Kön	Ålder	År 1	År 2	Ämneskombinationer och tjänst
Cia	K	43	1998	2004	svenska
Carina	K	46	2001	2001	svenska, historia
Camilla	K	63	1983	1986	svenska, engelska
Christel	K	45	2000	2000	svenska, engelska
Celina	K	48	1998	2001	svenska, tyska
Claire	K	60	1987	2001	svenska, engelska

### Irisskolan

	Kön	Ålder	År 1	År 2	Ämneskombinationer och tjänst
Ingrid	K	29	2014	2014	svenska, samhällskunskap
Isabelle	K	27	2014	2014	svenska, psykologi
Irma	K	45	2007	2014	svenska
Inez	K	42	2004	2010	svenska, spanska
Ida	K	48	1998	2007	svenska, historia, förstelärare

### Scleraskolan

	Kön	Ålder	År 1	År 2	Ämneskombinationer och tjänst
Siv	K	50	1992	1994	svenska
Signe	K	47	1995	1995	svenska, engelska
Stefan	M	40	2000	2000	svenska, engelska, förstelärare
Saga	K	47	1995	1996	svenska, samhällskunskap, förstelärare
Sofia	K	37	2006	2014	svenska
Sara	K	39	2006	2008	svenska, tyska

*Notera:* ”År 1” anger när läraren började arbeta som gymnasielärare medan ”År 2” anger när läraren började arbeta på skolan.



## Bilaga E: Transkriptionstecken

Följande tecken har använts vid transkription av samtal och intervjuer (delvis från Atkinson & Heritage, 1984):

(.)	Paus kortare än en sekund
(3)	Paus uttryckt i sekunder
<u>måste</u>	Starkt betonade ord och uttryck
för om –	Talstreck anger avbrutet tal
”själ å hjärta”	Citationstecken anger citat från elevtext
(mmm)	Parentes anger ickeverbal kommunikation eller kompletterande upplysningar
flytta [upp [flytta ja	Hakparenteser anger överlappande tal och ny rad ny tur eller uppehåll i tal

## Bilaga F: IR-kodning

Följande symboler, värden och definitioner för IR-kodning utgår från Linell (2011).

Symbol	Värde	Definition
>	6	Initiativ som med en fråga eller begäran om omedelbar handling explicit begär respons utan att anknyta till tidigare tur eller topik.
^	5	Samma som > men inte lika begärande initiativ.
..>	5	Initiativ som anknyter till topik från tidigare i samtalet och explicit begär respons.
:>	5	Initiativ som innebär en ”kringkommentar” om en perifer aspekt av ämnet och explicit begär respons.
<=>	5	Initiativ som anknyter till egen föregående tur och explicit begär respons.
<>	4	Initiativ som anknyter till huvudinnehåll för diskussionen i tidigare tur och explicit begär respons.
..^	4	Samma som ..> men utan att begära respons, upprepar något tidigare nämnt.
=>	4	Samma som <=> men utan att tillföra något mer, dvs. upprepar något.
:^	4	Samma som :> men utan att explicit begära respons, dvs. mer som ett konstaterande.
(>	3	Förberedande initiativ om ny topik.
<^	3	Samma som <> men utan att begära respons, dvs. mer som ett konstaterande så kallad expanderade respons.
->	2	Begäran om reparation för att upprepa eller förtydliga något, så kallad uppskjutande fråga.
<	2	Minimal respons utan eget initiativ, dvs. en uppbackning.
-	1	Minimal men otydlig respons.

