Peer-reviewed article

Är vi överens om det? Samspel och samstämmighet i svensklärares bedömningssamtal om gymnasieelevers skrivande

Per Blomqvist*
Institutionen för språkdidaktik, Stockholms Universitet

Sammanfattning
Artikeln beskriver svensklärares interaktion och samstämmighet i samtal om summativ bedömning av gymnasieelevers skrivande. I ett dialogiskt perspektiv (Linell, 2011) undersöks interaktionella mönster avseende dominans, dynamik och koherens i tre svensklärargruppers bedömningssamtal. Dessa interaktionella mönster ställs i relation till uttryck för samstämmighet inom och mellan lärargrupperna. Resultatet visar att lärarna uttrycker hög grad av samstämmighet i bedömningssamtalen men att de bedömningar som lärarna gör enskilt i anslutning till samtalen enbart till viss del överensstämmer med gruppens beslut. Lägst grad av samstämmighet uppvisar den lärargrupp som i samtalen framstod som mest samstämmig. Där kännetecknas beslutsprocesserna av låg intensitet genom att få alternativ till bedömningar prövas. I de lärargrupper där intensiteten är högre genom att flera lärare styr samtalet i olika riktningar och där lärarna tillsammans prövar flera olika förslag på bedömning överensstämmer även enskilda lärares bedömningar i högre grad med gruppens beslut. Bedömningsöverensstämmelsen mellan lärargrupperna är däremot låg.

Nyckelord: *Skrivbedömning; beslutsprocesser; summativ bedömning; sambedömning; samstämmighet*

Abstract
This article reports on a qualitative study of the interaction in Swedish teachers' discussions about summative assessment of students' writing. Interactional patterns of dominance, dynamics and coherens in the decision-making processes are described and analyzed, comparing the degree of assessment consistency between and within the groups. The analysis 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, a variety of suggestions for assessment are made by the teachers. 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.

*Korrespondanse: Per Blomqvist, Institutionen för språkdidaktik, Stockholms Universitet. E-mail: per.blomqvist@isd.su.se

©2018 Per Blomqvist. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), allowing third parties to copy and redistribute the material in any medium or format and to remix, transform, and build upon the material for any purpose, even commercially, provided the original work is properly cited and states its license.

Citation: Per Blomqvist. "Är vi överens om det? Samspel och samstämmighet i svensklärares bedömningssamtal om gymnasieelevers skrivande' Nordic Journal of Literacy Research, Vol. 4, 2018, pp. 1–21. http://dx.doi.org/10.23865/njlr.v4.1050

Keywords: *Writing assessment; social moderation; decision-making process; summative assessment*

Received: October, 2017; Accepted: February, 2018; Published: April, 2018

Inledning

Forskning om skrivbedömning har återkommande visat att lärares bedömningar av elevers texter varierar i samstämmighet (Baker, 2010; Fasting et al., 2009; McNamara, 1996). En åtgärd som har framförts för att komma till rätta med problemet är sambedömning mellan lärare (Adie et al., 2012; Black et al., 2011), något som återkommer i allmänna rekommendationer, råd och modeller till stöd för lärares arbete med bedömning och betygsättning i svensk skola (Skolverket, 2009a; 2013). När lärare tillsammans diskuterar hur styrdokument ska tolkas i relation till bedömningar av elevers prestationer förväntas de utveckla en gemensam förståelse för vad bedömningar ska gälla vilket kan stödja arbetet för likvärdig och rättvis bedömning (Klenowski & Wyatt-Smith, 2010).

Samtidigt råder ovisshet om vilken påverkan sambedömning egentligen har på samstämmighet. En del studier visar att sambedömning av elevers skrivande till viss del leder till ökad samstämmighet (t.ex. Brown et al., 2004; de Eca & Torres, 2005; Elder et al., 2007) andra inte (t.ex. Baird et al., 2004; Jølle, 2014). Antalet studier är dessutom få och de bygger huvudsakligen på sambedömning i expertpaneler eller mellan lärare som genomgått någon form av bedömarträning, inte sambedömning mellan lärare på skolor lokalt. Skar och Jølle (2017) visar exempelvis att lärare efter år av bedömarträning visserligen kan utveckla en ökad samstämmighet men att variationer i enskilda lärares summativa bedömningar av elevers skrivande kvarstår (se även Knoch, 2011; Skar, Thygesen & Evensen, 2017; Weigle, 1998).

I ett mål och kunskapsrelaterat betygssystem som det svenska ställs höga krav på lärares förmåga till samarbete i bedömningsfrågor. Men trots att lärare i svensk skola ägnar betydande tid åt just sambedömning, framför allt bedömning av nationella prov (Skolinspektionen, 2015; SOU, 2016), finns mycket begränsad kunskap om hur sambedömning tar form i samtal mellan lärare på skolor, särskilt avseende hur samstämmigheten påverkas av lärares diskussioner om summativ bedömning. Såväl forskare (t.ex. Adie et al., 2012; Tengberg et al., 2017) som myndigheter (Skolverket, 2013) efterlyser därför forskning som undersöker hur lärare motiverar och kommunicerar summativa bedömningar samt hur lärares gemensamma beslutsprocesser vid sambedömning ser ut. Sådana studier kan bli ett viktigt bidrag till att utveckla lärares bedömarkompetens och minska variationer i olika lärares skrivbedömningar.

I denna studie undersöks sambedömning mellan svensklärare som i grupp diskuterar summativ bedömning av gymnasieelevers skrivande i relation till kursen *svenska 1*. Studien ingår i en större undersökning om svensklärares skrivbedömningssamtal där tidigare resultat visar betydande variationer i *vad* lärare grundar summativa bedömningar av gymnasieelevers skrivande på (Blomqvist, *u.u.*). Fokus för denna studie är att undersöka *hur* lärare interagerar av elevers skrivande: att beskriva, analysera och i

viss mån även förklara lärares beslutsprocesser i relation till deras uttryck för samstämmighet. Forskningsfrågorna är: (1) Hur interagerar lärargrupper i samtal om summativ bedömning av elevers skrivande, (2) Hur ger lärare uttryck för samstämmighet i dessa samtal? (3) Hur samstämmiga är lärare inom och mellan grupperna?

Tidigare forskning

Ett flertal internationella studier har undersökt lärares gemensamma beslutsprocesser när de i par eller grupp diskuterar bedömning av elevers skrivande. Exempelvis undersöker Adie m.fl. (2012) lärares bedömningar som social praktik och hur lärares gemensamma arbete med att bedöma elevers skrivande kan utvecklas över tid. Särskild vikt läggs vid att utifrån sociokulturella perspektiv förklara vad som påverkar lärares bedömningar bortom kriterierna och synliggöra hur en ömsesidig förståelse gradvis kan växa fram genom att lärare kontinuerligt deltar i samtal om skrivbedömning. Även Colombini och McBride (2012) och Wyatt-Smith m.fl. (2010) studerar lärares skrivbedömningar som processer och hur lärargrupper gemensamt kan utveckla bedömarkompetens med hjälp av professionella samtal. Fokus i dessa studier är att analysera lärares strategier vid betygsättning med betoning på beslutsfattande processer.

Berge (2002) har synliggjort och problematiserat brister i de tolkningsgemenskaper som utmärker norsklärares skrivbedömningspraktiker, något som följts av flertalet norska studier av beslutsprocesser i lärares skrivbedömningar. Exempelvis undersöker Jølle (2014, 2015) vilka strategier lärare använder vid parbedömning av elevers skrivande och hur lärares bedömarkompetens kan utvecklas i professionella träningsprogram. Matre och Solheim (2015, 2016) undersöker norsklärares användning av bedömningsredskap samt hur lärare argumenterar och interagerar i samtal om skrivbedömning. Didaktiska implikationer av dessa studier är bl.a. att de betonar vikten av att lärare gemensamt utifrån konkreta elevtexter diskuterar beslut som rör bedömningars inriktning, genomförande och användning. På så sätt kan lärares olika uppfattningar om kvaliteter i elevers skrivande synliggöras, ett meningsutbyte som på sikt kan utveckla ett gemensamt ämnesspråk om skrivande, bedömning och lärande (Colombini & McBride, 2012; Evensen, 2012). Bedömning betraktas alltså som en i grunden normativ aktivitet där lärares möjligheter till att gemensamt problematisera skrivbedömning blir en förutsättning för likvärdighet (Berge, 2009; Skar et al., 2017).

Dialogiskt perspektiv

Den teoretiska utgångspunkten för denna studie är dialogisk. I korthet innebär det att individers meningsskapande betraktas som både dynamiska och sociala processer, något som sker *i* och *genom* interaktion med andra (Linell, 2001; Rommetveit, 1990). Med ett dialogiskt perspektiv på svensklärares samtal om skrivbedömning riktas intresset mot lärarnas yttranden i de kontexter som skapar och omger samtalen. Ett sådant perspektiv kan bidra till att synliggöra vad som utmärker de beslutsprocesser som svensklärare deltar i när de tillsammans med en ämnesgrupp på sin skola diskuterar summativ bedömning av elevers skrivande, något som kan bidra

med kunskaper om hur lärare motiverar och kommuniceras bedömning av elevers skrivande i relation till en specifik undervisningskontext. För att öka likvärdigheten i lärares summativa bedömningar kan sådana kunskaper ge nycklar till att säga något om vad sambedömning gör för samstämmigheten i lärares bedömningar.

Centrala begrepp för studien är *dominans, dynamik* och *koherens* (Linell, 2011; Linell & Gustavsson, 1987). Dominans avser hur initiativ och respons i lärarnas interaktion ger upphov till interaktionella styrkeförhållanden mellan lärarna medan dynamik beskriver själva rörelsen i samtalen: hur initiativ och respons styr samtalet i olika riktningar (Linell & Gustavsson, 1987). Koherens beskriver hur samtalen vävs samman genom att lärarna använder varandras yttranden i beslutsprocesserna vilket blir ett mått på hur väl det dialogiska samspelet fungerar i lärargruppernas bedömningssamtal (ibid.).

Empiri och metod

Empirin består av ljud- och videoupptagningar från tre svensklärargruppers bedömningssamtal samt ett frågeformulär som lärarna besvarade enskilt i anslutning till samtalen. Sammanlagt deltog 16 lärare från tre gymnasieskolor. Gemensamt för lärarna är att de undervisade i svenska innevarande läsår (2014–2015) och hade erfarenhet av att undervisa i kursen *svenska 1*. För övrigt skiljer sig lärargrupperna åt. Här följer en kortare presentation av deltagande svensklärare och gymnasieskolor.

Corneaskolan är en kommunal gymnasieskola, cirka 550 elever, med högskole- och yrkesförberedande program. Samtliga svensklärare är behöriga och den genomsnittliga yrkeserfarenheten uppgår till 19 år med ett intervall på 13 till 31 år. Ingen av lärarna är ämnesansvarig eller förstelärare.[1] Det pedagogiska arbetet på skolan sker i ämnesblandade programlag och svensklärarna träffas sällan för att gemensamt bedöma skrivande eller planera skrivundervisning. Detta sker däremot vid bedömning av nationella prov.

Irisskolan är en fristående gymnasieskola, cirka 450 elever, med högskoleförberedande program. Fyra av fem svensklärare är behöriga och den genomsnittliga yrkeserfarenheten uppgår till 7 år med ett intervall på 1 till 16 år. En av lärarna är ämnesansvarig och förstelärare. Det pedagogiska arbetet på skolan sker både i ämneslag och ämnesblandade programlag, och svensklärarna har ämnestid att gemensamt diskutera elevers skrivande och planera skrivundervisning.

Scleraskolan är även den en fristående gymnasieskola, cirka 550 elever, med högskoleförberedande program. Samtliga svensklärare är behöriga och den genomsnittliga yrkeserfarenheten uppgår till 16 år med ett intervall på 9 till 23 år. En av lärarna är ämnesansvarig och två förstelärare. Det pedagogiska arbetet på skolan sker framför allt i ämnesblandade programlag vilket innebär att svensklärarna sällan träffas som ämnesgrupp. Undantaget är vid sambedömning av nationella prov.

[1]Förstelärare är sedan 2011 en karriärtjänst för lärare i svensk skola. Tjänsten kan tilldelas legitimerade lärare som har undervisat i minst fyra år och som av huvudman bedöms ha intresse av att utveckla undervisningen (Svensk författningssamling, 2013:70, §3).

Sammantaget belyser dessa skillnader att svensklärare på en och samma skola inte kan ses som någon homogen grupp utan består av lärare med olika kunskaper och erfarenheter av att bedöma elevers skrivande. Dessa tre lärargruppers bedömningssamtal blir därför ett exempel på hur skrivbedömning kan ta form på en skola i relation till ett specifikt ämne och en specifik kurs.

Fokusgrupper och frågeformulär

Bedömningssamtalen ägde rum på respektive gymnasieskola och utformades som sessioner i fokusgrupper (Wibeck, 2010). Lärarnas uppgift var att utifrån enskild läsning av ett fokusmaterial bestående av tre elevers två texter, en utredande text och en litterär essä, tillsammans bedöma elevers skrivande summativt i relation till kursen *svenska 1*. Frågan som ställdes till lärarna var: ”Vilket betygsomdöme anser ni att den här elevens skrivande motsvarar?”.

Dessa bedömningssamtal liknar vad Adie m.fl. (2012) benämner ”konferensmodell” som *social moderation* (Wyatt-Smith et al., 2010) där lärare i dialog med varandra försöker komma överens om en bedömning så långt som möjligt. Något som skiljer dessa bedömningssamtal från vad som annars är vanligt vid sambedömning genom just *social moderation* är att lärarna inte enskilt bedömde elevernas skrivande *innan* samtalen påbörjades. Istället ställdes frågan om att summativt bedöma elevernas skrivande först när lärarna i grupp skulle diskutera elevernas texter. Fördelen med detta tillvägagångssätt är att lärarna inte bestämmer sig för ett betygsomdöme i förväg utan fokus hamnar på lärarnas resonemang när de tillsammans löser uppgiften. Nackdelen är att lärarna inte förbereder argument inför samtalet vilket kan göra argumentationen vassare. En annan skillnad är att lärarna inte hade tillgång till de styrdokument som reglerar bedömning av elevers skrivande på gymnasiet utan fick använda sig av sina kunskaper och erfarenheter från att undervisa i kursen *svenska 1*. Även om lärares tillgång till specifika kriterier vid bedömning kan bidra till att utveckla lärares bedömningsförmåga och öka graden av samstämmighet (Parr, 2007) visar studier att bedömningsredskap tenderar att styra lärares bedömningar i en mer instrumentell riktning där de fastnar i tolkningar av enskilda aspekter av texter istället för att se till elevers skrivförmåga i stort (Matre & Solheim, 2016).

Det bör påpekas att betygsomdöme inte är något vedertaget begrepp vid bedömning i svensk skola men används för att förtydliga lärarnas uppgift: att bedöma elevernas skrivande summativt i förhållande till de mål och kunskapskrav som behandlar skrivande i *svenska 1*. Inför samtalen poängterades att lärarna inte behövde komma överens om ett betygsomdöme men att de gärna fick redovisa hur långt de var överens. Ingen moderator ledde samtalen och lärarna avslutade själva samtalen när de uttryckte att de hade fullgjort uppgiften. I *Tabell 1* redovisas längden på bedömningssamtalen.

Det frågeformulär *(Bilaga 3)* som lärarna enskilt besvarade i direkt anslutning till bedömningssamtalen innebar att de dels bedömde elevernas skrivande, ”Vilket betygsomdöme anser *du* att de tre elevernas skrivande bäst motsvarar?” (ange *ett* betygsomdöme från A till F) dels utvärderade samtalen: ”I vilken utsträckning anser

Tabell 1. Samtalslängd per lärargrupp

Lärargrupp	Samtalslängd		
	Elev 1	Elev 2	Elev 3
Corneaskolan	4 m. 01 s.	3 m. 18 s.	4 m. 47 s.
Irisskolan	3 m. 42 s.	4 m. 27 s.	4 m. 40 s.
Scleraskolan	4 m. 56 s.	4 m. 12 s.	4 m. 36 s.
S:a	12 m. 39 s.	11 m. 57 s.	14 m. 03 s.

Notera: m. = minuter, s. = sekunder

du att du fick komma till tals under samtalet?" och "I vilken utsträckning fick samtalet dig att se nya perspektiv eller kvaliteter i elevernas skrivande?" (ange *en* siffra från *6=till stor del* till *1=inte alls*). Dessa frågor fyllde två funktioner. För det första kunde lärarnas enskilda bedömningar av elevernas skrivande synliggöra grad av samstämmighet inom lärargrupperna, för det andra kunde lärarnas uppfattningar om samtalens genomförande och verkan jämföras med mått på interaktionell dominans, dynamik och koherens i bedömningssamtalen.

Interaktionsanalys

Analysarbetet inleddes med att transkriptioner från en tidigare delstudie (Blomqvist, *u.u.*) förfinades avseende interaktionella drag som pausers längd, lärares intonation och gester.

Transkriptionerna analyserades sedan med *initiativ-respons*-analys, så kallad IR-kodning (Linell, 2011; Linell & Gustavsson, 1987). I korthet innebär det att yttranden i samtal poängsätts utifrån grad av initiativ och respons. Högst poäng (värde 6) får yttranden som för in nytt ämne i samtalet och explicit begär respons av övriga deltagare, ett så kallat *fritt begärande initiativ* (Linell, 2011), medan yttranden som enbart responderar genom att bekräfta tidigare, en så kallad *minimal adekvat respons* (ibid.) får lägst poäng (värde 1). Däremellan finns olika grader av interaktionell styrka i enskilda yttranden avseende kombinationer av initiativ och respons (värde 2–5). I *Bilaga 1* finns en förteckning över hur yttranden har kodats och poängsatts.

IR-kodningen ger olika kvantitativa mått som belyser interaktionell dominans, dynamik och koherens i lärargruppernas bedömningssamtal. Exempelvis synliggör lärarnas individuella IR-index, som är 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, som är summan av lärarnas IR-index, den genomsnittliga interaktionella dynamiken och koherensen i samtalet där ett högt IR-index innebär att samtalet byggs av flera högt poängsatta repliker, av initiativ som initierar nya ämnen och explicit begär respons från övriga deltagare (Linell & Gustavsson, 1987).

Linell (2011) använder sedan olika *koefficienter* för att beskriva ett samtals uppbyggnad varav två uppmärksammas i denna studie: S- och B-koefficienten. *S-koefficienten*

avser andelen styrande turer som för in nya ämnen och explicit begär respons från övriga deltagare, så kallade S-turer eller *directing moves* (Linell, 1990a). Genom att se till andelen styrande initiativ i samtalen ges ett mått på både dominans och dynamik i interaktionen. *B-koefficienten* avser i sin tur hur stor andel av turerna som kan anses vara balanserade genom att lärarna anknyter till varandras turer och accepterar varandras inlägg som relevanta, så kallade B-turer. Dessa ger tillsammans med andelen S-turer mått på koherens i interaktionen genom att de synliggör hur samtalen byggs av styrande och balanserande turer. Analysen av initiativ och respons i bedömningssamtalen knyts sedan till lärarnas uttryck för samstämmighet i beslutsprocesserna.

Analysarbetet innebar flera överväganden och ställningstaganden. Framför allt var det ibland svårt att avgöra den interaktionella styrkan i enskilda turer avseende om turer explicit begär respons från övriga lärare eller inte. Exempelvis kan ett yttrande som ”men källförteckningen” som anknyter till huvudinnehållet för diskussionen i tidigare tur kodas som antingen en tur som genom en fråga eller begäran vill ha respons (<>) eller en tur som inte begär respons (<^). Här har prosodin och responsen från övriga gruppdeltagare fått avgöra. Om de uppfattar uttrycket som en fråga ”men källförteckningen?” kodas det som ett begärande initiativ annars som ett konstaterande, ett icke-begärande initiativ. Att avgöra detta innebar flera omlyssningar av materialet vilket i sin tur gjorde transkriptionerna allt mer detaljerade.

Drygt 10% av de sammanlagt 431 replikerna har medkodats av en expert på området. Urvalet gjordes genom att tre partier à 15 turer slumpvis valdes ut från varje lärargrupps bedömningssamtal. Dessa partier inleddes med replik 26, 51, 76, eller 101 vilket gjorde att medkodningen kom att omfatta repliker från olika delar av bedömningssamtalen avseende olika elevers skrivande. Medkodningen resulterade i en reliabilitet på 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. Dessutom påpekar Linell och Gustavsson att variationer i kodning är 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 olika analyser leder till delvis olika alternativ innebär alltså inte att resultaten nödvändigtvis påverkas på ett genomgripande sätt. Med det sagt handlade variationerna här om mindre avvikelser som hade att göra med olika tolkningar av exempelvis hur explicit ett yttrande begär respons eller om introducerad topik kan anses vara ny eller knyta av till redan behandlad topik. Det senare kunde medkodaren inte avgöra eftersom det inte kunde utläsas enbart av urvalet repliker.

Resultat

Resultatet redovisas i tre steg. I ett första steg presenteras kvantitativa mått på interaktionell dominans i lärargruppernas bedömningssamtal, därefter mått på dynamik och koherens. I ett tredje steg presenteras sedan lärarnas uttryck för samstämmighet inom och mellan lärargrupperna.

Tabell 2. IR-profiler för Corneaskolans lärare

Lärare	Turer	S-turer	IR-index
Camilla	23	5	3,65
Carina	38	3	3,39
Celina	20	2	2,85
Christel	39	6	3,26
Cia	43	2	3,19
Claire	7	1	2,57
IR-differens			*1,08*

Steg 1: Interaktionell dominans

Corneaskolans bedömningssamtal omfattar 170 repliker. I *Tabell 2* redovisas hur dessa repliker fördelar sig mellan lärarna, antal styrande initiativ (S-turer) och IR-index.

Som *Tabell 2* visar domineras talutrymmet i Corneaskolans lärargrupp av framför allt tre lärare: Carina, Christel och Cia som bidrar med runt 40 turer var. Därefter följer Camilla med 23 turer, Celina med 20 och Claire som med sina sju turer, som dessutom är korta, knappt deltar alls i samtalet. Ser man till antalet styrande turer är Christel och Camilla de två lärare som genom att föreslå eller argumentera för ett visst betygsomdöme bidrar med flest initiativ som explicit kräver respons. Detta gör att det interaktionella styrkeförhållandet i gruppen delas relativt jämnt mellan majoriteten av lärarna. Camilla är även den lärare med högst IR-index i gruppen (3,65) följd av Carina (3,39), Christel (3,26) och Cia (3,19). Två lärare deltar mer sparsamt i samtalet och bidrar med färre styrande initiativ vilket gör att de har ett betydligt lägre IR-index än övriga: Celina (2,85) och Claire (2,57). Detta ger en IR-differens för gruppen på 1,08 vilket visar på skillnader mellan Camillas och Claires interaktionella dominans i samtalet. Mellan de fyra lärare som aktivt deltar i samtalet ligger dock IR-differensen på 0,46 vilket gör att styrkeförhållandena mellan majoriteten av lärarna i gruppen ändå kan betraktas som mer jämna. Dessa resultat kan betraktas i relation till lärarnas svar på frågan ”I vilken utsträckning anser du att du fick komma till tals under samtalet?” där samtliga lärare svarar *till stor del* (6 poäng på en skala från 1–6).

Irisskolans bedömningssamtal omfattar 124 repliker. I *Tabell 3* redovisas hur dessa repliker fördelar sig mellan lärarna, antal styrande initiativ (S-turer) och IR-index.

Av *Tabell 3* framgår att Irisskolans samtal i mycket hög grad domineras av två lärare: Irma och Ida. Dessa står för över 40 turer var vilket tillsammans är 71% av samtliga turer (88 av 124) och för 8 av 9 S-turer. Därefter följer Isabel som bidrar med 20 turer varav ett styrande initiativ medan Inez och Ingrid bidrar med betydligt färre turer och inga S-turer. Högst IR-index har Ida (3,36) följd av Irma (3,19) medan övriga har ett lägre eller betydligt lägre IR-index: Isabel (3,00), Ingrid som med fyra repliker knappt deltar i samtalet (2,75) och Inez (2,58). Detta ger en

Tabell 3. IR-profiler för Irisskolans lärare

Lärare	Turer	S-turer	IR-index
Ida	45	5	3,36
Inez	11	0	2,58
Ingrid	4	0	2,75
Irma	43	3	3,19
Isabel	20	1	3,00
IR-differens			*0,78*

Tabell 4. IR-profiler för Scleraskolans lärare

Lärare	Turer	S-turer	IR-index
Saga	35	10	3,83
Sara	27	1	3,15
Signe	30	0	2,93
Siv	17	1	3,24
Stefan	28	1	2,75
IR-differens			*1,08*

IR-differens för gruppen på 0,78 vilket förtydligar bilden av styrkeförhållandena mellan lärarna varierar. Dessa resultat kan betraktas i relation till lärarnas svar på frågan ”I vilken utsträckning anser du att du fick komma till tals under samtalet?” där samtliga lärare svarar *till stor del* (tre uppgav 6 poäng och två 5 poäng på en skala från 1–6).

Scleraskolans bedömningssamtal omfattar 137 repliker. I *Tabell 4* redovisas hur dessa repliker fördelar sig mellan lärarna, antal styrande initiativ (S-turer) och IR-index.

Av *Tabell 4* framgår att Scleraskolans samtliga lärare deltar aktivt i samtalet: fyra av fem lärare bidrar med mellan 27 och 35 turer var (Saga, Sara, Signe och Stefan) och Siv bidrar med 17 turer. Fördelningen av antalet styrande initiativ i samtalet är däremot mycket ojämnt där Saga bidrar med hela 10 av 13 S-turer. Utöver det bidrar Sara, Stefan och Siv med varsin S-tur medan Signe inte bidrar med något styrande initiativ alls trots att hon deltar aktivt i samtalet. Detta gör att Saga i betydande grad dominerar gruppens samtal interaktionellt vilket avspeglar sig i gruppens IR-differens på 1,08 där Saga har avsevärt högst IR-index på 3,83 medan övriga lärares IR-index ligger mellan 2,75 (Stefan) och 3,24 (Siv). Även dessa resultat kan betraktas i relation till lärarnas svar på frågan ”I vilken utsträckning anser du att du fick komma till tals under samtalet?” där fyra lärare svarar *till stor del* (tre lärare ger 6 poäng och en 5 poäng på en skala från 1 till 6) och en lärare något lägre (4 poäng).

För att illustrera hur interaktionell dominans kommer till uttryck i lärargruppernas samtal följer här ett exempel från Scleraskolan:

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 med 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 sin 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. Liknande uttryck för dominans återkommer i alla lärargruppers samtal: de lärare som bidrar med flest styrande turer använder sig även i hög grad av initiativ som anknyter till egen föregående tur.

Sammanfattningsvis kan sägas att graden av interaktionell dominans och hur den kommer till uttryck varierar mellan de tre lärargrupperna. På Corneaskolan delar majoriteten av lärarna på talutrymmet och flera lärare bidrar med ett antal styrande initiativ. På Irisskolan är det däremot två av fem lärare som äger talutrymmet och bidrar med majoriteten av gruppens styrande initiativ. På Scleraskolan är talutrymmet mest jämnt fördelat men samtidigt är det enbart en lärare som bidrar med i stort sett samtliga styrande initiativ.

Vilka lärare är det då som dominerar bedömningssamtalen? Tendensen är att lärare med lång yrkeserfarenhet, lärare med ämnesansvar och förstelärare dominerar samtalen interaktionellt. Detta blir särskilt tydligt på Irisskolan där skillnaden mellan de fem lärarnas yrkeserfarenhet är som störst. Ida och Irma har varit svensklärare i 16 respektive 10 år vilket är betydligt längre än de övriga lärarna, varav två är nyexaminerade; Ida är dessutom både ämnesansvarig och förstelärare. Att de interaktionella styrkeförhållandena är ojämna är något som Irisskolans lärare själva uppmärksammar; vid två tillfällen bjuder Irma explicit in resten av lärargruppen att delta aktivare i samtalet genom inbjudande initiativ: ”vad säger ni?” (tur 22) och ”är vi överens

om det?" (tur 59). Sådana initiativ förekommer inte på de två andra skolorna där de interaktionella styrkeförhållandena är mer jämna.

På Scleraskolan dominerar Saga samtalet interaktionellt. Hon är även en av tre lärare i gruppen med längst yrkeserfarenhet och förstelärare. På Corneaskolan där alla lärare har lång yrkeserfarenhet fördelas både talutrymme och andelen styrande initiativ mest jämnt mellan majoriteten av lärarna. Dessa resultat kan ställas i relation till att femton av sexton lärare i det efterföljande frågeformuläret ändå uppger att de *till stor del* fick komma till tals i samtalen.

Steg 2: Interaktionell dynamik och koherens

I *Tabell 5* redovisas balanserande och styrande turer i lärargruppernas bedömningssamtal samt IR-index på gruppnivå utifrån lärarnas IR-profiler (*Bilaga 2*).

Som *Tabell 5* visar byggs bedömningssamtalen framför allt av turer som är motpartsanknutna och som fokalt anknyter till föregående tur genom att explicit begära respons från övriga lärare (<>) eller genom att konstatera, kommentera och utveckla något men utan att begära respons (<^). Dessa balanserande turer (B-turer) uppgår till närmare hälften av samtliga turer i lärargruppernas samtal där lärarna prövar olika förhållningssätt i dialog med varandra istället för att med bestämdhet fastslå något eller abrupt bryta in med ett nytt ämne eller perspektiv och kräva respons på det. Ett exempelutdrag från Corneaskolans bedömningssamtal får illustrera det:

84 <^	Camilla:	nä det e ju årskurs ett
85 <^	Cia:	det e ju inte fel i genren eller –
86 <	Celina:	nä
87 =>	Cia:	eller i språket
88 <	Celina:	nä jag håller med dig
89 <^	Christel:	det e ett enormt drivet språk
90 <	Carina:	jag håller med

Förutom att lärarna genom minimal respons (<) på olika sätt bekräftar, accepterar och i viss mån utvecklar varandras yttranden, "nä jag håller med dig" (Celina, tur 88), och "jag håller med" (Carina, tur 90), är sättet på vilket ämnen initieras ofta glidande genom konstaterande som inte explicit begär respons, exempelvis "det e ju inte fel i genren eller – eller i språket" (Cia, tur 85–87). Sådana *expanderande responser* (Linell & Gustavsson, 1987) som kodas <^ bidrar till bilden av att lärarna använder varandra

Tabell 5. Balanserande och styrande turer

Lärargrupp	B-turer	S-turer	IR-index
Corneaskolan	80/170 (47%)	19/170 (11%)	3,25
Irisskolan	62/124 (50%)	9/124 (7%)	3,15
Scleraskolan	68/137 (50%)	13/137 (9%)	3,20

som resurser för att tillsammans lösa uppgiften. Om man dessutom ser till andelen turer som både uttrycker balans (<>, <^) och acceptans genom minimal respons (<) uppgår de till 71% av Corneaskolans repliker (121 av 170), 69% av Irisskolans (85 av 124) och 77% av Scleraskolans (105 av 137). Bedömningssamtalen kan framför allt beskrivas som ett reflekterande samspel lärare emellan. Här följer ytterligare ett exempel på det, från Scleraskolan:

29 <^	Stefan:	men
		(.) eller det kanske jag inte behöver veta
		(.) för jag tycker inte den når upp till nåt C-kriterie
30 <^	Siv:	jo innehållsmässigt på första texten tycker jag
31 <	Sara:	jag tycker också att den löser uppgiften på –
		((eh)) när det gäller framför allt innehåll
32 <	Stefan:	jo men det kan jag väl kanske hålla med om
33 <^	Siv:	men det gör den inte på andra texten så att då –
34 <	Saga:	jag känner som du *(vänd mot Sara)*
		(.) kanske
		(.) kanske C

Stefans inledande replik utgår från ett tidigare initierat förslag om att elevens skrivande motsvarar C: ”för jag tycker inte den når upp till nåt C-kriterie” (tur 29), vilket Siv bemöter med ”jo innehållsmässigt på första texten tycker jag” (tur 30) och Sara ”jag tycker också att den löser uppgiften på ((eh)) när det gäller framför allt innehåll” (tur 31). Stefan responderar att ”jo men det kan jag väl kanske hålla med om” (tur 32) varpå Siv påpekar att en av texterna inte håller samma nivå (tur 33) medan Saga backar upp Saras tur med: ”jag känner som du (.) kanske (.) kanske C” (tur 34). Återigen byggs interaktionen av expanderande responser som är motpartsanknutna där lärarna använder varandra som resurser för att tillsammans lösa uppgiften och nå ett gemensamt beslut. Lärarna framstår prestigelösa och öppna för att överväga andras åsikter vilket även återkommer i tentativa uttryck som: ”tycker jag” (tur 30), ”kan jag väl kanske hålla med om” (tur 32) och ”jag känner som du” (tur 34). Andelen styrande turer som abrupt för in ett nytt ämne i samtalen är däremot betydligt färre. I Corneaskolans samtal uppgår dessa till 19 (11% av samtliga turer). Detta är något fler än i Scleraskolans (13 st., 9%) och Irisskolans samtal (9 st., 7%).

För att illustrera hur dynamik och koherens tar sig konkreta uttryck i samtalen följer här en jämförelse mellan lärargruppernas beslutprocesser avseende samma elevs skrivande *(elev 1)* där turer som i förhandlingen explicit uttrycker ett betygsförslag har extraherats. Corneaskolans bedömningssamtal utmärks av att flera lärare bidrar med förslag på betygsomdöme innan beslut fattas: ”men den passerar för en E-nivå då?” (Camilla, tur 17), ”för vi förutsätts inte prata om B å D nu då eller?” (Cia, tur 19), ”nä men E tycker jag utan tvekan” (Carina, tur 20), ”å inte D på den här text sex?” (Christel,

tur 22), ”nä men E plus kanske” (Carina, tur 23), ”men alltså nästa nivå är väl C då?” (Camilla, tur 28), ”den når ju inte upp till C-nivå” (Camilla, tur 32), ”men D kan man ju prata om i betyg” (Cia, tur 35), ”om vi nu inte ska sätta ett D?” (Celina, tur 46), ”jag tänkte D spontant” (Christel, tur 52), ”inte uppe på C” (Cia, tur 53), ”om vi sätter E plus då?” (Carina, tur 58) och ”E plus” (Christel, tur 60). Liknande mönster återkommer i Scleraskolans bedömningssamtal: ”den får ett starkt D” (Signe, tur 2), ”jag skulle nog säga E” (Saga, tur 3), ”skulle kunna nå ett D” (Stefan, tur 11), ”D inom parentes” (Siv, tur 12), ”inte upp över E-nivå” (Saga, tur 17), ”övervägande del C-kriterier är det ju inte” (Saga, tur 26), ”jag tycker inte den når upp till nåt C-kriterie” (Stefan, tur 29) och ”jag känner som du (.) kanske (.) kanske C” (Saga, tur 34). Annorlunda är det däremot i Irisskolans bedömningssamtal där antalet förslag på betygsomdöme är betydligt färre och där lärarna utan större diskussion når fram till ett gemensamt beslut: ”den ligger ju som sagt på E” (Ida, tur 7), ”det e inte fel att sätta E på den” (Irma, tur 16), ”inte F i alla fall” (Isabel, tur 23) och ”nä jag tänker också E” (Ingrid, tur 27).

Dessa olika uttryck för dynamik och koherens i lärargruppernas bedömningssamtal kan ställas i relation till frågan ”I vilken utsträckning fick samtalet dig att se nya perspektiv eller kvaliteter i elevernas skrivande?”. Majoriteten av lärarna på de tre skolorna (13 av 16) angav att samtalen bidrog till det genom att ange 4–6 poäng på en skala om 6 poäng. Corneaskolan: fem av sex lärare; Irisskolan: tre av fem lärare; Scleraskolan: samtliga lärare. De tre lärare som angav att samtalet i liten utsträckning bidrog till det, genom att ange 1–3 poäng på en skala om 6 poäng, var lärare som i hög grad dominerade sina lärargruppers samtal och bidrog med flera styrande turer: Camilla, Ida och Irma.

Steg 3: Samstämmighet

Lärargruppernas uppgift bestod i att diskutera ”Vilket betygsomdöme anser ni att den här elevens skrivande motsvarar?”. Inför samtalen informerades lärarna om att de inte behövde komma överens om *ett* betygsomdöme men att de gärna fick redovisa vad de var överens om och vari oenighet bestod. Detta till trots uttrycker lärargrupperna genomgående något som kan liknas vid konsensusbeslut där lärarna i slutet av betygssamtalen ofta manifesterar sin enighet genom att gemensamt summera gruppens beslut. Här är ett exempel på det från Scleraskolans bedömningssamtal om elev 2:

63 <>	Signe:	nä men då blir det ju ett C
64 <	Saga:	absolut ett C minst sagt
65 <	Stefan:	ja
66 <^	Signe:	men hade kursen varit lite längre så känns ju det här som en elev som är på gång
67 <	Stefan:	absolut
68 =>	Signe:	som har det i sig väldigt nära
69 <^	Stefan:	B och på väg uppåt
70 <^	Signe:	ja för vissa passager e ju det –

Liknande mönster återkommer i majoriteten bedömningssamtal (sex av nio). Här är ytterligare ett exempel på det, hämtat från Corneaskolans samtal om elev 3:

174 =>	Carina:	så klarar den ju C
175 <>	Cia:	ett C då
176 =>	Carina:	C minus
177 <	Camilla:	ja
b	(mmm x 2)	
178 <	Celina:	ja

Däremot, i tre av bedömningssamtalen, uttrycks inte samstämmigheten lika explicit genom att lärarna gemensamt summerar gruppens beslut. Istället avslutar en lärare samtalet med att kort konstatera vad gruppens beslut blev, ett konstaterande som ingen lärare yttrar någon invändning emot. Ett exempel på det är när Scleraskolans lärare tidigt i samtalet om elev 3 uttrycker enighet om att betygsomdömet är C: ”det e ju i alla fall ett C” (Saga, tur 91), ”ja det e det ju” (Signe, tur 92) och ”ja det e det absolut tycker jag” (Sara, tur 93). I det fortsatta samtalet diskuteras sedan ytterligare alternativ på betygsomdömen, exempelvis genom att uttrycker Sara att ”man kan se tendenser på en A-nivå” (tur 126) och Signe ”ja B-nivå men med en potential” (tur 127) men lärargruppens gemensamma beslut blir slutligen ett C när Siv avrundar samtalet med ”nä C” (tur 138) vilket bemöts av uppbackningar genom nickningar utan att någon lärare opponerar sig.

Samma mönster återkommer i två av Irisskolans beslutsprocesser. I samtal om elev 2 uttrycker tre lärare tidigt i samtalet enighet om att betygsomdömet är D: ”jag säger D” (Irma, tur 110), ”ja D” (Inez, tur 111) och ”D tycker jag” (Ida, tur 112). Fler betygsförslag diskuteras inte utan samtalet handlar sedan mest om vilka svårigheter det innebär att bedöma elevers skrivande summativt. När Ida sedan avslutar samtalet genom att konstatera ”ja det ligger lite på D-känning där” (tur 125) är det ingen lärare som opponerar sig. I *Tabell 6* redovisas lärargruppernas beslut om betygsomdömen.

Som *Tabell 6* visar överensstämmer lärargruppernas summativa bedömningar avseende en elevs skrivande (elev 1). I övrigt skiljer sig deras summativa bedömningarna åt på mellan ett och fyra betygssteg: Corneaskolans A jämfört med Irisskolans D (elev 2) och Irisskolans B jämfört med Corneaskolans och Scleraskolans C (elev 3). Det finns däremot inget mönster som visar att någon lärargrupp genomgående sätter högre eller lägre betyg än övriga.

Tabell 6. Lärargruppernas betygsomdömen

	Corneaskolan	Irisskolan	Scleraskolan	Diff.
Elev 1	E	E	E	0
Elev 2	A	D	C	4
Elev 3	C	B	C	1

Notera: Diff. anger avsteg från betygssteg mellan lärargruppernas betygsomdömen.

Hur förhåller sig då lärargruppernas gemensamma beslut om betygsomdöme gentemot enskilda lärares bedömningar? I *Tabell 7–9* redovisas lärarnas bedömningar utifrån frågan ”Vilket betygsomdöme anser *du* att elevens skrivande motsvarar?”.

Som *Tabell 8* visar har Irisskolans lärargrupp störst variation i samstämmighet inom gruppen.Och är inte överens om något betyg. Deras betyg på elev 1 skiljer med ett betygssteg genom att en lärare sätter ett D, övriga E. Deras betyg på elev 2 skiljer med tre betygssteg genom att en lärare sätter D, två C och två B. Deras betyg på elev 3 skiljer med fyra betygssteg, genom att en lärare sätter D, två C, en B och en A. Få lärare på Irisskolan sätter det betyg som gruppen kom fram till.

Tabell 7. Lärargruppens och enskilda lärares beslut om betygsomdöme – Corneaskolan

	Gruppens betyg	Enskilda lärares betygsomdömen						
		Camilla	Carina	Celina	Christel	Cia	Claire	Diff. Grupp
Elev 1	E	E	E	C	E	E	E	2
Elev 2	A	B	A	A	A	A	A	1
Elev 3	C	C	C	C	C	C	C	0
Diff. ind.		1	0	2	0	0	0	(3)

Notera: Diff. anger avsteg från gruppens betygsomdöme, på gruppnivå och individnivå.

Tabell 8. Lärargruppens och enskilda lärares beslut om betygsomdöme – Irisskolan

	Gruppens betyg	Enskilda lärares betygsomdömen					
		Ida	Inez	Ingrid	Irma	Isabel	Diff. grupp
Elev 1	E	E	E	E	E	D	1
Elev 2	D	C	D	C	B	B	6
Elev 3	B	B	C	A	C	D	5
Diff. ind.		1	1	2	3	5	(12)

Notera: Diff. anger avsteg från gruppens bedömning, på gruppnivå och individnivå.

Tabell 9. Lärargruppens och enskilda lärares beslut om betygsomdöme – Scleraskolan

	Gruppens betyg	Enskilda lärares betygsomdömen					
		Saga	Sara	Signe	Siv	Stefan	Diff. grupp
Elev 1	E	E	E	E	D	E	1
Elev 2	C	C	C	C	C	C	0
Elev 3	C	B	B	C	C	C	2
Diff. ind.		1	1	0	1	0	(3)

Notera: Diff. anger avsteg från gruppens bedömning, på gruppnivå och individnivå

Betydligt högre samstämmighet redovisas för Corneaskolans och Scleraskolans lärargrupper där avvikelserna mellan enskilda lärares och lärargruppens beslut ligger på tre steg vardera (se *Tabell 7* och *Tabell 9*). Sammantaget är samstämmigheten lägst i de fall där lärarna inte explicit summerar sitt beslut om summativ bedömning: i Irisskolans samtal om elev 2 och elev 3 samt i Scleraskolans samtal om elev 3.

Diskussion

Denna studie ger inblickar i svensklärares interaktion och samstämmighet när de i grupp diskuterar summativ bedömning av gymnasieelevers skrivande. Resultatet visar att den lärargrupp som i samtalen framstod som mest samstämmig uppvisar lägst grad av samstämmighet. Beslutsprocesserna i dessa samtal kännetecknas av låg intensitet där få alternativ till bedömningar prövas. I de två lärargrupper där intensiteten är högre genom att flera lärare styr samtalet i olika riktningar och där lärarna tillsammans prövar en rad olika alternativ och förslag till bedömning överensstämmer även enskilda lärares bedömningar i högre grad med gruppens beslut. Dessa samtal kännetecknas av att mer utrymme ges åt att behandla och problematisera olika aspekter och förslag på bedömning, vilket i sin tur leder till att lärarna i högre grad riktar in samtalen på sådant som de är oeniga om än uppehåller sig vid sådant som de är eniga om (jfr Columbini & McBride, 2012; Matre & Solheim, 2016).

I likhet med tidigare forskning om hur sambedömning påverkar lärares samstämmighet visar alltså denna studie att sambedömning *kan* leda till en viss grad av samstämmighet men att den inte *behöver* göra det. Mycket av den kunskap som lärare bär med sig om bedömning har som all kunskap en tyst dimension (Wyatt-Smith et al., 2010) och vad som framstår som avgörande för om sambedömning ska leda till högre grad av samstämmighet är att lärare så explicit som möjligt uttrycker vad de grundar sina bedömningar av elevers skrivande på. På så sätt blir ett fungerande dialogiskt samspel nödvändigt för att synliggöra olika lärares uppfattningar om kvalitetsnivåer i elevers skrivande (Berge, 2002; Matre & Solheim, 2016).

Resultatet bör dock ställas i relation till de relativt korta bedömningssamtal som analyseras i denna studie. Det är möjligt att lärargruppernas diskussioner hade lett till högre grad av samstämmighet om besluten om summativ bedömning hade ägnats mer tid men samtidigt är det ingen lärargrupp som efterfrågar det. Möjligen motsvarar tidsramarna i någon mån den erfarenhet som lärarna själva uppger att de har från sambedömning, nämligen bedömning av nationella skrivprov i svenska. Utöver det bör resultatet även ställas i relation till lärares möjligheter att gemensamt med ämneskolleger på skolan öva sambedömning. Dessa möjligheter framstår som mycket begränsade på de tre gymnasieskolorna och det saknas rutiner för återkommande organiserade bedömningssamtal mellan ämneslärare. Detta innebär bland annat att yngre och mer oerfarna lärare inte ges någon möjlighet att skolas in i bedömning genom samverkan med äldre och mer erfarna kolleger (Baird et al., 2004;

Crisp, 2013; Little et al., 2003). Inte heller ges lärarna möjlighet att tillsammans didaktiskt utforma skrivuppgifter och kurser utifrån bedömning av elevers skrivande (Black et al., 2011; Tengberg et al., 2017). Den låga samstämmigheten inom och mellan de tre lärargrupperna framstår som en självklar konsekvens av allt detta.

Litteratur

Adie, L. E., Klenowski, V., & Wyatt Smith, C. (2012). Towards an understanding of teacher judgement in the context of social moderation. *Educational Review, 64,* 223–240.

Allal, L. (2013). Teachers' professional judgement in assessment: a cognitive act and a socially situated practice. *Assessment in Education: Principles, Policy & Practice, 20:1,* 20-34.

Baker, B. A. (2010). Playing with the stakes: A consideration of an aspect of the social context of a gatekeeping writing assessment. *Assessing Writing, 15(3),* 133–153.

Baird, J.-A., Greatorex, J., & Bell, J. F. (2004). What makes marking reliable? Experiments with UK examinations *Assessment in Education: Principles, Policy & Practice, 11(3),* 331–348.

Berge, K. L. (2002). Hidden Norms in Assessment of Students' Exam Essays in Norwegian Upper Secondary Schools. *Written Communication, 19(4),* 458–492.

Berge, K. L. (2009). Er tolkningsfellesskap mulig å oppnå i skriveprøver? I H. Otnes, F. Hertzberg, L. S. Evensen O. K. Haugaløkken (red.) (2009) *Tekstvurdering som didaktisk utfordring*. Oslo: Universitetsforlaget.

Black, P., Harrison, C., Hodgen, J., Marshall, B., & Serret, N. (2010). Validity in teachers' summative assessments. *Assessment in Education Principles, Policy & Practice*. *17*, 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?. *Assessment in Education Principles, Policy & Practice*, *18(4)*, 451–469.

Blomqvist, P., *u.u.* (2018) Vad uppmärksammar lärare i samtal om skrivbedömning? Svensklärares normer för beslut om summativ bedömning. *Nordisk tidskrift för allmän didaktik*.

Brown, G. T. L., Glasswell, K., & Harland, D. (2004). Accuracy in the scoring of writing: Studies of reliability and validity using a New Zealand writing assessment system. *Assessing Writing*, *9(2)*, 105–121.

Colombini, C. B., & McBride, M. (2012). "Storming and norming": Exploring the value of group development models in addressing conflict in communal writing assessment. *Assessing Writing*, *17(4)*, 191-207.

Connolly, S., Klenowski, V., & Wyatt-Smith, C. M. (2012). Moderation and consistency of teacher judgement: teachers' views. *British Educational Research Journal*, *38*, 593–614.

Crisp, V. (2013). Criteria, comparison and past experiences: how do teachers make judgements when marking coursework?. *Assessment in Education: Principles, Policy & Practice*, 20, 127–144.

Daly, A., Billington, L., Chamberlain, S., Meyer, L., Stringer, N., Taylor, M., & Tremain, K. (2011). *Principles of moderation of internal assessment*. Manchester, NH: AQA.

de Eca, P & Torres, M. T. (2005). Using portfolios for external assessment: how do teachers make judgements when marking coursework?. *Assessment in Education: Principles, Policy & Practise*, *20*, 127–144.

Elder, C., Barkhuizen, G., Knoch, U. & von Randow, J. (2007). Evaluating rater responses to an online training program for L2 writing assessment. *Language Testing*, *24(1)*, 37–64.

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. *Scandinavian Journal of Educational Research*, *53(6)*, 617–637.

Jølle, L. (2014). A dialogic approach for studying the development of rater competence. *Assessing Writing*, *20*, 37–52.

Jølle, L. (2015). Rater strategies for reaching agreement on pupil text quality. *Assessment in Education: Principles, Policy & Practice*, *22(4)*, 458–47.

Jönsson, A., & Thornberg, P. (2014). Samsyn eller samstämmighet? En diskussion om sambedömning som redskap för likvärdig bedömning i skolan. *Pedagogisk forskning i Sverige*, *19*, (nr. 4–5), 386–402.

Klenowski, V., & Wyatt-Smith, C. (2010). Standards, teacher judgement and moderation in contexts of national curriculum and assessment reform. *Assessment Matters*, 2, 107–131.

Knoch, U. (2011). Investigating the effectiveness of individualized feedback to rating behavior – a longitudinal study. *Language Testing*, *28(4)*, 179–200.

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: Universitetet i Linköping.

Linell, P., & Gustavsson, L. (1987). *Initiativ och respons. Om dialogens dynamik, dominans och koherens.* Linköping: Universitetet i Linköping.

Linell, P. (1990). The power of dialogue dynamics. I I. Markova and K. Foppa (red.). *The Dynamics of Dialogue*. 147–176. Hemel Hempstead: Harvester Wheatsheaf.

Little, J. W., Gearhart, M., Curry, M., & Kafka, J. (2003). Looking at student work for teacher learning, teacher community, and school reform. *Phi Delta Kappan*. *85*, 185-102.

Matre, S., & Solheim, R. (2015). Writing education and assessment in Norway: Towards shared understanding, shared language and shared responsibility. *Educational Studies in Language and Literature vol. 15*, 1–33.

Matre, S., & Solheim, R. (2016). Opening dialogic spaces: Teachers' metatalk on writing. *International Journal of Educational Research*, *80*, 188–203.

McNamara, T. F. (1996). *Measuring second language performance*. New York: Longman.

Rommetveit, R. (1990) On axiomatic features of a dialogical approach to language and mind. I I. Markova & Foppa (1990), 83–104.

Skar, G. B., Thygesen, R., & Evensen, L. S. (2017). Assessment for Learning and Standards: A Norwegian Strategy and Its Challenges. *Standard Setting in Education*, 225–241.

Skar, G. B., & Jølle, L. J. (2017). Teachers as raters: An investigation of a long-term writing assessment program. *L1-Educational Studies in Language and Literature*, *17*, 1–30.

Skolinspektionen (2015). *Ombedömning av nationella prov 2014: "Processerna spelar roll"*. Stockholm: Skolinspektionen.

Skolverket (2009). *Redovisning av regeringsuppdrag att ge förslag på hur det nationella provsystemet bör utvecklas och utformas.* Stockholm: Skolverket.

Skolverket (2013). *Sambedömning i skolan – exempel och forskning.* Stödmaterial. Nr 13:1395. Stockholm: Skolverket.

Smaill, E. (2012). Moderating New Zealand's national standards: teacher learning and assessment out-comes. *Assessment in Education: Principles, Policy & Practice,* 1–16.

SOU 2016:25 (2016). *Likvärdigt, rättssäkert och effektivt – ett nytt nationellt system för kunskapsutveckling.* Stockholm: Statens offentliga utredning.

Svensk författningssamling (2013). *Förordning (2013:70) om statsbidrag till skolhuvudmän som inrättar karriärsteg för lärare*. Sveriges riksdag: Utbildningsdepartementet.

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.

Weigle, S. C. (1998). Using FACETS to model rater training effects. *Language Testing*, *15(2)*, 263–287.

Wibeck, V. (2010). *Fokusgrupper. Om fokuserade gruppintervjuer som undersökningsmetod.* Lund: Studentlitteratur.

Wyatt-Smith, C., Alexander, C., Fishburn, D., & McMahon, P. (2017). Standards of practice to standards of evidence: developing assessment capable teachers. *Assessment in Education: Principles, Policy & Practice*, *24*, 250–270.

Wyatt-Smith, C., Klenowski, V., & Gunn, S. (2010). The centrality of teachers' judgementpractice in assessment: A study of standards in moderation. *Assessment in Education: Principles, Policy & Practice*, *17*, 59–75.

Bilagor

Bilaga 1. IR-koder och poängsättning (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, ofta genom att upprepa något.
:^	4	Samma som :> men utan att explicit begära respons, mer som ett konstaterande.
(>	3	Förberedande initiativ om ny topik.
<^	3	Samma som <> men utan att begära respons och 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, en så kallad *uppbackning*.
–	1	Minimal men otydlig respons.

Bilaga 2. IR-profiler
Corneaskolan

Symbol	**>**	**^**	**…>**	**:>**	**<=>**	**<>**	**..^**	**=>**	**:^**	**<^**	**(>**	**–>**	**<**	**–**
Poäng	**6**	**5**	**5**	**5**	**5**	**4**	**4**	**4**	**4**	**3**	**3**	**2**	**2**	**1**
Camilla	3	1	1		2			4		6		1	5	
Carina		3			5	5		2		14	1	-	8	
Celina			2			2				7		2	7	
Christel	1	3	2	1		2		1		20	2	2	5	
Cia	2					7	2	3		18	1	1	9	
Claire			1							1			5	
Antal	**6**	**7**	**6**	**1**	**7**	**16**	**2**	**10**	**0**	**66**	**4**	**6**	**39**	

Irisskolan

Symbol	**>**	**^**	**…>**	**:>**	**<=>**	**<>**	**..^**	**=>**	**:^**	**<^**	**(>**	**–>**	**<**	**–**
Poäng	**6**	**5**	**5**	**5**	**5**	**4**	**4**	**4**	**4**	**3**	**3**	**2**	**2**	**1**
Ida		5			3	2	2	6	1	17		3	4	2
Inez										7		1	4	
Ingrid						1				2			1	
Irma		3			1	5	1	2	1	19	2		9	
Isabel		1			2	3				6	1	2	5	
Antal	**0**	**9**	**0**	**0**	**6**	**11**	**3**	**8**	**2**	**51**	**3**	**6**	**23**	**2**

Scleraskolan

Symbol	**>**	**^**	**…>**	**:>**	**<=>**	**<>**	**..^**	**=>**	**:^**	**<^**	**(>**	**–>**	**<**	**–**
Poäng	**6**	**5**	**5**	**5**	**5**	**4**	**4**	**4**	**4**	**3**	**3**	**2**	**2**	**1**
Saga	2	6	2		3	5		2		9			6	
Sara		1			2	1	1	3		11	1		7	
Signe						3		3		17			6	1
Siv		1			1	5				4	1		5	
Stefan		1				3		1		10			13	
Antal	**2**	**9**	**2**	**0**	**6**	**17**	**1**	**9**	**0**	**51**	**2**	**0**	**37**	**1**

Bilaga 3. Frågeformulär

A. Lärarbedömning av elevtexterna

Vilket betygsomdöme anser *du* att de tre elevernas skrivande bäst motsvarar? Ringa in *ett* alternativ:

Elev 1 (text 1 och 6)	A	B	C	D	E	F
Elev 2 (text 2 och 5)	A	B	C	D	E	F
Elev 3 (text 3 och 4)	A	B	C	D	E	F

B. Utvärdering av gruppsamtalet

I vilken utsträckning anser du att du fick komma till tals under samtalet? Ringa in *ett* alternativ:

	6	5	4	3	2	1
Till stor del						*Inte alls*

I vilken utsträckning fick samtalet dig att se nya perspektiv eller kvaliteter i elevernas skrivande?

	6	5	4	3	2	1
Till stor del						*Inte alls*