

Per Sund is a docent in science education. Per's research interest is science education (SE) and environmental and sustainability education (ESE) from a teacher's perspective. He is involved in several research projects regarding assessment in science education. He is the link-convenor of the Environmental and Sustainability Education Research Network, ESER (no: 30) in the European Educational Research Association, EERA. The ESER network has more than 200 researchers from 32 countries. Per trains teachers and student teachers at national and at international level in global south in SE and ESE.

Louise Sund is an experienced secondary school teacher, teacher educator, and researcher at Mälardalen University's School of Education, Culture and Communication. She is a member of the research group SMED (Studies of Meaning-Making in Educational Discourses), a cross-university research group in the field of didactics and educational science. Her research interests include philosophical perspectives and approaches to education and sustainable development. She has a particular interest in poststructural criticism of universalistic approaches in environmental and sustainability education.

## PER SUND & LOUISE SUND

Mälardalen University, School of education, communication and culture, Sweden  
per.sund@mdh.se  
louise.sund@mdh.se

# ”Alla gör fel?!” – Hinder för lärares bedömning av elevers praktiska förmågor under ett nationellt prov

## *Abstract*

*Science teachers regard practical work as important which helps students to learn science. Besides theoretical knowledge practical work is considered to be an integral part of science education. This implies that in addition to theoretical science tests, there is a need to find new ways of assessing students' practical work. This empirical case study investigates two secondary school science teachers' assessments of students in three separate groups in the practical part of a Swedish national test in chemistry. Data is gathered using two video cameras and three pairs of spy camera glasses. The results show that individual and independent assessments of student's practical abilities are difficult due to the social interactions that take place and physical sources of error that occur in this type of laboratory setting. The results are discussed in relation to how teachers' assessments of students' practical abilities can be conducted and made more equal and fairer.*

## INTRODUKTION

Det nationella och systematiska testandet av elever har ökat markant i Sverige liksom i många andra delar av världen (Broadfoot & Black, 2004; Lundahl, Roman, & Riis, 2010). Syftet är att utvärdera elevers ämneskunskaper och förmågor men kan också utgöra en grund för en mer likvärdig nationell bedömning (Lundqvist & Sund, In Press; Sund, In Press). Ett sätt att försöka säkerställa att lärare bedömer elever på ett likvärdigt sätt är att utveckla nationella prov med praktiska moment. Inom detta moment ska lärare genomföra en given systematisk undersökning med sin klass, med utgångspunkt i detaljerade instruktioner, och bedöma elevernas prestationer med stöd av bedömningsanvisningar. I teorin verkar allt enhetligt och klart men i praktiken uppstår en mängd frågor: Klarar lärarna av att bedöma elevers praktiska förmågor? Kan lärare med hjälp av instruktionerna åstadkomma jämför-

bara laborativa provmiljöer och säkerställa likvärdigheten i bedömningarna? Vilka hinder finns för lärares bedömning av praktiska förmågor? Bedömning av praktiska förmågor är i sig komplext och det kompliceras ytterligare med praktiska provmoment inom ett nationella prov (Abrahams & Reiss, 2012; Harlen, 1999). Fokus i denna studie är olika slags hinder för lärares likvärdiga bedömning av elevers individuella praktiska förmågor vid en provsituation i en laboratoriemiljö.

Ämnesprov under samlingstermen "nationella prov" infördes i den svenska grundskolan i 1994 och har fram till 2009 genomförts i svenska, engelska och matematik. Från och med läsåret 2009/10 skriver alla elever i skolor 9 nationellt prov i ett av de naturorienterande ämnena (NO) fysik, kemi eller biologi. Enligt Skolverket är det huvudsakliga syftet med det nationella provsystemet att stödja lärarnas arbete med att genomföra likvärdig, rättvis bedömning och betygsättning av elever samt att ge underlag för analyser om i vilken utsträckning kunskapskraven uppfylls på skolnivå, på huvudmannanivå och på nationell nivå (Lundahl, 2010; Skolverket, 2010).

## BAKGRUND

Det råder en viss förvirring om definitionen av begreppet 'praktiskt arbete' inom undervisning i naturvetenskap (Dillon, 2008) och de flesta skolförsök i naturvetenskap följer en skriftlig instruktion som liknar ett recept (Abrahams & Millar, 2008). Skolverket kallar praktiskt arbete för en systematisk undersökning inom de nationella proven (Skolverket, 2010). Många lärare som undervisar i naturvetenskap anser att praktiskt arbete i skolämnesområdet NO är en viktig del av att lära sig naturvetenskap (Abrahams & Millar, 2008; Högström, Ottander, & Benckert, 2012; Watts & Wilson, 2013). Huvudargumenten är att eleverna lär sig den naturvetenskapliga metoden med tillhörande begrepp men också att elever ofta är positivt inställda till praktiskt arbete (Dillon, 2008; Högström, Ottander, & Benckert, 2010).

Trots att praktiska förmågor i naturvetenskap är högt skattade och beskrivna i forskningslitteraturen som viktiga råder det en brist på tydlighet om vad de står för och hur de kan bedömas (Reiss, Abrahams, & Sharpe, 2012). Bedömningen av hur en förmåga används är sammanvävd med kännedomen om ämnesområdet inom vilket förmågorna används. Av det skälet utvärderas alltid en kombination av förmågor och ämneskunskap (Harlen, 1999; Roberts & Gott, 2006). En utmaning för praktiska utvärderingar är att engagera elever att göra sitt bästa. Forskning visar att ju mer en intressant och viktig en uppgift uppfattas desto mindre variationer i elevers resultat kan förväntas (Roberts & Gott, 2006). Det finns fler faktorer att reflektera över i samband med bedömning. Laborationslokaler som provmiljö påverkar också utförandet eftersom den signalerar att det är en speciell typ av tänkande som är relevant under genomförandet av laborationer. Kopplingen mellan ett praktiskt prov och läroplanens mål är inte alltid helt given. I en tidigare studie framkom det exempelvis att den vetenskapliga metoden som är ett viktigt läroplansmål inte nämndes av någon deltagande lärare i samband med bedömning av deras elevers praktiska arbete (Ottander & Grelsson, 2006). Sammanfattningsvis kan man säga att bedömning av förmågor under praktiskt arbete i en laboratoriemiljö är svårt och komplicerat.

## SKOLVERKETS SYSTEMATISKA UNDERSÖKNING

Enligt det centrala innehållet i kursplanen för kemi i åk 9 ska eleverna kunna genomföra systematiska undersökningar och olika separations- och analysmetoder (Skolverket, 2011b). Preciseringsen av vilken typ av analyser elever ska lära sig framgår inte i kursplanerna. Skolverkets ambitionsnivå för att utvärdera det centrala innehållet med nationella proven i åk 9 är hög och erbjuder lärare möjligheter att bedöma elevers olika slags naturvetenskapliga kunskaper som omnämns i kursplanerna (Skolverket, 2011a). Det nationella provet består av ett teoretiskt delprov (Del A) och en systematisk praktisk undersökning (Del B). I det teoretiska provet får de tillfälle att tillämpa sina kemikunskaper både inom

det vetenskapliga fältet och i samhällsdiskussioner. Det handlar med andra ord om att använda sina kunskaper och förklara samband i både naturen, samhället och i elevernas vardagliga liv. Förutom att visa olika tillämpningar av teoretiska kunskaper erbjuds också elever att visa sina praktiska experimentella förmågor i en systematisk undersökning. Syftet med denna praktiska deluppgift B är att erbjuda lärare tillfälle att studera, eller *mäta* som det står i lärarinstruktionerna, elevers förmåga att genomföra en systematisk undersökning.

Det systematiska testet i Del B nationella provet i kemi för åk 9 är uppdelat i tre delar, planering (I), genomförande (II) och utvärdering (III) och den rekommenderade tiden för varje del är 30 min. Del I och III utvärderar naturvetenskapliga idéer och begrepp medan Del II observerar elevers utförande. I Del I *planerar* eleverna undersökningen. När eleverna planerat klart så *genomför* de undersökningen i Del II självständigt eller med hjälp av en instruktion med en ’öppen’ karaktär. I denna studie bedömer lärarna i samtliga fall att respektive elev ska använda den medföljande laborationsinstruktionen under genomförandet. De förväntas svara på några utredande frågor men i övrigt lämnas ett stort utrymme till elevens egen planering och initiativ. Del II är i fokus för denna studie. I Del III *utvärderar* de sin undersökning. De sammanställer resultaten i tabellform, och förklarar mer i detalj vad som hänt under försöket. Till sist ombeds de att ge förslag till förbättringar av undersökningen. Del II är den praktiska delen med flest bedömningsdetaljer i anvisningarna. Del I har färre bedömningsdetaljer än Del II medan Del III har färre detaljer men mer teoretiska bedömningsdetaljer om villkoren för naturvetenskapens kunskapsproduktion. Del I och Del III lämpar sig också väl att genomföras skriftligt i vanlig skolsal medan Del II ovillkorligen är av praktisk natur. I bedömningsanvisningarna finns ingen skrivning om de tre delarnas inbördes betydelse för den sammanlagda bedömningen av en elevs prestation. Den skriftliga informationen i Del I och III som elever lämnat in kan studeras av lärare i flera omgångar medan Del II är ett moment som endast kan bedömas under ett enstaka tillfälle. Det gör att läraren är speciellt fokuserad under Del II och antecknar mycket. Den kan därför få ett större värde i den sammanlagda bedömningen än de andra två delarna. Enligt anvisningarna rekommenderades lika mycket tid för alla tre delarna men genomförandet i denna studie visade att både elever och lärare ägnade lång tid till Del II och kortast tid till Del I.

De nationella proven har lång sekretesstid vilket innebär att de provuppgifter som förekommer i nationella prov inte får spridas eller användas utanför specifika provdatum. Denna sekretess gör det omöjligt för denna studie att i detalj redovisa den uppgift eleverna utför i samband med den aktuella systematiska undersökningen. Detta medför med nödvändighet att beskrivningarna i denna studie är vaga. Den systematiska undersökningen, Del II, går ut på att undersöka tre vätskor. Till sin hjälp har eleverna tillgång till lämpliga kemikalier och laborationsutrustning. Lärarna blandar tre stamlösningar enligt anvisningar och det är dessa som eleverna undersöker. De förväntas jämföra en egenskap hos vätskorna och därefter placera dem i en fallande skala enligt instruktionens frågeställning.

Instruktioner till läraren handlar om vilka moment som ingår samt tidsåtgång. Den beskriver säkerhetsutrustning, laborationsutrustning och kemikalier som ska finnas till hands och om hur stamlösningarna tillreds men ger inga instruktioner om hur kommunikationen mellan lärare och elever eller mellan olika elever i klassrummet förväntas gå till. I information till eleverna beskrivs att ’det nationella ämnesprovet ger dig möjlighet att visa vad du kan i kemi’. Därefter beskrivs tre olika betygsnivåer. Den svenska betygsskalan är i sin helhet fem-gradig A-E. Det nationella ämnesprovets instruktioner till eleverna utger att det erbjuder dem möjligheter att visa vad de kan på betygsnivåerna E, C och A, där A är högsta betyg. Kunskaper mellan A och C utgör betyget B samt mellan C och E ger betyget D. Bedömningsanvisningarna ger läraren detaljerade instruktioner om vad som ska bedömas i elevernas arbete. Lägsta betygsnivå handlar exempelvis om säkerhetsarbete och hur aktuella kärl och kemikalier hanteras. Nästa betygsnivå omfattar val av laborationsutrustning med exempelvis olika mättnogsnoggrannhet. På den högsta betygsnivå ska läraren exempelvis uppmärksamma på vilket sätt elever läser av skalor. För att läraren ska ha möjlighet att observera de högre betygsnivåerna behöver läraren vara mycket nära enskilda elever och i detalj följa deras arbete.

## SYFTE

Syftet med denna fallstudie är att undersöka förutsättningarna för lärare att göra individuella bedömningar av elevers laborativa förmågor under ett praktiskt moment. Olikheter i de tre gruppernas sätt att välja laborationsutrustning i inledningen av det praktiska provet inspirerade författarna till att studera förutsättningarna för lärares likvärdiga bedömning.

Studien besvarar forskningsfrågan:

- Vilka hinder finns för lärares likvärdiga bedömningar av elevers individuella prestationer?

Resultaten synliggör ett antal praktiska och sociala hinder för lärares möjligheter till en likvärdig bedömning. I diskussionen diskuteras hur dessa hinder i en bedömningspraktik kan undvikas eller om ambitionsnivå behöver ändras.

## METOD

Det ursprungliga syftet med att spela in elever som laborerar i NO i denna studie var att studera deras sätt att 'tala' naturvetenskap. Vid dörren klassrummet säger läraren att det är första dagen som den systematiska undersökningen inom det nationella provet ska genomföras. Nationella prov är sekretessbelagda men efter överväganden om att författaren har skrivit på sekretessförbindelser i ett tidigare forskningsprojekt om nationella prov påbörjades datainsamlingen.

Studieobjektet är lärare och elever som genomför en systematisk undersökning. Det praktiska delprovet utförs här av två skolklasser. Det är vanligt att i Sverige dela upp två skolklasser i tre laborationsgrupper i NO för att garantera säkerheten i laboratoriet. Totalt är det 38 elever som genomför delmomentet och det är två lärare, samma vid varje tillfälle, som handhar utrustning och observerar. Den ena läraren är klassens ordinarie lärare medan den andra läraren är NO-lärare på samma skola men som inte känner de aktuella eleverna särskilt väl. Studien följer alla tre laborationsgrupperna i en laborationslokal/NO sal under tre på varandra efterföljande dagar. Det innebär att villkoren för de tre grupperna är mycket lika. Datainsamlingen sker med två fasta videokameror och tre par spionglasögon i ca 60 min/grupp. Datamängden är totalt ca femton timmar. En fast videokamera står längst fram i klassrummet och filmar specifikt området där laborationsutrustning och kemikalier är placerade. En videokamera står längst bak och filmar mer översiktligt. Spionglasögon kan beskrivas som glasögon med vanligt glas. Mitt på skälmen ovanför näsan finns en liten kamera och mikrofon monterad. Spionglasögon fångar upp det eleven tittar på och skriver, med god bild- och ljudkvalitet. Tre par spionglasögon delas ut till elever och urvalet av dem handlar i första hand om deras frivillighet, men hänsyn tas också till kön och deras fysiska placering i klassrummet. Lärarna fick inga glasögon för det ursprungliga forskningssyftet för studien var att studera elevers språk i det naturvetenskapliga klassrummet (se ovan) men syftet ändrades till istället till att studera hinder för lärares bedömning i inledningen av provet då grupp två började laborera. Gruppernas olika val, inte enskilda elevers val, av laborationsutrustning indikerade starka och intressant sociala fenomen i situationen.

De fasta kamerorna fångar översiktbilder och spionglasögon enskilda detaljer och samtal. Under provsituationen förväntas eleverna vara tysta, men korta lågmälda konversationer förekommer och de fångas upp med spionglasögonen. Spionglasögonen fångar upp elevers individuella aktiviteter men även vad kamrater intill gör. Den samlade inspelningsutrustningen gör det möjligt att exakt se vad flera elever gör individuellt, i vilken ordning och hur de arbetar med utrustningen, vem de talar med och oftast vad också lärarna säger samt var de befinner sig i laborationssalen. Det innebär att det i samband med analysen är enkelt att kontrollera klassrumsobservationerna ur flera perspektiv. Det går i praktiken att kors-referera observationer med hjälp av inspelningar med olika kameror vid samma tidpunkt. Flera viktiga detaljer i bedömningsanvisningarna för respektive betygsnivå framkom tydligast med hjälp av spionglasögonen.

Eleverna som tilldelades glasögon upplystes noga om deras funktion och ombads att lämna glasögonen på bänken om de lämnade klassrummet. Ur ett forskningsetiskt perspektiv kan man tänka sig att forskaren kan få syn på detaljer av mer otillbörlig karaktär som exempelvis att blicken ’fastnar’ på detaljer av mer personlig karaktär. I detta material fanns inget som kunde uppfattas som integritetskränkande men om så varit fallet hade sådana sekvenser uteslutits ur datamaterialet. Ett annat etiskt spörsmål handlar om lämpligheten att låta elever ha spionglasögon under ett nationellt prov då eleverna kan uppfattas bli störda under själva provsituationen. Eftersom spionglasögon i praktiken liknar vanliga skyddsglasögon som alla elever har i laborations salen tog det inte mer än någon minut innan de aktuella eleverna glömt bort att de hade andra glasögon på sig.

### Analysmetod

Analysarbetet består av att studera all data upprepade gånger. Det är viktigt att jämföra data från flera olika inspelningsskällor i real tid. Till forskningsfrågan utvecklas tre specifika analysfrågor som används i analysen av data:

- 1) I vilka situationer blir elevernas sociala interaktioner synliga?
- 2) Vad begränsar lärares möjligheter till att utföra likvärdiga bedömningar?
- 3) Vilka fysiska faktorer hindrar en likvärdig och rättvis bedömning?

Data analyserades med en tolkande och iterativ metod (Kvale & Brinkmann 2009; Wolcott, 1994), där analysfrågorna utkristalliserade tre olika kategorier med respektive analysfråga. Analyscykeln upprepas tills kategoriseringarna ligger fast.

### RESULTAT

Forskningen i denna studie genomförs i en provsituation vilket innebär att eleverna inte förväntas tala med varandra eller läraren. Detta står inte uttryckligen i instruktionerna men det nationella provet är generellt ett underlag för en enskild summativ bedömning (Skolverket, 2010), vilket i praktiken innebär att eleverna bör vara tysta. Det innebär att resultatet i studien inte kan förmedlas som excerpt av lärares och elevers muntliga uttryck i första hand. Istället presenteras i det följande avsnittet tre klassrumsbeskrivningar av gruppernas provsituation. Beskrivningarna är något olika i sin karaktär i ett försök erbjuda läsaren ett rikt empiriskt material trots att provsekretessen har en begränsande inverkan på presentationen av studien. Beskrivningarna är från början en form av tolkningar av verkligheten och därför presenteras de som resultat. Ett sammanfattande resultatavsnitt presenteras efter beskrivningarna.

### Grupp ett (16 elever + 2 lärare)

Generellt sett är klassrumssituationen under detta provtillfälle ganska livlig där många elever diskuterar, sitter bredvid varandra (inte tillräckligt med bänkar) och tittar i varandras anteckningar. Lärarna hanterar inledningsvis huvudsakligen material och kemikalier, samt sanerar efter några spillolyckor och ett glaskross, men efterhand hinner de också med att utföra den planerade bedömningen.

Inledningsvis ger lärarna en kort information om vad som ska göras enligt instruktionerna. Eleverna skriver en planering (Del I) för laborationen men läraren bedömer att alla bör följa den färdiga laborationsanvisningen som följer med provet. Under tiden som eleverna tar på sig skyddsförkläden och skyddsglasögon ställer lärarna fram laborationsutrustning, nödvändiga kemikalier samt de lösningar som ska undersökas. Allt ställs fram på katedern och ett på ett rullbord intill. En flicka tar först av alla elever laborationsutrustning. Det var ett provrörställ med provrör. I denna grupp hämtar samtliga elever först provrör och provrörställ samt de kemikalierna som ska tillsättas, men provrören används inte senare mer än av ett fåtal elever. De flesta elever använder istället 100 ml bägare (små bägare) som blandningskärl för vätskorna.

Vid katedern och rullbordet står 6-8 elever samtidigt och häller upp kemikalier. Det är trångt och det pågår ett myller av aktiviteter under ett lågt sorl. Lärarna står intill vid tavlan en meter ifrån dem och tittar på. När eleverna hämtar provlösningarna (stamlösningar) häller alla elever upp vätskorna ur tre 3000 ml bägare (stora bägare) märkta A, B och C direkt i tre små bägare. En del elever märker sina bägare andra inte. Om de råkar hälla upp för mycket vätska i de små bägarna häller de tillbaka överskottet direkt i stamlösningarna A, B och C. Eleverna häller vätskor fram och tillbaka mellan kärlen och spiller lite ibland.

En elev med spionglasögon som väntar på sin tur att hälla upp vätskor frågar de båda lärarna vid tavlan om det är så här man ska göra. Han står intill dem vid tavlan och tittar ut över sina kamrater som står tätt runt bordet med kemikalier. Den ene läraren svarar att *"det går att göra så här"*. Eleven går förbi den första läraren, vänder sig och frågar den andra läraren om det är rätt sätt. Läraren svarar inte men eleven fortsätter att titta på läraren som börjar le när den ser att eleven håller kvar blicken och fortsätter titta. Eleven utbrister högt så att alla elever hör - *"Alla gör fel?!"* - han tittar först på läraren och sedan på sina kamrater på andra sidan bordet. De fortsätter sin aktivitet efter att ha snabbt kastat en blick mot lärarna vid tavlan.

Data från spionglasögonen visar att under tiden som eleverna hämtar kemikalier frågar de varandra oavbrutet olika praktiska frågor. Är det här A? Hur mycket ska man ha? Ska man ha natrium (natriumhydroxid)? Är bägarna slut? Det tar ca tio minuter för alla elever att hämta utrusning och hälla upp de flesta vätskorna, men också fortsättningsvis rör sig elever från sin bänk till och från katedern för att hämta glasstavar, dropprör, mer kemikalier. Inledningsvis när eleverna börja komma igång med att arbeta mer självständigt vid sina platser går lärarna runt och småpratar med enskilda elever. Eleverna har många frågor om vad som ska göras och lärarna svarar på saker som rör material men oftast inte själva uppgiften. Det är frågor om material, mängder men också av redovisningsnatur som var resultat och tabeller ska redovisas.

Två elever hämtar efter en stund mätglas som står intill de små bägarna på kemikaliebordet. De är de enda två som använder mätglas i hela gruppen. En av dem har spionglasögon och med hjälp av data från dessa går det att observera att den elev som uttryckte sin oro över att alla kanske gör fel är den enda som använder mätglas och dessutom gör det på ett sätt som enligt bedömningsanvisningarna är på högsta betygsnivå. Han mäter på nytt upp alla vätskevolymerna med god noggrannhet vid sin laborationsbänk.

En av lärarna har ett observationsprotokoll med sig i handen. Han går runt och tittar och småpratar lite och ibland går han åt sidan i ett hörn av klassrummet för att göra en del noteringar på sitt papper. Med jämna mellanrum går lärarna runt tillsammans och pratar med eleverna och sedan stannar lärarna upp och diskuterar med varandra 30-40 sekunder innan de fortsätter att gå runt. Eleverna blir färdiga med den praktiska delen i lite olika takt, skriver i resultat och börja diska efter sig.

### **Grupp två (12 elever + 2 lärare)**

I grupp två är det generellt en lugnare klassrumssituation jämfört med grupp ett. På katedern och på ett rullbord intill står all laborationsutrustning liksom provvätskorna som ska undersökas. Stamlösningarna i de stora bägarna är exakt de samma som grupp ett tidigare använt. Det går att avgöra genom märkningens utseende och de något lägre vätskenivåerna. Efter en viss väntan i hela gruppen reser sig en av pojkarna och tar tre bägare och ett dropprör, och börjar därefter överföra provvätskor ur de stora bägarna till de små bägarna. En grupp om 7-8 elever står runt bordet och småpratar med varandra när de hämtar olika lösningar. De använder ett dropprör för att föra över vätska i mindre bägare och det tar flera minuter. När nästan alla elever fyllt i sina bägare med dropprör tar en flicka initiativet och börjar hälla över provvätska direkt från den stora bägaren i de små bägarna. Samma händelser inträffar som i den första gruppen. De som häller över en alltför stor volym provvätska i en liten bägare häller tillbaka överskottet i den ursprungliga stamlösningen.

I denna grupp är det en pojke som tar laborationsmaterial först. Då han närmar sig katedern står bägarna lättast åtkomliga. Alla elever i gruppen använder bägare. En flicka tar ett mätglas och håller över vätska i den. Hon mäter inte upp någon volym utan använder den som transportkärl.

När eleverna börjar att laborera går lärarna runt en liten stund och småpratar med respektive elev. De stannar upp längst bak i klassrummet, och observerar, och utbyter erfarenheter med varandra under tiden som de blickar ut över klassrummet. Det är lugnt i klassen och det hörs glasklirr från omrörningen av vätskor. Eleverna observerar varandras blandade vätskor. En flicka säger till en lärare – *”jag förstår inte”*. *”Det gör du visst, du kan”* – svarar läraren. Läraren går fram till en pojke som sitter framför flickan. Han har liksom flickan hållt lika mycket vätska i alla tre provlösningarna (de bör var olika mängder anm.). Läraren frågar hur han vet det. Eleven svarar att han insett det när han hällde ur mätglaset. Flickan, som har spionglasögon, har gjort likadant och hört samtalet mellan läraren och pojken framför, hon frågar läraren – *”kan man göra om?”* Läraren ser på klockan och säger – *”det hinner du”*. Flickan och pojken diskar och gör om försöket. Lärarna står på nytt om diskuterar med varandra längst bak i klassrummet. De två elever som gör om försöket blir sist kvar i klassrummet.

### Grupp tre (10 elever + 2 lärare)

Inledningsvis visar den ena läraren gruppen var alla kemikalier finns och vilka de är. Den andre läraren uppmanar dem att läsa de utdelade instruktionerna noggrant innan de börjar laborera.

Generellt är klassrumssituationen lugn och vid katedern står allt material och provvätskorna (stamlösningarna), samma vätskor som grupp ett och två redan använt. Det kan man se på att vätskenivåerna i de stora bägarna som nu är ännu lägre än när grupp två började. Den första eleven som väljer laborationsutrusning är en pojke som tar ett mätglas vilket leder till att nästan alla senare tar med ett mätglas till sin arbetsplats. Ingen använder dem senare. Mätglaset var placerade i mitten och längst fram på kemikaliebordet och provrören står på rullbordet intill. Ingen i denna grupp hämtar eller använder provrör. Eleverna använder små bägare för stamlösningar och andra vätskor. Den andra eleven som kommer fram till katedern tar tre bägare. Han märker upp dem vilket leder till att andra elever som hämtar bägare uppmärksammar att det redan finns märkningar på flera bägare. De sorterar dem i bokstavsordning. En elev börjar ta upp stamlösning med dropprör men han ser att kamraten mitt emot på andra sidan bordet börjar hålla upp lösning i bägare, då slutar han att använda droppröret. Han håller direkt i små bägare och inom kort håller alla elever stamlösningarna direkt i små bägare med påföljd att eventuellt överskott liksom i tidigare grupper hålls tillbaka i de stora bägarna.

En grupp om fem elever laborerar och pratar med varandra på ett återhållsamt sätt. De visar varandra tyst sina lösningar och anteckningspapper. De tittar, jämför och fortsätter att diskutera lågmält. De kommunicerar utan att lärarna uppfattar det. En lärare med observationsprotokoll observerar dem på någon meters avstånd men studerar i första hand elevernas lösningar.

Efter en stund blir en del elever klara. En elev, som bär spionglasögon, upptäcker att han gjort fel. Lösningarna förändras inte som han tänkt sig. Han håller i mer och mer vätska men inga färgomslag sker. Det framgår att han från början har hållt i lika mycket vätska som stamlösning. En lärare kommer gående förbi honom och ser att eleven ser ut att söka efter något i laborationsanvisningarna. *”Vad letar du efter?”* – säger läraren. *”Efter vad jag gjort för fel”*, svarar eleven. *”Varför säger du det?”* – undrar läraren. Eleven svarar – *”jag har tillsatt för mycket vätska från början”*. Han berättar kort för läraren hur han gjort. *”Om du skulle göra om det här, hur skulle du göra det?”* – undrar läraren. Eleven förklarar kortfattat hur han skulle göra och frågar läraren om han hinner göra om, vilket han hinner. Läraren kommer fram till författarna som filmar i klassrummet och säger att det är svårt att inte be dem att göra om från början om de hamnat helt fel. Lärare säger att det känns som att man lurar dem om man inte är uppriktig.

## Resultatsammanfattning

Tre huvudkategorier av lärares hinder för likvärdiga bedömningar av elevers individuella prestationer utkristalliserades i data: 1) Elevers interaktioner med varandra, 2) Lärares interaktion med elever och 3) Systematiska fysiska felkällor

### 1) Elevers interaktioner med varandra

Vid tre olika moment under den systematiska undersökningen framstår elevers sociala interaktioner som särskilt tydliga hinder för en individuell bedömning av elever.

a) I samband med val av laborationsmaterial

I samtliga tre tillfällen uppstår gruppenomen i samband med valet av laborationsutrustning. Utrustningen som väljs först är olika i de tre grupperna. Den utrustning som den förste eleven väljer är den som de efterföljande kamraterna också väljer. Den elev väljer som väljer först tar den utrustning som står närmast placerad fysiskt i rummet. Eleven studerar inte hela materialet översiktligt för att kunna göra ett mer aktivt reflekterat val utan tar ofta snabbt tag i det närmaste föremålet.

b) Vid hanteringen av kemikalier vid katedern

Eleverna kommunicerar hela tiden med varandra på olika sätt trots att det är en provsituation. De överlägger med varandra om hur de ska göra när de håller upp vätskor. Eleverna trängs runt bordet med kemikalier och hanterar utrustningen på ett likartat sätt. Om en elev börjar använda ett dropprör för att föra över stamlösning till en liten bägare så gör allesammans det inom en kort stund. I en annan grupp håller en elev upp vätska direkt i en graderad bägare, vilket innebär att inom kort gör samtliga det runt kemikaliebordet.

c) Under det genomförandet vid arbetsplatsen

Ofta hamnar eleverna ungefär samtidigt i samma moment i laborationen. De gör små pauser och tittar upp, på varandra, och påbörjar nästa steg ungefär samtidigt. De droppar i lösningar och de rör om med en glasstav i bägarna. Det uppstår tidvis ett gemensamt 'glasstavs-klirrande' mot glasbägarnas väggar under de olika upprepade omgångarna av omrörning. Ett annat exempel som ökar samtidigtheten är eleverna tittar på varandras lösningar. I stora delar utför eleverna laborationens olika moment parallellt under stor påverkan av varandra.

### 2) Lärares interaktion med elever

Det framkommer tre olika slags hinder för lärarna att göra likvärdiga bedömningar av elevers praktiska förmågor. Den första är av mer fysisk karaktär medan de två efterföljande faktorerna handlar om mellanmänsklig interaktion mellan lärare och elev.

a) Det är svårt för lärare att se detaljer i elevers hantering av utrustningen

En laborationslokal är ofta trångt möblerad med höga bänkar och det är svårt att få en överblick över vad var och en av eleverna gör vid sin arbetsplats. Under genomförandet går elever fram och tillbaka mellan sin arbetsplats och katedern, kompletterar sitt material och kemikalier. Det är ständigt någon i rörelse. Det gör att läraren ofta blir skymd vilket gör det svårt att under en längre sammanhållen tid systematiskt observera en elev. Det är många delmoment i varje laboration som ska bedömas och de utförs ofta nära intill eleven rent fysiskt. Ofta sker olika moment nästan i 'händerna' på eleven och de observerar vätskenivåer och läser av sin utrustning på bråkdelar av en sekund. Detta gör det svårt för läraren att se vilket material eleven verkligen använder sig av och hur det går till. Det är besvärligt för läraren att se med vilken noggrannhet en elev utför exempelvis en volymmätning. Tidvis hinner lärarna att vara aktiva och fråga elever men läraren får dock endast reda på en liten del av det som sker under hela genomförandet.



Ett tydligt exempel på lärares bedömningssvårigheter är en specifik situation som kunde följas med hjälp av ett par spionglasögon. En elev med spionglasögon använder ett mätglas på ett sätt som beskrivs i den högsta betygsnivån i bedömningskriterierna. Ingen av de två lärarna har möjlighet att observera denna betygsgrundande händelse. En lärare är tillfälligt utanför klassrummet och den andra pratar med en elev på andra sidan klassrummet.

b) Det är svårt för lärare att undvika lotsning

När lärarna går runt i klassrummet för att observera hinner de ofta inte ställa mer utredande frågor till elever utan eleverna frågar istället läraren olika saker av praktisk karaktär. Det kan handla om var det finns fler bägare eller var de ska hålla överskottsvätskor. Ibland ställer elever frågor om provet och hur försiktigt och undvikande lärarna än svarar på elevernas frågor riskerar de att 'lotsa' dem i någon riktning. Ibland säger läraren korta fraser, ibland inga ord men ett leende eller blinkning med ögonen bekräftar för eleven om man är på rätt väg.

c) Det är svårt för lärare att undvika att göra in-situ bedömningar

Lärarna uttalar flera gånger att de hamnar i situationer där de förväntas att 'rätta uppgiften' genom återkoppling under utförandet. De tyckte själv vid ett par tillfällen att de sagt långt mer än de borde gjort. De insåg att de delvis 'rättat' uppgiften direkt på plats – in situ. Det gick ibland inte att svara på den aktuella elevfrågan i relation till vad eleverna i praktiken hade gjort. Det gick inte att endast 'justera' inriktningen på genomförandet, exempelvis genom att tillsätta mer vätska/indikator, eftersom eleverna ibland från början gjort helt fel. Lärarna sa att de känner sig nästan svarslösa. I praktiken uppfattar eleverna däremot lärarnas reaktion direkt och förstår att något är fel. Lärares – ”*ja, du hinner*” – bekräftar för eleven att den hamnat fel.

### 3) Systematiska fysiska felkällor

Tre kategorier fysiska systematiska felkällor uppmärksammades som hinder för en likvärdig och individuell bedömning. De två första är av handhavande-karaktär medan den tredje är en resursfråga som kan variera mellan skolor. Den första underkategorin kan utgöra ett hinder för likvärdig bedömning då den kan leda till felaktiga resultat. Den andra kategorin leder till att bedömningen inte blir individuell medan den tredje kategorin kan utgöra ett hinder för att upprätthålla likvärdigheten mellan olika skolors möjligheter att genomföra det praktiska provet.

a) Felaktiga analytiska handlingar minskar likvärdigheten

Instruktionerna till lärarna inför den systematiska undersökningen är detaljerad men omfattar naturligtvis inte allt. Det bör kanske ha stått att stamlösningarna ska förnyas efter varje provtillfälle, vilket också kan gälla andra vätskor. Stamlösningarna i detta fall är inte de samma efter att de tre grupperna laborerat. Samtliga grupper hällde tillbaka överskotten av stamlösningarna efter volymmätningen i små bägare i den stamlösningen lärarna tagit fram ur kemikalieförrådet. Stamlösningarna blir starkt kontaminerade redan under och efter första provtillfället med grupp ett.

b) Gammal märkning kvar på utrustning minskar den individuella bedömningen

Ett ytterligare exempel på systematiskt fel är att eleverna ska bedömas avseende hur de märker upp sina blandningskärl. I bedömningsanvisningen är en godkänd nivå att kärnen ska märkas tydligt. När grupp två och tre laborerade var alla små blandningsbägare redan uppmärka i förvaringslådan. Dessa elever riskerade inte att missa det specifika bedömningskriteriet. Elever kan också få vägledning av att se vilka kärl som var diskade (våta) efter tidigare gruppers användning.

c) Möjlig brist på utrustning minskar likvärdigheten mellan skolor

Den systemiska undersökningen ställer mycket stora krav i praktiken. Om alla elever ska ha möjligheten att lyckas genomföra undersökningen på ett sätt som motsvarar de högsta betygskriterierna skulle de behöva ha tillgång till betydligt mer laborationsutrustning än vad många skolor kan upp-

bringa. Det är väldigt trångt vid kemikaliebordet eftersom eleverna där försöker hämta vätskor och utföra volymmätningar samtidigt. Det är inte möjligt att vid kemikaliebordet mäta upp volymer av lösningar på ett analytiskt mer korrekt sätt på grund av utrymmesskäl och säkerhetsskäl. De missar därmed chansen att uppnå de högsta betygsnivåerna enligt kunskapskriterierna. Eleverna behöver fler bågare för att kunna ta med sig kemikalierna till sin arbetsplats för att där använda rätt volymmätningstrustning på ett avancerat sätt. Under genomförandet i de tre aktuella grupperna tar de små bågarna slut vid två tillfällen. Eleverna får leta upp bågare i diskstall och diska ur dem, eller hämta vätskor med andra kärl som exempelvis mätglas eller provrör. Det är således inte helt säkert att det i praktiken finns tillräckligt med glasvaror på en skola för att eleverna ska få möjligheter att visa sina förmågor för de högsta betygskvaliteterna.

## DISKUSSION

Denna studie omfattas av en sekretess som kan uppfattas innebära validitetsproblem för resultatpresentationen då antalet excerpt och beskrivningar med nödvändighet behöver begränsas. Det innebär att resultatet kan betraktas som författarnas observationer alternativt tolkningar av vad som sker i skollaboratoriet. Ett sätt att säkerställa resultatens validitet är att två författare var och en för sig avkodat data, och därefter diskuterat och sammanställt resultaten. Resultaten har därefter presenterats och diskuterats inom två olika forskargrupper. Datainsamlingsmetoden med både fasta kameror och spionglasögon möjliggör för forskarna att gå tillbaka till och korsreferera data från olika kameror för att på nytt betrakta en situations specifika hinder för lärarens bedömning. Två forskare som enskilt eller tillsammans kan återvända till avgörande händelser i klassrumspraktiken ökar validiteten för en fallstudie som denna och därmed ökar generaliserbarheten (Flyvbjerg, 2006). Det innebär att denna fallstudies resultat får därigenom ett högre allmän intresse för lärare och forskare genom att de kan känna igen sig i de aktuella situationerna och konsekvenserna av dem, för sig själva för sina elever. Det innebär att igenkänningen av resultaten medverkar till att en del av de urskilda hindren i denna studie kan reflekteras över av lärare och forskare och därigenom i viss mån kunna motverkas i bedömningspraktiken.

### Den praktiska provsituationens komplexitet

Det är lätt att tycka att lärarna kunde ha styrt upp hela provsituationen betydligt strängare och varit tydligare med att eleverna inte får prata med varandra. Av de tre områden som analysfrågorna berör finns ingen omnämnd i instruktionsmaterialet. Ett sätt att förstå det är att det förutsätts från provkonstruktörernas sida att det går att återskapa en 'skriftlig provsalsmiljö' också under praktiska omständigheter i ett laboratorium. Om lärarna ska fokusera på att skapa en formellare provmiljö skulle det uppta en stor del av deras tid som var avsatt för observationer och bedömningar av elevernas individuella praktiska förmågor.

Det är svårt att skapa en praktisk provsituation där kommunikation inte förekommer. I en laboratoriemiljö rör sig eleverna kontinuerligt och de gör dessutom liknande handlingar synliga för varandra ungefär samtidigt. Det begränsade utrymmet i laboratoriet tillsammans med den lilla gruppstorleken skapar tillsammans en känsla av samhörighet, och det innebär att det nästan är ofrånkomligt att de börjar socialt interagera och kommunicera med varandra. Även om lärarna skulle få tyst på eleverna så pågår en icke-verbal kommunikation genom exempelvis den sociala interaktionen som uppstår i samband med att välja lika laborationsutrustning och arbeta parallellt. Den kommunikationen kvarstår och utgör fortfarande hinder för lärares rättvisa bedömningar av elevers individuella förmågor.

I denna situation är det också svårt för lärare att undvika lotsning, det vill säga att eleverna får lite mer information än vad lärarna egentligen får delge eleverna. Ofta är det omedvetet men ibland uttrycker lärarna en form av omsorg och känsla för eleven. Det blir synligt då elever har kommit helt fel

i sin undersökning. Eleven märker att resultat uteblir och att kamraterna gör något annat. I det läget kan lärare hamna i en situation där eleven behöver få återkoppling om det gör rätt eller fel. Läraren ’rättar’ uppgiften akut genom att uppmuntra eleven att göra om undersökningen.

Tidigare forskning har visat att lärare ofta känner en omsorg om sina elever och vill gärna att det ska gå bra för dem (Sund & Wickman, 2008, 2011). Resultatet i denna fallstudie visar på en sådan omsorg vid ett par tillfällen i olika grupper vilket medförde en överenskommelse mellan elev och lärare att göra om undersökningen och lyckas nå ett godkänt resultat.

### Likvärdighet och rättvisa

Det finns flera sätt att förstå hur svårt det är att i praktiken använda laborativa undersökningar som betygsunderlag. Uppgiften att genomföra en systematisk undersökning som Del B i nationellt prov i NO kan ur ett lärarperspektiv beskrivas som ett omöjligt uppdrag. Trots att lärare följer de detaljerade och noggranna provinstruktionerna kan de inte bedöma elevernas arbete på ett likvärdigt sätt. Eleverna är för många och läraren hinner inte uppmärksamma allt som är av betydelse enligt bedömningsanvisningarna. En systemisk undersökning kan uppfattas som orättvist mot både eleverna och lärarna eftersom de utsätts för en omöjlig situation. Elevernas betygsgrundande individuella prestationer kan komma att helt eller delvis skymmas av dem själva eller andra elever som hämtar utrustning vid katedern.

De systemiska undersökningar som beskrivits ovan är på flera sätt inte likvärdiga. För det första genomförs undersökningen inom det nationella provet endast vid ett enstaka tillfälle där läraren inte hinner få ett helhetsgrepp om vad och hur elever verkligen gör. Många viktiga iakttagelser missas per se av läraren. För det andra ger inte undersökningssituationen eleverna möjligheter att arbeta helt självständigt och visa sin individuella förmågor. För att det ska vara möjligt att helt individualisera det praktiska provet krävs en lokal som liknar en ’språkstudio’ med en egen arbetsplats med sidovägar, egen utrustning och egna kemikalier. Det kanske skulle behövas till och med hörlurar med musik för att skapa en total social avskildhet.

Bedömningen är orättvis mot elever ligger också i att ett grundläggande kemiskt begrepp (anges ej på grund av sekretessen) som den systematiska undersökningen starkt relaterar till inte nämns i centralt innehåll för kemi i åk 9. De nämns för första gången explicit i gymnasiekursen Kemi 2. Det innebär att lärarna kanske inte ens har undervisat eleverna om detta praktiska begrepp. Relationen mellan aktuella ämnesbegrepp och den praktiska förmågorna är viktig. Det är viktigt att enbart bedöma praktiska förmågor i förhållande till innehåll där begreppsförståelsen inte blir ett hinder för användningen av praktiska förmågor (Harlen, 1999). Det är ganska uppenbart att dessa tre elevgrupper som genomförde provet inte fått lära sig att laborera på ett noggrant analytiskt sätt. Det står i kursplanen för åk 9 att de ska lära sig en analytisk metod, inte att de ska kunna arbeta analytiskt. Skolundervisningen bör i nationella prov utvärderas så långt det är möjligt i förhållande till den undervisning som skett (Harlen, 1999). Karaktären på de aktuella uppgiften i denna systematiska undersökning riskerar att missa möjligheten att kontrollera och bedöma måluppfyllelsen av den aktuella läroplanen.

Sammantaget kan man säga att det är svårt att bedöma elever på ett likvärdigt sätt inom det aktuella svenska upplägget av en systematisk undersökning. I denna studie för lärarna inga elaborerande samtal där de söker elevers djupare kunskaper om vad de gör och varför eftersom detta skulle riskera att sprida och delge nya kunskaper till kamrater som laborerar intill. Enligt internationell forskning är bedömning av praktiska förmågor svårt, tidskrävande och mycket kostsamt (Gott & Duggan, 2002). Det innebär att det svenska skolsystemet antingen behöver avsätta ytterligare resurser för att säkerställa bedömningen av den systematiska undersökningen eller att ta beslutet att lägga ned det aktuella upplägget då det är dyrt och osäkert. Det går att utvärdera praktiska förmågor med mer teoretiska verktyg som beskrivs nedan.

### **Den praktiska bedömningens möjliga genomförande**

Att bedöma praktiska förmågor behöver i sig inte vara ett problem om läraren får tillfälle att mer kontinuerligt arbeta med laborativa övningar. Under dessa mer vardagliga omständigheter hinner lärare bilda sig en mer korrekt och allsidig uppfattning om hur kunniga enskilda elever är trots att de ofta jobbar parvis eller i grupper. En erfaren lärare som studerar sina laborationsgrupper ser ofta vem eller vilka som leder arbetet och som har de kvalitativt bästa laborativa förmågorna.

Ett annat sätt att göra det mer naturligt för elever att inte kommunicera allt för mycket med varandra under provsituationen är att de utför olika undersökningar i ett och samma klassrum. Det får kommunicera en del öppet eftersom det är fortfarande fruktbart att kommunicera kemi med varandra men det är inte är någon större idé att snegla på varandras försöksupställningar. Att genomföra olika uppgifter inom klassrummet kan anpassas till både enskilda elever och elevgrupper.

Att lära sig att arbeta i grupp är generellt viktig i samhället. Många arbetsgivare värderar förmågor som att arbeta i grupp så det kan vara en poäng att värdera både individuella och kollektiva förmågor samtidigt (Reiss et al., 2012). Ett av huvudproblemen med att genomföra en praktisk övning är de tysta omständigheterna vilket implicerar att uppgiften istället kan utföras i gruppform. Gruppens medlemmar har rätt att samtala med varandra och resultatet ger en bild av hur gruppens/klassens förmågor som helhet står sig i jämförelse med andra skolor. Lärares diskussioner med studenter om deras metoder är särskilt fruktbara om de ställs så att de kräver någon form av reflektion (Harlen, 1999). Det blir också lättare för läraren som kan koncentrera sig på att samtala med grupper om 4-5 individer istället för ett tjugotal elever.

### **Vad tillför praktiska prov skolan?**

Lärarnas bedömning av elever står inte och faller med denna specifika provsituation. I praktiken kan lärarna i Sverige ge bra bild över elevers förmågor när de väger in sin samlade kunskap i bedömningen av elevers praktiska förmågor. Detta är möjligt att uppnå utan ett praktiskt delprov som denna systematiska undersökning. Resultatet i denna studie visar att trots Skolverkets höga ambitionsnivå med att inkludera en holistisk systemisk undersökning i de nationella NO-proven så är utbytet med avseende på att betygsunderlaget lågt. Det finns åtskilliga hinder för lärares bedömningar. Det är fullt möjligt att inom den ordinarie verksamheten nå goda bedömningar om elevers förmågor.

Det är inte helt enkelt att förstå vad den systematiska undersökningen tillför skolan i dess helhet. De kan säkerligen inspirera lärare att göra fler laborationer och kanske implementera en del nya undersökningar och arbetsätt. Detta behöver dock ställas emot de omfattande kostnader som det kostar att årligen genomföra nationella prov (Dillon, 2008; Roberts & Gott, 2006). Det är en mängd värdefulla skolresurser som tas i anspråk under vårterminen i åk 9. De nationella proven tar mycket tid och lärarresurser i anspråk och det kan vara så att det räcker med det teoretiska delprovet för en mer likvärdig och rättvis betygssättning i NO. Teoretiska nationella prov i kombination med att låta skolor/lärare bedöma elever under ett helt år i samband med ett tjugotal utvalda praktiska uppgifter (Ofqual, 2015) kan vara ett sätt att göra det på. Ett annat sätt att öka likvärdigheten i bedömning är att fortbilda lärare specifikt i bedömning.

Skolutveckling och alldeles särskilt utveckling i NO-undervisning går att närma sig på många sätt. Syftet med den systemiska undersökningen som betygsunderlag är gott men fungerar sämre i praktiken visar denna fallstudie. Dessa lärare och elever representerar inte hela landet men visar troligen på olika hinder som förekommer i olika utsträckning i olika provsituationer. En fallstudie som denna kan generellt visa att det finns anledning att anta att det förekommer hinder av olika slag som behöver ytterligare studeras (Flyvbjerg, 2006). Det är inte svårt att se att provsituation är arrangerad och konstlad. Det kommer fler mätningar i framtiden. Storskaliga prov blir allt fler och verkar vara här för att stanna. En fråga som kan vara nyttig att ställa regelbundet är: Ska vi värdera det vi kan mäta eller värdera det vi vill mäta (Biesta, 2009)?

Det är nödvändigt att med jämna mellanrum ställa kritiska frågor om storskaliga prov och den skärpa att mäta det som de utger sig för att göra. Det är många som vill mäta skolans verksamhet på olika sätt. Ytterst handlar det om en tilltro till skolans kompetenta lärare och att låta dem utforma förut-sättningarna för en likvärdig och rättvis bedömning av sina elever. Lärarna kan i sin professionella praktik ta sig förbi flera av de hinder för en likvärdig och rättvis bedömning som denna studie upp-märksammat.

## REFERENSER

- Abrahams, I., & Millar, R. (2008). Does Practical Work Really work? A study of the effectiveness of practical work as a teaching and learning method in school science. *International Journal of Science Education*, 30(14), 1945-1969.
- Abrahams, I., & Reiss, M. J. (2012). Practical Work: Its Effectiveness in Primary and Secondary Schools in England. *Journal of Research in Science Teaching*, 49(8), 1035-1055.
- Biesta, G. (2009). Good Education in an age of measurement: On the need to reconnect with the question of purpose in education. *Educational Assessment, Evaluation and Accountability*, 21(1), 33-46.
- Broadfoot, P., & Black, P. (2004). Redefining assessment? : The first ten years of Assessment in Education. *Assessment in Education: Principles, Policy & Practice*, 11, 7-27.
- Dillon, J. (2008). A Review of the Research on Practical Work in School Science. London: King's College.
- Flyvbjerg, B. (2006). Five Misunderstandings About Case-Study Research. *Qualitative Inquiry*, 12(2), 219-245.
- Gott, R., & Duggan, S. (2002). Problems with the Assessment of Performance in Practical Science: Which way now? *Cambridge Journal of Education*, 32(2), 183-201.
- Harlen, W. (1999). Purposes and Procedures for Assessing Science Process Skills. *Assessment in Education: Principles, Policy and Practice*, 6(1), 129-144.
- Högström, P., Ottander, C., & Benckert, S. (2010). Lab Work and Learning in Secondary School Chemistry: The Importance of Teacher and Student Interaction. *Research in Science Education*, 40, 505-523. DOI 10.1007/s11165-009-9131-3
- Högström, P., Ottander, C., & Benckert, S. (2012). Laborativt arbete i grundskolans senare år: Lärares perspektiv. . *Nordic Studies in Science Education*, 6(1), 80-91.
- Kvale, S., & Brinkmann, S. (2009). *InterViews: Learning the Craft of Qualitative Research Interviewing*. Thousand Oaks, CA: Sage.
- Lundahl, C. (2010). Nationella prov – ett redskap med tvetydiga syften. [National tests – a tool with ambiguous purposes]. In G. Lundahl & M. Folke-Fichtelius (Eds.), *Bedömning i och av skolan - praktik, principer, politik*. [Assessment in and of schools - practice, principles, policies] (pp. 223-242). Lund: Studentlitteratur.
- Lundahl, C., Roman, H., & Riis, U. (2010). Tidigt ute med sena betyg - sent ute med tidiga! Svensk betygspolitik i ljuset av internationell betygsforskning och betygssättning i Europa. [Early with late rating - late with early! Swedish assessment policies in light of international assessment research in Europe.] *Pedagogisk forskning i Uppsala 157*. Uppsala: Uppsala University.
- Lundqvist, E., & Sund, P. (In Press). Selective traditions in group discussions – teachers' views about good science and the possible obstacles when encountering a new topic. *Cultural Studies in Science Education (accepted Sept, 2015)*.
- Ofqual. (2015). *Assessment of Practical Work in New Science GCSEs - Summary*. [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/408513/2015-03-03-assessment-of-practical-work-in-new-science-gcses-summary.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/408513/2015-03-03-assessment-of-practical-work-in-new-science-gcses-summary.pdf).
- Ottander, C., & Grelsson, G. (2006). Laboratory work: the teachers' perspective. *Journal of Biological Education*, 40(3), 113-118.

- Reiss, M., Abrahams, I., & Sharpe, R. (2012). Improving the assessment of practical work in school science *Leading Education and Social Research*. London: Institute of Education University of London and The University of York.
- Roberts, R., & Gott, R. (2006). Assessment of performance in practical science and pupil attributes. *Assessment in Education: Principles, Policy & Practice*, 13(1), 45-67.
- Skolverket. (2010). Ämnesproven i biologi, fysik och kemi i årskurs 9. *En redovisning av utvärderingsomgången [Subject tests in biology, physics and chemistry in year 9. A report from the test set]*. Stockholm: Skolverket [National School Agency].
- Skolverket. (2011a). Kunskapsbedömning i skolan – praxis, begrepp, problem och möjligheter.
- Skolverket. (2011b). Läroplan för grundskolan, förskoleklassen och fritidshemmet.
- Sund, P. (2016). Discerning selective traditions in science education – A qualitative study of teachers' responses to what is important in science teaching. *Cultural Studies in Science Education*. DOI 10.1007/s11422-015-9666-8
- Sund, P., & Wickman, P.-O. (2008). Teachers' Objects of Responsibility - Something to Care about in Education for Sustainable Development? *Environmental Education Research*, 14(2), 145-163.
- Sund, P., & Wickman, P.-O. (2011). Socialization Content in Schools and Education for Sustainable Development - I. A study of Teachers' Selective Traditions. *Environmental Education Research*, 17(5), 599-624.
- Watts, A., & Wilson, F. (2013). The assessment of practical science: a literature review *Cambridge Assessment*. Cambridge: Research Division, Assessment Research and Development Cambridge Assessment.
- Wolcott, H. F. (1994). *Transforming qualitative data: description, analysis, and interpretation*. Thousand Oaks, California Sage Publications, Inc.