Formativ bedömning: En enkel match eller en svår utmaning?

Effekter av en kompetensutvecklingssatsning på lärarnas praktik och på elevernas prestationer i matematik

Erika Boström

Institutionen för naturvetenskapernas och matematikens didaktik
Umeå 2017

This work is protected by the Swedish Copyright Legislation (Act 1960:729)
ISBN: 978-91-7601-706-7
ISSN: 1652-5051
Omslagsfoto: Kerstin Östberg, omarbetat av Mathias Norqvist & Therese Ekström
Elektronisk version tillgänglig på http://umu.diva-portal.org/
Tryck: UmU tryckservice, Umeå universitet
Umeå, Sverige 2017

Till Nils och Ines

Innehåll

Abstract **ii**
Avhandlingens artiklar **iii**
1. Introduktion **1**
    1.1 Syfte och översikt 4
2. Bakgrund **6**
    2.1 Ramar för avhandlingsprojektet 6
    2.2 Formativ bedömning 10
    2.3 Formativ bedömning – effekter på elevers prestationer 14
    2.4 Kompetensutveckling och införlivandet av formativ bedömning 17
3. Forskningsdesign och metoder **27**
    3.1 Övergripande forskningsdesign 27
    3.2 Design av kompetensutvecklingen 31
    3.3 Datainsamling 33
    3.4 Analysmetoder 39
    3.5 Forskningsetiska överväganden 44
4. Resultat **46**
    4.1 Effekter av kompetensutvecklingssatsningen 46
    4.2 Varför gjorde lärarna dessa förändringar och vilka hinder mötte de? 50
5. Diskussion **55**
    5.1 Resultaten i relation till Modellen om lärares professionella utveckling 55
    5.2 Lärarna genomförde förändringar, men varför gjorde de inte än mer? 62
    5.3 Den uteblivna effekten på elevernas prestationer i matematik 65
    5.4 Lärdomar om kompetensutvecklingssatsningar (i formativ bedömning) 66
    5.5 Reflektioner kring avhandlingens kvalitet 68
    5.6 Avhandlingens bidrag till forskningsfältet och utbildningsväsendet 71
    5.7 En enkel match eller en svår utmaning? 72
6. Korta sammanfattningar av artiklarna **74**
    6.1 Artikel I 74
    6.2 Artikel II 75
    6.3 Artikel III 76
    6.4 Artikel IV 77
Efterord **79**
Referenser **81**

Abstract

Research reviews have shown that the use of formative assessment in classroom practice can substantially improve student achievement. However, a strong research base about how to support teachers' implementation of such formative classroom practice is lacking. In this thesis, I investigate the effects of a comprehensive professional development programme (PDP) in formative assessment on teachers' classroom practice and students' achievement in mathematics. In addition, I identify reasons for the changes made in the teachers' formative classroom practice. Fourteen randomly selected year - 7 mathematics teachers participated in the PDP. The teachers' formative classroom practice before and after attending the programme was analysed and described, and reasons for their change in practice were explored. The effect of the changes in formative classroom practice on students' mathematics achievement was examined using pre- and post-tests administered to both the intervention group and a control group. A mixed methods approach with classroom observations, teacher interviews, questionnaires and student achievement tests in mathematics was used in the studies included in the thesis.

The results show that the teachers used aspects of formative assessment in their classroom practice before the PDP, but that there was plenty of room for development towards a more effective formative assessment practice. Several possibilities for developing the practice were identified. After the PDP the teachers believed in the idea of formative assessment and were motivated to make changes towards a more formative classroom practice. The teachers included new formative assessment activities in their classroom practice, but in different ways and to different degrees. The characteristics of these changes were identified, and also the characteristics of the PDP that the teachers found to be influential for their development of the formative classroom practice. Results also show that the teachers' motivational beliefs held after the PDP was an explanatory factor for their changes in practice. However, the formative assessment practice the teachers implemented did not have a significant effect on their students' achievement compared to the control group. In addition, there was no correlation between the number of formative assessment activities implemented by the teachers and their students' achievement gains. Reasons for these non-effects on student achievement, and for the teachers' degree and type of implementation of formative assessment in the classroom practice, are discussed in the thesis.

Avhandlingens artiklar

Avhandlingen innehåller följande artiklar, där Artikel I är publicerad och Artikel II - IV är under förberedelse för att skicka till tidskrift.

Artikel I
Andersson, C., Boström, E. & Palm, T. (2017). Formative assessment in Swedish mathematics classroom practice. *Nordic Studies in Mathematics Education*, 22(1), 5-20.

Artikel II
Changes in teachers' formative classroom practice after a professional development programme, and important conditions for change
Författare: Erika Boström & Torulf Palm

Artikel III
Motivational beliefs as an explanation for the effect of professional development programs in formative assessment on teacher practice
Författare: Erika Boström & Torulf Palm

Artikel IV
The impact of a specific formative assessment practice, on student achievement in mathematics.
Författare: Erika Boström & Torulf Palm

Alla artiklarna är samförfattade med min handledare, Torulf Palm. Artikel I skrevs dessutom tillsammans med Catarina Andersson, som också var doktorand inom det Formativa projektet (projektet beskrivs i avsnitt 2.1). I Artikel I har jag och Catarina gjort merparten av arbetet när det gäller insamling och bearbetning av empiriskt material, och i Artikel II- IV har jag gjort merparten av arbetet när det gäller insamling och bearbetning av empiriskt material. Torulf har i egenskap av forskningsledare haft huvudansvaret för den övergripande planeringen av projektet och dess datainsamling, och jag har varit aktivt deltagande i diskussioner och beslut. Torulf (och Catarina i Artikel I) har liksom jag gett betydande intellektuellt bidrag i analys- och skrivfasen. Vid skrivandet av artiklarna har jag bidragit i processen och skrivandet av artiklarnas alla delar. Samtliga författare har läst igenom och stått bakom den slutgiltiga versionen av Artikel I som skickats till tidskrift. Artikel I dubbel-publiceras här med godkännande från den aktuella tidskriften.

1. Introduktion

Flera forskningsöversikter har lyft fram användningen av formativ bedömning i undervisningen som ett av de mest effektiva sätten att öka elevernas lärande (Black & Wiliam, 1998; Hattie, 2009, s. 297). Det har också varit ett ökat intresse för formativ bedömning bland forskare under de senaste åren (Baird, Hopfenbeck, Newton, Stobart & Steen-Utheim, 2014; Hirsh & Lindberg, 2015), och sedan Black och Wiliams (1998) inflytelserika forskningsöversikt har det i många länder genomförts olika reformer för att utveckla användningen av formativ bedömning (Black, 2015; Birenbaum m.fl., 2015). Att införliva formativ bedömning i undervisningen har dock visat sig vara svårt, då det för många lärare innebär en förändring av synen på lärarens roll i klassrummet:

"It calls for a set of pedagogical habits of mind and pedagogical content knowledge that most teachers do not have and that are not simple to acquire"
(Shoenfeld, 2014, p. 407).

"For many teachers, adopting formative assessment practices is difficult because it involves a radical change in the way in which they relate to their students and the ways they behave in the classroom. What is called for is nothing less than a change in the ways they perceive, and strive to implement, their role as teachers."
(Black, 2015, s.171)

Många forskare menar att idén med formativ bedömning måste förstås på djupet, för att sedan kunna införlivas på ett tillförlitligt sätt i undervisningen (Flórez & Sammons, 2013; Pedder & James, 2012). I allmänhet så hävdar förespråkare för formativ bedömning att det finns ett behov av större förståelse för komplexiteten i att införliva formativ bedömning i praktiken och att detta inkluderar att acceptera processen som en långsam process som måste ske gradvis, samt att lärare behöver mycket stöd i denna process (Flórez & Sammons, 2013; Leahy & Wiliam 2012; Black, 2015). När det gäller att möta de utmaningar som följer med att försöka öka användningen av formativ bedömning i undervisningspraktiken så spelar kompetensutvecklingsinsatser en viktig roll (Laveault, 2016). Dock så visar forskning om kompetens-utvecklingsinsatser i formativ bedömning att nationella reformer i olika länder inte fått det utfall som många hoppats på (Tierney, 2006; Bierenbaum m. fl., 2015) och beslutsfattare i många länder efterfrågar en mer omfattande användning av formativ bedömning i skolorna (Cizek, 2010; Smith, 2011; Stiggins, 2002). Försök att öka användningen av formativ bedömning i undervisningen har dock inte visat sig vara så framgångsrika i att åstadkomma en utvecklad formativ klassrumspraktik (De Lisle, 2015; Flórez & Sammons, 2013; Hume och Coll, 2009; James & McCormick, 2009; Ofstead, 2008; Schneider & Randel, 2010; Wylie & Lyon, 2015), i den utsträckning att det medför ökat lärande hos eleverna (Bell, Steinberg, Wiliam & Wylie, 2008; Jönsson, Lundahl & Holmgren, 2015, Randel, Apthorp, Beesley, Clark &

Wang, 2016; Schneider & Randel, 2010). Forskning har lyft och belyst svårigheter med att stödja lärare att införliva formativ bedömning i undervisningen i den utsträckning att det får en betydande inverkan på elevernas resultat. Shavelsons (2008) citat, där han beskriver sina erfarenheter av att skapa, införliva och studera effekter av formativ bedömning talar sitt tydliga språk:

"After five years of work, our euphoria developed into a reality that formative assessment, like so many other education reforms, has a long way to go before it can be wielded masterfully by a majority of teaches to positive ends"
(Shavelson, 2008, p.294)

Inom forskningsfältet om formativ bedömning har det också visat sig vara en stor variation i hur formativ bedömning konceptualiseras och införlivas, vilket har skapat kontroverser kring de belägg som finns kring ökat lärandehos elever som användningen av formativ bedömning sägs ge (Briggs, Ruiz-Primo, Furtak, Shepard, & Yin, 2012; Dunn & Mulvenon, 2009; Filsecker & Kerres, 2012; Kingston & Nash, 2011; McMillan, Venable, & Varier, 2013). Svårigheter i att identifiera 'best practice' gällande användningen av formativ bedömning i praktiken har också lyfts fram (Dunn & Mulvenon, 2009). Ett antal empiriska studier rapporterar dessutom om ett ytligt införlivande av formativ bedömning (Hume & Coll, 2009; James & McCormick, 2009; Marshall & Drummond, 2006, Torrance, 2007), vilket även förs fram i flera teoretiska reflektioner om forskningsfältet (bl.a. Klenowski, 2009; Swaffield, 2011; Torrance, 2012). Detta kan ses som en indikation på den komplexitet och de svårigheter som införlivandet av formativ bedömning verkar innebära. Dessutom visar den ovan nämnda forskningen på vikten av att forskningsstudier måste vara tydliga med och redovisa egenskaperna hos den formativa bedömningspraktiken som införlivas (och inte bara att formativ bedömning införlivas). Detta är även något som Bennett (2011) och Kingston och Nash (2011) efterlyser då de önskar fler studier för att avgöra vilka faktorer som inverkar på effektiviteten av formativ bedömning. Studier som undersöker effekter av kompetensutvecklingsinsatser i formativ bedömning på både lärarnas praktik och på elevernas prestationer har just börjat dyka upp inom fältet. Schneider och Randel (2010) betonar att sådana studier är viktiga för att avgöra om kompetensutvecklingsinsatser i formativ bedömning är effektiva avseende lärarnas förändrade praktik och elevernas ökade prestationer. De framhåller även att sådan forskning är värdefull för att bidra med förståelse om för vem kompetensutvecklingsinsatser i formativ bedömning är effektiva och under vilka förutsättningar. Dessutom konstaterar de att kompetensutvecklingssatsningar i formativ bedömning oftast genomförts med lärare över ämnesgränser, och att antagandet då är att lärare har de ämneskunskaper och de pedagogiska färdigheter som krävs för att genomföra et framgångsrikt införlivande av formativ bedömning i sin klassrumspraktik. Men de konstaterar också att lärare inte alltid har de ämneskunskaper och

pedagogiska färdigheter som behövs för att införliva formativ bedömning på ett effektivt sätt (Schneider & Randel 2010). Detta tyder på att det finns ett behov av kompetensutvecklingssatsningar i formativ bedömning där lärare även kan få stöd inom sitt ämnesområde.

Sammanfattningsvis, för att bidra med kunskap om hur vi kan stödja lärare att införliva formativ bedömning av hög kvalitet i sin klassrumspraktik så behövs fler studier som undersöker effekter av kompetensutvecklingssatsningar i formativ bedömning på både lärarnas praktik och på elevernas prestationer i matematik. Detta för att bidra med kunskap om vilka kompetensutvecklingsinsatser som är effektiva och för att bidra med kunskap om egenskaper i den förändrade praktiken som ger effekt på elevernas prestationer. Dessutom bör sådana studier även utröna viktiga egenskaper i kompetensutvecklingsinsatsen för att stödja införlivande av formativ bedömning, samt för vem och i vilken kontext sådana kompetensutvecklings satsningar är effektiva. Sådan forskning behövs för att bidra med mer kunskap om vilka egenskaper den formativ bedömning bör ha som ska inkluderas i kompetensutvecklingssatsningar, hur dessa ska utformas samt vilket stöd olika lärare behöver för att på bästa sätt kunna och vilja införliva formativ bedömning i sin klassrumspraktik.

I denna avhandling så har intentionen varit att möta många av dessa ovan beskrivna behov genom att genomföra en interventionsstudie av en omfattande kompetensutvecklingssatsning för matematiklärare i formativ bedömning. De ingående artiklarna studerar effekter av kompetensutvecklingssatsningen (interventionen) på lärarnas praktik och på elevernas prestationer i matematik, samt söker möjliga orsaker till utfallet. Då formativ bedömning visat sig ha stor potential för att öka elevers prestationer så har det varit aktuellt i Sverige liksom i många andra länder att utveckla användningen av formativ bedömning för att förbättra lärares och elevers lärande (Flórez & Sammons, 2013; Hirsh & Lindberg, 2015). Forskare behöver dessutom bli bättre på att bygga broar mellan forskning och praktik, det vill säga att de behöver bli bättre på att finna vägar för hur forskningsresultat kan komma till nytta och användas i praktiken, något som Paul Black, 2015 poängterar:

" Academics in education have to find ways to build fruitful interactions between their world and the world of practising teachers if they are ambitious to explore, and to learn how to implement, the potential benefits of their work."
(Black, 2015. s. 175).

Inom det svenska skolväsendet finns också en önskan om att ta beslut på vetenskaplig grund, och i skollagen uttrycks att utbildningen ska vila på vetenskaplig grund (och beprövad erfarenhet), vilket denna typ av forskning kan bidra till. Den här forskningen kan därmed vara värdefull för politiker,

rektorer och andra beslutsfattare som tar beslut om kompetensutvecklingssatsningar, såväl som för lärare och elever som är de viktigaste aktörerna inom skolväsendet. Denna forskning fyller därmed dubbla syften, och kan bidra både till forskningsfältet i formativ bedömning och till skolpraktiken.

1.1 Syfte och översikt

1.1.1 Avhandlingen som del i ett större projekt

Denna avhandling har sitt ursprung i och utgör en del av ett projekt som delvis finansierats av en kommun och med pengar från Skolverket. Den forskning som mynnade ut i de fyra ingående artiklarna i denna avhandling genomfördes inom ramen för projektet som fortsättningsvis kommer att kallas *Formativa projektet*. Det Formativa projektet skapade förutsättningar för att genomföra en omfattande kompetensutvecklingssatsning i formativ bedömning för matematiklärare, samtidigt som det erbjöd goda möjligheter till forskning. Detta samarbete möjliggjorde att cirka hälften av de matematiklärare som läsåret 2011/2012 skulle undervisa årskurs 4 eller årskurs 7 i kommunen kunde väljas genom slumpmässigt urval, för att sedan erbjudas att delta i den omfattande kompetensutvecklingssatsningen. Fortsättningsvis kommer lärarna som undervisar i årskurs 4 att kallas mellanstadielärare och lärarna som undervisar i årskurs 7 att kallas högstadielärare. Kompetensutvecklingen genomfördes sedan på liknande sätt för de två grupperna av lärare. Samarbetet med kommunen möjliggjorde också att alla kommunens elever fick göra tester i matematik, både före och efter införlivandet av formativ bedömning, vilket gjorde att vi kunde använda en experimentell design med interventionsgrupp och kontrollgrupp. I denna avhandling fokuseras högstadielärarna (som under läsåret 2011/2012 undervisade årskurs 7) i det Formativa projektet, som beskrivs mer utförligt i avsnitt 2.1

1.1.2 Syfte och innehåll

Det övergripande syftet med det Formativa projektet och även med denna avhandling är att bidra med kunskap och förståelse av viktiga faktorer när det gäller att stödja matematiklärares införlivande av formativ bedömning i sin klassrumspraktik. Mer specifikt är avhandlingens syfte att beskriva och bidra med kunskap om den (eventuella) effekt en genomförd kompetensutvecklingssatsning i formativ bedömning gav på högstadielärarnas formativa klassrumspraktik och på elevernas prestationer i matematik. Dessutom är syftet att bidra med kunskap om orsaker till att lärarna att förändrade sin klassrumspraktik på det sätt och i den omfattning som de gjorde.

I kappan kommer jag att beskriva resultaten i de fyra ingående artiklarna och även att göra en syntes. I syntesen är min avsikt att väga samman resultaten från artiklarna och diskutera dessa som en helhet för att möta syftet med

avhandlingen. Jag kommer specifikt att lyfta fram de resultat som bidrar till kunskap och förståelse av orsaker till lärarnas införlivande av formativ bedömning.

De fyra ingående artiklarna i denna avhandling belyser ovanstående syften utifrån olika perspektiv: Hur används formativ bedömning av ett slumpmässigt urval av högstadielärare (och även mellanstadielärare) i matematik i en kommun (Artikel I)? Hur förändrade högstadielärarna sin klassrumspraktik i matematik efter en omfattande kompetensutvecklingssatsning i formativ bedömning, vilka egenskaper i kompetensutvecklingen och andra förutsättningar i skolkontexten upplevde de som viktiga för förändringen, och vilka hinder för förändring upplevde de? Förekommer likheter och skillnader avseende dessa frågor jämfört med mellanstadielärarna? (Artikel II)? Kan högstadielärarnas förändring mot en mer formativ klassrumspraktik förklaras och förstås utifrån deras motivation? Kan skillnader mellan hur högstadielärare och mellanstadielärare förändrade sin klassrumspraktik förklaras och förstås utifrån skillnader i motivation (Artikel III)? Vilka effekter har den införlivade formativa bedömningspraktiken på elevernas prestationer i matematik och finns det något samband mellan antalet nya bedömningsaktiviteter som införlivas och elevernas lärande (Artikel IV)? Den första studien genomfördes före kompetensutvecklingssatsningen och inkluderar även mellanstadielärarna i det Formativa projektet. De andra tre studierna genomfördes under och efter kompetensutvecklingssatsnigen och i Artikel II III och IV görs även vissa jämförelser med motsvarande resultat i de parallella studierna som rör mellanstadielärarna i det Formativa projektet.

1.1.3 Avhandlingens disposition

Detta är en sammanläggningsavhandling med en inledande så kallad kappa och fyra artiklar. Kappan består av sex kapitel och i det första kapitlet ges en introduktion av forskningsfältet och avhandlingens syfte. I det andra kapitlet, bakgrund, ges en överblick av befintliga teoretiska och empiriska studier om formativ bedömning och om kompetensutveckling som är av särskild relevans för den här avhandlingen. I Kapitel 2 presenteras också de ramverk som används i avhandlingen. Det tredje kapitlet behandlar forskningsdesign och metoder och där återfinns även designen av den givna kompetensutvecklingen. I det fjärde kapitlet beskrivs avhandlingens resultat, där avsikten är att knyta samman och till viss del göra en syntes av resultaten. Det femte kapitlet innehåller en diskussion kring avhandlingens resultat och slutsatser, avhandlingens relevans och hur den relaterar till forskningsfältet. Det sista kapitlet består av korta sammanfattningar av de ingående artiklarna.

2. Bakgrund

I detta avsnitt beskriver jag inledningsvis ramarna för det projekt som denna avhandling utformats inom. Därefter ger jag en överblick av de teoretiska och empiriska studier om formativ bedömning och om kompetensutveckling som är av särskild relevans för den här avhandlingen samt de ramverk som används.

2.1 Ramar för avhandlingsprojektet

Detta avhandlingsprojekt genomfördes inom ramen för ett större projekt, det Formativa projektet som var ett samarbetsprojekt mellan en kommun och Umeå universitet. Kommunen sökte och fick pengar från Skolverkets matematiksatsning 2009 - 2011, en satsning på matematik där kommuner under 3 år kunde söka pengar för att utveckla matematikundervisningen. Kommunen kunde själv bestämma vad pengarna skulle satsas på genom att skriva fram det i sin ansökan. Inom utvecklingsprojektet satsade man på två spår. Det ena spåret innebar en allmän satsning på matematik för alla matematiklärare i kommunens grundskolor. Det andra spåret innebar att genomföra en omfattande kompetensutvecklingssatsning för ett antal matematiklärare i formativ bedömning, det Formativa projektet. Det Formativa projektet inleddes med en pilotstudie som inbegrep en omfattande kompetensutvecklingssatsning för några av kommunens högstadielärare under hösten 2010. Huvudstudien påbörjades under senare delen av hösten 2010, och under våren 2011 genomfördes två liknande omfattande kompetensutvecklingssatsningar för ett större antal mellanstadielärare och högstadielärare. Totalt deltog cirka 45 matematiklärare i det Formativa projektet, inom vilken detta avhandlingsprojekt genomfördes.

2.1.1 Ett samarbetsprojekt mellan Umeå universitet och en kommun

Forskargruppen från universitetets sida bestod av Torulf Palm, Catarina Andersson och jag själv. Dessutom så bidrog Peter Nyström vid planering och uppstarten av projektet och Tomas Bergqvist deltog i delar av data-insamlingen och planeringen av hur analysen av lärarnas praktik efter kompetensutvecklingen skulle genomföras. Torulf Palm var forskningsledare och även handledare till de två doktoranderna, där jag ansvarade för de studier som genomfördes på högstadienivå och den andra doktoranden (Catarina) ansvarade för de studier som genomfördes på mellanstadienivå. Forskargruppen ansvarade för designen och organisationen av forsknings-studierna såväl som designen och genomförandet av kompetensutvecklingen. Kommunen i sin tur ansvarade för den initiala kontakten med skolorna och bistod även med administrativt stöd i de delar av datainsamlingen som

omfattade elevdata. Kommunen bidrog genom att skapa kontakter med lärare och rektorer och information (och marknadsföring) om projektet kunde därmed på ett bra sätt administreras till alla berörda. Dessa informationsutskick handlade bland annat om tillvägagångssätt vid datainsamling och återkoppling. Vi i forskargruppen deltog i allt detta och genom samarbetet med kommunen kunde det ske genom kommunens kommunikationskanaler. Kommunen ansvarade också för finansieringen av att friställa de lärare som skulle delta i kompetensutvecklingen motsvarade 20% av en heltidstjänstunder en termin. De pengar kommunen fått från Skolverket bidrog till att kommunen hade ekonomi att genomföra detta projekt, vilket var mycket värdefullt.

Genom detta samarbete fick kommunen möjlighet att få (då forskargruppen ansvarade för kompetensutvecklingen) en omfattande kompetensutveckling för ett stort antal lärare, som dessutom följdes upp genom forskning. Kommunen fick också tillgång till de tester (som kan användas som diagnoser) som konstruerades särskilt för forskningsstudien. Forskargruppen å sin sida fick möjlighet att genomföra en storskalig forskningsstudie om en kompetensutvecklingssatsning i formativ bedömning med ett slumpmässigt urval av lärare. Detta ömsesidiga samarbete gynnade därmed båda parter. Innan projektet initierades genomfördes en kartläggning av aktuell kompetensutvecklingsforskning av forskningsledaren. Den kartläggningen (som också bidrog till kompetensutvecklingssatsningens design) och kommunikation med verksamma lärare stödde kommunen i deras beslut att satsa på detta projekt. Forskargruppen hade en kontaktperson på kommunen som drev projektet och ansvarade för ansökan av pengar från Skolverket. Alla de rektorer som var ansvariga för de lärare som deltog i kompetensutvecklingen hade i förväg gått med på att ordna det praktiska så att lärarna skulle kunde delta. Som tidigare nämnts så innebar det att lärarna skulle få nedsättning i sin tjänst motsvarande 20% av en heltidstjänst. Rektorerna skulle ombesörja att de deltagande lärarna inte skulle vara schemalagda med undervisning då träffarna på universitetet var planerade, helst en hel dag fri från undervisning. För högstadielärarna var kursträffarna på universitet förlagda en halvdag (4h) per vecka, alltid samma dag och tid under en termin.

2.1.2 Den svenska kontexten

Då studien genomfördes i Sverige, vill jag här ge en övergripande bild av den svenska kontexten som är viktig för att förstå förutsättningarna för det formativa projektets genomförande. Formativ bedömning och den positiva effekt det kan ha på elevers prestationer har uppmärksammats i Sverige, och detta har inte minst visat sig i beslut på nationell nivå. Mellan 2009 och 2011 genomfördes en stor matematiksatsning, varav det Formativa projektet utgjorde en del. Formativ bedömning hade börjat uppmärksammas i Sverige

innan det Formativa projektet genomfördes men det hade ännu inte börjat nå ut på bred front. Vidare har det svenska skolsystemet under senare år genomgått stora förändringar och många frågor rörande bedömning har varit i fokus i samband med en mängd reformer, detta särskilt i samband med skrivandet och implementering av den senaste läroplanen för grundskola (Lgr 11) och gymnasium (Gy 11). Lärare och skolledare har fått ta sig an frågor rörande bedömning, både med summativa och formativa syften, i större utsträckning (Hirsch & Lindberg, 2015). Skolverket har även propagerat för den formativa bedömningens plats i skolan genom allmänna råd och riktlinjer.

Formativ bedömning har fortsatt att ha en framträdande roll i Sverige också inom 'matematiklyftet', en stor nationell utvecklingssatsning i matematik som pågått 2012 - 2016 med syfte att utveckla matematikundervisningen i Sverige (Skolverket, 2012c). I flera av de undervisningsmoduler som ingår i den satsningen så är formativ bedömning inbäddat. En annan indikation är formativ bedömnings framträdande roll inom SKOLFORSK, Vetenskapsrådets projekt från 2014 med att kartlägga utbildningsvetenskaplig forskning på uppdrag av regeringen. Syftet med kartläggningen var att de skulle skapa en plattform av kunskapsunderlag för det nybildade Skolforskningsinstitutet att utgå från i deras fortsatta arbete med skolans vetenskapliga förankring. Inom projektet så genomfördes en kartläggning av aktuell forskning inom formativ bedömning (Hirsch & Lindberg, 2015). I en annan kartläggning, också genomförd på uppdrag av SKOLFORSK (Ryve m. fl. 2015), ansågs formativ bedömning vara ett av tre viktiga utvecklingsområden för utveckling av matematikundervisningen i Sverige. Det har också genomförts olika kompetensutvecklingsinsatser i Sverige där formativ bedömning varit i fokus, till exempel "Boråsprojektet" (Jönsson, m. fl., 2015). Detta tyder på att utveckling av matematikundervisning och användandet av formativ bedömning är högt prioriterade områden i Sverige. Vad gäller forskning kopplat till formativ bedömning så ökar antalet studier som rör formativ bedömning i Sverige, men är fortfarande relativt begränsat, vilket gäller både generellt (Hirsch & Lindberg, 2015) såsom i matematik (Ryve m.fl., 2015).

2.1.3 Översikt av forskningsprojektet

I det här avsnittet presenteras i Figur 1 en översikt av det totala forskningsprojektet. Forskningsprojektet inkluderar kompetensutvecklingssatsningen som genomfördes inom pilotstudien samt de två kompetens utvecklingssatsningarna, en för högstadielärare och en för mellanstadielärare, som genomfördes inom forskningsprojektet. Huvudstudien gällde både högstadielärare och mellanstadielärare, men i denna avhandling fokuseras högstadielärarna, dock så görs jämförelser med motsvarande studier för mellanstadielärarna. Fyra av de artiklar som blev resultatet från huvudstudien, de som berör högstadielärarna, ingår i avhandlingen (Artikel I-IV), se Figur 1.

<table>
<tr><td colspan="6">Formativa projektet</td></tr>
<tr><td colspan="6">Pilotstudie med kompetensutveckling för högstadielärare (n=8)</td></tr>
<tr><td colspan="6">Huvudstudie</td></tr>
<tr><td colspan="3">Högstadielärare (n=17)</td><td colspan="3">Mellanstadielärare (n=21)</td></tr>
<tr><td colspan="6">Lärarnas användning av formativ bedömning innan kompetensutvecklingen (n=38)
(Artikel I)</td></tr>
<tr><td colspan="3">Kompetensutveckling för högstadielärare</td><td colspan="3">Kompetensutveckling för mellanstadielärare</td></tr>
<tr><td>Effekter på lärarnas klassrumspraktik och orsaker till effekter på lärarnas praktik
(Artikel II)</td><td>Orsaker till effekter på lärarnas praktik utifrån ett motivations-perspektiv
(Artikel III)</td><td>Effekter på elevernas prestationer i matematik
(Artikel IV)</td><td>Artikel II</td><td>Artikel III</td><td>Artikel IV</td></tr>
</table>

Figur 1. Figuren presenterar en översikt över hela det Formativa projektet vilket inkluderar en pilotstudie och en huvudstudie. Dessutom presenteras kompetensutvecklingssatsningarna som genomfördes inom projektet och de fyra artiklar (gällande högstadielärarna) som blev resultatet av huvudstudien.

2.2 Formativ bedömning

I detta avsnitt presenteras först en redogörelse för vad formativ bedömning kan anses vara, och därefter beskrivs konceptualiseringen[1] av formativ bedömning som används i denna avhandling.

2.2.1 Vad är formativ bedömning?

En bedömning som sker under lärprocessen för att föra lärandet framåt, och där informationen som genereras i bedömningen används för att stödja elevers lärande är en formativ bedömning. Begreppet summativ bedömning används om informationen som genereras i bedömningen inte används för att stödja elevers lärande utan enbart sammanfattar det lärande som skett. Formativ bedömning ställs ofta mot summativ bedömning, och diskussioner har förts om nödvändigheten i att separera dessa. Att använda dessa begrepp för att skilja på den funktion som bedömningen fyller, anser dock de flesta vara användbart. Bedömningen kan därmed ske på samma sätt men hur informationen sedan används kan vara olika, det vill säga den funktion bedömningen fyller kan skilja sig åt. Samma bedömning kan därmed användas både för formativa och summativa syften vilket medför att det ibland kan vara problematiskt att göra en absolut distinktion mellan dem (Bennett, 2011).

Termerna formativ bedömning och Bedömning för lärande (BfL) används ibland med olika innebörd och ibland synonymt inom forskningsfältet om formativ bedömning (Swaffield, 2011). Några som varit tongivande i utvecklandet av begreppet formativ bedömning är Paul Black och Dylan Wiliam med kollegor. Den definition av formativ bedömning som Black och Wiliam använde i sin forskningsöversikt 1998, utvecklades senare till en mer detaljerad och förfinad definition i Black & Wiliam (2009). Den senare kan anses inkludera många av förekommande definitioner av formativ bedömning och BfL, och är den definition som används i denna avhandling och i de ingående artiklarna. Dock poängterar Black & Wiliam (2009) i sin definition att bedömningen måste åtföljas av handling (dvs, någon måste agera på bedömningen) och definitionen kan därmed inte anses inkludera de definitioner av formativ bedömning och Bedömning för lärande som bara inkluderar skrivningar om bedömningens syfte istället för bedömningens faktiska funktion. Definitionen av Black & Wiliam (2009) presenteras nedan (på nästa sida) och därefter presenteras också en svensk översättning av definitionen:

[1] Med konceptualisering avser jag hur hela ”begreppsapparaten” runt formativ bedömning förstås och används, dvs termen konceptualisering kan därmed innefatta både definition och operationalisering av definitionen

"Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited."
(Black & Wiliam, 2009, p. 9).

"Bedömning fungerar formativt när bevis för elevens prestation tas fram, tolkas och används av lärare, elever eller deras kamrater för att besluta om nästa steg i undervisningen som förmodligen blir bättre, eller bättre grundade, än de beslut de skulle ha fattat om bevis inte funnits"
(Wiliam, 2013 s. 58)

Definitionen innefattar flera olika sätt att genomföra formativ bedömning. En variant fokuserar att läraren använder test för att samla information om elevers lärande vilket följs av en justering av undervisningen (t.ex. Yeh, 2009). I detta sammanhang är det vanligt att datorprogram används för att testa eleverna (t.ex. Burns Klingbeil & Ysseldyke, 2010; Koedinger, McLaughlin & Heffernan, 2010). Andra fokuserar på lärarens feedback som en respons på den insamlade informationen om elevernas lärande (t.ex. Rakoczy m. fl., 2013). Andra varianter fokuserar elevernas aktiva deltagande i den formativa bedömningsprocessen (t.ex. Marchis, 2012). Det kan vara elever som självreglerare av sitt lärande, vilket inkluderar självbedömning och efterföljande agerande för att nå lärandemålen (Zimmerman, 2002). Elevernas aktiva deltagande i den formativa bedömningsprocessen kan också handla om att stödja varandras lärande. Detta kan ske på olika sätt, till exempel genom kamratbedömning och efterföljande förslag till kamraten om hur de kan agera för att nå lärandemålen. Vissa forskare som beforskar dessa aspekter av formativ bedömning använder det termen för det specifika fokus som studeras, t.ex. feedback, medan andra använder mer övergripande termer såsom formativ bedömning och BfL. Då de olika varianterna av formativ bedömning kan fungera kompletterade (t.ex. lärares feedback kan mottas på ett bättre sätt av elever som är effektiva självreglerare av sitt lärande), kan man anta att det finns potential för ökade lärandemöjligheter genom att integrera dessa strategier till en helhet. Det skulle då kunna leda till ytterligare förbättringar i elevers prestationer. Detta är också något som bland annat Hawe och Parr (2014) argumenterar för: De hävdar att löftet som formativ bedömning (eller BfL) för med sig endast kan nås när strategier är antagna på sätt som återspeglar dess enhetliga natur, främjar kvalitativa utfall, och ger eleverna en central roll i sitt lärande. Det förekommer olika men liknade förslag på hur en sådan praktik kan konceptualiseras och operationaliseras (Arter, 2009, Wiliam & Thompson, 2008). Wiliam och Thompson (2008) operationaliserade en idé om att involvera flera olika strategier för formativ bedömning, som alla kretsar kring den *stora idén* att framkalla belägg om elevers lärande och sedan att informationen används för att justera

undervisning (och lärande) för att bättre möta elevernas lärandebehov. Konceptualiseringen av formativ bedömning som använts i denna avhandling inbegriper Black och Wiliams definition (2009) och Wiliam & Thompsons (2008) operationalisering och beskrivs mer ingående i nästa avsnitt.

2.2.2 Formativ bedömning - definition och ramverk

Definition av Black och Wiliam (2009), se ovanstående avsnitt, är den som används i de ingående studierna och i denna avhandling. Definitionen betonar formativ bedömning som en naturlig inneboende del i klassrumspraktiken. Den utgör alltså en övergripande beskrivning av idén med formativ bedömning som en enhetlig praktik av integrerade strategier. I definitionen inkluderas i begreppet ”*undervisning*” (eng. instruction) kombinationen av undervisning och lärande, all verksamhet avsedd att skapa lärande. Definitionen operationaliseras av Wiliam och Thompson (2008) i ett ramverk som underlättar förståelsen och den praktiska användningen av formativ bedömning i klassrumspraktiken. Ramverket av Wiliam & Thompson (2008), bestående av en *stor idé* och fem nyckelstrategier, utgjorde basen för kompetensutvecklingen i denna studie och strukturerade analysen av lärarnas formativa klassrumspraktik (Artikel I & II). I ramverket operationaliseras formativ bedömning som en klassrumspraktik som baseras på ”trohet” gentemot en *”stor ide”*, att belägg för lärande ska användas för att anpassa undervisningen för att bättre möta elevernas lärandebehov (med andra ord att undervisningen ska *anpassas* efter elevernas lärandebehov), och en kompetent användning av fem nyckelstrategier, se Figur 2 (Wiliam & Thompson, 2008). Enligt dem kan formativ bedömning förstås som att det involverar tre huvudprocesser i undervisning; (1) att identifiera var eleverna befinner sig i sitt lärande (2) var de är på väg i sitt lärande, och (3) hur de ska ta sig dit (ursprungligen från Ramaprasad, 1983; Sadler, 1989). De poängterar också att detta är något som kan utövas av tre aktörer i klassrummet (läraren, eleven och klasskamraterna), vilket resulterar i en matris bestående av nio celler som kan organiseras som ”fem nyckelstrategier” av formativ bedömning utifrån aktörernas roll i respektive process, se Figur 2. Nyckelstrategierna som presenterats i Figur 2 är som följer (från Wiliam, 2013, s. 61):

1) Klargöra, delge och skapa förståelse för lärandemål och kriterier för framsteg
2) Genomföra effektiva diskussioner, aktiviteter och lärandeuppgifter och som tar fram belägg för lärande
3) Ge feedback som för lärandet framåt
4) Aktivera eleverna som läranderesurser för varandra
5) Aktivera eleverna till att äga sitt eget lärande

<table>
<tr><th></th><th>Vart eleven är på väg</th><th>Vart eleven befinner sig just nu</th><th>Vägen dit</th></tr>
<tr><td rowspan="2">Lärare</td><td rowspan="2">NS 1.
Klargöra, delge och skapa förståelse för lärandemål och kriterier för framsteg</td><td>NS 2.
Skapa och leda effektiva klassrumsdiskussioner, lärandeuppgifter och aktiviteter som tar fram belägg för elevers lärande</td><td>NS 3.
Ge feedback som för lärandet framåt</td></tr>
<tr><td rowspan="2">Kamrat</td><td rowspan="2">Förstå och dela lärandemål och kriterier för framsteg</td><td colspan="2">NS 4.
Aktivera eleverna som läranderesurser för varandra</td></tr>
<tr><td>Elev</td><td>Förstå och dela lärandemål och kriterier för framsteg</td><td colspan="2">NS 5.
Aktivera eleverna till att äga sitt eget lärande</td></tr>
</table>

Figur 2. Förhållandet mellan nyckelstrategierna, undervisningsprocesserna och aktörerna i klassrummet (utifrån figurer i Wiliam & Thompson, (2008, s. 63) och Wiliam, (2011, s. 61))

De tre processerna (att identifiera var eleverna befinner sig i sitt lärande, var de är på väg i sitt lärande, och hur de ska ta sig dit) utgör den definierande egenskapen i formativ bedömning i definitionen likväl som inom den *stora idén* i ramverket. Detta då de utöver att tydliggöra vikten av att målen för lärandet är tydliga fokuserar på framkallandet, tolkandet och användandet av information för att ta beslut om nästa steg i undervisningen (inkluderat lärandet) för att nå dessa mål. Figur 2 visualiserar hur de fem nyckelstrategierna kan kopplas till den definierande egenskapen av formativ bedömning likväl som hur de relaterar till varandra och integreras i ett enhetligt ramverk för formativ bedömning. Den *stora idén* kan genomsyra allt arbete med formativ bedömning, och klargör att belägg om elevernas lärande bör samlas in för att sedan användas av aktörerna i klassrummet (lärare, elever och (klass)kamrater) för att ta beslut om nästa steg i undervisningen – med andra ord är undervisningen *följsam* mot elevernas behov. För att detta arbete ska vara effektivt, betonar den första nyckelstrategin vikten av att lärare och elever når en gemensam förståelse av lärandemålen. Den andra nyckelstrategin handlar om att samla belägg för elevers lärande. De insamlade beläggen (informationen) kan sedan användas för att ge feedback som möter elevernas lärandebehov (Nyckelstrategi 3) eller till att ta bättre grundade beslut om hur undervisningen ska fortsätta. Nyckelstrategi 4 och 5 klargör att alla aktörerna i klassrummet, läraren, eleverna och deras klasskamrater, (och inte bara läraren, kan (och bör) vara aktiva aktörer i dessa processer. Eleverna kan till exempel själva bedöma sina prestationer för att avgöra hur de ska gå vidare och ta nästa steg i sitt lärande i linje med den stora idén. Vidare kan läraren stödja elevernas motivation och färdigheter att ta rollen som kamrat- och självreglerande lärande. För att vara effektiva bör dessa strategier införlivas kontinuerligt och dagligen i klassrumspraktiken och många av dessa strategierna har enskilt visat sig vara effektiva för elevernas lärande (Wiliam,

2007). Dessutom, som tidigare nämnts, när dessa strategier används tillsammans i klassrumspraktiken, som en *integrerad helhet*, finns potential att de ytterligare stödjer varandra för att förbättra elevernas engagemang och lärande. Detta då lärares feedback och justerad undervisning, baserat på insamlad information om elevers lärande, kan ses som kompletterande komponenter till elevers självreglerande lärande och kamratlärande (Andersson & Palm, 2017a). Läraren och eleverna arbetar tillsammans för att stödja lärandet genom interaktion under alla nyckelprocesser.

Black och Wiliam (2009) klargör också att *"formative assessment is concerned with the creation of, and capitalization upon, 'moments of contingency' in instruction for the purpose of the regulation of learning processes"* (Black and Wiliam, 2009, s. 10). En aspekt som skiljer mellan olika sätt att genomföra formativ bedömning rör tidsramen för justeringscykeln, vilket inkluderar längden av cykeln och frekvensen av cykler. Längden av justeringscyklen handlar om den tid som går mellan att information om elevers lärande framkallas till dess att informationen används i form av feedback eller justerad undervisning (Wiliam & Thompson, 2008). Längden av cykeln kan variera från sekunder till mer än ett år, och Wiliam delar utifrån detta in cykellängden som Kort-cykel (inom eller mellan lektioner, 5 sekunder till 48 timmar), medium-cykel (inom och mellan undervisningsmoment, en till fyra veckor) och lång - cykel (mellan betygsperioder, terminer och år, fyra veckor till ett år) (Wiliam & Thompson, 2008; Wiliam, 2010). En lärare kan till exempel agera på information om elevers lärande genom att justera undervisningen samma lektion som informationen samlades in eller genom att ge feedback till eleven som sedan direkt agerar på detta genom att till exempel ändra sitt inlärningsbeteende. I dessa fall skulle längden av justeringscykeln vara ett antal sekunder eller minuter. Läraren kan också vänta med att agera på den insamlade informationen till nästa lektion, till nästa moment eller till nästa termin. I dessa fall skulle cykellängden beskrivas i termer av dagar eller månader. Frekvensen av justeringscykeln, innebär hur ofta undervisning eller lärande justeras, baserat på den framkallade informationen om elevernas lärande (Andersson & Palm, 2017a). Längden av cykeln, likväl som frekvensen av cykler kan variera, både för individuella elever och för hela klassen.

2.3 Formativ bedömning – effekter på elevers prestationer

Som tidigare nämnt lyfter Black & Wiliam (1998) i sin forskningsöversikt fram att stora effekter på elevernas prestationer är möjliga om formativ bedömning används i klassrumspraktiken. Den definition av formativ bedömning som de då använde (likväl som definitionen från 2009) inkluderar flera olika sätt som formativ bedömning kan genomföras på. Detta innebar att deras forskningsöversikt inkluderade publikationer som undersökte effekten av olika strategier för formativ bedömning. Potentialen att förbättra elevers lärande för var

och en av dessa strategier bekräftas också i olika forskningsöversikter. Dessa forskningsöversikter visar starka samband mellan elevers prestationer och formativa bedömningsstrategier såsom feedback (Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Shute, 2008), självreglerat lärande (vilket inkluderar självbedömning och efterföljande agerande för att nå lärandemålen) (Dignath & Büttner, 2008; Ross, 2006) och kamratlärande inkluderande kamratbedömning och efterföljande feedback (Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003; Van Zundert, Sluijsmans & Merrienboer, 2010). Att användandet av bedömningsmatriser kan ha potential att inverka positivt på elevers lärande konstaterats i en forskningsöversikt av Pandero och Jönsson (2013) med fokus på användning av matriser för formativa syften.

Forskningsöversikter med ett särskilt fokus på matematik har också visat samband mellan elevers prestationer och lärares justering av undervisningen, baserat på den insamlade informationen om elevernas lärande (National mathematics Advisory Panel, 2008; Yeh, 2009), och självreglerat lärande (Dignath & Büttner, 2008). En forskningsöversikt som inkluderar olika varianter av formativ bedömning som alla delar den gemensamma egenskapen att aktörer i klassrummet (lärare, elever och klasskamrater) samlar information om elevers lärande och baserat på den informationen justerar undervisningen visar också på en positiv relation mellan formativ bedömning och elevers prestationer i matematik (Palm, Andersson, Boström & Vingsle, accepterad för publikation). Tre varianter av formativ bedömning studerades i forskningsöversikten och alla tre indikerar en positiv relation till elevernas prestationer. I den första varianten av formativ bedömning ligger fokus på att läraren samlar information ofta om elevernas kunskaper för att justera undervisningen (vilket också kan ske genom feedback). Många av de ingående studierna kopplade till denna variant använde datorprogram för att framkalla information. Forskningsöversikten visar även att den andra varianten, som fokuserar egenskaper hos den feedback läraren ger, kan ha en positiv effekt på elevernas prestationer. Typen av feedback är dock avgörande vilket är i enlighet med tidigare forskning om feedback (t.ex. Hattie & Timperley, 2007). I översikten konstaterar de också en stark relation mellan varianten av formativ bedömning som handlar om självbedömning och efterföljande agerande.

Utifrån ovanstående kan konstateras att forskningsbasen för att formativ bedömning har stor potential att gynna elevers lärande kan sägas vara stark. Dock har de stora effektstorlekar som ofta görs gällande som vedertagna i forskningslitteraturen kritiserats. Det förekommer stor variation i storleken på den effekt som olika studier visat att formativ bedömning har på elevernas prestationer, och vissa interventioner med formativ bedömning har inte gett någon effekt på elevernas prestationer (Bennet, 2011; Dunn & Mulvenon,

2009; Kingston & Nash, 2011). Det är också en stor variation i hur formativ bedömning konceptualieras (som tidigare nämnts) och införlivas, vilket har orsakat kontroverser kring de belägg som finns som hävdar stora vinster avseende elevernas prestationer när formativ bedömning används (Briggs, Ruiz-Primo, Furtak, Shepard, och Yin, 2012; Dunn & Mulvenon, 2009; Filsecker & Kerres 2012, Kingston & Nash, 2011; McMillan, Venable & Varier, 2013). I en forskningsöversikt genomförd av U.S. Department of Education (National Mathematics Advisory Panel, 2008) som inkluderade en undersökning av effekten av lärares användning av information från frekvent administrerade små test för att ge differentierad undervisning för elever drogs slutsatsen att formativ bedömning i matematik har stor potential att öka elevers prestationer. Men de konstaterade också att mer forskning behövs för att kunna dra reliabla slutsatser om denna effekt.

Då formativ bedömning kan genomföras på många olika sätt kan därmed även effekten på elevernas prestationer variera stort från en specifik införlivande av formativ bedömning till en annan. Vissa generella egenskaper för effektiv formativ bedömning kan anses vara kända från till exempel Black och Wiliams (1998) forskningsöversikt och andra forskningsöversikter om till exempel feedback (Shute, 2008; Hattie & Timperley, 2007). Men det är inte klarlagt vad som skiljer effektiva formativa bedömningspraktiker från mindre effektiva (Jönsson m.fl., 2015), och bland annat Bennett (2011) frågar efter studier för att avgöra vilka faktorer som påverkar effekten av formativ bedömning. För att flytta fram forskningsfältet om formativ bedömning efterfrågar Bennet (2011) '*a theory of action*'. Han anser att det behövs studier som utvärderar de underliggande mekanismerna som är tänkta att orsaka de önskade effekterna av formativ bedömning då vi annars inte kan förutsäga under vilka förutsättningar eller för vilka grupper som formativ bedömning fungerar.

Endast ett fåtal empiriska studier har undersökt effekten av en formativ bedömningspraktik som integrerar flera strategier på elevernas prestationer, och effekten av en sådan formativ bedömningspraktik på elevers prestationer är hittills inte så välkänd (Flórez & Sammons, 2013). Det finns dock några exempel på studier där effekten på elevers prestationer av några kompetensutvecklingsinitiativ i formativ bedömning undersöks (t. ex. Bell, Steinberg, Wiliam, & Wylie, 2008; Randel m. fl., 2016). Det finns också exempel på några studier där effekten på kompetensutvecklingsinitiativ på lärarnas praktik har undersökts (t.ex. Jönsson m. fl., 2015; Robinson, Myran, Strauss, & Reed, 2014; Wylie & Lyon, 2015). I studierna av Jönsson m. fl. (2015) och Robinson m. fl. (2014) fanns även elevernas prestationer med som ett effektmått. I Jönsson m. fl. (2015) användes ett mått på elevernas medelbetyg för att diskutera möjliga utfall av elevernas prestationer. I Robinson m. fl. (2014)

baserades slutsatserna på ett begränsat antal självrapporter från lärarna, vilket i sin tur begränsade deras resultat som de inte kunde koppla till elevernas resultat på ett meningsfullt sätt (något som författarna själva konstaterar). Avseende studierna av Bell m. fl. (2008) och Randel m. fl. (2016), finns indikationer på att både omfattningen och ”troheten” mot kompetensutvecklingen var begränsad, och dessa studier visar inte heller någon effekt på elevernas prestationer. Detta medför att de två senare studierna inte bidrar med forskningsbelägg för formativ bedömnings effekter på elevers prestationer. De visar däremot på svårigheten med att utforma och genomföra kompetensutvecklingssatsningar i formativ bedömning som ger effekt på lärarnas klassrumspraktik och på elevers prestationer, särskilt då formativ bedömning konceptualiseras som en integrerad helhet av strategier.

De två studier jag funnit som använde en design med interventionsgrupp och kontrollgrupp, (Wiliam, Lee, Harrison & Black, 2004; Andersson & Palm, 2017a), visade positiva resultat avseende elevernas prestationer. I “The Kings’s Medway- Oxfordshire Formative Assessment Project (KMOFAP)”, rapporterade Wiliam med kollegor (Black, Harrison, Lee, & Marshall, 2003; Black & Wiliam, 2003; Wiliam m. fl., 2004) positiva resultat på elevernas prestationer då lärarna i projektet utvecklade sin praktik inom fyra områden som är inkluderade i ramverket av Wiliam and Thompson (2008). Den andra studien, som ingår i det Formativa projektet, använde Wiliam & Thompsons (2008) ramverk och inkluderar 22 slumpmässigt valda mellanstadielärare (Andersson & Palm, 2017a). Resultaten visar att klasserna som undervisades av lärare i interventions gruppen, de lärare som utvecklat sin formativa klassrumspraktik, presterade bättre på ett test efter ett skolår jämfört med klasser som undervisats av lärare i kontrollgruppen, när hänsyn tagits till ett förtest.

Sammanfattningsvis kan konstateras att då det förekommer en stor variation i hur formativ bedömning konceptualiceras och införlivas behövs det mer forskning om vilken typ av formativ bedömning som ger effekt på elevers prestationer. Det finns således ett behov av att urskilja vilka egenskaper av formativ bedömning som har vissa effekter, samt för vem och i vilken kontext som denna formativa klassrumspraktik är effektiv. Dessutom finns det få studier som undersöker formativ bedömning som en integrerad helhet av flera strategier och dess effekt på elevers prestationer.

2.4 Kompetensutveckling och införlivandet av formativ bedömning

Under de senaste åren har formativ bedömning förordats av beslutsfattare i skolfrågor runt om i världen (Birenbaum m. fl., 2015). Trots att forsknings-basen för formativ bedömning är relativt väletablerad och accepterad i de

olika länderna har utbildningspolicyn inte fullt ut antagits på ett sådant sätt att det lett till en signifikant förändring i lärarnas klassrumspraktik (Birenbaum m. fl., 2015). Formativ bedömning förespråkats i många länder (OECD, 2005) men ansträngningar att sprida användningen av formativ bedömning har inte medfört att formativ bedömning, som genomförs med hög kvalitet, observerats frekvent i praktiken (t.ex. De Lisle, 2015; Hume & Coll, 2009; Carless, 2005; Flórez & Sammons, 2013; Ofsted, 2008; Stiggins, 2002).

Försök att främja användningen av formativ bedömning har ofta inte varit framgångsrika avseende att utveckla en bestående formativ bedömningspraktik (De Lisle, 2015; Hume och Coll, 2009; James & McCormick, 2009; Schneider & Randel 2010). Lärare behöver på djupet förstå den komplexa idén som formativ bedömning utgör för att kunna överföra den förståelsen till praktiken på ett tillfredställande sätt (Pedder & James, 2012; Robinson m. fl., 2014). Kompetensutvecklingsinsatser spelar då en viktig roll i att möta de utmaningar införlivandet av en formativ bedömningspraktik verkar innebära (Laveault, 2016). Interventioner med kompetensutvecklingsinsatser i formativ bedömning har mött på svårigheter, och ofta inte resulterat i en substantiellt förbättrad formativ bedömningspraktik (t.ex. James & McCormick, 2009; Jönsson m. fl. 2015; Randel m. fl., 2016; Wylie & Lyon, 2015). Dessutom behöver man investera även på andra nivåer i skolsystemet för att möjliggöra en mer spridd och högkvalitativ formativ klassrumspraktik. Skolledare behöver stöd och utbildning i anslutning till att formativa praktiker ska införlivas, och broar mellan forskning, policy och praktik måste stärkas och byggas (Hirsch & Lindberg, 2015). Många forskare har konstaterat att lärare behöver mycket stöd för att införliva formativ bedömning på ett tillfredställande sätt, och det har inom forskningsfältet konstaterats att man vet för lite om hur man på bästa sätt kan stödja lärare (t.ex. Wiliam, 2010; Schneider & Randel, 2010). Nästa avsnitt behandlar kompetensutveckling av lärare och därefter följer ett avsnitt som handlar om viktiga förutsättningar för, och svårigheter med, en framgångsrik införlivning av formativ bedömning.

2.4.1 Kompetensutveckling generellt

Inom forskningsfältet kring kompetensutveckling har fokus länge legat på att identifiera viktiga designegenskaper i kompetensutvecklingssatsningar (Kennedy, 2016). Inom fältet finns nu en samsyn om egenskaper i kompetensutvecklingssatsningar som är viktiga för att åstadkomma en förändring i lärares praktik och ökade elevprestationer (Desimone, 2009; Timperley, Wilson, Barrar & Fung, 2007). Dessa egenskaper kan summeras som ämnesfokus, aktivt lärande (t.ex. praktiska övningar för att nå djupinlärning), koherens (t.ex. mellan det som fokuseras i kompetensutvecklingen och lärarnas föreställningar eller vidare trender i policys), utsträckning av och

tillräckligt med tid, samt möjligheter till samarbete och diskussioner med kollegor. Dessutom fann Timperley m. fl. (2007) att engagemang från skolledare och tillgång till extern expertis var viktiga för att kompetens-utvecklingar skulle ha inverkan på elevernas prestationer. Schneider & Randel (2010) som i sin forskningsöversikt kopplar litteratur om kompetens-utveckling generellt till litteratur om formativ bedömning identifierade sju viktiga faktorer kring kompetensutveckling generellt som troligen är särskilt viktiga vad gäller kompetensutveckling i formativ bedömning. Dessa faktorer var: administrativt stöd, individualisering av lärarnas mål, kunskap om innehållet, tid, samarbete, koherens och aktivt lärande. En nyligen genomförd forskningsöversikt med fokus på viktiga förutsättningar för införlivande av formativ bedömning identifierade också liknande viktiga egenskaper avseende kompetensutveckling och stöd till lärare (Heitink, Van der Kleij, Veldkamp, Schildkamp, & Kippers, 2016). Ovanstående indikerar att forskning kring kompetensutveckling generellt kan bidra och utgöra en bas för forskning om kompetensutveckling i formativ bedömning.

Dock har man fått olika utfall av kompetensutvecklingsinsatser både när de kända egenskaperna är uppfyllda och inte, vilket medfört att dessa egenskaper inte fungerar för att förutsäga lärarnas lärande i det enskilda fallet (Kennedy, 2016; Opfer & Pedder, 2011). Varje egenskap i designen hos en kompetens-utveckling kan skilja på viktiga sätt, påverka med olika intensitet och uppträda på olika sätt i olika kontexter, och dessa skillnader kan påverka lärares lärande (Kennedy, 2016; Opfer & Pedder, 2011). Vidare argumenterar Opfer och Pedder (2011) att ett fokus på närvaro eller frånvaro av de olika egenskaperna begränsar våra möjligheter att förstå och förutsäga varför vissa utfall av kompetensutvecklingssatsningar kommit till stånd. De argumenterar också för att professionellt lärande för lärare utgör ett komplext system där variabler interagerar och kombineras på olika sätt och med olika intensitet och påverkar lärares lärande. Den tid som behövs för att lärare ska förstå nya idéer och överföra denna förståelse till praktiken kommer till exempel att variera beroende på tidigare kunskap och föreställningar och på stöd från extern expertis. Att utveckla en bättre förståelse för varför och hur lärare lär och använder kompetensutvecklingssatsningar för att utveckla sin praktik har därför blivit ett viktigt område att utforska (Opfer & Pedder, 2011).

Opfer och Pedder (2011) föreslår att forskning om kompetensutveckling bör byta fokus från en 'orsak och effekt approach' mot ett fokus på orsaksförklaringar, det vill säga till exempel studier som inte bara visar att mycket tid i kompetensutvecklingen är bra utan även visar varför mycket tid är bra. Sådana studier kan bidra med förståelse av under vilka förutsättningar, varför och hur lärare lär (Opfer & Pedder, 2011. Kennedy (2016) instämmer och föreslår forskningsdesigner och kompetensutvecklingsmodeller som är

mer knutna till underliggande teorier om lärares motivation och lärares lärande. Hon lyfter att alltid då man studerar lärande, oavsett om det gäller elevers eller lärares lärande utgör motivation en viktig faktor. Då motivation är den drivande kraften bakom mänskligt beteende (Ryan & Deci, 2000) kan ett motivations perspektiv vara användbart i strävan att förstå lärares val då de engagerar sig i kompetensutvecklingssatsningar för att lära och införliva nya idéer i klassrumspraktiken. Många misslyckanden med kompetens-utvecklingsinitiativ, i till exempel formativ bedömning, verkar vara relaterade till att lärare inte upprätthåller och fortsätter sina ansträngningar gentemot kompetensutvecklingsinsatsen och dess införlivande (Bell m. fl., 2008; Randel m. fl., 2016; Schneider & Randel, 2010).

För att få bättre förståelse för de val lärarna gör i kompetensutvecklings-satsningen, antas i Artikel III ett motivationsperspektiv och *förväntan-värde-teorin* (eng. Expectancy-value theory of achievement motivation) som är utvecklad av Eccles och Wigfield (2000) användes i analysen. Teorin beskrivs i avsnitt 2.4.4. För att få bättre förståelse för hur kompetensutvecklings-satsningen påverkar lärarnas motivation att förändra sin praktik, lärarnas lärande och lärarnas praktik används en modell för lärares förändring (eller professionella utveckling). *Modellen om lärares professionella utveckling* av Clarke och Hollingworth (2002) finns beskriven i avsnitt 2. 4. 3. Modellen används i Artikel III och i avhandlingen som helhet för att knyta samman och diskutera resultaten. Nästa avsnitt ger en översiktlig bakgrund kring forskning om införlivande av, och kända svårigheter med, lärares införlivande av formativ bedömning.

2.4.2 Införlivande och svårigheter med införlivande av formativ bedömning

Black & Wiliam konstaterade redan 1998, att införlivandet av formativ bedömning inte på något sätt är enkelt eller rättframt (Black & Wiliam, 1998b). Flertalet studier har visat på svårigheter med att införliva formativ bedömning. Dessa utmaningar och svårigheter som uppkommit har inom forskningsfältet kopplats till olika orsaker och några av dessa kommer att tas upp i detta och avsnitt. Dessutom beskrivs forskning som belyser olika egenskaper i den formativ bedömning införlivats.

Yttre faktorer, såsom skolkontexten och trender av ansvarsskyldighet påverkar införlivande av formativ bedömning (Birenbaum m. fl. 2016; Hirsh, Lindberg, 2015). I många länder i västvärlden har den rådande testkulturen med många summativa bedömningar lyfts som en orsak till att införlivande av en bestående formativ klassrumspraktik av hög kvalitet varit svårt att få till stånd (Birenbaum m. fl. 2015; Hirsh & Lindberg, 2015). Även om det är möjligt för summativ bedömning och formativ bedömning att samexistera

(Black, Harrison, Hodgen, Marshall & Serret, 2010; Brookhart, 2010) blir det problematiskt när det i hög grad är summativa bedömningar som räknas inom utbildningssystemet (Bennet, 2011). Likaså i länder där testkulturen inte är så stor, som till exempel i Norge, verkar man ha mött svårigheter med att införliva formativ bedömning i stor omfattning (Hopfenbeck, 2015).

Komplexiteten av en formativ klassrumspraktik (Vingsle, 2014; Andersson & Palm, 2017b) i sig själv kan vara en bidragande orsak till att införlivandet medför stora utmaningar. Nära knuten till denna orsak, som tidigare nämnts, är att formativ bedömning måste förstås på djupet för att sedan kunna införlivas på ett tillfredställande sätt i klassrumspraktiken (Flórez & Sammons, 2013; Pedder & James, 2012; Robinson m. fl., 2014). En annan svårighet är kopplad till att undervisning generellt utgör en komplex situation, som därmed är svår att förändra (Leahy & Wiliam, 2012; Pedder & James, 2012; Kennedy, 2016). Då undervisningssituationen i sig själv är komplex kan det vara särskilt svårt att införliva något så komplext som formativ bedömning, och för många lärare innebär det en stor förändring i deras sätt att agera i klassrummet (Black, 2015). Komplexiteten kan därmed utgöras av både att en formativ klassrumspraktik är komplex (särskilt om vi ser på formativ bedömning som en *integrerad helhet*) och att det är komplext att förändra sin klassrumspraktik (bland annat då det innebär att läraren måste förändra sin invanda syn på lärande och agerande i klassrummet).

Hur begreppet formativ bedömning har kommit att tolkas och (misstolkas) har utgjort ett problem. I till exempel England lanserades 2008 en nationell strategi för formativ bedömning (de använde termen ”Assessment for Learning” Bedömning för lärande). Forskare har dock kritiserat reformen för att huvudidéerna med formativ bedömning misstolkats, och att det i sin tur bidragit till införlivandet av en förvrängd syn av konceptet, där eleverna till exempel inte fokuseras som aktiva aktörer i lärandeprocessen (Swaffield, 2011). Klenowski (2009) för en diskussion kring detta för i sitt positionerande paper kring hur tidigare definitioner av begreppet *Bedömning för Lärande* successivt misstolkats och hur förståelsen av begreppet därmed börjat ändra innebörd. Ett exempel som nämns är hur ”deciding where the learner are in their learning, where they need to go, and how to best get there” (Klenowski, 2009 s. 263) har tolkats som en uppmaning till lärare att summativt testa sina elever ofta för att bedöma uppnådda nivåer i enlighet med föreskrivna nationella mål och kunskapskrav. Något som enligt forskare (t.ex. Swaffield, 2011; Torrance, 2007; 2012) fått till följd att begreppet kan förstås på ett mer instrumentellt sätt. En sådan tendens anar även (Hirsh & Lindberg, 2015) i sin genomgång av empiriska studier om formativ bedömning, och då särskilt i studier som fokuserar IKT (informations och kommunikationsteknik) som formativt verktyg. Att begreppet förstås på ett instrumentellt sätt fick till följd

att man inom vissa forskningskretsar införde begreppet 'andra generationens Bedömning för Lärande' (Klenowski, 2009). Detta för att försöka flytta fokus mot elevens delaktighet och roll i den formativa processen samt att det skulle ske dagligen i klassrumspraktiken.

Många fallstudier (t.ex. Hume & Coll, 2009) har också påvisat en instrumentell förståelse av formativ bedömning som snarare skulle kunde benämnas "teaching for the test". Dessa författare väver samman sina argument med det som Torrance (2007, s. 282) framför då han beskriver förändringen av begreppet som "assessment for learning has become assessment as learning". Vidare har forskning visat att införlivandet av formativ bedömning ibland resulterat i mekaniska och ritualiserade metoder som läggs ovanpå lärares ordinarie undervisningsmetoder och används på ett oreflekterat sätt (Hirsh & Lindberg, 2015) och många studier som visat på ytligt införlivande förekommer (t.ex. Hume & Coll, 2009; Marshall & Drummond, 2006; James & McCormick, 2009). I samma anda konstateras att många lärares införlivande av formativ bedömning ofta är mer "by the letter" än "by the spirit" (Marshall & Drummond, 2006). En följd av detta är att den viktiga roll som eleverna spelar, och deras eget ansvar för sitt lärande, inte beaktas som sig bör (Swaffield, 2011).

Några utmaningar avseende med vilken kvalitet formativ införlivas (och vad som innebär hög kvalitet) då formativ bedömning införlivas kommer att lyftas i detta avsnitt. En aspekt som kan kopplas till kvaliteten av den införlivade formativa praktiken gäller det uttrycks av bland annat Bennet (2011). Han poängterar att det inte räcker att samla information, utan att informationen också måste användas på ett genomtänkt sätt. Heritage, Kim, Vendlinski, & Herman (2009) fann till exempel att det var lättare för lärare att samla belägg för elevernas kunnande än att tolka informationen och sedan agera på ett fördelaktigt sätt för elevernas lärande. Att tolka och använda information om elevernas kunnande är ett steg som kräver mycket av läraren, vilket bland annat Vingsle (2014) visar i sin fallstudie. Forskare betonar också vikten av, och hur, de tre processerna (att samla belägg om elevernas lärande, att tolka informationen och att agera på den informationen i undervisningen) som ingår i den formativa bedömningscykeln är ömsesidigt beroende av varandra för att det ska resultera i en formativ bedömningspraktik av hög kvalitet (Heritage & Heritage, 2013; Schneider & Gowan, 2013). Vidare betonas att en brist i ett av dessa steg kan få stor inverkan på hur effektiv den formativa bedömningscykeln är. Det finns också studier som visar att hela den formativa bedömningscykeln inte genomförs fullt ut i undervisningen i stor utsträckning. I en studie visar Wylie & Lyon (2015) till exempel att en stor andel av lärarna inte genomförde hela den formativa bedömningscykeln i någon stor utsträckning efter en relativt omfattande kompetensutvecklingssatsning. En

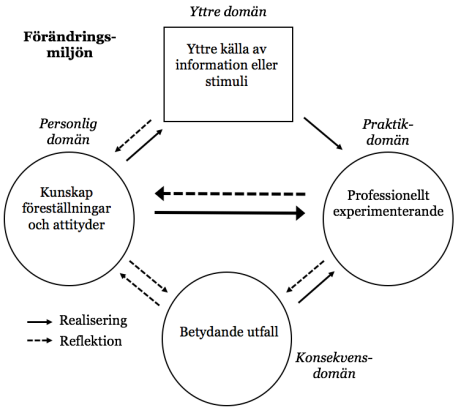
annan aspekt, som tidigare nämnts, handlar om att elevens centrala roll inom formativ bedömning inte fått det fokus som många önskat, till exempel konstateras att elevens roll inom formativ bedömning i USA ännu inte är väletablerad (Birenbaum m. fl. 2015). Storskaliga longitudinella studier i andra länder, där man försökt införliva formativ bedömning på senare år, där formativ bedömning ses som en integrerad helhet av flera strategier, har visat att Nyckelstrategi 4 (elever som lärande resurser för varandra) och Nyckelstrategi 5 (självreglerat lärande) inte är närvarande i undervisningen i så hög grad (t.ex. Lysaght & O'Leary, 2013). Andra exempel där lärarna deltagit i relativt omfattande kompetensutvecklingssatsningar visar på liknande resultat (t. ex. Jönsson m. fl. 2015). En stor utmaning gäller därmed att involvera eleverna mer i den formativa bedömningsprocessen.

I ovanstående avsnitt (2.4.2) har jag belyst några av de svårigheter med införlivandet av formativ bedömning som är aktuella inom forskningsfältet. I avhandlingen har två ramverk använts för att belysa och förstå orsaker till lärarnas införlivande av formativ bedömning och beskrivs i nästkommande avsnitt. I avsnitt 2.4.3 beskrivs en *modell för lärares professionella utveck*ling. Därefter beskrivs i avsnitt 2.4.4 den motivationsteori som använts för att belysa varför lärarna införlivade formativ bedömning.

2.4.3 Modellen om lärares professionella utveckling

Modellen om lärares professionella utveckling (eng. The interconnected model of professional growth), se Figur 3, utgör en komplex modell för lärares förändring eller professionella utveckling och presenterades av Clarke & Hollingsworth (2002). Modellen kan beskrivas som bestående av en bakgrund, *förändringsmiljön*, fyra *domäner* (kvadrat och cirklar i figuren) samt de processer (pilar i figuren) som medierar en förändring mellan domänerna. Två typer av domäner finns representerade i modellen. Den *yttre domänen*, i figuren illustrerad av en kvadrat, är placerad utanför lärarens personliga värld. Den yttre domänen utgörs av källor till information, stimuli eller stöd. De andra tre domänerna *praktikdomänen* , *personliga domänen* och *konsekvensdomänen* och är illustrerade som cirklar i figuren och utgör i kombination den individuella lärarens professionella värld. Lärarens professionella värld innefattar lärarens professionella ageranden, de konsekvenser som framträder av dessa ageranden, och de kunskaper och *föreställningar* (eng. beliefs) som föranledde och svarade på dessa ageranden. En förändring kan ske i någon av modellens fyra domäner och typen av förändring återspeglar den specifika domänen. Att experimentera med en ny undervisningsstrategi skulle till exempel tillhöra praktikdomänen, medan ny kunskap eller nya föreställningar (t. ex. föreställningar som påverkar motivation) skulle tillhöra den personliga domänen. En förändrad uppfattning av vad som för den individuella läraren uppfattas som *betydande*

utfall (eng. Salient Outcomes) relaterat till klassrumspraktiken skulle tillhöra konsekvensdomänen. Vidare så *medieras* en förändring i en domän till förändring i en annan domän genom processerna *reflektio*n och *realisering*[2] *(eng. enactment)*. Processen av *reflektion* ska förstås som aktivt ihållande och noggrann eftertanke. *Realisering* refererar till "putting into action of a new idea or a new belief or a newly encountered practice" (ibid., s. 953), vilket är skiljt från att bara agera (då ageranden sker i Praktikdomänen). Ett agerande representerar realisering av någonting som läraren vet, tror eller har upplevt.



Figur 3. Modellen om lärares professionella utveckling (efter Clarke & Hollingsworth, (2002, s. 951) – min översättning)

I modellen skiljs de fyra domänerna från förändringsmiljön och från de medierande processerna reflektion och realisering. Förändringsmiljön för lärare utgörs av den kontext i vilken lärare arbetar, det vill säga skolkontexten. Aspekter i skolkontexten kan verka genom att stödja eller motverka lärares professionella utveckling, vilket innebär att förändring i varje domän och effekten av varje medierande process gynnas eller hindras av det stöd och/eller de motverkande faktorer som förekommer i varje enskild lärares skolkontext. Skolkontexten inkluderar till exempel tillgång till kompetens-utvecklingen och den tid som kompetensutvecklingen tar i anspråk, möjligheter att testa nya idéer i klassrumspraktiken och rektorers upp-muntran att experimentera med nya undervisningstekniker.

Hur en förändring sker kan med hjälp av modellen förstås som *förändringssekvenser*. En förändringssekvens består av två eller fler domäner tillsammans med pilarna (dvs. reflektion och realisering) som kopplar

[2] Enactment har översatts till realisering, då jag inte fann något annat lämpligt begrepp och är det begrepp som jag fortsättningsvis kommer att använda.

samman domänerna. Förändring i en domän leder inte alltid till förändring i en annan domän, men när det gör det så används begreppet förändrings-sekvens. En sådan förändring kan vara flyktig, ett enda tillfälle av experimentering som snabbt överges, eller den kan vara mer varaktig och då används begreppet *utveckling*. Clarke & Hollingsworth (2002) gör alltså en skillnad i när förändringen är mer tillfällig, förändring i en domän påverkar till exempel på förändring i en annan domän, vilket då utgör en till två pilar i modellen jämfört med när förändringen är varaktig (fler pilar i modellen) och benämner det utveckling. Förekomsten av förändringar som är mer varaktiga används för att identifiera *professionell utveckling*. En förändringssekvens som är associerad med sådan professionell utveckling benämns *utvecklings nätverk*. Ett exempel på lärares professionella utveckling skulle då kunna beskrivas på följande sätt med hjälp av modellen: Läraren deltar i en kompetensutveckling och en föreläsning fungerar som yttre stimuli vilket leder till att läraren genom realisering testar en ny aktivitet i sin klassrums-praktik. Läraren tyckte att det fungerade bra att använda aktiviteten och att eleverna blev mer engagerade då hen använde aktiviteten. Detta skulle i modellen beskrivas som att läraren reflekterar över hur det har fungerat att genomföra aktiviteten i praktiken vilket leder till en förändring av lärarens kunskaper och föreställningar. Läraren reflekterar och drar även nya slutsatser om vilka betydande utfall som genomförandet av aktiviteten i praktiken medförde, vilket genom reflektion medför att läraren ändrar sina kunskaper och föreställningar. Lärarens förändrade kunskaper och före-ställningar påverkar i sin tur, så att läraren reflekterar och genom realisering testar aktiviteten igen i klassrumspraktiken. Dessutom påverkar lärarens förändrade kunskaper och föreställningar på lärarens reflektion och realisering av andra och nya idéer som presenteras i kompetensutvecklingen, samt över vilka slutsatser om de betydande utfall som användningen av aktiviteten i praktiken medför och så vidare.

2.4.4 Förväntan- värde-teorin

Den andra teoretiska utgångspunkten som spelar en stor roll i analysen är *förväntan–värde-teorin* (eng. expectancy value theory of achievement motivation) som utvecklats av Eccles och Wigfield (Eccles m. fl. 1983; Eccles & Wigfield 2002; Wigfield & Eccles, 2000). Teorin har ofta använts för att förklara och förstå elevers motivation. Teorin användes i analysen av data för att förstå de val lärarna gjorde när de deltog i kompetensutvecklingen och sedan införlivade formativ bedömning i sin klassrumspraktik. Enligt förväntan-värde-teorin är *förväntningar av framgång* (eng. expectancy of success) och de *värden* (eng. values) individen ser i att lära eller genomföra en aktivitet de huvudsakliga faktorerna för motivation och förväntas direkt påverka prestationsval, ansträngning och uthållighet (Wigfield & Eccles, 2000). De *föreställningar* (eng. beliefs) som inverkar på individens

motivation, det vill säga individens *motivationsföreställningar,* är enligt teorin alltså förväntningar av framgång och de värden läraren ser i till exempel att genomföra en aktivitet. Förväntningar av framgång kan definieras som individens föreställningar om hur väl de kommer att klara en kommande uppgift, antigen direkt eller i en framtid (Wigfield & Eccles, 2000 s. 70). Den förväntade framgången av att lära och införliva formativ bedömning i sin klassrumspraktik kan inkluderas i sådana föreställningar. Olika komponenter av värden är *personligt värde* (eng. attainment value), *inre värde* (eng. intrinsic value), *nyttovärde* (eng. utility value) och *kostnad* (eng. cost). Personligt värde avser vikten av att genomföra en aktivitet väl och är kopplat till personens identitet eller självbild. När en aktivitet bekräftar en själv som person känns den viktig. Det personliga värdet är alltså kopplat till relevansen av att engagera sig i en aktivitet för att bekräfta eller inte bekräfta betydande aspekter av ens ideala eller faktiska 'självschema', såsom kompetens inom ett område (Wigfield & Eccles, 1992). Därigenom kan sådana självscheman involvera idéer om vad en bra lärare är. Inre värde handlar om den tillfredställelse som man får av att genomföra aktiviteten och handlar alltså om det intresse, de positiva känslor eller den glädje som en aktivitet lockar fram. Sådana känslor kan till exempel framkallas av tillfredställelsen av att lära sig och använda nya undervisningsmetoder såsom formativ bedömning. Nyttovärdet handlar om användbarheten och nyttan av aktiviteten och hur en aktivitet passar in i en individs framtida planer. Förbättrat lärande för elever eller ökat elevengagemang i klassrummet kan till exempel anses vara nyttovärden. Kostnad, som är ett negativt värde, handlar om vad det kostar någon att utföra en viss aktivitet. Det hänvisar till hur beslut om att engagera sig i en aktivitet (till exempel lära sig att införliva formativ bedömning) begränsar tillgången till andra aktiviteter (till exempel planering av vanliga klassrumsaktiviteter). Det hänvisar också till hur mycket ansträngning som krävs för att utföra aktiviteten och den känslomässiga kostnaden. I Artikel III fokuseras lärarnas förväntningar av framgång och att de kan se värden i att lära och införliva formativ bedömning i sin klassrumspraktik.

3. Forskningsdesign och metoder

De fyra studierna i denna avhandling är alla inom ramen för det tidigare beskrivna forskningsprojektet. Designen av projektet kan sägas vara problemorienterad, dvs vi har använt de metoder som vi tror kan ge data för att kunna belysa de problemställningar vi haft. Specifika beskrivningar av metoderna finns i de ingående artiklarna, i detta kapitel försöker jag därför att upprepa så lite som möjligt och snarare komplettera med detaljer som inte fanns med i artiklarna. Kapitlet inleds med den övergripande forskningsdesignen för avhandlingsprojektet. Därefter presenteras designen av kompetensutvecklingen, följt av datainsamling, analysmetoder och etiska överväganden.

3.1 Övergripande forskningsdesign

Forskningsdesign behandlar vilka frågor man ställer och hur forskningsprocessen planeras och genomförs, dvs insamling och analys av data (Bryman, 2011). Forskningen som presenteras i denna avhandling kan beskrivas som en interventionsstudie. Effekten av interventionen (kompetensutvecklingen i formativ bedömning för matematiklärare) undersöktes både på lärarnivå och på elevnivå (analys gjordes på klassnivå). Huvudkomponenten i designen är interventionen och en mix av metoder användes med både kvantitativa och kvalitativa inslag. Data samlades in vid flera tillfällen och på flera olika sätt, se Figur 4 och Tabell 1.

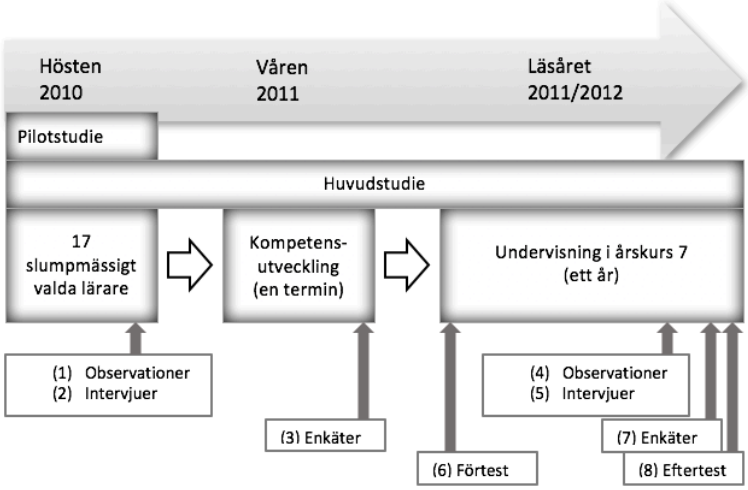
3.1.1 Mixed methods design

En kombination av både kvantitativa och kvalitativa metoder definieras som *mixed methods design* (Cohen, Manion & Morrison, 2011). En mix av metoder användes för att undersöka lärarnas formativa klassrumspraktik (Artikel I och II) och för att undersöka lärarnas motivation för att förändra sin formativa klassrumspraktik (Artikel III), där kvalitativa analyser mixades med kvantitativa analyser. Artikel IV, där elevernas prestationer i matematik undersöktes, hade däremot en i huvudsak kvantitativ karaktär. Forskare har kritiserat en skarp kontrast mellan kvalitativa och kvantitativa metoder (Bryman, 2011), och Cohen m. fl. (2011) argumenterar för att forskning som utnyttjar flera metoder är den nya forskningsparadigmen inom utbildningsvetenskap. Sådan forskning kan utnyttja de kompletterande kvaliteterna hos de två forskningsstilarna och kombinera kvantitativa och kvalitativa dataanalyser för att få en bred bild av det beforskade objektet. Ett brett spektrum av empiriska data samlades i avhandlingsprojektet och fyra metoder för datainsamling användes (dock gäller tre metoder på lärarnivå och en metod elevnivå), vilket möjliggör triangulering av data. Triangulering kan definieras som användandet av två eller flera metoder för datainsamling av någon aspekt av mänskligt beteende (Cohen m. fl., 2011). Tekniker som

innefattar triangulering används för att försöka kartlägga eller mer utförligt förklara rikedomen och komplexiteten av mänskligt beteende genom att både kvantitativa och kvalitativa data används och genom att studera det från mer än en utgångspunkt (Cohen m. fl., 2011). En annan förtjänst med triangulering av data argumenterar Wellington (2000) för, som menar att den mest värdefulla aspekten av triangulering är dess användning för validering - för att validera påståenden, intervjuuttalanden, transskript, eller data från enkät-undersökningar. Att använda en mix av metoder har bidragit till att resultat och slutsatser i de ingående artiklarna och i avhandlingen som helhet har blivit mer fylliga och dessutom har möjligheten att triangulera data kunnat ökat trovärdigheten i slutsatserna som dras.

3.1.2 Översikt över avhandlingsprojektet och dess datainsamling

I detta avsnitt presenteras en översikt över hela avhandlingsprojektet. Hur olika personer i forskargruppen genomförde arbetet och det externa stöd som användes kommer också att beskrivas. Hela projektet planerades runt den omfattande kompetensutvecklingssatsning som genomfördes. Kompetensut-vecklingen designades särskilt för detta forskningsprojekt och dess design beskrivs utförligt i avsnitt 3.2. Figur 4 presenterar datainsamlingen inom avhandlingsprojektet och när datainsamlingen skedde i tid och i förhållande till kompetensutvecklingssatsningen. I Tabell 1 presenteras en översikt över avhandlingen, bestående av pilotstudien och huvudstudien. De olika datakällorna är numrerade i Figur 4 och denna numrering återkommer i Tabell 1 i kolumnen 'Empiri', för att tydliggöra vilka datakällor som använts i för de olika forskningsfrågorna. I Tabell 1 presenteras forskningsfrågor, empiri, dataanalys och den artikel som detta presenteras i.



Figur 4. Datainsamling inom avhandlingsprojektet samt när den skedde i förhållande till kompetensutvecklingssatsningen.

Tabell 1. Avhandlingsöversikt

Forskningsfrågor	Empiri	Dataanalys	Artikel
	Pilotstudien		
Hur använder matematiklärarna formativ bedömning i sin klassrumspraktik?	8 högstadielärare Intervjuer Observationer	Kategorisering av klassrumspraktik och intervjusvar	
	Huvudstudien		
Använder matematiklärare i kommunen formativ bedömning och om de gör det, hur använder de formativ bedömning?	17 högstadielärare och 21 mellanstadielärare Observationer (1) Intervjuer (2)	Kategorisering av klassrumspraktik och intervjusvar	Artikel I
	Kompetensutvecklingen		
På vilka sätt förändras högstadielärarnas formativa klassrumspraktik efter kompetens-utvecklingssatsningen? Hur är dessa förändringar i jämförelse med de som mellanstadielärarna gjorde?			
Vilka förutsättningar för införlivande av formativ bedömning i sin praktik upplevde högstadielärarna som viktiga, och vilka hinder för förändring upplevde dom? Hur är dessa förutsättningar och hinder i jämförelse med de som mellanstadielärarna upplevde?	14* högstadielärare (22 mellanstadie-lärare) Observationer (4) Intervjuer (5) Enkäter (3 & 6)	Kategorisering av klassrumspraktik, intervjusvar och enkätsvar	Artikel II
Hur kan högstadielärarnas förändring mot en mer formativ klassrumspraktik förklaras av deras motivationsföreställningar?		Kategorisering av intervjusvar och enkätsvar	
Hur kan skillnaden i antalet införlivade formativa bedömnings-aktiviteter för högstadie- och mellanstadielärarna förklaras av skillnader i deras motivationsföreställningar?	14* högstadielärare (22 mellanstadie-lärare) Intervjuer (5) Enkäter (3 & 6)	Statistisk analys (medelvärden och signifikanstest) av enkätsvar	Artikel III
Har den formativa klassrumspraktiken som införlivades av högstadielärarna, efter kompetensutvecklingssatsningen, någon påverkan på deras elevers prestationer i matematik?	291 elever i 14 klasser som undervisades av lärare i interventions-gruppen och 275 elever i 15 klasser som undervisades av kontrollgruppslärare	Statistisk analys (ANCOVA)	
Finns det en korrelation mellan antalet nya formativa bedömningsaktiviteter som införlivades av lärarna och elevernas prestationsvinster i matematik?	För- och eftertest för elever (7 & 8)	Partiell korrelation	Artikel IV

* 3 lärare kunde av olika skäl (såsom flytt och byte av jobb) inte följas upp efter kompetensutvecklingen och därför är det 14 högstadielärare som ingår i Artikel II-IV.

Jag och forskningsledaren genomförde pilotstudien och forskningsledaren ansvarade för design och genomförande av kompetensutvecklingen i pilotstudien likväl som i huvudstudien. Jag deltog vid kompetensutvecklingen i pilotstudien och i huvudstudien och förde minnesanteckningar under träffarna med lärarna. Efter varje träff diskuterade och reflekterade forskningsledaren och jag kring träffen och kring nästa träffs upplägg och innehåll. Forskargruppen planerade insamling och analys av data och datainsamling och dataanalys genomfördes av doktoranderna, med stöd från forskningsledaren (som även var handledare för oss doktorander). Jag ansvarade för högstadielärarna (som också är i fokus i denna avhandling) och min doktorandkollega ansvarade för mellanstadielärarna. Vid datainsamlingen för Artikel II bistod också en annan forskare som genomförde intervjuer och cirka hälften av observationerna. Dessutom hyrdes några pensionerade matematiklärare in och bistod med rättning av testerna som eleverna gjorde.

3.1.3 Pilotstudien

Pilotstudien beskrivs kort här för att ge information om hur den genomfördes och vad den bidrog med, eftersom pilotstudien inte presenteras i sin helhet i något annat avsnitt i avhandlingen. De fyra artiklarna som ingår i avhandlingen presenteras däremot inte separat här utan sammanfattas i Kapitel 6. Pilotstudien startade under senare delen av våren 2010. Rektorer kontaktades och informerades om den kommande kompetensutvecklingssatsningen som skulle genomföras under hösten 2010. Rektorerna gav sedan förslag på åtta högstadielärare som visat intresse av att delta i kompetensutvecklingssatsningen i formativ bedömning. Klassrumsobservationer och intervjuer genomfördes därefter med de åtta lärare som nästkommande termin (hösten 2010) skulle delta i kompetensutvecklingssatsningen. Under hösten 2010 genomfördes sedan kompetensutvecklingen. Pilotstudien bidrog till utveckling och design av kompetensutvecklingen i huvudstudien, samt utveckling och utformning av datainsamlingen (observationer och intervjuer). Alla intervjuer transkriberades, och tre av de åtta lärarnas klassrumspraktik analyserades genom narrativ. Narrativen strukturerades enligt den stora idén och de fem nyckelstrategierna av respektive lärares klassrumspraktik. Struktureringen utifrån ramverket gjordes för att utröna om vi från observationer och intervju fick sådan information att vi sedan utifrån dessa narrativ skulle kunna göra de analyser vi avsåg göra. Några justeringar av intervjuguide och observationsschema gjordes inför huvudstudien. Pilotstudien genomfördes av mig och forskningsledaren. Forskningsledaren ansvarade för kompetensutvecklingen, jag genomförde klassrumsobservationerna, och lärarintervjuerna genomfördes av oss tillsammans. Analysen i form av narrativen genomfördes av mig med stöd från forskningsledaren.

3.2 Design av kompetensutvecklingen

Designen av kompetensutvecklingen baserades på tidigare forskning om kompetensutveckling (t.ex. Timperley m. fl., 2007; Desimone, 2009; Villegas Reimers, 2003), se även avsnitt 2.4. Som tidigare nämnts har tidigare forskning visat på svårigheter med att införliva formativ bedömning i klassrumspraktiken, och lärare behöver därför mycket stöd i den processen (t.ex. Flórez & Sammons). Detta låg till grund för beslutet att genomföra en omfattande kompetensutvecklingssatsning innehållande de flesta egenskaper som framgångsrika kompetensutvecklingar visat sig ha enligt tidigare forskning. Kompetensutvecklingen var processorienterad, hade en formativ karaktär och var organiserad och leddes av forskningsledaren (som även kommer att kallas kursledaren). Innehållet i kompetensutvecklingen var formativ bedömning som en *integrerad helhet* (se avsnitt 2.2.2) i enlighet med Wiliam och Thompsons (2008) ramverk. Kompetensutvecklingen inkluderade att lärarna fick gott om tid (ca 216 timmar totalt), vilket var möjligt då lärarna fick nedsättning i sin tjänst motsvarande 20% av en heltidstjänst under den termin då kompetensutvecklingen pågick. Kompetensutvecklingen startade i januari 2011 och pågick under vårterminen 2011. På grund av organisatoriska skäl flyttades två kursträffar till september istället för att enligt tidigare planering enbart genomföras under vårterminen. Detta var dock något som i efterhand uppskattades av lärarna då det gav dem en bra möjlighet att inspireras och planera för det kommande införlivandet av formativ bedömning under läsåret 2011/2012. Kompetensutvecklingen bestod av kursträffar på universitet, då de deltagarna träffade kursledaren (forskningsledaren) i fyra timmar en gång i veckan under en termin (96 timmar totalt). Dessutom hade lärarna utöver dessa träffar 120 timmar till förfogande. Denna tid var avsedd till att läsa kurslitteratur, till reflektion, till samarbete med övriga kursdeltagare, till att planera för införlivandet av nya formativa bedömningsaktiviteter i sin klassrumspraktik och till att reflektera över utfallet.

En vanlig kursträff bestod av föreläsningar som presenterade teorin om formativ bedömning och dess forskningsbas. I linje med slutsatserna från en forskningsöversikt om kompetensutveckling av Kennedy (2016) blev lärarna inte ålagda exakt hur, och vilka, formativa bedömningsaktiviteter de skulle införliva. Däremot fick de vid föreläsningarna en presentation av ramverket av formativ bedömning av Wiliam & Thompson (2008). Lärarna presenterades också med forskningsbelägg som visade på värdet med att använda formativ bedömning. Utöver detta presenterades konkreta aktiviteter som exempel på hur den stora idén och nyckelstrategierna kan komma till uttryck i praktiken. Vid föreläsningarna lades stor vikt på att deltagarna skulle skapa en djup förståelse av syftet med formativ bedömning, den stora idén och de fem nyckelstrategierna (se avsnitt 2.2.2). Det kompletterades som tidigare

nämnts, med exempel på konkreta formativa bedömningsaktiviteter att använda i klassrummet. En annan viktig komponent i kompetensutvecklingen var att lärarna förväntades omsätta teori till praktik. Mellan kursträffarna förväntades lärarna testa de nya formativa bedömningsaktiviteter som introducerats och diskuterats vid kursträffarna i sin klassrumspraktik. Kursträffarna inkluderade även diskussioner i liten och stor grupp om formativ bedömning och införlivandet av formativ bedömning. Diskussionerna gav lärarna möjligheter att diskutera erfarenheter från förra veckans införlivning. Diskussionerna gav även lärarna möjlighet att utvärdera hur införlivandet av formativa bedömningsaktiviteter fungerat och att dela upplevelser av framgång likväl som misslyckanden. De kunde även vid dessa diskussioner hjälpa varandra att komma över hinder och utveckla och modifiera användningen av de nya formativa bedömningsaktiviteterna. Kursledaren stödde dessa diskussioner när det bedömdes lämpligt, och gav till exempel förtydliganden kring teorin om formativ bedömning eller gav förslag till förändringar av aktiviteter som inte fungerat bra i praktiken.

Lärarna förväntades också ta ett aktivt ansvar i designen av kursträffarna. De deltog till exempel i beslut om vad olika kursträffar skulle fokusera, tidsramar och vilka formativa bedömningsaktiviteter de skulle testa i sin klassrumspraktik inför nästa kursträff. De flesta formativa bedömningsaktiviteterna som introducerades i kompetensutvecklingen diskuterades och utarbetades under träffarna och utifrån lärarnas behov fick detta ta olika lång tid. Användningen av så kallade små skrivtavlor (eng. *miniwhiteboards*) diskuterades till exempel återkommande under flera kursträffar, då lärarna succesivt överkom praktiska hinder vid användning av dessa i klassrummet. Små skrivtavlor kan användas som ett svarssystem, vid informationsinsamling om elevernas kunskaper, så att alla elever får ge svar på de frågor läraren ställer. Vid den sista kursträffen gavs lärarna tid och möjlighet att lägga upp en plan för hur de planerade att införliva formativ bedömning i sin klassrumspraktik. En plan kan utgöra ett stöd för lärarna, samtidigt var det något som helt lämnades över till dem, då vi inte ville att lärarna skulle känna något krav på om, vad och hur mycket de avsåg att införliva under det kommande läsåret. Lärarna blev ombedda att inte prata med sina kollegor om vad de lärt under kompetensutvecklingen. Detta för att i största möjliga mån undvika överspillningseffekter till kontrollgruppen.

Mellanstadielärarna i den parallella studien deltog i en liknande kompetensutvecklingssatsning. Det var dock några små skillnader i upplägget. Alla mellanstadielärare var fria från undervisning en hel dag. Den dagen var förlagd på universitetet med kursträffar som varade i sex timmar, till skillnad från högstadielärarna vars kursträffar på universitetet utgjordes av en halv dag (fyra timmar). Att kompetensutvecklingens upplägg skiljde sig för de två

lärargrupperna berodde främst på att några av högstadielärarnas rektorer hade svårigheter att friställa dem en hel dag från undervisning. Tanken var också att många av högstadielärarna hade kollegor på sina skolor som de skulle kunna samarbeta med genom att de arbetade på så stora skolor att flera lärare från samma skola deltog i kompetensutvecklingssatsningen. Dock var det bara cirka hälften av de deltagande högstadielärarna som hade kollegor på sin skola att samarbeta med, och trots att mellanstadielärarna skulle arbeta utan stöd från kursledaren under de två ”extra” timmarna fanns han ofta tillgänglig. Detta gav varje vecka mellanstadielärarna mer tid med en expert och också bättre förutsättningar för samarbete och diskussion med de andra kursdeltagarna.

3.3 Datainsamling

Data samlades in via lärarintervjuer, lärarenkäter, klassrumsobservationer och matematiktest för elever. Datainsamlingen beskrivs under avsnitten 3.3.2 - 3.3.5 och urvalet av studiens deltagare presenteras i avsnitt 3.3.1.

3.3.1 Deltagare

Deltagare i studien var lärare (Artikel I-III), tillsammans med deras elever (Artikel IV), från en medelstor kommun i Sverige. Lärarna som deltog i kompetensutvecklingssatsningen valdes genom slumpmässigt urval från de lärare i kommunen som nästkommande skolår var ålagda att undervisa matematik i årskurs 7. Urvalsproceduren skedde genom stratifierat urval på skolnivå. Tjugo högstadielärare erbjöds att delta i kompetensutvecklingen, två lärare tackade nej och en bytte stadium att undervisa på innan kompetensutvecklingen startade. Sjutton högstadielärare deltog därför i delstudien som genomfördes före kompetensutvecklingen (Artikel I). Av olika anledningar kunde tre av dessa inte följas genom hela huvudstudien (en flyttade, en började undervisa på ett annat stadium och fick då ingå i den parallella studien för mellanstadielärare och en bytte arbete). I delstudierna efter kompetensutvecklingen (Artikel II - IV) studerades därför 14 högstadielärare. I Artikel I och III ingick även mellanstadielärare (21 stycken i Artikel I och 22 stycken i Artikel III), där urvalet skett på motsvarande slumpmässiga sätt. Vad vi kunde bedöma visade inte bortfallsgruppen några särskilda tendenser, till exempel att de skulle vara mindre motiverade eller dylikt, utan skälen till att inte delta var av annan personlig eller organisatorisk karaktär. Värt att notera var dock att det ursprungliga bortfallet var något större för mellanstadielärarna, där sju lärare valde att inte delta i studien, medan motsvarande antal för högstadielärarna endast var två lärare (som sedan i ett tidigt skede i studien blev tre).

I Artikel IV ingick även de klasser, med elever i årskurs 7, som lärarna undervisade och en motsvarande kontrollgrupp. Interventionsgruppen bestod av de

14 lärare som deltagit i kompetensutvecklingssatsningen och deras elever. Kontrollgruppen bestod av de andra lärarna som undervisade i matematik i årskurs 7 det aktuella läsåret och deras elever. De lärare som deltagit i pilotstudien skulle inte undervisa en klass i årskurs 7 det aktuella läsåret, men det blev så för några lärare och därför uteslöts deras klasser från studien. Under läsåret 2011/2012 visade det sig att en lärare som deltagit i kompetensutvecklingen undervisade i matematik i två klasser (årskurs 7). Lotten fick då avgöra och en av dessa klasser fick ingå i interventionsgruppen och den andra klassen uteslöts ur studien. I Artikel IV ingick slutligen tidigare nämnda 14 lärare som deltagit i kompetensutvecklingen och deras elever, samt kontrollgruppen som bestod av 15 lärare och deras elever. Klasserna som undervisades av lärarna i interventionsgruppen (de *formativa lärarna*) bestod av totalt 291 elever och klasserna i kontrollgruppen bestod av 275 elever. Bakgrunden hos dessa elever var varierande avseende socioekonomiska och kulturella aspekter. En fördel med slumpmässigt urval av lärare är att vi får ett representativt urval (så bra som möjligt) av lärare i kommunen som undervisar en klass i matematik i årskurs 7. Detta borde innebära att vi fick lärare med olika bakgrund, kvalifikationer, lärarerfarenhet, intresse för formativ bedömning och så vidare.

3.3.2 Intervjuer

De två intervjuerna, före och efter kompetensutvecklingen, var semistrukturerade och en intervjuguide användes. En fördel med att använda en intervjuguide är att den bidar till att göra datainsamlingen relativt systematisk för varje respondent samtidigt som intervjun påminner om en vanlig konversation (Patton, 2002). En annan fördel är att intervjuaren kan styra över hur den ofta begränsade tiden som finns tillgänglig vid intervjusituationen fördelas (Patton, 2002). Detta var värdefullt då intervjuerna, särskilt den andra intervjun, var omfattande. Under intervjun tillämpades ett tillvägagångssätt med sonderings och uppföljningsfrågor för att fånga det som önskades och för att hålla fokus. Att använda sonderingsfrågor är enligt Patton (2002) “a skill that comes from knowing what to look for in the interview, listening carefully to what is said and what is not said, and being sensitive to the feedback needs of the person being interviewed” (Patton, 2002, s. 374). Genomgående genom intervjun ombads respondenten (läraren) vid behov att exemplifiera och ge förtydliganden. I den första intervjun var syftet att identifiera hur undervisningen genomfördes med avseende på formativ bedömning innan kompetensutvecklingen. Intervjuguiden var strukturerad enligt de fem nyckelstrategierna och den stora idén (Wiliam & Thompson, 2008) och innehöll stödfrågor såsom ”Hur används resultat från det diagnostiska testet?”. Lärarna uppmanades att använda sin egen terminologi för att beskriva sin undervisning då de inte kunde förväntas

vara bekanta med den formativa bedömningsterminologin såsom ”nyckelstrategier”.

Den andra intervjun handlade om undervisningen efter kompetensutvecklingen och orsaker till hur lärarna förändrade sin praktik. Då fokuserades de fem nyckelstrategierna och den stora idén, vilket inkluderade de formativa aktiviteter som lärarna använde i klassrummet, och orsaker till införlivandet av formativ bedömning. Intervjuguiden var indelad i tre delar. I den första inledande delen av intervjun ombads läraren att ge en beskrivning av sin undervisning efter kursen och de förändringar hen hade gjort avseende formativ bedömning. För att minska risken att någon stor förändring inte skulle komma fram i intervjun fick läraren då möjlighet att prata fritt om de förändringar hen hade gjort. Den andra delen av intervjun följde strukturen av de fem nyckelstrategierna och den stora idén och läraren ombads att utveckla sina uttalanden angående de förändringar hen gjort kopplat till varje nyckelstrategi. Dessutom ställdes i denna del av intervjun frågor som sökte orsaker till införlivandet av olika aktiviteter. Sådana frågor var till exempel: ”Vad trodde du värdet skulle vara med att använda den aktiviteten? Hur fungerade det? Var det svårt eller enkelt?” I den tredje delen ombads läraren att mer generellt ge orsaker till sina val av klassrumspraktik efter kompetensutvecklingen. Tre typer av frågor användes: (1) frågor om varför specifika ändringar genomförts och inte andra (”Vad har påverkat dig att göra ändringar i den här omfattningen? Varför gjorde du allt detta och inte mindre? Varför gjorde du inte mer?”) (2) Frågor om det stöd de hade på skolan som har påverkat införlivandet av formativ bedömning (”Är det något stöd som du har fått på skolan som har påverkat införlivandet positivt? Är det något stöd/förutsättning på skolan som du saknar för att kunna genomföra den formativa klassrumspraktik du skulle vilja genomföra? Vilket stöd hade du önskat att du hade fått från till exempel din rektor?”), och (3) frågor om lärarens generella upplevelser från införlivandet av formativ bedömning (t.ex. ”Vad är dina upplevelser av det här läsåret med undervisning som innehöll nya formativa bedömningsaktiviteter? Vilka fördelar och nackdelar har du upplevt med detta sätt att undervisa?”). Den andra omgången intervjuer, efter kompetensutvecklingen, genomfördes av en annan forskare som var väl insatt i ämnet och intervjuguiden.

3.3.3 Observationer

Vid observationerna fördes fältanteckningar och ett observationsschema användes. Två observationer per lärare genomfördes före och efter kompetensutvecklingen. Observationerna fyllde flera syften. Ett syfte var att bidra med en referensram till, och att stödja kommunikationen i, intervjun så intervjuaren kunde relatera till konkreta situationer från observationerna. Ett annat syfte var att via triangulering stödja intervjudata vid identifiering av

formativa bedömningsaktiviteter. För att underlätta identifiering av dessa aktiviteter var observationsschemat strukturerat i linje med den stora idén och de fem nyckelstrategierna. Observationsschemat bidrog med struktur och stödjande frågor såsom ”Hur presenterar läraren lärandemålen?”. Detta exempel på stödfråga relaterar till Nyckelstrategi 1. Min inställning under observationerna var att samla så mycket information som möjligt, med fokus på indikationer av formativ bedömning i klassrummet. Utifrån observationernas syfte, att vi ville få en bild av deras ordinarie klassrumspraktik, ville vi göra oannonserade observationer. Alla lärare gav sitt samtycke till detta och vi kom överens om att jag skulle meddela dem under vilken tidsperiod på cirka sex veckor som observationerna skulle ske, och att jag sedan oanmäld kunde ansluta vid lektionsstart. Lärarna ombads även meddela mig om det skedde några förändringar i deras ordinarie schema, så att jag fick veta om en lektion var inställd eller om den inte var lämplig att besöka (t. ex. om eleverna skrev prov). Endast någon enstaka lärare tog sådan kontakt som alltså inte var så omfattande, och den kontakten skedde via mail. Endast vid några få tillfällen kunde jag inte genomföra observationerna som planerat. Vid den första omgången observationer (före kompetensutvecklingen) hade jag ett tight schema för att hinna genomföra de planerade observationerna och ett något annorlunda upplägg var därför nödvändigt. Jag meddelade då läraren kvällen innan, via mail eller sms, att jag planerade besöka deras lektion nästkommande dag. Lärarna hade då möjlighet att meddela mig om lektionen var inställd, och vi stämde träff på skolan för att jag enkelt skulle hitta till lektionssalen. Vid observationstillfället informerades klassen om vem jag var, och att mitt fokus var att observera läraren och inte dem (eleverna). De informerades även om att min avsikt var att inte störa eller påverka undervisningen. Det var ett sätt att hantera ”observatörseffekten” som enligt Wellington (2000) är ett ”unnatural behaviour by the subjects of research due to the presence of an outsider” (Wellington, 2000 s. 66). Min upplevelse var att eleverna var vana vid att ha besökare i klassrummet och att de inte brydde sig nämnvärt om min närvaro (vilket förvånade mig lite). Endast vid något enstaka tillfälle tog någon eller några elever kontakt. Observationerna före kompetensutvecklingen genomfördes i årskurs 7 till årskurs 9.

3.3.4 Enkäter

De två enkäterna inkluderade frågor med fem svarsalternativ och öppna frågor. Valet att använda frågor med båda dessa varianter av svarsalternativ gjordes utifrån de analyser som planerades. Kvantitativa analyser (frågor med fem svarsalternativ) användes för att få ett mått på lärarnas grad av motivation och för de öppna frågorna användes mer kvalitativa analyser. Svarsalternativen i flervalsfrågorna tilldelades nummer från 1 – 5 (där 1 = liten utsträckning eller liten förändring, och 5 = stor utsträckning eller stor förändring). Båda enkäterna fokuserade lärarnas grad av motivation och de

motivationsföreställningar som enligt *förväntan-värde-teorin* påverkar motivationen. De nyss nämnda motivationsföreställningarna var förväntad framgång (den första enkäten) och värde (personligt värde, nyttovärde och kostnad). Den första enkäten fokuserade på lärarnas upplevelser av kompetensutvecklingssatsningen och deras planer och förväntningar på det kommande införlivandet av formativ bedömning. Den andra enkäten bestod av frågor om lärarnas upplevelser av sin förändrade praktik, orsaker till dessa förändringar och deras planer och förväntningar på sin framtida klassrumspraktik. De frågor som var relevanta för studierna finns återgivna i Artikel II och Artikel III.

3.3.5 Testen

Lärarna fick i slutet av våren 2011 information om att alla kommunens årskurs 7 klasser skulle göra ett test (vi kallade det då en diagnos) i början på nästkommande läsår. Före eleverna hade anlänt i augusti fick lärarna den utförliga informationen om testet, bland annat att testet bestod av två delar som skulle genomföras under två på varandra följande lektioner med start den tredje matematiklektionen. Förtestet bestod av två delar, där eleverna fick 40 minuter till förfogande på varje del och där miniräknare endast var tillåten på del två. Den totala poängen på testet var 60 poäng. Den första delen inkluderade flervalsfrågor, ”fyll i luckor uppgifter” och uppgifter där eleverna förväntades ge ett kort svar. Del två bestod huvudsakligen av uppgifter som krävde att eleverna redovisade sina lösningar. Förtestet var utformat för att mäta elevernas uppnående av lärandemålen för årskurs 6 i de nationella kursplanerna vilket utöver det centrala innehållet inkluderade bland annat resonemangsförmåga, problemlösningsförmåga och kommunikationsförmåga.

I slutet av våren 2012, på bestämda dagar, genomfördes det andra testet av kommunens alla elever i årskurs 7. I början av terminen fick lärarna information om de aktuella testdatumen, men testet tillsammans med instruktioner om genomförandet distribuerades till lärarna först några dagar innan det skulle genomföras. Eftertestet var utformat för att testa innehåll i de nationella kursplanerna och som täcktes av läroböckerna i matematik som användes i kommunen. Eftertestet bestod också av två delar, där eleverna fick 40 min till förfogande på varje del och där miniräknare endast var tillåtet på del 2. Den maximala poängen på eftertestet var 58 poäng. Eftertestet hade ett liknande upplägg som förtestet.

Testen var särskilt konstruerade för detta projekt. Forskargruppen utformade testen tillsammans med provutvecklare av nationella prov samt erfarna högstadielärare som fungerade som en referensgrupp. Testen var utformade för att testa den matematik som är specificerad i de nationella ämnesplanerna

och kursplanerna. Förtestet var utformat för att testa innehållet (enligt kursplaner) upp till årskurs 6 och eftertestet var utformat för att testa innehållet under årskurs 7. Valet att göra två olika test (och inte två lika) var framförallt för att vi ville mäta det som det var tänkt att eleverna skulle lära sig under årskurs 7. Dessutom ville vi att testresultaten skulle vara direkt användbara för lärarna i sin formativa klassrumspraktik. Testen fyllde därmed dubbla syften, då de var lämpliga för studien samtidigt som de kunde vara användbara för läraren. Att använda samma test vid olika tillfällen, kan ha sina fördelar när man vill undersöka elevernas förbättrade prestationer, men det kan också ha sina nackdelar när eleverna redan är vana vid testuppgifterna. För att försäkra oss om att testen var lämpliga för eleverna i alla klasser inkluderades bara innehåll som var specificerat i kursplanen och som dessutom fanns med i alla läroböcker som användes vid den tidpunkten i kommunen. Testen utprövades först i en annan kommun och därefter gjordes vissa ändringar. Provutvecklingsgruppen konstruerade också detaljerade rättningsmallar som för vissa uppgifter även innehöll autentiska och poängsatta elevlösningar. Lärarna ombesörjde genomförandet av testen och de fick noggranna instruktioner om alla steg i denna process, till exempel hur provet skulle introduceras, tidsramar och hur lärarna sedan skulle leverera proven tillbaka till forskargruppen. Förtestet rättades av medlemmar i forskargruppen och en grupp erfarna pensionerade matematiklärare. Eftertestet rättades av samma grupp erfarna pensionerade matematiklärare. Rättarna visste inte vilka klasser som tillhörde interventionsgruppen och vilka klasser som tillhörde kontrollgruppen.

Alla data behandlades konfidentiellt och överfördes till statistikprogram, Excel och SPSS. Alla kommunens matematiklärare som undervisade i årskurs 7 fick därefter en fil med sina elevers resultat, där resultatet på varje uppgift redovisades samt information om vad de olika uppgifterna avsåg att mäta. Lärarna fick därmed information om deras elevers ämneskunskaper i matematik som de sedan kunde använda för att informera och planera undervisningen. Lärarna informerades också om att deras elevers resultat inte skulle rapporteras till någon annan. Dessa åtgärder vidtogs då vi ville främja en positiv inställning till testen. Lärarna fick därmed, med minimal arbetsinsats, väl utformade test som de inte behövde rätta själva och vars resultat som endast rapporterades tillbaka till dem (och till forskarna, men för andra syften). Att få lärarna positivt inställda till testerna fyllde även det dubbla syften. Antagandet var att om lärarna hade en positiv inställning gentemot testen skulle de även följa instruktioner bättre. Detta skulle då höja kvaliteten i mätningen och därmed också kvaliteteten av studien. Vi ville också att både forskare och lärare skulle dra fördel av projektet och ville därför att lärarna varken skulle känna sig, eller bli, utnyttjade.

Som ett mått på reliabiliteten uppskattades testens interna konsistens med hjälp av Cronbach's alfa. Cronbach's alfa för förtestet beräknades till 0,88 och för eftertestet till 0,92, vilket indikerar hög reliabilitet. En alfakoefficient mellan 0,8 - 0,9 indikerar hög reliabilitet för testet och en alfakoefficient >0,9 indikerar väldigt hög reliabilitet för testet (Cohen m. fl., 2011). En uppgiftsanalys för testerna genomfördes också. Analysen visade att de flesta uppgifterna diskriminerade på ett bra sätt. Några få uppgifter diskriminerade inte fullt så bra, men de inkluderades för att inleda testen med några "uppvärmningsuppgifter", och de hade bara en marginell effekt på den totala reliabiliteten för testet.

3.4 Analysmetoder

I det här avsnittet beskrivs hur analysen av data har skett för att besvara de forskningsfrågor som ingår avhandlingens artiklar. Avsnittet innehåller också till viss del motiveringar till valet av analysmetod som valdes.

3.4.1 Lärarnas användning av formativa bedömningsaktiviteter

I analysen användes intervju och observationsdata för att identifiera de formativa bedömningsaktiviteter som lärarna använde regelbundet i klassrumspraktiken innan kompetensutvecklingen (Artikel I), samt de nya aktiviteter som användes regelbundet efter kompetensutvecklingen (Artikel II). Analysen baserades på ramverket av Wiliam och Thompson (2008), bestående av fem nyckelstrategier och en stor idé. Ramverket i sig specificerar inte särskilda aktiviteter som ska användas för att uppnå syftet med den stora idén och nyckelstrategierna. En formativ bedömningsaktivitet definierades som: aktiviteter som används i klassrumspraktiken som har potential att bidra till uppfyllandet av målet med åtminstone en nyckelstrategi och den *stora idén* med formativ bedömning. Många olika aktiviteter kan användas för att nå dessa syften, och några kan anses mer framgångsrika i att bidra till att uppnå till exempel en gemensam förståelse av lärandemålen (Nyckelstrategi 1). Alla fem nyckelstrategier kan bidra till den stora idén, att använda belägg för elevers lärande för att justera undervisningen till att bättre möta elevernas lärandebehov. Att samla belägg om elevers lärande (Nyckelstrategi 2) är till exempel nödvändigt för att man ska kunna göra justeringar som bättre möter elevernas lärandebehov. Bortsett från lärares feedback (Nyckelstrategi 3), så är lärarens justering av undervisningen inte inkluderad i någon av strategierna. I analysen kompletterades därför nyckelstrategierna med den delen av den *stora idén*, som vi kallade justerad lärarundervisning (eng. "adjusted teacher instruction (ATI)". Den justerade lärarundervisningen beskriver vilken typ av justering läraren gör som en respons på den insamlade informationen.

Analysen för Artikel I genomfördes på följande sätt. Fältanteckningar från observationerna strukturerades enligt de fem nyckelstrategierna och den *stora idén*. Fältanteckningarna innehöll anteckningar om allt som jag som observatör uppfattat skulle kunna anses vara en indikation på formativ bedömning. Alla intervjuer spelades in (ljud) och transkriberades. Lärarnas intervjuuttalanden användes sedan tillsammans med fältanteckningar från observationerna för att skapa beskrivningar, som inkluderade citat från båda datakällorna, av varje lärares undervisning med fokus på formativ bedömning. Om det förekom någon diskrepans mellan de två datakällorna noterades det, men det var väldigt ovanligt. Dessa beskrivningar strukturerades också enligt de fem nyckelstrategierna och den *stora iden*. Lärarna fick därefter möjlighet att läsa beskrivningarna och kommentera den tolkning som gjorts av deras klassrumspraktik. I vissa fall bad jag dem också förtydliga vissa oklarheter. Beskrivningarna justerades sedan till en version som både jag som forskare och läraren var överens om, en procedur som är en form av deltagarvalidering (Kvale, 2014). Dessa beskrivningar användes sedan som data i den fortsatta analysen. I beskrivningarna söktes efter aktiviteter som ansågs tillhöra till de fem nyckelstrategierna eller justerad lärarundervisning. Frekvent använda formativa bedömningsaktiviteter identifierades enligt en trestegsprocess och de aktiviteter som uppfyllde de uppsatta kriterierna för dessa bedömdes som att läraren använde aktiviteten regelbundet i sin klassrumspraktik. I Steg 1 identifierades alla formativa bedömningsaktiviteter som läraren använde. En särskild utmaning när det gällde kategoriseringen i Steg 1 rörde framkallande av information om elevers lärande (Nyckelstrategi 2), då dessa belägg inte alltid användes för formativa syften (till exempel kunde prov användas enbart för summativa syften). För att en aktivitet skulle kategoriseras som att den ansågs höra till Nyckelstrategi 2 krävde vi att informationen skulle användas för att justera undervisningen för åtminstone några elever. I Steg 2 avgjordes om aktiviteten kunde anses användas regelbundet av läraren. Vad som ska räknas som regelbundet är självklart en bedömningsfråga och kan skilja mellan olika aktiviteter. Intervjuuttalanden användes för att avgöra om aktiviteten användes regelbundet. Om läraren till exempel uttryckte att hen använder aktiviteten ”ofta” eller ”brukar använda”, alternativt angav hur ofta, varvid minst varannan vecka ansågs som ofta (förutom för vissa aktiviteter såsom prov som inte kan förväntas ges så ofta) så uppfylldes kriteriet. Slutligen, för att passera Steg 3 krävde vi att klassificeringen av en aktivitet som formativ skulle vara välgrundad. Kriteriet var att läraren gav en ”rik” beskrivning av aktiviteten eller att intervjudata kompletterades med observationsdata. Observationsdata användes därmed för att komplettera och verifiera intervjudata. En ”rik” beskrivning innebar att läraren exemplifierade hur hen använde aktiviteten eller gav en utförlig beskrivning av aktiviteten. Det räckte alltså inte att läraren bara sa att hen använde aktiviteten.

Analysen för Artikel II genomfördes på liknande sätt som för Artikel I, med den skillnaden att det var förändringar i den formativa klassrumspraktiken som analyserades. Analysen genomfördes i två steg. I det första steget användes intervjutranskript och fältanteckningar från observationerna för att identifiera de nya formativa bedömningsaktiviteter (förändringar) som läraren införlivat i sin klassrumspraktik. En aktivitet ansågs som en ny aktivitet i lärarens praktik om aktiviteten inte använts tidigare eller om den använts i mindre omfattning än tidigare. En aktivitet kunde också anses vara en ny aktivitet om den nu användes på ett modifierat sätt. Nya formativa bedömningsaktiviteter kunde alltså vara aktiviteter som presenterats under kompetensutvecklingen och som införlivats som de presenterats där eller i en ny form. Det kunde också vara gamla aktiviteter som modifierats eller som nu användes mer frekvent. Dessutom krävdes att aktiviteten användes regelbundet som ett vanligt inslag i lärarens klassrumspraktik. För att avgöra om aktiviteten användes regelbundet användes intervjuuttalanden som indikerade hur ofta en aktivitet användes. Uttalanden som till exempel ”mycket”, ”ofta”, “vid varje kapitel” eller “varje vecka”, tolkades som att de beskrev tillräcklig regelbundenhet. Gällde det en aktivitet som läraren redan använde före kompetensutvecklingen krävdes att lärarna angav att aktiviteten nu användes ”oftare” eller ”mycket mer” eller ett liknande uttalande. Uttalanden som ”sällan”, ”ibland” eller ”bara testat”, tolkades å andra sidan som att aktiviteten inte användas regelbundet och ansågs då inte vara en förändring av klassrumspraktiken. Vid behov gjordes även jämförelser med de formativa aktiviteter som läraren använde före kompetensutvecklingen (Artikel I). Ingen deltagarvalidering skedde denna gång av tidsskäl och praktiska skäl. En jämförelse med motsvarande resultat avseende mellanstadielärarnas införlivande av nya formativa aktiviteter gjordes också och några framträdande likheter och skillnader identifierades.

Tillförlitligheten i kategoriseringen ombesörjdes genom att alla beslut om vilka aktiviteter som skulle anses vara formativa bedömningsaktiviteter skedde gemensamt i forskargruppen. En utgångspunkt för indelningen av aktiviteter i kategorier var också att varje formativ bedömningsaktivitet enbart skulle tillhöra en kategori, för att undvika dubbelkategorisering. Detta sätt att dela in aktiviteter blir då lite konstruerat då lärarens användning av aktiviteten kan fylla flera syften. Det huvudsyfte som läraren hade med att använda aktiviteten fick avgöra den slutliga kategoriseringen. Eftersom vi ville få ett mått på antalet aktiviteter som lärarna använde före kompetensutvecklingen (Artikel I) och ett mått på de nya aktiviteter lärarna använde (Artikel II) valde vi dock att göra kategoriseringen på detta sätt. Alla tveksamma fall kring kategoriseringen diskuterades också i forskargruppen så att vi nådde en samsyn kring kategoriseringen av dessa (vilket också bör innebära att vi nådde relativt god samsyn kring övrig kategorisering).

3.4.2 Viktiga förutsättningar och hinder för införlivandet av formativ bedömning

I analysen användes enkät och intervjudata för att undersöka vilka förutsättningar som lärarna upplevde var viktiga för att införliva formativ bedömning i sin klassrumspraktik och vilka och hinder de upplevde för förändring (Artikel II). Den övergripande utgångspunkten i analysen var de tre huvudkomponenterna för kvalitativ dataanalys enligt Miles och Huberman (1994): datareduktion, datadisplay och dra och verifiera slutsatser. Först valdes, ur intervjutranskript och svaren på de öppna frågorna i enkäterna, data ut data som var relevanta för forskningsfrågan. Data ordnades och strukturerades därefter i kategorier i Excel. Dessa kategorier utvecklades successivt då jag började dra slutsatser kring viktiga egenskaper i kompetensutvecklingen och kring andra förutsättningar och hinder utifrån enkätsvar och intervjuuttalanden.

Analysen genomfördes i två steg. Först söktes svar utifrån vilka egenskaper i kompetensutvecklingen som lärarna ansåg varit viktiga för deras möjligheter att lära och införliva formativ bedömning. Lärarnas svar på de öppna frågorna i Enkät 1 och i intervjun analyserades utifrån egenskaper i kompetensutvecklingen i kategorier. Utgångspunkten var de sex kategorier som identifierats i den parallella studien för mellanstadielärare som deltagit i en liknande kompetensutveckling, men jag var även öppen för andra viktiga egenskaper (och därmed nya kategorier) som kunde dyka upp i data. De frågor som användes i Enkät 1 var frågorna A-D nedan, där A-C var följdfrågorna till följande två frågor som besvarats på en skala 1 - 5: I vilken utsträckning tycker du att du har förvärvat de kunskaper i formativ bedömning du behöver för att känna dig redo att bedriva en sådan undervisning på ett effektivt sätt? I vilken utsträckning tycker du att du har haft goda förutsättningar att förvärva dessa kunskaper? (A) Vilka förutsättningar är det? (B) Vilka förutsättningar har du inte haft som skulle ha underlättat förvärvandet av dessa kunskaper? (C) Vilka faktorer i kompetensutvecklingen tror du är avgörande för om den ska leda till de undervisningsförändringar du som lärare vill göra? (D) Ge feedback på kompetensutvecklingen och dess genomförande enligt principen ”2 stjärnor och ett obegränsat antal förslag till förbättringar”. Det var främst del 3 i intervjun, då lärarna fick besvara frågor om orsaker till de förändringar de gjort, varför de gjorde dessa förändringar och inte färre och varför de inte gjorde mer förändringar som användes i analysen. Beskrivningar av de identifierade kategorierna och exempel och klassificeringar av uttalanden från lärarna återfinns i resultatavsnittet i Artikel II. Antalet lärare som nämnde egenskapen i varje kategori gav ett mått på hur vanlig varje kategori var.

I steg två söktes svar till vad lärarna ansåg vara viktiga förutsättningar och hinder för att införliva formativ bedömning, som relaterade till andra

förutsättningar än egenskaper i själva kompetensutvecklingen. Lärarnas svar på öppna frågor i enkäterna och i intervjun, främst från del 3, var de data som analyserades och sorterades i kategorier. De enkätfrågor som användes i Enkät 1 var följdfrågan till ”I vilken utsträckning tror du att du skulle behöva andra förutsättningar än du har för att kunna genomföra de förändringar i undervisningen du vill?” som besvarades som flervalsfråga (1 – 5) och följdes av ”Vilka förutsättningar är det?” samt ”Känner du att du under läsåret skulle behöva fortsatt stöd för att kunna göra de modifieringar mot en mer formativ klassrumspraktik du vill göra? Vilket stöd skulle i så fall det vara?” I Enkät 2 användes frågan ”Vilket stöd/vilka förutsättningar är viktigast för att du fortsättningsvis kommer att använda dig av formativ bedömning i klassrummet?” Antalet lärare som nämnde varje förutsättning eller hinder gav ett mått på hur vanlig varje förutsättning/hinder var. De frågor som var relevanta för analys återfinns också i resultatavsnittet i Artikel II. En jämförelse med motsvarande resultat avseende de förutsättningar och hinder mellanstadielärarna upplevde gjordes och några framträdande likheter och skillnader identifierades.

3.4.3 Lärarnas motivationsföreställningar

I analysen användes främst enkätdata men också intervjudata för att undersöka hur högstadielärarnas förändringar mot en mer formativ klassrumspraktik kan förklaras av deras motivationsföreställningar (Artikel III). Som ett mått på lärarnas grad av motivation och deras motivationsföreställningar (dvs. till exempel förväntad framgång av att införliva formativa aktiviteter och olika värde som kan uppkomma av att införliva aktiviteter) beräknades medelvärden av lärarnas svar på enkätfrågor med fem svarsalternativ och som relaterade till detta. Frågorna finns presenterade i resultatavsnittet i Artikel III. De orsaker som lärarna gav för införlivandet eller inte införlivandet av formativ bedömning i de öppna enkät och intervjufrågorna avseende deras motivationsföreställningar kategoriserades också. Antalet lärare som gav svar som klassificerades i varje kategori kunde då bidra som ett ytterligare mått på lärarnas motivationsföreställningar. Detta mått användes för triangulering av data och för eventuell förstärkning av säkerheten i resultaten om lärarnas motivationsföreställningar som baserats på frågor med fem svarsalternativ. Lärarnas svar till de öppna frågorna bidrog också med ytterligare förståelse och en rikare beskrivning av de data som var mer kvantitativa. Baserat på dessa analyser användes därefter *förväntan-värde-teorin* för att dra slutsatser om varför lärarna införlivade formativa bedömningsaktiviteter. Dessutom gjordes en jämförelse av högstadielärarnas och mellanstadielärarnas motivation och motivationsföreställningar avseende deras svar på flervalsfrågorna. Jämförelsen gjordes för att bidra med ytterligare förståelse för hur lärarnas motivation och motivationsföreställ-

ningar påverkar införlivandet av formativ bedömning. För att testa om skillnaderna mellan grupperna (högstadielärare och mellanstadielärare) var signifikanta så genomfördes Mann- Whitney U test.

3.4.4 Elevernas prestationer i matematik och relationen med införlivade formativa bedömningsaktiviteter

I analysen användes testdata (dvs. elevernas resultat på förtestet och eftertestet) för att avgöra om den införlivade formativa bedömningspraktiken hade någon effekt på elevernas prestationer i matematik (Artikel IV). En envägs mellangruppsanalys av kovarians (ANCOVA) genomfördes. Vi valde att göra analysen på lärarnivå (eller egentligen klassnivå), då lärarna var de som var valda genom slumpmässigt urval. Att välja att göra analysen på lärarnivå är ett analytiskt tillvägagångssätt som är konceptuellt korrekt då det är läraren som utsatts för ”behandlingen”, det vill säga kompetens-utvecklingen (Kennedy, 2016). Lärarna användes därmed som analysenhet i studien om den införlivade formativa bedömningspraktiken hade någon påverkan på elevernas prestationer i matematik. Som ett mått på klassrumspraktikens effekter användes medelvärdet av elevernas resultat på eftertestet då hänsyn tagits till förtestet för eleverna i lärarens klass. Analysen med ANCOVA användes för att testa för signifikanta skillnader mellan testresultaten i klasserna i interventionsgruppen och i kontrollgruppen. För att säkerställa antaganden för ANCOVA; normalitet och homogenitet av varianser, användes Shapiro Wilk’s test and Levene’s test. För att kontrollera linearitet och homogenitet av regressionslutningar gjordes grafiska kontroller.

För att avgöra om det fanns någon relation mellan antalet nya införlivade formativa bedömningsaktiviteter av läraren och elevernas prestationer i matematik i lärarens klass genomfördes en partiell korrelation. En partiell korrelation är liknande en Pearson produkt-moment korrelation, förutom att den tillåter kontroll för ytterligare en variabel (i detta fall förtestet).

3.5 Forskningsetiska överväganden

I Sverige tillhandahåller Vetenskapsrådet etiska regler och riktlinjer för utbildningsforskning (Vetenskapsrådet, 2002; 2011) vilka följdes genom hela forskningsprocessen. När forskningsprojektet planerades togs hänsyn till etikprövningslagen, med senare tillägg (SFS 2008:192). I projektet är det lärarnas klassrumspraktik i förhållande till formativ bedömning som är det studerade fenomenet (dessutom gjorde eleverna test i matematik, något som jag återkommer till senare). Lärarnas agerande, upplevelser och uppfattningar är det som fokuseras. Då studien inte berör sådant som kan benämnas som känsliga uppgifter gjordes bedömningen att studien inte var av sådan karaktär att en formell etikprövning var nödvändig. När studien påbörjades hösten 2010, var det vetenskapsrådets forskningsetiska principer för

humanistisk-samhällsvetenskaplig forskning som främst var gällande (Vetenskapsrådet, 2002 (http://codex.vr.se/texts/HSFR.pdf)). Under projektets gång har jag fortsatt att förhålla mig till de nya dokument och riktlinjer som vetenskapsrådet tillhandahåller.

Vid mötet med lärarna innan kompetensutvecklingen startade fick de information om de studier som var planerade i forskningsprojektet. Samtliga lärare gav sitt samtycke till att delta i forskningsprojektet då de tackade ja till att delta i kompetensutvecklingen. Den information de fått tidigare kompletterades då med ytterligare information om syftet med forskningsprojektet, deras deltagande i studien, datainsamling och hur dessa data sedan skulle hanteras och deras möjlighet att närhelst kliva av forskningsprojektet men ändå fortsätta deltagande i kompetensutvecklingen osv. Denna information diskuterades och informationen förtydligades dessutom vid flertalet tillfällen under kompetensutvecklingen. Till grund för denna information låg de grundläggande forskningsetiska principerna: informationskravet, samtyckeskravet, konfidentialitetskravet samt nyttjandekravet (Vetenskapsrådet, 2002; 2011). Vi ansåg att inget samtycke behövdes för elevernas del, då den datainsamling vi gjorde i klassrummet rörde lärarnas ageranden och uttalanden. Data från elevernas matematiktester administrerades av kommunen och ansågs vara ett vanligt tillfälle för utvärdering, som en kommun kan välja att göra. Elevernas testresultat hanterades i Excel och de avidentifierades innan vi fick resultaten från kommunen.

4. Resultat

Resultaten från de artiklar som ingår i avhandlingen är presenterade i sin helhet i respektive artikel. Här presenteras inte resultaten artikelvis utan istället presenteras en sammanfattning och syntes av resultaten från Artikel I – IV. Dessutom presenteras några opublicerade data, vilka bidrar med en något utförligare exemplifiering av några av resultaten då utrymmet i artiklarna är begränsat. Resultatkapitlet är upplagt så att först presenteras effekter av kompetensutvecklingen i avsnitt 4.1, och därefter presenteras några möjliga orsaker och förklaringar till utfallet i avsnitt 4.2.

4.1 Effekter av kompetensutvecklingssatsningen

De lärare som ingår som deltagare i huvudstudien var slumpmässigt utvalda, vilket innebär att det bör vara en god spridning avseende lärarnas undervisningserfarenhet, deras intresse för formativ bedömning och så vidare. För att kunna studera hur lärarna förändrade sin klassrumspraktik, så undersöktes först hur lärarnas klassrumspraktik såg ut före kompetensutvecklingssatsningen med avseende på formativ bedömning. Lärarnas formativa klassrumspraktik före kompetensutvecklingen kunde då utgöra en 'baslinje' att utgå från i studien om hur formativ bedömning införlivades. Det gav dessutom en bild av hur matematiklärares formativa klassrumspraktik ser ut för lärare i en kommun. I detta avsnitt presenteras därför först lärarnas formativa klassrumspraktik före kompetensutvecklingen (avsnitt 4.1.1), följt av de förändringar lärarna gjorde (avsnitt 4.1.2) och sist presenteras effekten av den förändrade undervisningen på elevernas prestationer i matematik (avsnitt 4.1.3).

4.1.1. Hur såg lärarnas formativa bedömningspraktik ut före kompetensutvecklingen?

Resultaten från Artikel I visar att ett slumpmässigt urval av högstadielärare (och även mellanstadielärare) i en kommun i Sverige till viss del använder formativ bedömning i sin klassrumspraktik, men på olika sätt och i varierad grad. Lärarna samlade till exempel information om elevernas kunskaper (Nyckelstrategi 2) och utifrån den informationen gav de eleverna feedback (Nyckelstrategi 3) och justerade undervisningen. Lärarna beskrev lärandemålen (Nyckelstrategi 1), och då ofta de mål som fanns presenterade i läroboken. Lärarnas sätt att arbeta med nyckelstrategi 4 och 5 handlade i huvudsak om att de uppmuntrade sina elever att hjälpa varandra (Nyckelstrategi 4) och att ta större ansvar för sitt eget lärande (Nyckelstrategi 5). Även om resultaten visar på att formativ bedömning används i matematikundervisningen, visar de också att det finns stor potential att utveckla den formativa klassrumspraktiken (Artikel I). Ett exempel handlar om lärarnas insamling av information om elevernas kunskaper med påföljande justering av undervisningen. Lärarna samlade ofta information om elevernas kunnande

då de gick runt och hjälpte eleverna som räknade enskilt i sin matematikbok. Informationen använde lärarna sedan (förutom till individuell feedback) till att justera den planerade undervisningen. Justeringen av undervisningen kunde ske direkt under samma lektion, till exempel genom att gå igenom och diskutera en uppgift med hela klassen vid tavlan eller genom att planera om helklassgenomgången inför nästa lektion. Dessa justeringar baserades dock ofta på den information läraren fått från ett fåtal studenter och främst de elever som läraren noterat hade svårigheter med något under lektionen. När lärarna gjorde justeringar efter informationsinsamlingar från många elever, t.ex. efter kapitelprov, så gällde justeringen även här oftast bara ett fåtal elever - De som inte klarat provet. Här fanns följaktligen potential att utveckla både informationsinsamlingen så att den oftare gäller alla elever och justeringen, så att den blir bättre anpassad till alla, eller åtminstone till fler elevers individuella lärandebehov. Ett annat exempel var att det oftast var de elever som räckte upp handen och besvarade lärarens frågor vid helklassgenomgångar som bidrog till informationsinsamlingen och följaktligen fick läraren främst information om vad dessa elever kunde. Detta skulle kunna förbättras om läraren i högre grad använder sig av ett system då alla elever får svara. Dessutom var det främsta syftet med att ställa frågor, för flertalet av lärarna, att skapa engagemang. Att locka fram information om elevernas kunnande var ofta ett sekundärt syfte. Men informationen som lärarna fick från några få elever, det vill säga de elever som räckte upp handen och svarade på frågor, användes ändå ofta av läraren som belägg för hur läraren sedan valde att fortsätta med sin helklassgenomgång.

Ytterligare ett exempel utgörs av lärarnas sätt att arbeta med Nyckelstrategi 4 (att aktivera eleverna som resurser för varandra) och Nyckelstrategi 5 (att stödja dem i att bli mer självreglerande av sitt lärande). Lärarnas sätt att arbeta med dessa strategier handlade ofta om att uppmuntra eleverna till att hjälpa och ta hjälp av sina kamrater, men de beskrev sällan hur detta görs på ett bra sätt. De uppmuntrade även eleverna att ta eget ansvar, till exempel genom att lyfta hur viktigt det är att de använder lektionstiden till matematik, men de hjälpte dem inte att till exempel övervaka hur de använder sin tid. Lärarna använde sällan aktiviteter som särskilt hjälpte eleverna med *hur* de skulle ta en aktiv roll i dessa nyckelstrategier (Artikel I). Resultaten i Artikel I visade också att de två lärargruppernas (högstadielärare och mellanstadielärare) formativa klassrumspraktik före kompetensutvecklingssatsningen var relativt lika. Båda grupperna använde ungefär lika många formativa bedömningsaktiviteter i sin klassrumspraktik (median för högstadielärarna var 15 och för mellanstadielärarna 14). Några små skillnader identifierades dock, till exempel så var det endast nio av de 21 mellanstadielärarna som använde information från prov på ett formativt sätt, det vill säga använde informationen för att justera undervisningen för åtminstone några

elever, medan motsvarande siffra för högstadielärarna var 16 av 17 lärare (Artikel I).

4.1.2. Vilka förändringar mot en mer formativ klassrumspraktik gjorde lärarna?

Resultaten visar att alla lärare förändrade sin klassrumspraktik mot att bli mer formativ och de införlivade nya formativa bedömningsaktiviteter som de använde regelbundet i sin klassrumspraktik, dock i varierad grad och på olika sätt (Artikel II). Lärarna började använda nya formativa bedömningsaktiviteter som stärkte en klassrumspraktik i linje med den stora idén om att framkalla belägg om elevers lärande för att justera undervisningen så att den bättre möter elevernas lärandebehov. Lärarna införlivade 3 - 19 nya formativa bedömningsaktiviteter och medianen för antalet nya aktiviteter var 11,5. Sex lärare införlivade få (3 – 7) nya formativa bedömningsaktiviteter, åtta lärare införlivade 11 - 19 nya formativa bedömningsaktiviteter och ingen lärare införlivade 8 - 10 aktiviteter. De största förändringarna i antal införlivade aktiviteter på gruppnivå kan relateras till nyckelstrategierna i följande ordning: (1) störst förändringar relateras till Nyckelstrategi 2, (45 nya aktiviteter), (2) medelstora förändringar relateras till Nyckelstrategi 1, 3 och 5 samt lärarens justering av undervisningen, (23 – 26 nya aktiviteter) och (3) små förändringar relateras till Nyckelstrategi 4, (8 nya aktiviteter). Vid en jämförelse med mellanstadielärarna i den parallella studien, som bedrev en liknande formativ klassrumspraktik före kompetensutvecklingen, införlivade mellanstadielärarna betydligt fler formativa bedömningsaktiviteter på gruppnivå än vad högstadielärarna gjorde. För mellanstadielärarna var medianen för antal nya införlivade formativa bedömningsaktiviteter 20 och för högstadielärarna var medianen 10 (Andersson & Palm, 2017b; Artikel II).

De vanligaste och mest frekventa förändringar som lärarna gjorde var att de oftare lockade fram belägg för elevernas lärande (aktiviteter som kan kopplas till Nyckelstrategi 2) med syfte att sedan justera undervisningen. Lärarna införlivade nya sätt att göra små och snabba bedömningsaktiviteter regelbundet i sin undervisning, till exempel genom att ställa frågor där alla elever fick svara med hjälp av så kallade utgångspass[3] eller små skrivtavlor[4]. Genom användningen av dessa formativa bedömningsaktiviteter (kopplade till Nyckelstrategi 2) samlar de flesta lärarna oftare och på ett strukturerat sätt information om fler, och ofta alla, elevers lärande. En annan aktivitet som också kan kopplas till Nyckelstrategi 2 är användningen av ett system för att

[3] Ett så kallat utgångspass (engelska: exitpass) är då eleverna besvarar en eller flera frågor ställda av läraren på ett papper då de lämnar klassrummet. Det är ett svarsystem som används för att läraren ska få svar från alla elever.

[4] Liten skrivtavla (engelska: miniwhiteboard) är ett svarssystem som används för att läraren ska få svar från alla elever på en ställd fråga. Eleverna skriver då sitt svar på skrivtavlan och håller upp dem så att läraren snabbt kan bedöma alla elevers svar.

slumpmässigt fördela frågor under helklassgenomgångar, vilket cirka hälften av lärarna började använda. Att använda ett slumpmässigt system för att fördela frågor kan bidra till större elevengagemang och tänkande hos fler och ibland alla elever, genom att alla elever förväntas tänka och förbereda sig på att besvara frågan. Det kan även bidra med bättre information genom att informationsinsamlingen inbegriper fler elever eftersom elever som vanligtvis inte brukar räcka upp handen förväntas och får svara. Lärarna uttryckte att den frekventa informationsinsamlingen medfört att de vet mer om elevernas lärandebehov. De menade även att den frekventa insamlingen gett dem informationen tidigare, vilket bidragit till att de kan hjälpa sina elever på ett bättre sätt (Artikel II). Då lärarna använder dessa nya sätt att samla information från alla elever vid samma tillfälle och då alla elever får besvara samma fråga, får de information från alla elever om samma sak (t.ex. samma matematiska innehåll eller samma missuppfattning). De tolkningar lärarna kan göra om elevernas kunskaper kan därmed baseras på information från fler elever om till exempel samma matematiska begrepp. Detta medför att lärarna får en större ”säkerhet” i informationen om alla elevers lärande i klassen jämfört med tidigare, då det oftast var några få elever som besvarade frågorna som ställdes av läraren (eller frågade/bad om hjälp). Alla utom en lärare stärkte sitt arbete med att justera undervisningen utifrån den insamlade informationen om elevernas lärandebehov. Lärarna använde sig av de justeringar som de gjort tidigare men dessa justeringar genomfördes oftare och för några lärare på ett modifierat sätt. Hälften av lärarna började också att använda nya sätt att justera undervisningen. Dessutom bidrar lärarnas nya sätt att ofta samla information om elevernas kunskaper till att justeringarna som lärarna gör är mer välgrundade. Den frekventa informationsinsamlingen möjliggör också mer frekventa justeringar av undervisningen och följaktligen kortas tiden mellan justeringscyklerna. Lärarnas användning av vägvals-frågor[5] som besvaras under lektionen med hjälp av små skrivtavlor och i slutet av lektionen med hjälp av utgångspass, med efterföljande agerande, kan därmed bidra till att läraren kontinuerligt kan förse eleverna med lärande-aktiviteter som är bättre anpassade mot deras lärandebehov (Artikel II).

Inom några av de identifierade utvecklingsområdena, till exempel gällande lärarnas insamling av information om elevernas kunskaper (Nyckelstrategi 2) och efterföljande justering av undervisningen, kan därmed slutsatsen dras att lärarna som grupp utvecklat sin formativa klassrumspraktik relativt mycket. Inom ett annat område, där det också fanns stor utvecklingspotential, gjorde dock lärarna relativt små förändringar. Det gäller hur de andra aktörerna i

[5] En vägvalsfråga är en fråga som ställs av läraren för att få information om alla elevernas kunskaper. Informationen används därefter av läraren för att ta beslut om den fortsatta undervisningen.

klassrummet (dvs. eleverna) involveras i den formativa bedömningsprocessen (Artikel II). Lärarnas sätt att arbeta med Nyckelstrategi 4 (att aktivera eleverna som resurser för varandra) och Nyckelstrategi 5 (att stödja dem i att bli mer självreglerande av sitt lärande), förändrades endast i liten grad på gruppnivå efter kompetensutvecklingssatsningen. Läraren var följaktligen även fortsättningsvis den som hade det stora ansvaret för den formativa klassrumspraktiken (Artikel II).

4.1.3 Effekter på elevernas prestationer i matematik

Resultaten visar att den förändrade formativa klassrumspraktiken inte gav någon effekt på elevernas prestationer i matematik (Artikel IV). Resultaten har som tidigare nämnts (Artikel II) visat att kompetensutvecklingssatsningen gav de slumpmässigt valda lärarna tillräckligt stöd så att lärarna förändrade klassrumspraktiken till att bli mer formativ. Dock kan konstateras att den förändrade formativa klassrumspraktiken inte var av sådant slag att den bidrog till att eleverna i dessa lärares klasser presterade bättre på ett test, då hänsyn tagit till förtestet, än vad de klasser som ingick i kontrollgruppen gjorde (Artikel IV). En jämförelse med den parallella studien för mellanstadielärare visar däremot på skillnader avseende hur den förändrade formativa klassrumspraktiken påverkade elevernas prestationer i matematik. De klasser i årskurs 4 som undervisades av lärare som deltagit i kompetensutvecklingssatsningen, som förändrat sin klassrumspraktik mot en mer formativ klassrumspraktik, presterade signifikant bättre på ett test i matematik, då hänsyn tagits till förtest, i jämförelse med de klasser som ingick i kontrollgruppen (Andersson & Palm, 2017a).

Resultaten indikerar även att antalet införlivade nya formativa bedömningsaktiviteter inte har något samband med elevernas resultat. Vi fann ingen korrelation mellan antalet införlivade aktiviteter per lärare och klassernas resultat på ett eftertest i matematik, då hänsyn tagits till resultaten på förtestet (Artikel IV). Detta indikerar att det är vilka aktiviteter som införlivas och med vilken kvalitet dessa aktiviteter införlivas, snarare än kvantiteten (antal införlivade aktiviteter) som inverkar på den formativa bedömningspraktikens påverkan på elevernas prestationer i matematik. Detta är något som också tas upp i diskussionen (Artikel IV).

4.2 Varför gjorde lärarna dessa förändringar och vilka hinder mötte de?

Ett syfte med avhandlingen är att bidra med kunskap och förståelse av det stöd lärare behöver för att införliva formativ bedömning i sin klassrumspraktik. Att undersöka vad som bidrog till och vad som hindrade högstadielärarnas införlivande kan bidra med ökad förståelse för vilket stöd lärare behöver. En jämförelse med resultaten i de parallella studierna som

genomfördes med mellanstadielärare inom det *Formativa projekte*t kan också bidra med ytterligare förståelse kring möjliga orsaker till utfallet för högstadielärarna. Detta då skillnader identifierades avseende de två lärargruppernas motivation (Artikel III), antalet införlivade formativa bedömningsaktiviteter (Artikel II) och resultaten på elevnivå (Artikel IV). Varför högstadielärarna förändrade sin klassrumspraktik på det sätt och i den omfattning som de gjorde samt möjliga förklaringar till detta söktes i egenskaper hos kompetensutvecklingssatsningen (avsnitt 4.2.1), i de hinder lärarna mötte under året av införlivande ute på skolorna (avsnitt 4.2.2) och i lärarnas motivation (avsnitt 4.2.3).

4.2.1 Design av Kompetensutvecklingen

Designen av kompetensutvecklingssatsningen har förmodligen påverkat att högstadielärarna gjorde förändringar i den omfattning som de gjorde. De egenskaper i kompetensutvecklingen som lärarna upplevde som viktiga för att kunna genomföra de förändringar som de ville göra, identifierades i Artikel II. De identifierade egenskaperna överensstämmer i stort med egenskaper som forskningsöversikter identifierat som viktiga för framgångsrika kompetensutvecklingsinsatser generellt (se bakgrund avsnitt 2.4.1) och var följande: (1) Aktiviteter att direkt använda i klassrummet, (2) Att kunna testa och få uppleva värdet av att använda formativa bedömningsaktiviteter, (3) Att kompetensutvecklingen knyter ihop teori och praktik, (4) Stöd från expert, (5) Tid, både mängd och utsträckning, 6) Diskussioner och samarbete med kurskollegor, (7) Formativ och processorienterad kompetensutveckling (Artikel II). Motsvarande viktiga egenskaper i kompetensutvecklingen identifierades också för mellanstadielärarna (Andersson & Palm, 2015). Det var dock en skillnad i designen av kompetensutvecklingen för mellanstadielärarna och för högstadielärarna. Mellanstadielärarna var frilagda från undervisning på skolan en hel dag och deras kursträffar på universitetet pågick i sex timmar. Högstadielärarnas kursträffar pågick i fyra timmar. Under dessa två extra timmar fick mellanstadielärarna därmed mer tid tillsammans med en expert och mer tid att tillsammans med de andra kursdeltagarna delta i diskussioner kring kursinnehållet och kring utfallet av experimenteringen med formativa aktiviteter i praktiken. Då mellanstadielärarna fick sammanhängande tid till kompetensutvecklingen, så fick de dessutom bättre förutsättningar för eget arbete och reflektion (Artikel II). Många av högstadielärarna uttryckte att de önskat mer sammanhängande tid (t.ex. 2 - 3 timmar i rad), då den frilagda arbetstiden som de fick (motsvarande 5 timmar) utöver kursträffarna ofta var utspridd över dagar och upplevdes som splittrad. Deras arbetssituation på skolan påverkade därigenom möjligheten att ägna dessa fem timmar åt kompetensutvecklingen. Detta indikerar att det i designen av kompetensutvecklingen är viktigt att ombesörja att den tid lärarna får också schemaläggs rent fysiskt, så att de ges möjligheter att samarbeta och diskutera

med kurskollegor. Just samarbete och diskussion med kollegor och kurskollegor var också något som högstadielärarna önskade mer av när de gav förslag på förbättringar i kompetensutvecklingen (Artikel II). De flesta högstadielärarna undervisade årskurs 9 i matematik då kompetensutvecklingen genomfördes. Deras elever skulle då samma termin förberedas för att skriva det nationella provet i matematik, något som de upplevde försvårade deras möjligheter att testa formativa bedömningsaktiviteter i klassrummet. De ansåg därför att det hade varit bättre om kompetensutvecklingen hade varit förlagd på hösten, alternativt att de undervisat årskurs 7 eller årskurs 8 istället (Artikel II). Denna skillnad i designen av kompetensutvecklingen kan ha påverkat de två lärargruppernas möjligheter att lära och införliva formativ bedömning, vilket kommer att diskuteras i Kapitel 5.

4.2.2 Hinder i praktiken

De hinder som lärarna upplevde för införlivande av formativ bedömning under läsåret efter kompetensutvecklingssatsningen, då de var tillbaka på skolan på full tid igen, identifierades också i Artikel II. Det var främst fyra hinder som lärarna refererade till och de var (1) tid, (2) att de samtidigt som införlivningen av formativ bedömning skulle implementera en ny kursplan, (3) sårigheter med att ändra sitt invanda sätt att undervisa, (4) situationen i klassen. Ett hinder var brist på tid, då de upplevde en hög arbetsbelastning. Bristen på tid för införlivande av formativ bedömning förstärktes ytterligare då de samtidigt skulle implementera en ny kursplan (Lgr11), vilket bidrog till en högre arbetsbelastning och därmed också mindre tid att planera och införliva formativ bedömning i sin klassrumspraktik. Lärarna hänvisade även till hur svårt det var att förändra sitt invanda sätt att bedriva undervisningen och att det tar tid att göra förändringar, särskilt i ett initialt skede, samt att arbetssituationen i klassen inverkade. Att få stöd från sin rektor framkom också som en viktig faktor och det stöd lärarna upplevde att de fått av sina rektorer varierade, där några lärare inte upplevde att de fått något stöd alls (Artikel II). Detta trots att alla rektorerna var väl införstådda i kompetensutvecklingssatsningen och bland annat var de som ombesörjde lärarnas nedsättning i tjänst som en förutsättning för deras deltagande i satsningen. Variationen bland högstadielärarna var som tidigare nämnts stor avseende det antal formativa bedömningsaktiviteter som de införlivade regelbundet i sin klassrumspraktik. Resultaten indikerar att de lärare som införlivade få (3 – 7) formativa bedömningsaktiviteter också upplevde större hinder för införlivande då de hade lite 'extraordinära förhållanden' avseende arbetssituationen i klassen, till exempel större klasser än tidigare (Artikel II). De två lärargrupperna (mellanstadielärarna och högstadielärarna) mötte i allmänhet samma hinder för förändring. Dessa var arbetssituationen i klassen och tid, men även vissa skillnader i hur mycket de upplevde att dessa hinder påverkade införlivandet av formativ bedömning kunde identifieras

(Andersson & Palm, 2015; Artikel II). Generellt upplevde högstadielärarna mindre stöd från sina rektorer och även att det varit svårare att frigöra tid för utvecklingsarbete, införlivandet av formativ bedömning. Dessutom poängterade flera högstadielärare att implementeringen av ny kursplan (Lgr 11) inverkat på deras möjligheter att införliva formativ bedömning, något som inte lyftes som ett hinder bland mellanstadielärarna (Andersson & Palm, 2015; Artikel II).

4.2.3 Lärarnas motivation

Att högstadielärarna förändrade sin formativa bedömningspraktik efter kompetensutvecklingssatsningen kan delvis förklaras av deras motivationsföreställningar (Artikel III). Resultaten visar att alla högstadielärare var relativt högt motiverade att förändra sin klassrumspraktik direkt efter kompetensutvecklingen (Artikel III), något som de också gjorde i varierad grad (Artikel II). Efter kompetensutvecklingssatsningen hade lärarna relativt höga värden på de motivationsföreställningar som enlig *Förväntan-värde-teorin* är de viktigaste faktorerna som styr motivation. Lärarna hade direkt efter kompetensutvecklingen höga förväntningar på att lyckas med införlivandet av formativ bedömning (*förväntad framgång*) och rapporterade dessutom relativt höga värden för att göra det (Artikel III). På frågor med fem svarsalternativ där 1 innebar liten utsträckning (eller små förändringar) och 5 innebar stor utsträckning (eller stora förändringar), rapporterade lärarna i medeltal 3,6 för förväntad framgång och 4,2 för motivation att införliva formativ bedömning. Dessutom rapporterade de generellt värden över 3 i medeltal avseende sina motivationsföreställningar för införlivande av formativ bedömning. Efter ett år av införlivande (läsåret 2011/2012) ansåg lärarna fortfarande införlivande av formativ bedömning som viktigt att genomföra och rapporterade höga personliga värden (i medeltal 3,9på motsvarande skala som ovan). Deras rapporterade värden för nyttan *(nyttovärden)* med att införliva formativ bedömning i sin klassrumspraktik minskade något vid en jämförelse direkt efter kompetensutvecklingen och efter ett år av införlivande av formativ bedömning. Lärarnas förväntade och upplevda kostnad var måttlig vid båda tillfällena (Artikel III).

En möjlig förklaring till att mellanstadielärarna införlivade fler formativa bedömningsaktiviteter kan härledas till att de generellt var mer motiverade än högstadielärarna. Mellanstadielärarna rapporterade, efter deltagande i kompetensutvecklingen, generellt högre värden avseende sin motivation att införliva formativ bedömning (Andersson & Palm 2015; Artikel III). Till exempel identifierades signifikanta skillnader i hur viktigt de ansåg det var att prioritera införlivande av formativ bedömning under perioder då deras arbetsbelastning upplevdes som (för) hög. På en fråga med fem svarsalternativ

(1 = liten utsträckning och 5 = stor utsträckning) rapporterade högstadielärarna i medeltal 3,1 och mellanstadielärarna 4,1 för hur viktigt de ansåg det vara att införliva formativ bedömning under perioder med hög arbetsbelastning (Artikel III). Resultaten indikerar också (men ej signifikant) att den grupp högstadielärare som införlivade få (3 – 7) formativa bedömningsaktiviteter generellt var mindre motiverade i jämförelse med de högstadielärare som införlivade fler (10 – 19) aktiviteter i sin klassrumspraktik (Artikel III). Vilka formativa bedömningsaktiviteter som lärarna valde att införliva kan också delvis förklaras med deras föreställningar (som påverkar deras motivation) avseende dessa aktiviteter efter kompetensutvecklingen (Artikel III). Resultaten visar att lärarna i hög grad valde att införliva formativa bedömningsaktiviteter som de upplevde bidrog med stor nytta (nyttovärde) och aktiviteter som inte innebar höga kostnader. Dessutom indikerar resultaten att lärarna valde att införliva aktiviteter som de kände att de behärskar att genomföra på ett bra sätt och som de kände sig trygga med att genomföra (förväntad framgång).

5. Diskussion

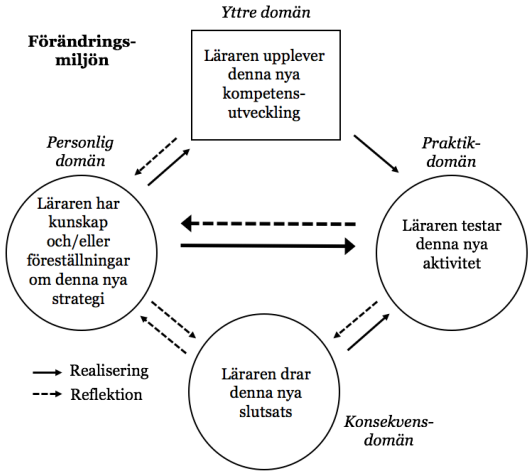
Avhandlingens syfte är att beskriva och bidra med kunskap om den (eventuella) effekt en genomförd kompetensutvecklingssatsning i formativ bedömning gav på högstadielärarnas formativa klassrumspraktik och på elevernas prestationer i matematik. Dessutom är syftet att bidra med kunskap om orsaker till att lärarna att förändrade sin klassrumspraktik på det sätt och i den omfattning som de gjorde. I detta kapitel förs en diskussion gällande studiens resultat. Kapitlet inleds med en syntes där avhandlingens resultat relateras till *Modellen om lärares professionella utveckling* (Clarke & Hollingsworth, 2002), då modellen kan utgöra ett stöd för att beskriva och förstå orsaker till lärarnas förändring. Den fortsatta diskussionen knyter också an till detta och fokuserar orsaker till lärarnas förändringar. Därefter diskuteras den uteblivna effekten på elevernas prestationer i matematik följt av reflektioner kring avhandlingens kvalitet. Efterföljande avsnitt behandlar avhandlingens bidrag till forskningsfältet och utbildningsväsendet. Kapitlet avslutas med reflektioner som knyter an till avhandlingens titel.

5.1 Resultaten i relation till Modellen om lärares professionella utveckling

Clark och Hollingsworths modell kan utgöra ett stöd för att beskriva och förstå orsaker till lärarnas förändring och finns tidigare beskriven i avsnitt 2.4.4. I Figur 5 visas en variant på modellen som explicit fokuserar läraren. Den kommer att användas för att diskutera hur kompetensutvecklingssatsningen och lärarnas arbetssituation på skolan kan ha påverkat lärarnas förändring (dvs. professionella utveckling). Viss koppling till lärarnas motivationsföreställningar görs också då förväntan - värde – teorin används i Artikel III.

Av de fyra domäner som ingår i modellen har lärarna framförallt studerats i förhållande till den yttre domänen, den personliga domänen och praktikdomänen i denna avhandling. I Artikel II studerades de förändringar lärarna gjorde mot en mer formativ klassrumspraktik, i relation till deras klassrumspraktik före kompetensutvecklingen (Artikel I). Detta kan beskrivas som en förändring i praktikdomänen. Lärarnas motivation (efter kompetensutvecklingssatsningen), för att införliva formativ bedömning studerades i Artikel III. Lärarnas motivationsföreställningar kan relateras till den personliga domänen, men ingen förändring har studerats. Vidare går vissa egenskaper i kompetensutvecklingssatsningen att relatera till den yttre domänen, se nästa avsnitt.

Enligt modellen kan lärares förändring starta i någon av domänerna som därefter kan förmedlas till en annan domän genom processerna *reflektion* och *realisering* (vilka utgörs av pilar i figuren). På vilket sätt lärares förändring inträffar är individuellt men kan ske genom olika *förändringssekvenser*. För att förändringen ska bli bestående krävs ett nätverk av förändringssekvenser för att kunna generera utveckling hos lärare. Som modellen tydliggör är det mycket som kan förändras och det är många aspekter som kan ha en påverkan på lärarnas förändring. Den process som framförallt studerats i avhandlingen är hur kompetensutvecklingen påverkat på lärarnas *realisering* av nya formativa bedömningsaktiviteter i klassrummet. Den process som framförallt studerats i denna avhandling är hur kompetensutvecklingen påverkat på lärarnas *realisering* av nya formativa bedömningsaktiviteter i klassrummet. Modellen synliggör att det inte är tillräckligt att enbart testa en aktivitet, för att det ska bli ett bestående inslag i lärarnas klassrumspraktik. Dessutom har jag indikationer på att många andra processer skett. Modellen kommer att användas för att diskutera hur denna förändring av klassrumspraktiken kan ha skett. Data gällande egenskaper i kompetensutvecklingen, skolkontexten samt information kring de förmedlande processerna *reflektion* och *realisering* kommer att användas för att diskutera lärarnas förändring i relation till modellen.



Figur 5. Modellen där varje domän operationaliserats så att ett explicit fokus på läraren framhålls (anpassad från Clarke & Hollingsworth (2002, s. 957) – min översättning).

I tillägg till de resultat som redovisats i avhandlingen finns indikationer från diskussioner vid kursträffarna hur processerna reflektion och realisering skett och därmed förändring skett. Den generella källan till stimuli (input från den yttre domänen) utgjordes av kompetensutvecklingen. Detta innefattade träffarna på universitet och läsningen av litteratur. Vid träffarna tog lärarna

del av föreläsningar och fick lära sig om formativ bedömning. Utöver detta fick de även konkreta förslag på formativa aktiviteter som de sedan kunde testa i sin egen klassrumspraktik (att testa hör till praktikdomänen). Vid träffarna deltog lärarna i föreläsningar och fick lära om formativ bedömning men de fick också konkreta förslag på formativa aktiviteter att testa i sin praktik (att testa hör till praktikdomänen). Exempel på sådana aktiviteter var att ställa frågor (vägvalsfrågor), som sedan skulle besvaras av alla elever. Frågorna ställs för att samla information om alla elevernas kunnande för att läraren därefter ska kunna anpassa undervisningen utifrån informationen och kan till exempel besvaras med hjälp av små skrivtavlor eller med hjälp av utgångspass. Vid en kursträff fick lärarna höra om syftet med att använda vägvalsfrågor som besvaras med utgångspass. De fick även aktiviteten exemplifierad under träffen och förväntades sedan testa utgångspass i praktiken innan nästkommande träff. Lärarna fick därmed ett yttre stimuli, som bidrog till deras realisering av utgångspass sin egen praktik. Vid nästa träff diskuterades utfallet av lärarnas experimenterande. De reflekterade bland annat över hur användningen av utgångspass fungerat. Det var enligt lärarna enkelt och inte så tidskrävande att använda sig utav utgångspass, vilket i sin tur bidrog till en förändring av deras föreställningar (fortsättningsvis i texten kommer jag att skriva föreställningar, men det kan även innefatta lärarnas kunskaper och attityder) om aktiviteten. Vidare reflekterade lärarna över vad användningen av utgångspass medfört avseende betydande utfall (konsekvensdomänen). Lärarna upplevde att de fick bra information om alla elevers kunskaper och de reflekterade och drog nya slutsatser om att använda utgångspass. Deras reflekterande kring detta medierade sedan en förändring av deras förställningar om att använda utgångspass. Mer specifikt förändrades lärarnas föreställningar om nyttan (nyttovärde), då de själva fått uppleva (och inte bara fått höra det berättas) att de fått bra information om alla elevernas kunskaper genom att använda utgångspass. När lärarna själva hade testat och upplevt att aktiviteten enkelt kunde genomföras, förändrades deras föreställningar (t.ex. motivationsföreställningar) om deras förmåga att klara av att genomföra aktiviteten på ett bra sätt (förväntad framgång). Lärarnas förändrade föreställningar om utgångspass medförde att de reflekterade på andra sätt och drog nya slutsatser om användandet av utgångspass vilket därmed bidrog till att de realiserade aktiviteten i praktiken igen. Lärarnas förändrade föreställningar inverkade också på lärarnas reflektion och realisering av nya och andra idéer och aktiviteter, såsom användningen av små skrivtavlor för samma syfte, som presenterades i kompetensutvecklingen. Sättet som lärarna pratade om dessa aktiviteter under kursträffarna indikerade detta. Kompetensutvecklingens upplägg, där lärarna vid diskussionerna under träffarna fick berätta om erfarenheter av experimenterandet i praktiken och reflektera tillsammans med en expert över hur det fungerat då de testat att använda utgångspass i sin

praktik bidrog till att lärarna fick ytterligare stimuli. Det bidrog till att lärarna såg vinsterna med att använda utgångspass för att samla information om alla elevernas kunskaper. De lärare som kanske inledningsvis inte såg nyttan med att använda utgångspass, eller som inte testat direkt som de uppmanats att göra, fick höra berättelser och andra lärares reflektioner. De fick då höra att det var snabbt och enkelt att genomföra och att det gav dem mer information om alla elevers kunnande. De flesta av lärarna trodde på idén med att använda utgångspass direkt efter att de första gången fått stimuli från kompetensutvecklingen och realiserade aktiviteten i praktiken, medan några lärare realiserade utgångspass i praktiken först efter att de fått höra av andra lärare hur bra de ansåg det var. Alla utom två lärare började efter kompetensutvecklingen att använda utgångspass regelbundet i sin klassrumspraktik, vilket tyder på en förändring i praktikdomänen.

Vid träffarna presenterades lärarna för ytterligare ett tillvägagångssätt för att låta alla elever få möjlighet att svara på lärarens frågor, vilket var genom att använda sig utav små skrivtavlor under helklassgenomgångar. Men endast cirka hälften av lärarna använde små skrivtavlor som svarssystem då de ställde frågor till eleverna efter ett år av införlivande av formativ bedömning. Att använda små skrivtavlor för att samla belägg om elevernas kunskaper kan utöver informationsinsamlingen också bidra till ökat elevengagemang och tänkande hos eleverna. Vid experimenterandet med små skrivtavlor upplevde många lärare praktiska hinder såsom exempelvis att eleverna ritade på de små skrivtavlorna, att det blev stökigt i klassrummet, att det tog lång tid att dela ut dem och att de avsedda pennorna för skrivtavlorna hade en tendens att försvinna. Några lärare tyckte därför att det var svårt att få användandet av små skrivtavlor att fungera på ett smidigt sätt i klassrummet. Lärarna var dock i diskussionerna vid träffarna överens om de betydande positiva utfall som användning av små skrivtavlor kunde medföra och medförde (bättre information om alla elevers kunskaper och ökat elev engagemang och tänkande). Vid diskussioner framkom förslag på hur hinder kunde överkommas, såsom att låta eleverna rita fritt på de små skrivtavlorna första gången så att det inte blir lika roligt nästa gång. Diskussionerna bidrog också till lärarnas reflektioner kring de betydande utfall som användandet av små skrivtavlor medför och de slutsatser lärarna drog om dessa. Ett vanligt argument som många lärare gav var att eleverna tyckte att det var roligt att få svara på lärarnas frågor genom att skriva på en liten skrivtavla och att lärarna upplevde större engagemang hos eleverna. Det i sin tur ledde till ytterligare reflektion hos lärarna, och förändrade föreställningar hos de flesta av lärarna om hur bra det faktiskt är att använda små skrivtavlor. Detta medförde att många av lärarna realiserade små skrivtavlor i praktiken igen. Lärarna som använt de små skrivtavlorna, föreslog att dessa skulle användas flera gånger för att komma förbi hindret med att eleverna ritade på dem. De hade upplevt att när eleverna

blev vana vid att använda dem så slutade de även att rita på dem. Att lärarna i diskussioner fick höra om alla fördelar och hur andra lärare kommit på lösningar för hindrande faktorer med användandet av små skrivtavlor bidrog till lärarnas reflektion och därmed förändrade föreställningar om att klara av att använda små skrivtavlor på ett smidigt och effektivt sätt i praktiken. I diskussionerna framkom även argument mot att använda små skrivtavlor. Ett sådant argument var att det kunde vara känsligt för eleverna att visa sina svar för de andra eleverna i klassen. De flesta argumenten som framkom mot små skrivtavlor rörde dock det praktiska i klassrummet, såsom att det blev stökigt och rörigt i klassrummet. Dessa upplevda negativa effekter som lyftes i diskussionerna påverkade också hur lärarnas reflekterade över användandet av små skrivtavlor och de nya slutsatser de drog om att realisera små skrivtavlor i praktiken. För vissa lärare kan det ha bidragit till att de att inte realiserade små skrivtavlor i praktiken igen. Några lärare kom därmed inte över dessa praktiska hinder med att använda små skrivtavlor trots att de fick höra många andra lärares ”framgångssagor” och lösningar på praktiska problem. Möjligheten att diskutera tillsammans bidrog till lärarnas reflektion kring användningen av små skrivtavlor i praktiken, vilket medierade en förändring i föreställningar hos lärarna kring deras förutsättningar att realisera små skrivtavlor i sin praktik. Några av lärarna experimenterade med små skrivtavlor i praktiken först efter att de i diskussioner hört hur andra lärare hanterat hinder och hört deras reflektioner kring betydande utfall då de använt små skrivtavlor. Indikationer i data tyder på att de praktiska hindren kring användandet av små skrivtavlor var en försvårande faktor för de flesta av lärarna som inte lyckades införliva små skrivtavlor regelbundet i praktiken efter kompetensutvecklingen. Ett par av dessa lärare började dock använda dem i andra klasser där de inte upplevde att de mötte samma hinder i klassrummet. Om experimenterandet under kompetensutvecklingen fungerat ännu bättre hade dessa lärare troligtvis varit bättre förberedda för de hinder som de sedan mötte i praktiken under året av införlivande. Som tidigare nämnts, valde de flesta lärarna att regelbundet använda utgångspass i klassrumspraktiken. Lärarna upplevde inte några avgörande praktiska hinder kring att använda utgångspass och de fick samtidigt bra information om alla elevers kunskaper. De ytterligare fördelar som följer av användandet av små skrivtavlor är ökat elevengagemang under helklassdiskussioner och informationsinsamling under lektionen, vilket också möjliggör tätare anpassningar av undervisningen bidrar utgångspass dock inte till på samma sätt.

Kompetensutvecklingens upplägg, där yttre stimuli bland annat utgjordes av föreläsningar med input från en extern expert, men också av diskussioner med andra deltagare i både liten grupp och stor grupp bidrog till den förändring som skedde. En viktig egenskap i kompetensutvecklingen var att lärarna fick gott om tid till att delta i kompetensutvecklingen samt även att den var

utsträckt över en lång tid (en termin). Att kompetensutvecklingen var förlagd över en lång tid bidrog till att lärarna fick möjligheter till *reflektion* och *realisering*. Det vill säga, det var gynnsamt för de medierande processerna mellan modellens domäner. Det faktum att lärarna under kursträffarna fick förslag på aktiviteter att själva testa kan ha bidragit till lärarnas *realisering* av nya idéer och aktiveter i klassrumspraktiken. Dessa processer stärktes ytterligare genom att lärarnas experimenterande i klassrummet diskuterades vid nästkommande kursträff tillsammans med en expert. De egenskaper i kompetensutvecklingen som lärarna ansåg vara betydelsefulla för deras möjligheter att lära och införliva formativ bedömning var bland annat tillgång till expert, aktiviteter att testa i klassrummet, diskussioner vid kursträffar, teori kopplat till praktik samt tid (Artikel II).

Olika aspekter av skolkontexten, som i modellen utgör förändringsmiljön, verkar antingen stödjande eller motverkande för förändring i de olika domänerna och för de medierande processerna. Ett exempel på detta är att många av de lärare som undervisade i årskurs 9, upplevde det som svårt att i realiteten kunna testa de aktiviteter som föreslagits samtidigt som kompetensutvecklingen pågick. Vikten av att testa aktiviteter i praktiken kan förstås genom dess påverkan på de andra domänerna. Uteblivna möjligheter att testa aktiviteter i praktiken innebär även uteblivna möjligheter att reflektera över experimenterandet i praktiken. Detta kan i sin tur påverka på lärares kunskaper och föreställningar och så vidare. Det är därför av stor vikt att ombesörja lärarnas möjligheter att testa aktiviteter i klassrummet. Detta var något som högstadielärarna upplevde att de kunde haft bättre förutsättningar för, om de inte hade undervisat årskurs 9 (vilket de flesta gjorde) då kompetensutvecklingen pågick. Mellanstadielärarna uttryckte inte att de önskat att kursen hade varit förlagd på något annat sätt (Andersson, 2015). Flera av de högstadielärare som införlivade relativt få aktiviteter var också de som upplevde lite stöd från sin rektor och dessutom upplevde särskilda försvårande omständigheter i klassrummet (Artikel II), faktorer som kan kopplas till skolkontexten. Det stöd (till exempel uppmuntran eller visat intresse för förändringar i praktiken) som lärare får från sina rektorer, kan till exempel påverka hur läraren reflekterar över sitt experimenterande i klassrummet och hur viktigt de anser det är att göra förändringar (personligt värde) samt vad läraren värderar som betydande utfall.

Att många av rektorerna inte friställde lärarna en hel dag från undervisning medförde att högstadielärarna enbart fick fyra timmar till de förlagda träffarna på universitetet istället för de sex timmar som mellanstadielärarna fick. De två extra timmarna som mellanstadielärarna fick under träffarna på universitet kan ha påverkat lärarnas möjligheter att lära och införliva formativ bedömning på följande sätt: Mer tid i kursträffar innebär också mer yttre

stimuli både från kurskollegor och genom diskussioner och från den externa experten. Detta i sin tur kan påverka lärarnas realisering av de idéer som presenteras under träffarna. Det påverkar också på deras reflektion över sin egen och andra lärares experimentering i praktiken och betydande utfall. Detta kan leda till en förändring i deras kunskaper och föreställningar, vilket i sin tur påverkar deras reflektion över innehållet i kursträffarna och vilka idéer som de realiserar i klassrummet och så vidare.

Inga indikationer framkom i data på att lärarna hade några betydande kunskaper om formativ bedömning vid studien av lärarnas praktik före kompetensutvecklingen (Artikel I). Efter kompetensutvecklingen trodde alla lärare på idén med formativ bedömning (Artikel II) och de ansåg att de hade den kunskap de behövde för att införliva formativ bedömning i sin praktik (Artikel III). Utifrån detta kan vi dra slutsatsen att det under kompetensutvecklingen skett en förändring i den personliga domänen. Lärarnas *personliga värden* indikerar också att deras syn på vad som utgör kompetenta lärares klassrumspraktik stämmer väl överens med en formativ klassrumspraktik. Lärarna rapporterade också relativt höga värden (på en 5-gradig skala) avseende deras motivation att införliva formativ bedömning efter kompetensutvecklingen (Artikel III). Detta kan tolkas som indikationer på att en förändring i den personliga domänen har skett. Alla lärare som deltog i kompetensutvecklingen gjorde varaktiga förändringar i sin formativa klassrumspraktik (regelbundet använda formativa bedömningsaktiviteter identifierades i Artikel II) vilket visar på en förändring i praktikdomänen. Praktikdomänen utgörs av lärarnas professionella experimenterande under kompetensutvecklingen, men också deras fortsatta experimenterande och införlivande under det kommande läsåret. De medierande processerna är en förutsättning för lärarnas fortsatta förändring, där varje individuell lärares förändring kan starta i någon av de olika domänerna och sedan genom reflektion och realisering påverka till en förändring i en annan domän. Detta kan i sin tur resultera i ett nätverk av förändringssekvenser. Clarke & Hollingsworth (2002) gör en skillnad i när förändringen är mer tillfällig (till exempel en förändring i en domän påverkar på en förändring i en annan domän vilket då utgör en till två pilar, dvs. förmedlande processer i modellen) jämfört med när förändringen är varaktig (fler pilar i modellen) som de kallar *utveckling*. Enligt Clarke & Hollingsworth (2002), krävs empiriska belägg för en varaktig förändring i lärarens praktik, kunskap eller föreställningar för att utveckling ska anses vara identifierad. Studien har med empiriska belägg visat att förändring har skett i de olika domänerna samt att de medierade processerna reflektion och realisering har stött denna förändring. Då kompetensutvecklingens upplägg fungerat stödjande för reflektion och realisering (många pilar i modellen) kan vi förstå att det bidragit till effekterna på lärarnas klassrumspraktik. Det har skett förändringar i alla domäner och

pilarna har gått åt alla håll i modellen, men på olika sätt för olika lärare. Det framkommer också hur viktigt det är att experimentera med aktiviteter i klassrummet för att skapa sig erfarenheter, och att det kan vara en bra start till lärarens förändringar av deras föreställningar om undervisningen. Det blir också synligt att förändringar i praktikdomänen kan skapa goda möjligheter för läraren att ”vandra vidare” i modellen.

Genom att diskutera avhandlingens resultat i relation till modellen synliggörs den komplexitet som lärares förändring utgör, och hur avgörande det är att se till lärarens situation under kompetensutvecklingen och deras arbetssituation på skolan (skolkontexten). Modellen tydliggör att det är mycket som kan ändras och att det är många aspekter som påverkar lärarnas förändring, likväl hur förändring i en domän bidrar till förändring i en annan domän. Det är många egenskaper i kompetensutvecklingen och egenskaper hos skolkontexten som påverkar lärares förändring och hur dessa påverkar varandra är komplext. Genom att använda modellen för att beskriva delar av avhandlingens resultat så har olika egenskapers påverkan på lärarnas förändring kunnat synliggöras. Hur processerna reflektion och realisering i modellen medierar förändring åt många håll och hur man genom att stödja dessa processer i kompetensutvecklingen kan åstadkomma att de sker och medierar förändring i de olika domänerna. Dessutom indikerar resultaten att om man får många processer att ske bidrar det till en mer varaktig förändring, dessutom framträder att vägarna till förändring (lärande) är många. Avhandlingens resultat bidrar därmed genom att synliggöra, på gruppnivå, hur en relativt stor grupp lärare efter en kompetensutvecklingssatsning gjort förändringar i sin formativa klassrumspraktik som blev varaktiga under det efterföljande året. Den påverkan de skillnader som förelåg i kompetensutvecklingssatsningens design, för högstadielärare och för mellanstadielärarna, kan också förstås utifrån modellen. Vad som kan upplevas som små skillnader i förutsättningar, till exempel den absoluta tiden i träffar på universitetet, kan få relativt stort genomslag avseende lärarnas förändring. Ett annat exempel är hur skillnader i skolkontexten, till exempel lärarnas möjligheter att testa aktiviteter i praktiken, kan få stor påverkan på lärarnas förändring.

5.2 Lärarna genomförde förändringar, men varför gjorde de inte än mer?

Alla de slumpvis utvalda lärarna gjorde förändringar i sin klassrumspraktik till att bli mer formativ, men på olika sätt och i varierad grad. För att få bättre förståelse för det stöd lärare behöver vid införlivande av formativ bedömning söker jag därmed att förstå möjliga orsaker till att lärarna införlivade formativ bedömning på det sätt som de gjorde, och orsaker till att de inte införlivade än mer formativ bedömning i sin klassrumspraktik. Kompetensutvecklingssatsningar riktar sig mot verksamma lärare som redan har utvecklat ett sätt

att hantera problem och balansera de många konkurrerande utmaningar och ideal som de möter i sin strävan att hjälpa elever att lära, och enligt Kennedy (2016) handlar det då inte om att bara erbjuda nya idéer om undervisning utan snarare om att erbjuda dem andra idéer än de som har guidat dem tidigare. En svårighet som lärarna mötte när de var tillbaka på skolan och ”utsattes” för en hög arbetsbelastning var att det var lätt att falla tillbaka i gamla vanor (Artikel II). Att för lärarna ägna tid och ansträngning till att ersätta och komplettera befintliga idéer och sätt att bedriva undervisning på som kan uppfattas som redan tillfredsställande med något annat kräver att de har motivation för att göra det, då motivation är den drivande kraften för människor att starta upp en aktivitet och också stanna kvar i aktiviteten (Ryan & Deci, 2000). Dessutom har tidigare studier visat att användande av formativ bedömning i klassrumspraktiken ställer höga krav på lärarens kunskaper och färdigheter (Vingsle, 2014), och läraren måste förändra sina invanda *'habits of mind'* (Schoenfeld, 2014), vilket visar på vikten av att lärarna verkligen är motiverade till att göra en förändring för att införlivandet av formativ bedömning ska bli framgångsrikt.

Kompetensutvecklingssatsningarna fick olika utfall för högstadielärarna och mellanstadielärarna trots att deltog i liknande kompetensutvecklingssatsningar där de allmänna egenskaperna, såsom tillgång till expert och mycket tid (både avseende direkt tid och utsträckt över lång tid) var lika. Utfallet skiljde sig både avseende lärarnas förändrade praktik och elevernas prestationer i matematik. Att de allmänna egenskaperna i kompetensutvecklingssatsningar inte räcker för att förutsäga utfallet konstateras av (Opfer & Pedder, 2011), något som även denna studie visar. Därför behövs studier som fokuserar på orsaksförklaringar, och inte enbart orsak – effekt (Opfer & Pedder, 2011). Med andra ord efterfrågar de till exempel studier som inte bara visar att mycket tid i kompetensutvecklingen är bra utan även visar varför mycket tid är bra. Enligt Kennedy (2016) behövs också studier som knyter dessa förklaringar till modeller för motivation och lärares lärande. Tid är till exempel identifierat som en viktig egenskap i framgångsrika kompetensutvecklingar, då mycket tid visat sig ge positiv effekt på utfallet av kompetensutvecklingar (t.ex. Tierney, 2006). Det är däremot mindre känt varför och hur tid påverkar på utfallet (dvs. orsaksförklaringar).

En jämförelse med mellanstadielärarnas förutsättningar avseende vissa aspekter i kompetensutvecklingen och i skolkontexten som också relateras till deras motivation kan därmed bidra med bättre förståelse kring högstadielärarnas införlivande av formativ bedömning. Några skillnader har redan lyfts i avsnitt 5.1, såsom skillnaden i tid i kursträffarna på universitet och möjligheter att testa aktiviteter i klassrummet och hur detta kan ha bidragit med bättre förutsättningar för mellanstadielärarna. Ett önskemål hos många av

högstadielärarna var en mer sammanhängande tid för att jobba med kompetensutvecklingen gällande reflektion och planering av införlivande av formativ bedömning. Denna önskan bottnade sig i att den frilagda arbetstiden utöver kursträffarna (5h) som de fick, ofta var utspridd över flera dagar. Detta skiljde sig från mellanstadielärarna, som var friställda en hel dag. Högstadielärarnas arbetssituation på skolan påverkade därigenom deras möjligheter att ägna dessa fem timmar åt kursen. Andra skillnader som framkom var att högstadielärarna upplevde en mer pressad arbetssituation i och med den samtidiga implementeringen av en ny kursplan (Lgr 11), något som inte framkom som ett upplevt hinder för mellanstadielärarna. Det verkar då i kompetensutvecklingen vara viktigt att ombesörja att den tid lärarna får också schemaläggs rent fysiskt, så att de att de ges möjligheter att samarbeta och diskutera med kurskollegor.

Slutsatser som kan dras från ovanstående diskussion samt i avsnitt 5.1, är den betydande vikten av att se till helheten - att förstå lärares förändring innebär förståelse för att det är ett komplext system, något som även betonas av bland annat Kennedy (2016) och Opfer & Pedder (2011). Att för lärare förändra sin klassrumspraktik är svårt, särskilt när det gäller att införliva något så komplext som formativ bedömning. För att få effekt krävs det att alla steg i kompetensutvecklingssatsningen ombesörjs, annars är risken att det inte blir någon större effekt. Resultaten indikerar att små skillnader i några av de viktiga egenskaperna i kompetensutvecklingssatsningen och i skolkontexten kan ha inverkat lärarnas (upplevda) möjligheter att införliva formativ bedömning i sin klassrumspraktik. Det blir därför viktigt att ta ett större grepp och ombesörja stöd på alla nivåer i skolsystemet. Dessutom blir det särskilt viktigt hur det stöd lärarna får fördelas. Det räcker till exempel inte med att enbart ge lärarna tid, i form av nedsättning i arbetstid. Gällande lärarnas arbetssituation, upplever de flesta sig ha en hög arbetsbelastning. I detta fallet när en ny kursplan samtidigt skulle implementeras blir det särskilt viktigt att ombesörja hur tiden faktiskt används. Det blir då viktigt att den tilldelade tiden används till aktiviteter som har visat sig vara gynnsamma för lärares lärande och att dessa aktiviteter ”schemaläggs” för att de ska bli av. Hur den absoluta tiden påverkar på lärarnas förändring måste också ställas i relation de andra aspekterna i det komplexa system som lärares lärande utgör. Framtida forskning om hur olika variabler påverkar lärares motivation och hur lärare lär kan ytterligare bidra till förståelse för hur vi kan stödja lärare att anta och genomföra den stora utmaning som införlivandet av en formativ bedömnings praktik innebär, likväl som när andra komplexa kompetensutvecklingsinsatser ska genomföras.

5.3 Den uteblivna effekten på elevernas prestationer i matematik

Formativ bedömning kan genomföras på många olika sätt och följaktligen kan effekten på elevers prestationer variera kraftigt från en specifik införlivad formativ klassrumspraktik till en annan. Tidigare forskning har också visat att detta är fallet, se avsnitt 2.2. Resultaten i denna avhandling indikerar att den typ av formativ bedömning som dessa lärare införlivade inte gav något resultat på elevernas prestationer i matematik. De krävs med andra ord större förändringar i praktiken för att det ska ge effekt på elevernas prestationer vilket jämförelsen med mellanstadielärarna visar på, som gjorde större förändringar. Dock indikerar resultaten i Artikel IV att antalet införlivade formativa bedömningsaktiviteter inte är avgörande utan att det snarare är egenskaperna hos aktiviteterna och hur aktiviteterna genomförs som är av vikt. En möjlig orsak till den uteblivna effekten skulle kunna härledas till den införlivade praktikens kvalitet, samt att vissa av de införlivade aktiviteterna inte är så effektiva för att uppnå syftet med nyckelstrategierna och den stora idén. En aspekt av kvalitet som vi och andra (t.ex. Tillema, Leenknecht, & Segers, 2011; Wylie & Lyon, 2015) tror kan vara avgörande är genomförandet av kärnan i formativ bedömning, den stora idén. Lärarna gjorde förändringar avseende den stora idén, att samla information om elevernas lärande som sedan används för att justera undervisningen så att den bättre möter elevernas lärandebehov. De gjorde stora förändringar avseende oftare, och strukturerad, insamling av information om elevernas prestationer vilket följdes av en justering av undervisningen. Därmed ökade frekvensen av de formativa justeringscyklerna (Artikel II). Men även om de samlar in mer information om alla studenters lärande och justerar undervisningen oftare, så råder en viss tveksamhet gällande hur väl lärarnas anpassningar faktiskt uppfyller alla elevernas individuella lärandebehov. I förlängningen är frågan vilken effekt det skulle ha fått på elevernas lärande. En kvalitetsaspekt kopplat till detta berör de frågor som lärarna ställer. Indikationer i resultaten finns på att många av lärarna oftast vid dessa nya typer av informationsinsamlingar ställer frågor som testar elevernas ”baskunskaper”, vilket då kan få följder avseende vilka justeringar de kan göra. Om lärarna i huvudsak testar elevernas ”baskunskaper”, kommer justeringarna endast att möta lärandebehoven hos de elever som ej förstått grunderna. Det finns då en potentiell risk att lärarna mestadels bara möter några elevers individuella lärandebehov, och att de fortfarande inte justerar utifrån många elevers lärandebehov i klassen. Detta kan vara en bidragande orsak till att det inte blev några större förändringar avseende klassens resultat. Fler kvalitetsaspekter diskuteras i Artikel IV.

5.4 Lärdomar om kompetensutvecklingssatsningar (i formativ bedömning)

I detta avsnitt förs en diskussion kring avhandlingens resultat och några lärdomar som kan dras om kompetensutvecklingsinsatser i formativ bedömning. Ett formativt angreppssätt kan vara användbart för andra utvecklingsområden, och inte bara avseende att elevers lärande ska stödjas på ett effektivt sätt. Därför antar jag i detta avsnitt ett formativt angreppssätt för att visa på de lärdomar om kompetensutvecklingssatsningar i formativ bedömning som denna avhandling kan bidra med. Avsnittet är upplagt, och struktureras, utifrån tre frågor som guidar de viktiga processerna i formativ bedömning, *vart ska vi (målen), var är vi (nuläget)* och *hur ska vi ta oss dit* (se avsnitt 2.2.2)?

Vart ska vi? Formativ bedömning har visat sig vara ett av kraftfullt sätt att förbättra elevers lärande. Därför är ett mål att fler lärare får kunskaper om formativ bedömning och införlivar mer formativ bedömning i sin klassrumspraktik på ett sådant sätt att det ger positiva utfall för elevernas resultat. Tidigare forskning har indikerat vikten av att lärare förstår idén med formativ bedömning och lär på djupet (t. ex. Flórez & Sammons, 2013), vilket är viktigt för att förändringar i deras formativa praktik ska bli bestående. Utifrån det Formativa projektet kan vi dra slutsatsen att då formativ bedömning konceptualiseras som en integrerad helhet av strategier (Wiliam & Thompson, 2008) finns stor potential för förbättrade prestationer hos eleverna (Andersson & Palm, 2017a). Ett mål kan då vara att fokusera på formativ bedömning som en integrerad helhet av strategier, då en sådan syn på formativ bedömning bör ha potential att ge än större lärandevinster än om enskilda strategier fokuseras (Hawe & Parr, 2014). De deltagande högstadielärarna trodde också efter kompetensutvecklingssatsningen på idén med formativ bedömning (Artikel II). De ansåg också det viktigt att göra förändringar i sin praktik (Artikel III), något som stödjer att innehållet i kompetensutvecklingssatsningen var något som lärarna trodde att de skulle ha nytta av i sin undervisning (Artikel III). Sist men inte minst, målet är att stödja lärarna till att införliva en formativ klassrumspraktik med hög kvalitet, och några kvalitetsaspekter har lyfts i Artikel IV.

Var är vi? Resultaten i Artikel I kan ge en indikation om var vi är idag och vilka utvecklingsmöjligheter som finns då lärarnas formativa klassrumspraktik karakteriserades och utvecklingsmöjligheter identifierades. Genom att både beskriva vad lärarna gör (egentligen vad de gjorde då, innan kompetensutvecklingen, vilket kan ha ändrats lite under de år som gått) och inte gör, ger det en indikation om var vi är idag. De deltagande lärarna använde till viss del formativ bedömning i sin klassrumspraktik, men det fanns också utvecklingspotential (Artikel I). Det fanns utvecklingspotential

inom alla de ingående strategierna och den *stora idén* (Artikel I). Eftersom lärarna var slumpmässigt valda kan resultaten indikera att denna utvecklingspotential gällde för övriga högstadielärare som undervisade i årskurs 7 i kommunen.

Hur ska vi ta oss dit? Resultaten i Artikel II, III och IV kan bidra med kunskaper om hur vi kan ta oss närmare målet, då de undersöker effekter och orsaker till effekter av en kompetensutvecklingssatsning i formativ bedömning. Den kunskapen kan vara viktig för att lära om hur vi kan ta oss närmare målet och detta avsnitt baseras på några förändringar som skulle kunna bidra med ett annat (bättre) utfall av kompetensutvecklingssatsningen för högstadielärarna. Vikten av att se helheten har poängterats tidigare, till exempel stöd från rektor, att ombesörja möjligheter att testa aktiviteter i klassrummet och så vidare. Även fördelningen av tid, då många lärare efterfrågade sammanhängande tid och även en omfördelning av tid så att lärarna får mer tid under själva införlivandet, så att inte all tid fördelas under själva kompetensutvecklingen. Kompetensutvecklingssatsningar bör i ännu högre grad anpassas efter den lärargrupp man möter, så att kompetensutvecklingen inkluderar en bättre förberedelse för de eventuella hinder lärarna kommer att möta ute på skolorna och i klassrummen vid införlivningen av formativ bedömning. Ett sådant exempel är att högstadielärarna upplevde sin arbetssituation på skolan mer hindrande än vad mellanstadielärarna gjorde (Artikel II), vilket kan betyda att de behöver ytterligare eller en annan typ av stöd. De efterfrågade även sammanhållen tid, vilket kan härledas till att de inte fått en undervisningsfri dag som frigjord tid för deltagande i kompetensutvecklingen. En ytterligare bidragande orsak till att de upplevde den tid de hade till förfogande som splittrad kan vara att högstadielärarnas arbetssituation på skolan ofta är mer splittrad, då de möter fler elever och byter salar oftare än vad mellanstadielärare vanligtvis gör. Byte av sal var något som vissa lärare upplevde som förhindrande gällande införlivandet av vissa formativa bedömningsaktiviteter, då det innebar att de behövde bära med sig material. Ett annat exempel gäller hur de ska hantera stökiga klasser och klasser med stor spridning i kunskaper, något som framkom som ett hinder för många lärare, men särskilt för några av de högstadielärare som införlivade få formativa bedömningsaktiviteter. Genom att ta upp och diskutera sedan tidigare kända möjliga hinder under kompetensutvecklingen kan läraren då förekomma vissa problem och situationer som är vanliga att de uppstår. Lärarnas möjligheter att tillsammans reflektera och diskutera kan bidra till en förändring i deras kunskaper och föreställningar (t.ex. deras tro på att de klarar av att införliva formativ bedömning, dvs. deras *förväntningar på framgång).* När kompetensutvecklingen fokuserar på formativ bedömning som en helhet av integrerade strategier, indikerar resultaten att vissa strategier kan behövas

fokuseras mer. Något som var tydligt i denna studie var att få lärare arbetade på ett aktivt sätt med Nyckelstrategi 4 (att stödja eleverna till att fungera som läranderesurser för varandra) och något mer med Nyckelstrategi 5 (att stödja eleverna att bli mer självreglerande av sitt lärande) vilket innebär att det fortfarande är läraren som tar det största ansvaret för den formativa klassrumspraktiken. Detta kan förstås mot tidigare studier av andra relativt storskaliga insatser, dock med mindre stöd till lärarna, där konceptualiseringen av formativ bedömning skett på ett liknande sätt som i denna studie. De studierna har också visat att förändringar kopplade till Nyckelstrategi 4 (t.ex. Jönsson m. fl., 2015) och 5 (t. ex. Wylie & Lyon 2015) inte fått stor genomslagskraft. Större vikt bör därför läggas på att stödja lärarna i hur de kan utveckla sin praktik i relation till dessa nyckelstrategier i framtida kompetensutvecklingsinsatser. Många forskare (t.ex. Swaffield, 2011; Klenowski, 2009) har också uppmärksammat detta och konstaterar att man tappar en stor del av den potential som formativ bedömning har om man inte involverar alla aktörer i klassrummet. Att stödja eleverna i att bli självständiga, ansvarstagande och effektiva lärande är en viktig pusselbit för att formativ bedömning ska nå den potential som det har. En annan viktig pusselbit som är viktig att fokusera i kompetensutvecklingsinsatser är att stödja hur lärarens justering av undervisningen kan genomföras på ett effektivare sätt, något som tidigare forskning visat vara svårt (Heritage m. fl., 2009; Schneider & Gowan, 2013). Resultaten i Artikel III indikerar också att lärarna i hög grad införlivade aktiviteter som innebär stor nytta och liten kostnad att genomföra. Aktiviteter som innebär både hög nytta (eller fördröjd nytta) verkar då väljas bort av lärarna. Det framstår då som viktigt att i kompetensutvecklingssatsningar stödja lärare att testa och införliva aktiviteter, trots att de innebär hög kostnad (och hög nytta). Ofta är kostnaden högre i det initiala skedet, och läraren kan därmed få stöd till att ta det steget och till att få uppleva nyttan med aktiviteten, vilket kan bidra till en fortsatt utveckling.

5.5 Reflektioner kring avhandlingens kvalitet

Vid bedömning av kvaliteten hos kvantitativ forskning används ofta begrepp som validitet och reliabilitet. Vid kvalitativa ansatser menar till exempel Bryman (2011) att det finns andra begrepp som bättre beskriver relevanta kvalitetsaspekter. En studies tillförlitlighet stärks bland annat genom tydliga och transparenta beskrivningar av hur man gjort (Guba och Lincoln, 1994). Jag har valt att använda Schoenfelds (2007) tre kriterier för kvalitet i forskning för att diskutera kvaliteten i min avhandling. Schoenfelds kriterier passar både kvalitativa och kvantitativa ansatser vilket därmed är lämpligt för denna avhandling. De kriterier han föreslår är *Generaliserbarhet* - Vilka situationer och kontexter är forskningen tillämplig på? *Trovärdighet* - Varför ska man tro på det författaren säger? *Viktighet* - Varför ska man bry sig? Det

sistnämnda av Schoenfelds (2007) kriterier, *viktighet* knyter jag även an till i avsnitt 5.6.

Generaliserbarhet handlar om vilka situationer och kontexter forskningen är tillämplig på. I min studie dras slutsatser om lärarnas formativa klassrumspraktik före kompetensutvecklingssatsningen, deras förändrade formativa klassrumspraktik efter kompetensutvecklingssatsningen och orsaker till denna förändring. Då lärarna var slumpmässigt utvalda, kan dessa slutsatser generaliseras och antas gälla för gruppen högstadielärare i matematik i kommunen. Det kan också finnas andra populationer i liknande kontexter som resultaten skulle kunna gälla för. En fördel med ett slumpmässigt urval av de deltagande lärarna var att vi kunde anta att lärarnas ingående egenskaper, såsom intresse för formativ bedömning och deras lärarerfarenhet, inte skiljer sig nämnvärt mellan interventionsgruppen (de lärare som deltog i kompetensutvecklingssatsningen) och kontrollgruppen. De deltagande lärarna kan också anses vara ”vanliga” lärare och kompetensutvecklingssatsningens effekter på dessa vanliga lärares formativa klassrumspraktik, är det som undersöks. Detta kan bidra med kunskap om hur kompetensutvecklingssatsningar kan fungera för stora grupper av lärare, där till exempel inte enbart motiverade entusiaster som anmäler sig frivilligt är de som data baseras på. Storleken på urvalet påverkar de slutsatser man kan dra. En önskan är därför ofta att ha ett stort urval med litet bortfall. Bortfallet avseende denna studie för högstadielärarna kan dock anses vara litet inledningsvis, då det endast var två lärare som tackade nej till att delta i kompetensutvecklingen. Under studiens gång föll dock ytterligare fyra lärare bort. Jag har dock inga indikationer på att dessa lärare uppvisade några särskilda egenskaper som skulle skilja dem från de andra deltagarna, utifrån de skäl de angav för att inte delta. Bortfallet för mellanstadielärarna var lite större, något som kan ha inverkat på resultatet, men vilket det inte heller finns några indikationer på (Andersson & Palm, 2017a). Avseende resultaten på elevnivå så innebär ett större urval, det vill säga större grupper av lärare, att sannolikheten ökar för att små skillnader i resultat mellan grupperna blir signifikanta. Vid en jämförelse utgör mellanstadielärarna ett större antal, vilket innebär att vi skulle kunna ställa oss frågan: Om vi hade haft fler högstadielärare i urvalet av högstadielärare, hade skillnaden mellan interventionsgruppen och kontrollgruppen då varit signifikant? Mest troligt hade det inte gjort någon skillnad eftersom skillnaden mellan grupperna (interventionsgruppen och kontrollgruppen) var så pass liten.

Studiens *trovärdighet* handlar om varför man ska tro på resultaten och slutsatserna i studien. En aspekt som rör de slutsatser vi kan dra berör noggrannheten i mätningen av elevernas resultat. Testen var noggrant utformade. De testade enbart innehållet som fanns i kursplaner och dessutom ingick i de

matematikböcker som användes i kommunen. Dessutom utprövades testen, och utöver provutvecklargruppen (bestående av bland annat nationella prov utvecklare) var det referensgrupper med erfarna lärare som bedömde att testens innehåll var lämpligt för årskursen. Detta indikerar att testen var utformade på ett sådant sätt att de väl mätte det som de var avsedda för att mäta. Noggranna och utförliga bedömningsanvisningar, att en grupp av erfarna lärare rättade proven (och dessutom satt tillsammans och kunde diskutera) och höga värden på Cronbach's alfa bidrar till mätningens trovärdighet. Med testen följde dessutom noggranna instruktioner till lärarna om hur de skulle genomföras och hur de skulle distribueras tillbaka till kommunen (till exempel fick de adresserade kuvert att skicka tillbaka proven i så att det var enkelt och så att testen inte behövde "passera" deras rektorer). Detta för att lärarna skulle ombesörja att genomförandet av proven skedde på samma sätt i alla klassrum och att de till exempel skulle veta hur de skulle svara på frågor från eleverna så att liknande stöd gavs till alla elever. En annan aspekt kopplat till mätningen av elevernas resultat gäller om det kan ha förekommit "överspillningseffekter" mellan lärarna i interventionsgruppen och kontrollgruppen. Inga indikationer fanns på att det skulle ha skett. Lärarna som deltog i kompetensutvecklingen uppmanades att inte diskutera formativ bedömning och kompetensutvecklingens innehåll med kollegor under det år då kompetensutvecklingen och studien pågick (under året av införlivande). En diskussion kring detta fördes också med lärarna i slutet av året av införlivande av formativ bedömning och då framkom också att de inte kände till någon annan kompetensutvecklingsinsats kring formativ bedömning och formativ bedömning var inte heller lika aktuellt i Sverige då som det är idag. De slutsatser som kan dras från detta är att trovärdigheten är relativt hög avseende mätningen av elevernas resultat. Metod och analys avseende lärarnas formativa klassrumspraktik, deras motivation och egenskaper i kompetensutvecklingen och skolkkontexten som verkat stödjande och hindrande genomfördes också på ett omsorgsfullt och noggrant sätt. Datainsamlingen omfattade flera olika typer av data, som användes för triangulering, vilket ger stöd åt de slutsatser som har dragits. I enkäten användes till exempel både slutna och öppna frågor, vilket också kan bidra till ökad trovärdighet. Dessutom genomfördes i Artikel I en deltagarvalidering som indikerade att tolkningen som gjorts gav en god bild av lärarnas praktik. Vid beskrivning av datainsamling och analys har transparens eftersträvats.

Studiens *viktighet* handlar om varför man ska bry sig om de resultat som framkommit. Studien undersöker flera områden av forskning som har beskrivits som viktiga. Ett viktigt forskningsområde är enligt Schneider och Randel (2010) att utröna vilken typ av formativ bedömning som ger effekter på elevernas prestationer, och fler studier behövs för att avgöra vilka faktorer som inverkar på effektiviteten av formativ bedömning (Bennett, 2011;

Kingston & Nash, 2011). Ett annat forskningsområde rör kompetensutveckling inom formativ bedömning, där fler studier som undersöker effekter av kompetensutvecklingsinsatser i formativ bedömning på både lärarnas praktik och på elevernas prestationer efterfrågas (Schneider & Randel, 2010). Sådana studier är viktiga för att avgöra om kompetensutvecklingsinsatser i formativ bedömning är effektiva avseende lärarnas förändrade praktik och elevernas ökade prestationer (Schneider & Randel, 2010). Schneider och Randel (2010) framhåller även att sådan forskning är värdefull för att bidra med förståelse om för vem kompetensutvecklingsinsatser i formativ bedömning är effektiva och under vilka förutsättningar (Schneider & Randel, 2010 s. 268). Många forskare har också konstaterat att man vet lite om hur man på bästa sätt kan stödja lärare till att införliva formativ bedömning på ett tillfredställande sätt (t.ex. Wiliam, 2010; Schneider & Randel, 2010). Den här avhandlingen bidrar genom att ge ett konkret exempel på en typ av formativ klassrumspraktik och dess effekt på elevernas prestationer. Dessutom bidrar avhandlingen till diskussionen kring viktiga faktorer för att stödja lärares införlivande av formativ bedömning likväl som orsaker som verkat hindrande för lärares införlivning. En jämförelse mellan utfallen för de två lärargrupperna, mellanstadie- och högstadielärare, i det Formativa projektet har också bidragit till en mer nyanserad bild av hur egenskaper i kompetensutvecklingssatsningen och kontexten inverkade lärarnas lärande och införlivande av formativ bedömning i klassrumspraktiken. Mer om avhandlingens ”viktighet” beskrivs i nästa avsnitt.

5.6 Avhandlingens bidrag till forskningsfältet och utbildningsväsendet

Att utforma kompetensutvecklingssatsningar som erbjuder sådant stöd till lärarna att alla, och inte bara entusiasterna, utvecklar sin formativa klassrumspraktik är något som är av intresse för forskningsfältet såväl som för utbildningsväsendet. Särskilt då många kompetensutvecklingsinsatser i formativ bedömning inte resulterat i det utfall som man hoppats på (Birenbaum, m. fl. 2015). Avhandlingens resultat har bidragit med kunskap om det stöd lärarna behöver och kring viktiga egenskaper kompetensutvecklingssatsningar för lärarnas införlivande av formativ bedömning i klassrumspraktiken. Avhandlingens resultat har också bidragit med kunskap kring egenskaper i den formativa klassrumspraktik som införlivas under specifika förutsättningar. Forskningsresultaten från det Formativa projektet visar att det är möjligt att genom en kompetensutvecklingssatsning få ett slumpvis urval av lärare att förändra sin formativa klassrumspraktik på ett sådant sätt att det ger effekt på elevernas prestationer i matematik (Andersson & Palm, 2017a). Denna avhandling visar att en formativ bedömningspraktik med de egenskaper som vi karakteriserat för högstadielärarna i Artikel II inte gav effekt på elevernas prestationer i matematik efter ett år av införlivande.

Resultaten har därmed även bidragit med indikationer om viktiga kvalitetsaspekter i den införlivade formativa bedömningspraktiken. Resultaten ger också indikationer på hur den givna kompetensutvecklingssatsningen skulle kunna utvecklas ytterligare för att ge önskvärda resultat (såsom bättre resultat för eleverna) även i andra kontexter (t.ex. för högstadielärare). I kompetensutvecklingen skulle det till exempel vara bra att fokusera mer på att stödja lärarna avseende hur de kan engagera eleverna i den formativa bedömningen. Ett annat exempel gäller stöd till lärarna att tolka information om elevernas kunnande och efterföljande justering av klassrumspraktiken. Ytterligare ett exempel är att ombesörja att den tid lärarna får till förfogande för kompetensutvecklingen verkligen används till aktiviteter som är gynnsamma för lärarnas lärande, såsom att testa aktiviteter i klassrummet och att diskutera och reflektera över utfallet tillsamman med en expert. Avhandlingen bidrar därmed med erfarenheter och viktiga kunskaper om designegenskaper hos kompetensutvecklingssatsningar som kan vara värdefulla för att få liknande kompetensutvecklingssatsningar i formativ bedömning att fungera (än) bättre för olika lärare i olika kontexter. Flera av de egenskaper som diskuterats är också av generell karaktär, såsom hur olika egenskaper i kompetensutvecklingssatsningen och i skolkontexten kan påverka lärarnas motivationsföreställningar och därmed lärarnas motivation att lära och förändra sin praktik. Insikterna från denna avhandling avseende kompetensutvecklingssatsningar i formativ bedömning kan därmed även bidra med insikter gällande design av kompetensutvecklingssatsningar inom andra områden.

5.7 En enkel match eller en svår utmaning?

Denna avhandling visar på att skolväsendet har en svår utmaning framför sig, något som redan Black och Wiliam konstaterade 1998 vilket följande citat antyder:

> *"Thus, the improvement of formative assessment cannot be a simple matter. There is no 'quick fix' that can be added to existing practice with promise of rapid reward. On the contrary, if the substantial rewards of which the evidence holds out promise are to be secured, this will only come about if each teacher finds his or her own ways of incorporating the lessons and ideas that are set out above into his or her own patterns of classroom work. This can only happen relatively slowly, and through sustained programmes of professional development and support. This does not weaken the message here – indeed, it should be a sign of its authenticity, for lasting and fundamental improvements in teaching and learning can only happen in this way." (Black & Wiliam, 1998b, s. 15)*

Att designa kompetensutvecklingssatsningar som stödjer olika lärare i olika kontexter till att utveckla sin formativa bedömningspraktik är en svår utmaning. Vi har tagit ett steg i rätt riktning, då vi genom det Formativa projektet visat att det är möjligt att få ett slumpvist urval av lärare (och inte bara entusiasterna) att förändra sin formativa klassrumspraktik på ett sådant sätt att det ger effekt på elevernas resultat, men vi blev också varse om att det

kan stanna vid en förändring på lärarnivå. Vi har också fått indikationer på det kan behövas kompetensutvecklingssatsningar liknande denna (med bland annat gott om tid och mycket stöd) för att den stora massan av lärare ska få det stöd de behöver för att utveckla en mer framgångsrik formativ klassrumspraktik. Ett sådant stöd bör inkludera bland annat kompetenta utbildare (så att lärarna får stöd i att lära vad formativ bedömning innebär på djupet), möjlighet att testa aktiviteter i klassrummet och möjlighet till att diskussioner och stöd från kolleger, stöd från rektorer samt tid (och då särskilt träff/mötestid frigjord till detta utvecklingsarbete). Tyvärr sker liknande omfattande kompetensutvecklingssatsningar sällan då de är resurskrävande. Hinder för införlivande som lärarna mött har belysts vilket tillsammans med andra viktiga designegenskaper i kompetensutvecklingssatsningen kan bidra till beslut om hur befintliga resurser bör prioriteras (till exempel genom att fokusera lärares kompetensutvecklingstid på få saker över lite längre tid.) Schneider och Andrade (2013) beskriver väldigt målande vad som bör strävas mot:

"In order to achieve the desired increases in student achievement to which we all aspire, the education community must be just as nimble in supporting teacher learning as we expect teachers to be in supporting student learning and administrators to be in supporting teacher learning." (Schneider & Andrade, 2013, s. 162)

Trots det faktum att formativ bedömning har visat sig ha potential att ge effekt på elevers prestationer, är det ingen enkel match att erbjuda ett sådant omfattande stöd, att lärare förändrar sin formativa klassrumspraktik på ett sådant sätt att det förbättrar elevers prestationer. Denna avhandling kan bidra till ökad förståelse om viktiga designegenskaper att beakta i framtida kompetensutvecklingssatsningar i formativ bedömning som siktar mot framgångsrika resultat både avseende lärarnas lärande och elevernas lärande. Avhandlingen bidrar med viktiga lärdomar som kan behövas för att anta den svåra utmaningen i att utforma kompetensutvecklingssatsningar i formativ bedömning som ger resultat både avseende lärarnas lärande, som kommer till uttryck i deras klassrumspraktik, och sist men inte minst som dessutom bidrar till positiva utfall för deras elever.

6. Korta sammanfattningar av artiklarna

I detta kapitel presenteras korta sammanfattningar av de ingående artiklarna i avhandlingen (Artikel I – Artikel IV). Artiklarna finns att läsa i sin helhet i slutet av avhandlingen.

6.1 Artikel I

Forskning har visat att elevers skolprestationer kan förbättras avsevärt genom användning av formativ bedömning i klassrummet. Dock vet vi relativt litet om svenska matematiklärares användning av formativ bedömning, och därmed också relativt litet om värdet av kompetensutvecklingsinsatser i formativ bedömning. Syftet med studien var att karaktärisera lärarnas formativa klassrumspraktik och besvara följande forskningsfråga: Använder matematiklärare i en kommun formativ bedömning och om de gör det, hur använder de formativ bedömning? Trettioåtta slumpvis utvalda mellanstadie- och högstadielärare i en medelstor kommun i Sverige deltog i studien. Lärarintervjuer och klassrumsobservationer användes för att undersöka lärarnas formativa klassrumspraktik under matematiklektioner. Ett ramverk för formativ bedömning som innefattar en stor idé och fem nyckelstrategier (Wiliam & Thompson, 2008) strukturerade analysen.

Resultaten visar att lärarna regelbundet använder formativa bedömningsaktiviteter i sin klassrumspraktik, men också att det finns stort utrymme för en mer utvecklad formativ klassrumspraktik. På gruppnivå använde lärarna formativa bedömningsaktiviteter inom alla fem nyckelstrategierna och använde olika metoder för att justera undervisningen baserat på information om elevernas lärande. På en generell nivå använde de två grupperna av lärare (mellanstadie- och högstadielärare) liknande formativa bedömningsaktiviteter, men på en mer detaljerad nivå identifierades vissa skillnader. En skillnad var till exempel att mellanstadielärarna rättade elevernas räknehäften för att få information om deras elevers lärande medan prov användes i högre grad av högstadielärarna. Ett exempel på potentialen för vidareutveckling av den formativa klassrumspraktiken grundar sig på att även om lärarna samlar information om elevernas kunnande på flera sätt i det dagliga arbetet, samlar de oftast bara belägg från några få individuella elever och inte från hela klassen. Många anpassningar av undervisningen till klassens lärandebehov baseras då på information från några få elever. Det var även vanligt att anpassningen av undervisningen genomförs med endast en elev åt gången.

6.2 Artikel II

Forskningslitteraturen innehåller många exempel på försök att öka användningen av formativ bedömning i undervisningen som inte visat sig vara särskilt framgångsrika i att åstadkomma en utvecklad formativ klassrumspraktik. Studien handlar om ett slumpmässigt urval av högstadielärare i matematik som har genomgått en omfattande kompetensutvecklingssatsning i formativ bedömning. De deltagande lärarnas formativa klassrumspraktik studerades före kompetensutvecklingen (Artikel I). Kompetensutvecklingen baserades på en konceptualisering av formativ bedömning i enlighet med Wiliam och Thompsons (2008) ramverk, som en stor idé och fem nyckelstrategier. Syftet med studien var att karakterisera de förändringar lärarna gjort mot en mer formativ klassrumspraktik efter kompetensutvecklingen samt att söka orsaker till deras förändringar. Ett annat syfte var att identifiera skillnader avseende högstadielärarnas och mellanstadielärarnas (som deltagit i en liknande kompetensutvecklingssatsning) införlivande av formativ bedömning samt skillnader i orsaker till dessa förändringar. Studien skulle besvara följande forskningsfrågor:

- På vilka sätt förändras högstadielärarnas formativa klassrumspraktik efter kompetensutvecklingssatsningen? Hur är dessa förändringar i jämförelse med de som mellanstadielärarna gjorde?
- Vilka förutsättningar för införlivande av formativ bedömning i sin praktik upplevde högstadielärarna som viktiga, och vilka hinder för förändring upplevde dom? Hur är dessa förutsättningar och hinder i jämförelse med de som mellanstadielärarna upplevde?

Klassrumsobservationer och lärarintervjuer användes för att undersöka lärarnas förändringar mot en mer formativ klassrumspraktik efter kompetensutvecklingssatsningen. Ramverket av Wiliam och Thompson (2008) strukturerade analysen. För att undersöka orsaker till lärarnas förändringar av klassrumspraktiken användes lärarintervjuer och öppna frågor i enkäter.

Resultaten visar att alla lärare förändrade sin klassrumspraktik till att bli mer formativ. De införlivade nya formativa bedömningsaktiviteter som de använde regelbundet i sin klassrumspraktik, dock i varierad grad och på olika sätt. Lärarna började använda nya formativa bedömningsaktiviteter som stärkte en klassrumspraktik i linje med den stora idén av att framkalla belägg om elevers lärande för att justera undervisningen så att den bättre möter elevernas lärandebehov. Lärarna införlivade 3 - 19 nya formativa bedömningsaktiviteter och medianen för antalet nya aktiviteter var 11,5. Den stora skillnaden avseende högstadie- och mellanstadielärarnas införlivande av formativ bedömning var att mellanstadielärarna införlivade fler formativa bedömningsaktiviteter i sin praktik. Resultaten visar att de egenskaper i kompetensutvecklingen som högstadielärarna ansåg vara viktiga för att kunna utveckla sin formativa klassrumspraktik var: 1) att den innehöll aktiviteter

som direkt kunde användas i klassrummet, 2) att de fick testa och uppleva värdet av att använda formativa bedömningsaktiviteter, 3) att kompetensutvecklingen knöt ihop teori och praktik, 4) att de fick stöd av en expert, 5) att de hade tid att ägna åt fortbildningen, 6) att de fick diskutera och samarbeta med kurskollegor, och 7) att kompetensutvecklingen var processorienterad och formativ. De egenskaper i kompetensutvecklingen som högstadielärarna ansåg som viktiga för införlivande av formativ bedömning var i stort desamma som för mellanstadielärarna. Detsamma gällde de hinder för införlivande som de båda lärargrupperna upplevde. De hinder som lärarna tog upp var främst kopplade till deras arbetssituation på skolan, såsom brist på tid då de upplevde en hög arbetsbelastning och situationen i klassen. Högstadielärarna upplevde även att implementeringen av en ny kursplan som skulle ske samtidigt bidrog till ytterligare arbetsbelastning och brist på tid, något som inte framkom som ett hinder för mellanstadielärarna.

6.3 Artikel III

Forskning har identifierat egenskaper hos kompetensutvecklingar som är viktiga för att åstadkomma förändringar i lärarnas praktik och i elevernas prestationer. Utfallet av en kompetensutvecklingsinsats verkar dock inte gå att förutsäga baserat på om dessa designegenskaper är närvarande eller ej. I artikel III antogs ett motivationsperspektiv i en strävan att förstå de val lärare gör i kompetensutvecklingar. Som en tolkande lins i analysen användes förväntan-värde-teorin (Wigfield & Eccles, 2000) och enligt teorin är det två motivationsföreställningar; förväntningar av framgång och de värden som individen ser i att lära eller genomföra en aktivitet som direkt påverkar individens motivation (se avsnitt 2.4.4). Syftet med studien var att undersöka varför det slumpmässiga urvalet av högstadielärare, efter att de deltagit i en kompetensutvecklingssatsning i formativ bedömning, förändrade sin praktik på det sätt som identifierats i artikel II. Dessutom var syftet att undersöka varför högstadielärarnas förändringar var annorlunda jämfört med de förändringar som mellanstadielärarna gjorde, som deltog i den parallella studien där utfallet avseende lärarnas praktik och elevernas prestationer i matematik var ett annat. I studien fokuseras följande forskningsfrågor: (1) Hur kan högstadielärarnas förändring mot en mer formativ bedömningspraktik förklaras av deras motivationsföreställningar?" (2) Hur kan skillnaden i antalet införlivade formativa bedömningsaktiviteter för högstadielärarna och mellanstadielärarna förklaras av skillnader i deras motivationsföreställningar?" Baserat på två enkäter och lärarintervjuer undersöktes lärarnas motivationsföreställningar kopplade till införlivande av formativ bedömning i klassrumspraktiken. Enkäterna innehöll öppna frågor och frågor som skulle besvaras på en femgradig skala. I jämförelsen med mellanstadielärarnas motivationsföreställningar användes Mann-Whitney U test för att testa för signifikanta skillnader mellan grupperna.

Resultaten visar att lärarnas förändringar mot en mer formativ klassrumspraktik kan förklaras av deras motivationsföreställningar efter kompetensutvecklingen. Detta gäller antalet formativa bedömningsaktiviteter som lärarna införlivade likväl som vilka aktiviteter de valde att införliva. Högstadielärarna rapporterade relativt höga värden avseende deras föreställningar om förväntad framgång och värden för införlivande av formativ bedömning. De förväntade sig samtidigt måttliga kostnader av att införliva formativ bedömning. Det större antal formativa bedömningsaktiviteter som mellanstadielärarna införlivade stämmer överens med dessa lärares motivationsföreställningar. Detta då mellanstadielärarna rapporterade högre värden avseende deras föreställningar om förväntad framgång och värden för införlivande av formativ bedömning än högstadielärarna, utan att deras föreställningar om kostnader med införlivandet var högre. I artikeln diskuteras hur dessa föreställningar kan ha utvecklats genom att koppla nyanser av egenskaper hos läraren, skolkontexten och kompetensutvecklingen till mekanismer identifierade i en modell av lärares professionella utveckling (Clarke & Hollingsworth, 2002).

6.4 Artikel IV

Ett antal forskningsöversikter har visat att formativ bedömning kan öka elevers prestationer. Men även om man inom forskningsfältet är relativt överens om formativ bedömnings potential att påverka elevers prestationer, förekommer stor variation i hur formativ bedömning konceptualiseras och införlivas och stor variation i effekter på elevprestationer. Ett slumpmässigt urval av 14 högstadielärare (som det aktuella året undervisade årskurs 7) i matematik deltog i en omfattande kompetensutvecklingssatsning kring formativ bedömning, med utgångspunkt i ett ramverk från Wiliam och Thompson (2008). Efter kompetensutvecklingssatsningen införlivade lärarna ett antal olika formativa bedömningsaktiviteter regelbundet i sin undervisning. Syftet med studien var att undersöka om de formativa bedömningsaktiviteter som de deltagande lärarna införlivade hade någon effekt på elevernas prestationer i matematik jämfört med en kontrollgrupp. Ett annat syfte var att undersöka om det fanns en relation mellan antalet införlivade nya formativa bedömningsaktiviteter (som identifierats och beskrivs i Artikel II), och elevers prestationsökningar i matematik. I studien fokuseras följande forskningsfrågor: (1) Har den formativa klassrumspraktiken som införlivades av högstadielärarna, efter kompetensutvecklingssatsningen, någon inverkan på deras elevers prestationer i matematik? (2) Finns det en korrelation mellan antalet nya formativa bedömningsaktiviteter som införlivades av lärarna och elevernas prestationsvinster i matematik? Alla elever i kommunen gjorde det aktuella läsåret ett förtest och ett eftertest särskilt konstruerade för denna studie. De klasser som undervisades av lärarna som deltagit i kompetensutvecklingen utgjorde interventionsgruppen och övriga klasser i kommunen

utgjorde kontrollgruppen. Som ett mått på klassrumspraktikens effekter användes medelvärdet av elevernas resultat på eftertestet då hänsyn tagits till förtestet för eleverna i lärarens klass. Analys med ANCOVA användes för att testa för signifikanta skillnader i elevernas testresultat mellan interventionsgruppen och kontrollgruppen. För att undersöka om det fanns en relation mellan antalet införlivade nya formativa bedömningsaktiviteter och elevers prestationsökningar i matematik genomfördes en partiell korrelation, där hänsyn togs till resultatet på förtestet.

Resultaten av ANCOVA visar att efter justering av poängen på förtestet så var det ingen signifikant skillnad i poäng på eftertestet mellan interventionsgruppen och kontrollgruppen ($F(1,26) = 0{,}005$, $p = 0{,}946$). Detta innebär att det inte finns något stöd i dessa data för att de elever som undervisats av lärare som förändrat sin formativa klassrumspraktik, efter deltagande i kompetensutvecklingen, har utvecklat sina matematikkunskaper mer efter ett läsår än de elever som undervisats av andra lärare. Vi fann heller ingen relation mellan antalet nya införlivade formativa bedömningsaktiviteter av lärarna och elevernas prestationssökningar i matematik. Den partiella korrelationen visade en svag, men inte signifikant, positiv korrelation mellan antalet nya införlivade formativa bedömningsaktiviteter av lärarna och elevernas resultat på eftertestet, då hänsyn tagits till elevernas resultat på förtestet ($r = 0{,}154$, $n = 14$, $p = 0{,}616$). Att det inte blev någon effekt på elevernas lärande kan bero på egenskaperna i den formativa klassrumspraktik som lärarna införlivade. Ett exempel där den införlivade formativa klassrumspraktiken gav positivt utfall är den parallella studien med mellanstadielärare, vars införlivade formativa klassrumspraktik gav ökade prestationsvinster för eleverna. Några möjliga förklaringar till skillnader i utfallet på elevprestationer mellan elevgrupperna diskuteras med ett fokus på frågor som rör egenskaper i den införlivade formativa klassrumspraktiken.

Efterord

Tänkte knyta an till titeln, som likväl skulle kunna syfta på vad det innebär att vara doktorand. Att vara Doktorand: En enkel match eller en svår utmaning? Många brukar beskriva doktorandtiden som en resa och jag kan inte annat än att hålla med. Inte har det varit någon enkel match, utan minst sagt en svår utmaning. Den som fick mig att anta utmaningen var Peter, som pushade mig till att söka och börja som licentiand. Men också Torulf som sedan uppmuntrade mig att fortsätta som doktorand. Jag visste inte vad jag gav mig in på, och det visst nog inte ni heller ;-). Stort Tack! Utan er hade detta inte blivit av! Tack för ert stora stöd och alla lärorika samtal. Ni har verkligen utmanat mig i mina tankar, fått mig att tänka kritiskt och också uppmuntrat mig. Utmaningen har pågått och jag har haft bättre och sämre avbrott under doktorandtiden, de bättre har mina barn stått för, vad vore livet utan er? De sämre stod akilles för, men nu är den matchen tack och lov över. Under resans gång har jag tagit ett rejält kliv, från en 5-poängs uppsats inom lärarprogrammet (som jag dessutom skrev tillsammans med en kursare) till denna avhandling. Vet inte om jag kan rekommendera andra att ta sådana kliv… i sann formativ anda är det bra både för måluppfyllelse och för motivationen att ta sig fram med hjälp av delmål. Jag valde och antog utmaningen själv, men som tur är har jag haft ett helt lag med mig. Det har inte varit som det brukar vara då jag antar utmaningar i beachvolley, då man bara är två spelare, utan i det här laget har det varit många. Så utöver mina handledare, huvudcoach Torulf och biträdande coach Peter, är det många fler jag vill tacka som varit med i mitt lag.

Jag vill tacka alla de fantastiska kompetenta lärare som lät mig ta del av deras arbete i skolan, deras tankar och utveckling. Det har också varit roligt att forska om något som man verkligen tror på. En lärares citat som jag burit med mig och som även visar vad jag tycker om formativ bedömning är:

> *”Man kan ju inte bortse från detta nu när man är invigd”*

Jag vill också sända ett tack till forskarskolan fontD, och alla de som jag träffade vid sammankomsterna i Norrköping. Dessutom vill jag tacka USE, då jag deltog i några av deras kurser och alla jag träffade i det sammanhanget. Sedan vill jag såklart tacka alla trevliga, roliga, kunniga kollegor som jag fått träffa och lära känna på NMD. Ett stort tack också till UFM, det har varit en förmån att få delta i den miljön som UFM utgör med alla kunniga, hjälpsamma forskare som ni alla är. Jag vill särskilt tacka Tomas, som ryckte in under min föräldraledighet och genomförde intervjuer och observationer. Jag vill också särskilt tacka Magnus, som var min slutläsare och gav mig värdefull feedback.

Dessutom vill jag tacka alla doktorandkollegor. Då man hållit på så länge om jag gjort, är det ju några som ”passerat”. Några av er har varit särskilt viktiga under olika perioder av min doktorandtid. Lena, du hjälpte mig verkligen i starten, att finna min plats och alla diskussioner kring allt vad det innebar att vara doktorand. Du bidrog också med många goda skratt och trevliga samtal. Maria, ditt stöd under de sista åren har varit superbt, vad gjorde jag utan våra kaffestunder, och långa tisdagsarbetskvällar, (och ikearesor ☺)? Lotta, alla givande lärorika samtal, både som doktorandkollega, lärarkollega och som vän, och alla (gap) skratt. Såå skönt att vi fick dela slutspurten tillsammans! Sofie, som doktorandkollega, Fofu-kollega och spurtkollega, om någon dag är det din tur och det kommer att gå galant. Jag vill verkligen tacka er alla doktorander som funnits med under olika perioder av min tid som doktorand (om jag nämner fler namn så kommer jag att glömma någon). TACK!! Ett dunder tack också till alla er som hjälpte mig med stort och smått under slutspurten, Maria, Mathias, Ingela, Lena, Lotta! Vill också passa på att tacka Kerstin, vars fina foto ni ser på framsidan av avhandlingen. Stort tack till alla mina kära vänner, som fått mig att göra annat än att jobba (även om jag varit svårlockad på slutet) och som förgyller min tillvaro med beach, skratt, skidåkning och allt annat skoj. Stor kram och tack till min stora familj både här i Umeå, Bålsta och Göteborg. Som stöttat, uppmuntrat och fått mig att tänka på annat. Tack för att ni finns! Ni är bara bäst! Och ett särskilt stort TACK mamma för allt stöd och uppmuntran jag fått genom livet och att du alltid funnits vid min sida. En särskild tanke sänder jag också till pappa, som också alltid finns med mig. Jag vill också tacka min moster Marika. Åsså min lilla familj, Jonas - för att du är en klippa, som tror på mig och som stöttar i vått och torrt! Jag fick det bäst ;-). Nils och Ines - all lycka och glädje som ni ger mig, och bara med er närvaro så påminner ni mig om vad som är viktigt i livet. Tack för att ni finns och för att ni är ni!

Nu är denna utmaning snart över och det känns som att jag vill vänta lite innan jag antar någon ny. Men jag vet ju hur det kan vara, dyker det upp något, så vet man aldrig... ☺

Umeå, en sen kväll på NMD i maj

Referenser

Andersson, C. & Palm, T. (2015). Reasons for teachers' successful development of a formative assessment practice through professional development a motivation perspective. I C. Andersson, *Professional development in formative assessment: effects on teacher classroom practice and student achievement*. Umeå, Sweden: Umeå University (Doctoral thesis, Department of Science and Mathematics Education).

Andersson, C., & Palm, T. (2017a). The impact of formative assessment on student achievement: a study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learning and Instruction*, *49*, 92-102.

Andersson, C. & Palm, T. (2017b). Characteristics of improved formative assessment practice. *Education Inquiry*.

Andersson, C., Boström, E. & Palm, T. (2017). Formative assessment in Swedish mathematics classroom practice. *Nordic Studies in Mathematics Education*, *22*(1), 5-20.

Arter, J., (2009, april). *Classroom assessment for student learning (CASL) perspective on the JCSEE student evaluation standards*. Paper presented at the annual meeting of the American Educational Research Association, in the Division H symposium *JCSEE National Conference on Benchmarking Student Evaluation Practices*. Paper retriewed at http://downloads.pearsonassessments.com/ati/downloads/JCSEE%20Student%20Eval%20Practices%20Symp%20Paper%202009.pdf

Baird, J., Hopfenbeck, T., Newton, P., Stobart, G. & Steen-Utheim, A. (2014). *State of the field review: Assessment and learning*. Report for the Norwegian Knowledge Centre for Education, case number 13/4697. Retrieved from http://forskningsradet.no

Bell, C., Steinberg, J., Wiliam, D., & Wylie, C. (2008). *Formative assessment and teacher achievement: Two years of implementation of the Keeping learning on track program*. Paper presented at the annual meeting of the National Council on Measurement in Education, New York.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5-25.

Birenbaum, M., Kimron, H., & Shilton, H. (2011). Nested contexts that shape assessment for learning: Schoolbased professional learning community and classroom culture. *Studies in Educational Evaluation*, 37 (1), 35-48.

Birenbaum, M., DeLuca, C., Earl, L., Heritage, M., Klenowski, V., Looney, A., ... & Wyatt-Smith, C. (2015). International trends in the implementation of assessment for learning: Implications for policy and practice. *Policy Futures in Education*, *13*(1), 117-140.

Black, P. (2015). Formative assessment–an optimistic but incomplete vision. *Assessment in Education: Principles, Policy & Practice, 22*(1), 161-177.
Black, P., Harrison, C., Hodgen, J., Marshall, B., & Serret, N. (2010). Validity in teachers' summative assessments. *Assessment in Education:Principles, Policy & Practice*, 17(2), 215-232.
Black, P., Harrison, C., Lee, C., & Marshall, B. (2003). *Assessment for learning: Putting it into practice*. Buckingham: Open University Press.
Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in education, 5*(1), 7-74.
Black, P. & Wiliam, D. (1998b). *Inside the black box: Raising standards through classroom assessment*. Phi Delta Kappan, *80*(2), 139-148
Black, P., & Wiliam, D. (2003). 'In praise of educational research': Formative assessment. *British Educational Research Journal*, *29*(5), 623-637.
Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5-31.
Brookhart, S. (2010). Mixing it up: Combining sources of classroom achievement information for formative and summative purposes. In H. L. Andrade & G. J. Cizek (Eds.) *Handbook of formative assessment*, (pp. 279-296). Abingdon: Routledge.
Briggs, D. C., Ruiz-Primo, M. A., Furtak, E., Shepard, L., & Yin, Y. (2012). Meta-analytic methodology and inferences about the efficacy of formative assessment. *Educational Measurement: Issues and Practice, 31*(4), 13-17.
Bryman, A. (2011). Samhällsvetenskapliga metoder. (2., [rev.] uppl.) Malmö: Liber.
Burns, M. K., Klingbeil, D. A., & Ysseldyke, J. (2010). The effects of technology-enhanced formative evaluation on student performance on state accountability math tests. *Psychology in the Schools*, *47*(6), 582-591.
Carless, D. (2005). Prospects for the implementation of assessment for learning. *Assessment in Education Principles Policy and Practice, 12*(1), 39-54. doi:10.1080/0969594042000333904
Cizek, G. J. (2010). An introduction to formative assessment: History, characteristics, and challenges. I H. L. Andrade & G. J. Cizek (red.), *Handbook of formative assessment* (s. 3-17). Abingdon: Routledge.
Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education, 18*(8), 947-967.
Cohen, L., Manion, L. & Morrison, K. (2011). *Research methods in education*. (7. uppl.). New York, NY: Routledge.
De Lisle, J. (2015). The promise and reality of formative assessment practice in a continuous assessment scheme: the case of Trinidad and Tobago. *Assessment in Education: Principles, Policy & Practice, 22*(1), 79-103.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181-199.
Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, *3*(3), 231-264.
Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The Limited Scientific Evidence of the Impact of Formative Assessment in Education. *Practical Assessment, Research & Evaluation, 14*(7), 1-11.
Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies values and academic behaviours. I J. T. Spence. (red.), *Achievement and Achivement Motives: Psychological Approaches* (s. 75-146). San Francisco, CA: W H Freeman.
Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*(1), 109-132.
Filsecker, M., & Kerres, M. (2012). Repositioning formative assessment from an educational assessment perspective: A response to Dunn & Mulvenon (2009). *Practical Assessment, Research & Evaluation, 17*(16).
Flórez, M. T., & Sammons, P. (2013). *Assessment for learning: Effects and impact*. Reading: CfBT Education Trust.
Guba, E. & Lincoln, Y. (1994) Competing paradigm in qualitative research. I N., Denzin, & Y. Lincoln (red.), *Handbook of qualitative research* (s.105-117). Thousand Oaks, CA: Sage.
Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.
Hawe, E., & Parr, J. (2014). Assessment for learning in the writing classroom: An incomplete realisation. *Curriculum Journal, 25*(2), 210-237.
Heitink, M., Van der Kleij, F., Veldkamp, B., Schildkamp, K., & Kippers, W. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational Research Review, 17*, 50-62.
Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practice, 28*(3), 24-31.
Hirsh, Å., & Lindberg, V. (2015). Formativ bedömning på 2000-talet – en översikt av svensk och internationell forskning *SKOLFORSK delstudie XI*: Vetenskapsrådet. Vetenskapsrådets rapporter, 2015.
Hopfenbeck, T. N., Flórez Petour, M. T., & Tolo, A. (2015). Balancing tensions in educational policy reforms: large-scale implementation of Assessment for Learning in Norway. *Assessment in Education: Principles, Policy & Practice, 22*(1), 44-60.

Hume, A., & Coll, R. (2009). Assessment of learning, for learning, and as learning: New Zealand case studies. *Assessment in Education: Principles, Policy & Practice, 16*(3), 269-290.
James, M., & McCormick, R. (2009). Teachers learning how to learn. *Teaching and Teacher Education: An International Journal of Research and Studies, 25*(7), 973-982.
Jönsson, A., Lundahl, C., & Holmgren, A. (2015). Evaluating a large-scale implementation of Assessment for Learning in Sweden. *Assessment in Education: Principles, Policy & Practice, 22*(1), 104-121.
Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research, 86*(4), 945-980.
Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice, 30*(4), 28-37.
Klenowski, V. (2009). Assessment for learning revisited: An Asia-Pacific perspective. *Assessment in Education: Principles, Policy & Practice, 16*(3), 263-268.
Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284.
Koedinger, K. R., McLaughlin, E. A., & Heffernan, N. T. (2010). A quasi-experimental evaluation of an on-line formative assessment and tutoring system. *Journal of Educational Computing Research*, *43*(4), 489-510.
Kvale, S. & Brinkmann, S. (2014). *Den kvalitativa forskningsintervjun*. (3. [rev.] uppl.) Lund: Studentlitteratur.
Laveault, D. (2016). Building capacity: Professional development and collaborative learning about assessment. I D. Laveault and L. Allal (red.), *Assessment for Learning: Meeting the Challenge of Implementation* (s. 131-143). Switzerland: Springer International Publishing.
Leahy, S. & Wiliam, D. (2012). From teachers to schools: scaling up professional development for formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 49–71). London: Sage Publications.
Lysaght, Z., & O'Leary, M. (2013). An instrument to audit teachers' use of assessment for learning. *Irish Educational Studies*, *32*(2), 217-232.
Marchis, I. (2012). Self-Regulated Learning and Mathematical Problem Solving. *New Educational Review, 27*(1), 195-208.
Marshall, B., & Jane Drummond, M. (2006). How teachers engage with Assessment for Learning: Lessons from the classroom. *Research Papers in Education, 21*(2), 133-149.
McMillan, J. H., Venable, J. C., & Varier, D. (2013). Studies of the effect of formative assessment on student achievement: So much more is needed. *Practical Assessment, Research & Evaluation, 18*(2), 2.
Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: an expanded sourcebook* (2. uppl.). Thousand Oaks, CA: Sage.

National Mathematics Advisory Panel, (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*: US Department of Education.
Organisation for economic co-operation and development [OECD]. (2005). *Formative assessment: improving learning in secondary classrooms*. Paris: OECD.
Ofsted (2008). Assessment for learning: the impact of National Strategy support. London: Ofsted.
Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research, 81*(3), 376-407.
Panadero, E., & Jönsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review, 9*, 129-144.
Patton, M. Q. (2002). Qualitative research & evaluation methods (3. uppl.). London : Sage.
Pedder, D. & James, M. (2012). Professional learning as a condition for assessment for learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 33–48). London: Sage Publications.
Palm, T., Andersson, C., Boström,E., & Vingsle, C. (Accepted). A review of the impact of formative assessment achievement in mathematics. Nordic Studies in Mathematics Education.
Rakoczy, K., Harks, B., Klieme, E., Blum, W., & Hochweber, J. (2013). Written feedback in mathematics: Mediated by students' perception, moderated by goal orientation. *Learning and Instruction, 27*, 63-73.
Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science, 28*(1), 4-13.
Randel, B., Apthorp, H., Beesley, A. D., Clark, T. F., & Wang, X. (2016). Impacts of professional development in classroom assessment on teacher and student outcomes. *The Journal of Educational Research*, *109*(5), 491-502.
Robinson, J., Myran, S., Strauss, R., & Reed, W. (2014). The impact of an alternative professional development model on teacher practices in formative assessment and student learning. *Teacher Development, 18*(2), 141-162.
Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school studies: A meta-analytic review. *Journal of Educational Psychology, 95*(2), 240-257.
Ross, J.A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, and Evaluation, 11*(10), 1-13.
Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology, 25*(1), 54-67.
Ryve, A., Nilsson, P., Palm, T., Van Steenbrugge, H., Andersson, C., Bergwall, A., ... & Vingsle, L. (2015). Kartläggning av forskning om formativ bedömning, klassrumsundervisning och läromedel i matematik. Stockholm: Vetenskapsrådet.
Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119-144.

Schneider, M. C., & Andrade, H. (2013). Teachers' and administrators' use of evidence of student learning to take action: Conclusions drawn from a special issue on formative assessment. *Applied Measurement in Education*, 26(3), 159-162.

Schneider, M. C., & Randel, B. (2010). Research on characteristics of effective professional development programs for enhancing educators' skills in formative assessment. I H. L. Andrade & G. J. Cizek (red.), *Handbook of formative assessment* (s. 251-276). Abingdon: Routledge.

Schneider, M. C., & Gowan, P. (2013). Investigating teachers' skills in interpreting evidence of student learning. *Applied Measurement in Education*, *26*(3), 191-204.

Schoenfeld, A. (2007). Method. I F. K. Lester, Jr. (red.), *Second handbook of research on mathematics teaching and learning: a project of the National Council of Teachers of Mathematics* (s. 69–107). Charlotte, NC: Information Age Publishing.

Schoenfeld, A. H. (2014). What makes for powerful classrooms, and how can we support teachers in creating them? A Story of Research and Practice, Productively Intertwined. *Educational Researcher, 43*(8), 404-412.

SFS 2008:192 Lag om ändring i lagen (2003:460) om etikprövning av forskning som avser människor. Stockholm: Utbildningsdepartementet. Från http://www.lagboken.se/Views/Pages/GetFile.ashx?portalId=56&cat=27526&docId=181354&propId=5

Shavelson, R. J. (2008). Guest editor's introduction. *Applied Measurement in Education, 21*(4), 293-294.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Skolverket. (2012). *Tid för matematik: erfarenheter från matematiksatsningen 2009-2011*. Stockholm: Fritzes.

Smith, K. (2011). Professional development of teachers - A prerequisite for AfL to be successfully implemented in the classroom. *Studies In Educational Evaluation, 37*(1), 55-61.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment FOR learning. *Phi Delta Kappan, 83*(10), 758-765.

Swaffield, S. (2011). Getting to the heart of authentic assessment for learning. *Assessment in Education: Principles, Policy & Practice, 18*(4), 433-449.

Tierney, R. D. (2006). Changing practices: Influences on classroom assessment. *Assessment in Education: Principles, Policy and Practice, 13*(3), 239-264.

Tillema, H., Leenknecht, M., & Segers, M. (2011). Assessing assessment quality: Criteria for quality assurance in design of (peer) assessment for learning - A review of research studies. *Studies In Educational Evaluation, 37*(1), 25-34.

Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher Professional Learning and Development: Best Evidence Synthesis Iteration*. Wellington, New Zealand: Ministry of Education.
Torrance, H. (2007). Assessment "as" learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice, 14*(3), 281-294.
Torrance, H. (2012). Formative assessment at the crossroads: conformative, deformative and transformative assessment. *Oxford Review of Education, 38*(3), 323-342.
Van Zundert, M., Sluijsmans, D., & Van Merriënboer, J. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction*, *20*(4), 270-279.
Vetenskapsrådet (2002). *Forskningsetiska principer inom humanistisk-samhällsvetenskaplig forskning*. Stockholm: Vetenskapsrådet.
Vetenskapsrådet (2011). God forskningssed. Stockholm: Vetenskapsrådet.
Villegas-Reimers, E. (2003). *Teacher professional development: an international review of the literature*. Paris: International Institute for Educational Planning.
Vingsle, C. (2014) Formative assessment: Teacher knowledge and skills make it happen. Retrieved from DIVA. (Licenciatuppsats, Licenciate avhandlingar i pedagogiskt arbete ;15) Umeå universitet
Wellington, J. (2000). *Educational research: contemporary issues and practical approaches*. London: Continuum.
Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review, 12*(3), 265-310.
Wigfield, A., & Eccles, J. S. (2000). Expectancy - Value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68-81.
Wiliam, D. (2007). Keeping learning on track: classroom assessment and the regulation of learning. I F. K. Lester (red.), *Second handbook of research on mathematics teaching and learning* (s. 1053-1098). Charlotte, NC: Information Age Publishing.
Wiliam, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. I H. L. Andrade & G. J. Cizek (red.), *Handbook of formative assessment* (s.18-40). Abingdon: Routledge.
Wiliam, D. (2013). *Att följa lärande: formativ bedömning i praktiken*. (1. uppl.) Lund: Studentlitteratur.
Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy & Practice, 11*(1), 49-65.
Wiliam, D., & Thompson, M. (2008). Integrating assessment with learning: what will it take to make it work? I C. A. Dwyer (red.), *The Future of Assessment: Shaping Teaching and Learning (s. 53-82)*. Mahwah, NJ: Lawrence Erlbaum Associates.

Wylie, E. C., & Lyon, C. J. (2015). The fidelity of formative assessment implementation: issues of breadth and quality. *Assessment in Education: Principles, Policy & Practice, 22*(1), 140-160.

Yeh, S. S. (2009). Class size reduction or rapid formative assessment?: A comparison of cost-effectiveness. *Educational Research Review, 4*(1), 7-15.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*(2), 64-70.