



HÖGSKOLAN
DALARNA

Examensarbete för ämneslärarexamen

Grundnivå 2

Lärares resonemang vid provkonstruktion i Företagsekonomi

Författare: Tim Svanberg
Handledare: Steven Hunter Lindqvist
Examinator: Jan Morawski
Ämne/huvudområde: Pedagogiskt arbete
Kurskod: PG2066
Poäng: 15
Examinationsdatum: 180110

Vid Högskolan Dalarna finns möjlighet att publicera examensarbetet i fulltext i DiVA. Publiceringen sker open access, vilket innebär att arbetet blir fritt tillgängligt att läsa och ladda ned på nätet. Därmed ökar spridningen och synligheten av examensarbetet.

Open access är på väg att bli norm för att sprida vetenskaplig information på nätet. Högskolan Dalarna rekommenderar såväl forskare som studenter att publicera sina arbeten open access.

Jag/vi medger publicering i fulltext (fritt tillgänglig på nätet, open access):

Ja

Nej

Abstract:

The aim of this essay is to gain more knowledge of how teachers reason when they construct tests in accounting in the business program in upper secondary school in Sweden. The study uses a qualitative method of inquiry interviewing four business teachers in the business program. The teachers interviewed in this essay say that they are highly influenced by the student's textbooks and what these bring up. According to the teachers interviewed in this study the content of the tests is designed in accordance with the curricula, however it is difficult to design tests that differentiate students' knowledge, and as such grades, as described in the curriculum for the business course.

Nyckelord: Bedömning, Bokföring, Företagsekonomi, Gymnasium, Gymnasieskola, Kunskapskrav Prov, Provkonstruktion

Innehållsförteckning

| | |
|---|-----------|
| 1. INLEDNING | 4 |
| 1.1 SYFTE OCH FORSKNINGSFRÅGOR | 5 |
| 2. BAKGRUND..... | 6 |
| 2.1 STYRDOKUMENT | 6 |
| 2.2 DEFINITIONER | 7 |
| 2.3 PROVANVÄNDNING I SVENSKA SKOLAN | 7 |
| 3. TEORI OCH TIDIGARE FORSKNING..... | 9 |
| 3.1 TEORETISKT RAMVERK | 9 |
| 3.1.1 Validitet | 9 |
| 3.1.2 Kanes validitetsmodell..... | 11 |
| 3.1.3 Reliabilitet..... | 11 |
| 3.1.4 Under- och överrepresentation..... | 13 |
| 3.1.5 Objektiva prov och alternativa (subjektiva) bedömningar | 13 |
| 3.1.6 Alignment | 14 |
| 3.1.7 Konstruktionsregler | 15 |
| 3.2 TIDIGARE FORSKNING | 15 |
| 4. METOD..... | 17 |
| 4.1 VAL AV METOD FÖR UNDERSÖKNINGEN | 17 |
| 4.2 URVAL | 17 |
| 4.3 GENOMFÖRANDET | 18 |
| 4.4 ANALYS AV EMPIRI | 18 |
| 4.5 ETISKA ÖVERVÄGANDEN..... | 19 |
| 4.6 METODDISKUSSION..... | 19 |
| 5. RESULTAT | 20 |
| 5.1 LÄRARNAS PROVKONSTRUKTION | 20 |
| 5.2 INNEHÅLL PÅ PROVEN | 21 |
| 5.3 KONSTRUKTION AV PROV FÖR BETYGSÄTTNING OCH BEDÖMNING..... | 23 |
| 6. DISKUSSION..... | 27 |
| 6.1 VALIDITET..... | 27 |
| 6.2 KANES VALIDITETSMODELL | 28 |
| 6.3 RELIABILITET | 29 |
| 6.4 UNDER- OCH ÖVERREPRESENTATION | 30 |
| 6.5 OBJEKTIVA PROV OCH ALTERNATIVA (SUBJEKTIVA) BEDÖMNINGAR | 31 |
| 6.6 ALIGNMENT | 33 |
| 6.7 KONSTRUKTIONSREGLER | 33 |
| 6.8 ANALYS AV RESULTAT MOT TIDIGARE FORSKNING | 34 |
| 7. SLUTSATSER..... | 36 |
| 8. FÖRSLAG TILL VIDARE FORSKNING..... | 37 |
| REFERENSER..... | 38 |
| BILAGOR..... | 40 |
| BILAGA 1. INFORMATIONSBREV | 40 |
| BILAGA 2. FRÅGEFORMULÄR – SEMISTRUKTURERAD INTERVJU | 41 |

1. Inledning

Att förstå hur lärarna resonerar kring provkonstruktion är viktigt pga Skolverket (2017b) kräver att läraren ska sätta betyg efter avslutad kurs. Skolverket anger att betyg ska sättas utifrån elevernas förmåga att uppnå kunskapskraven. Kunskapskraven innehåller ett antal värdeord, eller progressionsord, som läraren behöver ta hänsyn till i sin betygsättning. Det är då värdefullt att se hur lärarna i sitt resonemang hanterar kunskapskraven när de konstruerar och bedömer provet. Om lärarna ska bedöma elevernas kunskaper utifrån vissa kriterier är det därför intressant att undersöka *vilka* kunskaper som lärarna tycker är viktiga att mäta och *hur* de väljer ut de kunskaper de vill mäta.

Jag fick idén till detta arbete under min andra VFU-period. Jag fick då möjlighet att undervisa i arbetsområdet Bokföring och även avsluta arbetsområdet med ett skriftligt prov. Jag upplevde då att det var mycket svårt att veta vad som förväntas av eleverna för att de ska nå upp till en viss betygsnivå. Det var också svårt att veta vilka kunskaper som ska testas. Hur vet jag om det exempelvis är bokföring av löner och momsredovisningar eller någonting annat som jag förväntas testa i provet och hur vet jag på vilken nivå eleverna förväntas befinna sig i dessa moment för att nå kravnivån för E, C eller A?

En undersökning av tidigare forskningen har inte visat studier på just detta område, vilket visar på behovet av ytterligare studier. Provkonstruktion är en del av testteorin och för detta forskningsområde har däremot teori gått att finna. Jag är intresserad av att undersöka och fördjupa mig i lärares resonemang. Denna studie tar del av tidigare forskning och teorier kring provkonstruktion för att få mera kunskap om lärares resonemang vid just provkonstruktion.

Christina Wikström (2013, ss. 36-37) menar att samhället har en stark tilltro till skolvärldens prov och att dessa säger någonting om elevens prestation, men det finns gott om faktorer som kan störa provens tillförlitlighet. Utifrån bristen på tidigare forskning och pga. vikten av att förstå hur lärarna resonerar när de konstruerar prov är denna studie intressant. Den här uppsatsen undersöker lärares resonemang vid provkonstruktion i bokföring genom intervjuer med lärare som undervisar i företagsekonomi 2.

1.1 Syfte och forskningsfrågor

Syftet med detta arbete är att få mera kunskap om hur lärare resonerar kring sitt konstruerande av prov i bokföring i ämnet företagsekonomi.

Forskningsfrågorna som detta arbete svarar på är följande:

1. Vilka kunskaper anser lärarna att de vill mäta med bokföringsprov i Företagsekonomi 2?
2. Hur säger sig lärarna välja ut de kunskaper de vill mäta med bokföringsprov i Företagsekonomi 2?
3. Hur resonerar lärare kring konstruktionen av bokföringsprov då det gäller att mäta elevernas olika kunskapsnivåer kopplat till kunskapskraven för kursen Företagsekonomi 2.

2. Bakgrund

I detta avsnitt kommer en bakgrund till studien att ges. Först sker en genomgång av de styrdokument från skolverket som styr hur betygsättning ska ske. Därefter redogörs för några begrepp som inte vidare förklaras i den löpande texten för att läsaren ska förstå dessa. Slutligen görs en kort historisk överblick över prov för att visa hur dessa har använts genom historien samt hur prov används för bedömning i nutid.

2.1 Styrdokument

Enligt Skolverket (2011, ss. 14-15) styrs betygsättning och bedömning genom *Läroplan, examensmål och gymnasiegemensamma ämnen för gymnasieskola 2011*. Bedömningen kopplat till betyg har sin grund i kunskapskraven som är graderade från A-F. Grunden för hur bedömningen görs är hur eleverna visar att de når upp till kunskapskraven för respektive betygsnivå. Om eleven ska få E, C eller A i betyg så behöver hon eller han till fullo ha uppfyllt kunskapskraven för just den nivån. Men om eleven exempelvis når upp till samtliga kunskapskrav för E och till övervägande del för C så ska eleven få betyget D. Analogt används samma resonemang för betyget B (Statens skolverk, 2011b, s. 12).

Ämnet företagsekonomi ska enligt *Läroplan, examensmål och gymnasiegemensamma ämnen för gymnasieskola 2011* ge eleverna:

[...]Möjlighet att utveckla kunskaper i företagsekonomiska teorier, modeller och metoder[...] (Skolverket, 2011, s. 1)

Vidare stipulerar kursplanen för Företagsekonomi 2 att undervisningen bland annat ska behandla bokföring (Skolverket, 2011, s. 9).

Bokföring är en företagsekonomisk metod. I *läroplan, examensmål och gymnasiegemensamma ämnen för gymnasieskolan 2011* (Skolverket, 2011) finns följande krav som är hänförliga till bokföringsområdet:

För betyget E:

Eleven redogör översiktligt för innebörden av företagsekonomiska begrepp. Vidare använder eleven med viss säkerhet enkla företagsekonomiska metoder. Med hjälp av enkla teorier och modeller förklarar eleven företagsekonomiska företeelser och sammanhang

För betyget C:

Eleven redogör utförligt för innebörden av företagsekonomiska begrepp. Vidare använder eleven med viss säkerhet företagsekonomiska metoder. Med hjälp av teorier och modeller förklarar eleven företeelser och sammanhang. Eleven värderar med enkla omdömen teorierna och modellernas användbarhet

För betyget A:

Eleven redogör utförligt och nyanserat för innebörden av företagsekonomiska begrepp. Vidare väljer och använder eleven med säkerhet företagsekonomiska metoder. Med hjälp av komplexa teorier och modeller förklarar eleven företagsekonomiska företeelser och sammanhang. Eleven värderar med nyanserade omdömen teoriernas och modellernas användbarhet (Skolverket, 2011, ss. 9-10)

Om lärarens bedömning ska fungera är det därför viktigt att läraren grundar sin bedömning i de statliga styrdokument som finns (ibid sid 6). Wikström (2013, s.35) betonar att prov som är rätt konstruerade kan ge den information som läraren behöver för att göra sin bedömning av elevens kunskaper men att ett felkonstruerat prov riskerar att ge utrymme för feltolkningar vid bedömningen.

2.2 Definitioner

I den här delen ges definitioner av viktiga begrepp som behövs för att förstå studien.

Företagsekonomi – Den här studien använder skolverkets definition av ämnet företagsekonomi.

Ämnet företagsekonomi behandlar företagande i vid bemärkelse och belyser såväl ekonomiska som sociala och miljömässiga aspekter. I ämnet ingår marknadsföring, ledarskap och organisation, redovisning och kalkylering. Ämnet omfattar även entreprenörskap, som inbegriper metoder och förhållningssätt för att kunna starta, driva och vidareutveckla verksamheter (Skolverket, 2011, s. 1).

Prov – Studien använder Wikströms definition av prov som ”kunskapsmätning i utbildningssammanhang” Wikström, 2013, s.13).

Bokföring -

Bokföring är det praktiska arbetet med att löpande registrera olika händelser, t ex inköp och försäljning på olika konton. Det är ett system för att ordna och behandla ekonomiska händelser (affärshändelser) i ett företag (Bolagsverket, 2017).

2.3 Provanvändning i svenska skolan

Elevers kunskaper har testats i prov genom historien och denna mätmetod kan spåras tillbaka till Mesopotamien (4000-5000 år sedan). Det äldsta provet i Europa kan dateras till 1100-tal då det användes vid universitet som examinerande instrument (Wikström, 2013, s. 20). Proven i Sverige är inte fullt så gamla. Husförhören kan sägas vara den äldsta formen av prov vi använt oss av. Ur ett konstruktionsperspektiv fanns det brister i dessa förhör då de som skötte förhören hade en varierande grad av kompetens. Först när folkskolan infördes kan vi tala om den typ av prov som vi är vana vid i dagen skola (Ibid s.21).

Sverige är enligt Wikström (2013, s. 19) ett av få länder som inte ger eleverna examinerande prov när de avslutar sina grundskole- och gymnasiekurser. Detta innebär att de prov vi har i Sverige inte ges samma tyngd och allvar som vid den tidigare studentexamen eller som i många andra länder. Detta innebär vidare att de

svenska proven också är mindre genomlysta eftersom det är upp till varje lärare att själva bestämma hur deras prov ska se ut. Undantaget i sammanhanget är de nationella proven som skolverket beställer av svenska universitet och högskolor och som tar år att producera (Skolverket, 2017; Skolverket, Statens skolverk, 2016). Wikström (2013, sid 19) menar dock att provens betydelse har ökat på senare år till följd av ett större resultat- och målfokus i den svenska skolan. Vidare används dessa prov i hög utsträckning i ett summativt syfte. Summativ bedömning kan definieras som

Bedömningar, som sammanfattar (summerar) elevens kunskapsnivå i förhållande till något kriterium, till exempel en betygsskala, ett statistiskt värde eller en norm (Lundahl C. , 2016, s. 11).

Således används alltså proven till stor del för att summera den kunskapsnivå som eleven når upp till.

Prov är en form av formell bedömning till vilken man brukar räkna såväl muntliga redovisningar som skriftliga prov (Statens skolverk , 2011b, ss. 10-11). Denna typ av uppgift ska leda fram till att läraren gör en bedömning av elevens prestation och ligger till grund för betygsättning. Det bedöms då som viktigt att läraren bottenar i hur bedömningen ska gå till och på vilket vis denna bedömning korresponderar med kunskapskraven. Informell bedömning är å andra sidan det som läraren gör under hand i klassrumssituationer eller i andra möten med eleverna och där läraren gör observationer av elevers färdigheter (Ibid, sid 11).

3. Teori och tidigare forskning

I det här kapitlet presenteras de teorier som finns kopplade till testteori och provkonstruktion. Enligt skolverket (2016) finns det några begrepp som är centrala vid provkonstruktion. Det är dessa som används i det teoretiska ramverket. Först beskrivs provvaliditet generellt och sedan Kanes validitetsmodell mer specifikt. Validitet avser enligt Wikström (2013, s. 44) i vilken utsträckning ett prov mäter det som det är avsett att göra. Kanes validitetsmodell är en specifik metod för att undersöka validiteten i ett bedömningsinstrument, såsom prov. Därefter sker en genomgång av begreppet reliabilitet som berättar hur pass tillförlitligt provet är. Eftersom läraren enligt Wikström (2013, s. 40) inte kan testa precis all kunskap som läraren vill mäta kommer provet behöva vara ett urval av denna kunskap. Därför är det enligt Wikström (ibid s. 41) viktigt att läraren resonerar kring hur deras konstruktion av provet påverkar reliabiliteten. Reliabiliteten säger dessutom i vilken utsträckning läraren, utifrån sitt konstruerade prov riskerar att bedöma eleven felaktigt så att elevens betyg inte stämmer överens med resultatet på provet (Jönsson, 2013, s. 68). Efter detta beskrivs över- och underrepresentation. Detta innebär att lärarna testar kunskaper som är irrelevanta för kunskapskraven eller att de missar att testa kunskaper som de vill mäta (Wikström, 2013, ss. 90-91). I vilken utsträckning resonerar lärarna kring detta när de väljer ut vilken kunskap de vill mäta? Sedan belyses vad objektiva prov och alternativa bedömningar är samt vilka konsekvenser lärarens val av provtyp får för bedömningen. Valet som lärarna gör där avgör enligt Frederiksen och Allan (1989, s. 28) hur lätt det är att göra en bedömning av elevens prestation. Efter det beskrivs graden av samstämmighet med läroplanerna. Om lärarna ska kunna bedöma proven efter kunskapskraven är det enligt Näsström (2008, s. 66) viktigt att de är samstämmiga med läroplanen. Sedan presenteras Wikströms (2013, ss. 103-104) rekommendation kring hur prov bör vara utformade för att hantera ovanstående aspekter. Slutligen redogörs för den tidigare forskning som finns på området.

3.1 Teoretiskt ramverk

3.1.1 Validitet

Inom testteorin handlar validitet enligt Wikström (2013, s. 44) om i vilken utsträckning ett prov mäter det som det är tänkt att mäta men också om vilken hänsyn som tas till de konsekvenser själva mätningen får. Det finns enligt Wikström (ibid, s. 44) olika diskurser kring hur begreppet validitet ska tolkas. Den tidigare diskursen innebar att validitet avser själva provinstrumentet, dvs hur väl provet mäter det som det är avsett att mäta. Numera avses snarare att tolkningen av det resultat som eleverna får på provet ska ha hög validitet. Messick (1989, ss. 5-6) är en av de som ansluter sig till denna tolkning och menar att provkonstruktören måste ta syftet och relevansen med provet samt användbarheten av resultatet i beaktande. Argumentet från Messick är att oavsett om provfrågorna är konstruerade på ett sådant vis att de mäter det de avser att mäta har det ingen betydelse om tolkningen av elevernas resultat skiljer sig åt.

Enligt Wikström (2013, ss. 44-45) finns det två olika sätt att försäkra sig om validiteten i ett mätinstrument. Det första är att göra en direkt mätning som innebär att en person visar att vederbörande har en viss förmåga genom att utföra någonting, exempelvis förmågan att köra bil. Ibland är det dock inte möjligt av praktiska skäl.

Läraren behöver då på annan väg skaffa den information som vederbörande hade fått om eleven hade utfört uppgiften direkt. Detta kallas för indirekt mätning. I ett provkonstruktionsperspektiv innebär detta att provkonstruktören behöver hitta en modell för att mäta de kunskaper som denne vill mäta. Detta kan också beskrivas som att hitta indikatorer på kunskaper som konstruktören vill mäta (ibid, s. 45). En indirekt mätning sker genom att på omvägar mäta kunskaper. Abstrakta mätföremål, såsom egenskaper och tankar behöver enligt Wikström (ibid, s. 45) i regel mätas med hjälp av en indirekt mätning. Om konstruktören tar omvägar för att mäta en kunskap behöver vederbörande enligt Wikström (ibid, ss. 45-46) försäkra sig om att validiteten inte försämras. Validitet i testorisammanhang kan enligt (Cronbach & Meehl, 1955, ss. 281-282, Wikström 2013, ss47-49) traditionellt delas in i tre delar:

1. Innehållsvaliditet: Detta är den mest centrala delen av validiteten enligt Wikström (ibid, s. 47) och ger svar på om läraren i provet får med den kunskap som läraren hade tänkt att få med. Detta innebär att de frågor som valts ut till provet behöver representera den kunskap som läraren vill mäta.
2. Kriterierelaterad validitet: Denna typ av validitet ger svar på frågan om läraren genom provet får den information som denne efterfrågar. I regel sätts det i förväg upp vilka kriterier som gäller för den kunskap som här ska fångas. Om läraren med hjälp av provet kan dra samma slutsatser om dessa kunskaper hos eleverna som när de bedöms med andra bedömningsinstrument så kan slutsatsen enligt Wikström (ibid, s. 47) dras att den kriterierelaterade validiteten är god. I sin tur kan denna validitet delas upp i a) *samtidig validitet* och b) *prediktiv validitet*. Om den samtida validiteten är god skulle elevens kunskap kunna testas med ett annat bedömningsinstrument samtidigt och få samma resultat. Om läraren däremot vill försäkra sig om att det som testas i provet också kan säga någonting om hur eleven kommer klara någonting i framtiden, pratar vi om prediktiv validitet.
3. Begreppsvaliditet: Wikström (ibid, s. 47) menar att denna validitet är den mest komplicerade men också den som är minst applicerbar på kunskapsprov eftersom den mest används inom psykometrin. Även begreppsvaliditet ger svar på om provet fångar upp det läraren avser att mäta men här på ett mer abstrakt plan.

Goldstein (2015, ss. 193-194) menar att validitet i testsammanhang inte låter sig mätas kvantitativ utan snarare att själva konstruktet - innehållet behöver sättas i förhållande till resultatet. Detta behöver göras av varje provkonstruktör eller provanvändare. Provvaliditet blir en subjektiv bedömning. Ändå anser en del forskare att det visst går att validitetstesta prov. Denna forskning sammanfattas av Tengberg (2016b, s. 25). Validitetstesten sker genom att eliminera de faktorer som kan skada validiteten. Dessa är konstrukt-irrelevans och under-representation. I korthet går konstrukt-irrelevans ut på att saker mäts i provet som inte är relevant medan under-representation avser att en del av det som ska mätas inte finns representerat i provet. Dessa begrepp presenteras närmare i 3.1.4.

Teorierna om validitet används i diskussionen för få ytterligare kunskap om hur lärarna väljer ut de kunskaper de vill mäta med bokföringsprov. Fångar lärarna kunskaperna i en direkt mätning eller försöker de hitta indikatorer på kunskapen de vill mäta? Innehållsvaliditet kan ge ytterligare information om vilka kunskaper lärarna anser att de vill mäta med provet. Kriterierelaterad validitet kan slutligen

användas för att addera kunskap om hur lärarna resonerar kring konstruktionen av sina bokföringsprov då det gäller att mäta elevernas olika kunskapsnivåer kopplat till kunskapskraven. Detta genom att undersöka om lärarna resonerar kring om samma bedömning av provresultatet skulle fås med andra bedömningsinstrument eller om lärarna anser att provet kan säga något om hur eleverna kommer prestera i framtiden.

3.1.2 Kanes validitetsmodell

I avsnittet ovan presenterades begreppet validitet och i detta avsnitt presenteras en specifik modell, av Kane, Crooks & Cohen (1999) för bedömning av provvaliditet. Kane, Crooks & Cohen (1999, s. 5) har lanserat en modell för tolkningar av resultat på exempelvis ett prov. Denna tolkning sker i flera steg. Först observerar läraren elevens prestation. Det kan då handla om elevens svar på provfrågorna. Svaret som eleven ger på provet leder fram till ett omdöme från läraren och en koppling till kunskapskraven för kursen. I detta led gör alltså den enskilda läraren en tolkning av elevens resultat. Hur pass trovärdig tolkningen är beror på hur tydliga bedömningsinstruktionerna är och hur pass styrd bedömande lärare är av dessa. För den som konstruerar provet är det alltså viktigt att göra en bedömningsinstruktion. I nästa steg ska läraren göra en bedömning huruvida den enskilda uppgiften som eleven svarat på också kan sägas ge svar på hur eleven skulle klara liknande uppgifter. Det sker alltså en generalisering av elevens svar. Jönsson (2013, s. 61) menar att om provet innehåller flera uppgifter där samma kunskaper testas, så är det lättare för bedömaren att dra slutsatser om elevens förmåga att klara likartade uppgifter. I sista steget ska bedömaren enligt Kane, Crooks & Cohen (1999, s.5) göra en extrapolering. Det innebär att läraren ska bedöma huruvida provsvaren kan gälla i situationer som inte har testats. I fallet om bokföring innebär detta att läraren ska göra en bedömning om de svar på bokföringsprovet som eleven ger också innebär att eleven har förmåga att faktiskt bokföra. Ju närmare uppgiften är den förmåga som ska testas desto lättare är det att extrapolera (Jönsson, 2013, ss. 61-62). Det åligger provkonstruktören enligt Jönsson (ibid, s. 63) att ta hänsyn till detta när provet konstrueras. Om frågorna ligger på en alltför abstrakt nivå eller om de är för grundläggande blir det enligt Jönsson (ibid, s.63) svårare att extrapolera. Kanes validitetsmodell används i diskussions-delen för att få mer kunskap om hur lärarna resonerar kring konstruktionen av sina bokföringsprov då det gäller att mäta elevernas olika kunskapsnivåer kopplat till kunskapskraven. Hur resonerar lärarna kring generalisering och extrapolering av provresultaten för att översätta resultatet från provet till betyg enligt kunskapskraven? Modellen är dessutom intressant för att analysera hur lärarna säger att de väljer ut de kunskaper som de vill mäta eftersom ju närmare provens uppgifter ligger själva metoden bokföring, desto lättare är det för läraren att extrapolera elevens kunskap.

3.1.3 Reliabilitet

Enligt Wikström (2013, s.39) är reliabilitet ett annat centralt begrepp inom testteorin, som ger svar på hur tillförlitligt ett prov är. En hög reliabilitet innebär att eleven skulle få samma resultat på provet oavsett när detta görs, om antagandet görs att eleven innehar samma kunskaper vid de separata provtillfällena. Då handlar det om att försöka få bort de fel som inte uppkommit till följd av slumpen. Om eleven behöver chansa för att de inte förstått frågan så har provet brister i reliabiliteten (Wikström, 2013, ss. 39-40). Men det kan också handla om helt yttre

omständigheter. En elev som sovit dåligt försämrar reliabiliteten i provet som mätinstrument. Ett sätt att hantera behovet av att ta bort slumpmässiga fel är enligt Wikström (2013, s.40) att lägga till ytterligare frågor. En för liten mängd frågor riskerar att försämma reliabiliteten. Provet kan enligt Wikström (ibid, s.40) ses som ett stickprov. All kunskap som finns inom ett område kan enligt Wikström (2013, s. 40) omöjligen testas på ett prov. Läraren måste då göra ett urval av vilka delar som kan representera hela kunskapsstoffet. Men detta kan föra problem med sig genom att frågorna inte blir representativa. Ju fler frågor provet har ju större är sannolikheten att läraren kan uttala sig om elevernas kunskapsnivå. Det blir en avvägningsfråga då det faller på sin egna orimlighet att uppnå full reliabilitet genom att lägga till provfrågor i det oändliga (ibid s.41-41). Även Wells & Wollack (2003, s. 2) lyfter upp denna problematik och menar att det är omöjligt att uppnå full reliabilitet. Om samma elever testas vid två olika tillfällen med samma prov så kommer ändå resultatet vara olika, trots att de har samma kunskap vid bägge tillfällen. Yttre omständigheter som elevernas mående, ljusstyrka i klassrummet etc. kommer bidra till försämrad reliabilitet. Det tycks råda samstämmighet i att det är svårt att få en hög grad av reliabilitet. Crocker & Algina (1986, s. 105) menar att eftersom vi vet att samma prov upprepat inte ger samma resultat så måste provkonstruktören eller provanvändaren fundera över vad detta får för konsekvenser för bedömningen. Detta hanteras enligt Wikström (2013, ss. 40-41) genom att tillse att provet består av tillräckligt många frågor. Detta i sin tur för med sig problem i form av att eleverna kan bli trötta, provtiden kan bli för lång och att omfattningen av själva provet blir alldeles för stort. Teorierna kring reliabilitet kan användas för att addera kunskap kring vilka kunskaper lärarna säger att de vill mäta med sina bokföringsprov och hur lärarna säger sig välja ut de kunskaper de vill mäta. Eftersom provet är att anse som ett stickprov är det intressant att titta närmare på hur lärarna väljer ut de kunskaper de vill mäta med detta stickprov och vilken kunskap detta stickprov då mäter.

En speciell form av reliabilitet är interbedömarreliabilitet. Denna form av reliabilitet anger enligt Lundahl (2016, s. 38) i vilken grad andra lärare skulle dra samma slutsatser och göra samma bedömning av elevens svar som provkonstruktören gör. I betygssystem med färre antal steg kommer interbedömarreliabiliteten att bli sämre då små avvikelser ger större genomslag (ibid, s.38). Vid ett värde på 1 i interbedömarkoefficient kommer alla lärare att göra samma bedömning av provet. De nationella proven har högre koefficient än lärarkonstruerade. De nationella proven har en koefficient på 0,8. Det innebär att var femte elev riskerar att bedömas på en felaktig nivå enligt (Jönsson, 2013, s. 68). Jönsson menar att det finns två alternativa vägar för att förbättra interbedömarreliabiliteten. Det första är att göra provet så pass smalt och insnävat att läraren i flera olika frågor kan testa samma förmåga. Det andra alternativet är att lägga till ytterligare frågor för att försäkra sig om att varje förmåga testas ordentligt. För att förbättra koefficienten 0,05 behöver läraren fördubbla antalet frågor. Jönsson visar på att ett prov skulle behöva vara 30 timmar långt för att få en nära nog helt tillförlitlig interbedömarreliabilitet. I rapporten *Kunskapsbedömning i skolan* (Statens skolverk, 2011b, s. 35) påtalas att ett annat sätt att hantera interbedömarreliabiliteten är att läraren bedömer provet tillsammans med en kollega. Detta höjer interbedömarreliabiliteten genom att konsensus nås kring bedömningen. Interbedömarreliabilitet kan användas för att få mer information om hur lärarna resonerar kring konstruktionen av sina bokföringsprov då det gäller att mäta elevernas olika kunskapsnivåer kopplat till

kunskapskraven. Sker sambedömning eller hanteras reliabiliteten genom att snäva in provområdet?

3.1.4 Under- och överrepresentation

Under- och överrepresentation är enligt Wikström (2013, ss. 90-91) en form av validitetsbrist i provet och vid provkonstruktion är detta en av fallgroparna som konstruktören kan falla i. Om provkonstruktören ställer för få frågor och därmed inte fångar alla kunskapsområden som ska testas lider provet av underrepresentation. Men om provet istället omfattar kunskaper som ligger utanför det som ska bedömas så lider provet av överrepresentation eller konstruktirrelevans. Provkonstruktören måste enligt Wikström (ibid, s.91) därför vara noga med att definiera och avgränsa det område som ska testas. Om eleven ska testas på bokföring och frågorna omges av en stor mängd svårtolkad text kommer provet att testa även läsförmågan och då mäter provet saker som är irrelevant för området. Den som innehar goda kunskaper i bokföring men har lässvårigheter riskerar att missgynnas av ett sådant prov. Detta skulle också kunna uppstå då läraren i ett bokföringsprov lägger in provfrågor om exempelvis affärsidéer. Tengberg (2016b, s. 25) understryker att läraren behöver ta fram en konstruktdefinition, vilket innebär att läraren behöver definiera den kunskap som ska mätas så att provet inte glider utanför denna kunskap eller mäter för lite. Provkonstruktören behöver fundera över följande:

Vad är det jag vill veta, varför vill jag veta detta och hur har jag tänkt att använda den information jag får genom provet? Svaren visar syftet med provet, och syftet styr allt från provets planering och konstruktion till användning och tolkning av dess resultat (Wikström, 2013, s. 92)

Teorin kring under- och överrepresentation kan användas för att ge ytterligare kunskaper kring vilka kunskaper lärarna vill mäta med sina bokföringsprov samt hur lärarna säger sig välja ut de kunskaper de vill mäta. Begreppet konstruktdefinition adderar information om vilka kunskaper som lärarna väljer ut att mäta på provet samt hur lärarna säger sig välja ut de kunskaper de vill mäta. Huruvida lärarna resonerar kring om de på provet testar irrelevanta kunskaper eller att de testat för lite, adderar kunskap kring hur lärarna säger sig välja ut de kunskaper de vill mäta med bokföringsprovet.

3.1.5 Objektiva prov och alternativa (subjektiva) bedömningar

Frederiksen J & Allan (1989, s.28) menar att prov i stora drag kan delas in två typer. De objektiva proven använder sig av enkla räkneoperationer för att läraren med enkelhet ska kunna summera elevens poäng och därefter sätta ett omdöme. Dessa är ofta prov där eleven ska kryssa i rätt alternativ, välja rätt/fel bland olika alternativ eller svara med ett enkelt ord. Det finns då enligt Frederiksen J & Allan (1989, s. 28) inte särskilt stort utrymme för tolkningar kring vad som är ett rätt svar eller inte. Objektiva prov har i regel hög reliabilitet och denna typ av prov minskar risken för bias men fungerar inte lika bra för att mäta djupare kunskaper. Metakognitiv kunskap och problemlösning är i regel svårt att dra slutsatser om utifrån ett objektiva prov. Det påtalas också att objektiva prov ger en indirekt mätning av kunskaper eftersom det ligger längre ifrån förmågan som ska testas (ibid s.29). Vidare

konstateras här också att användandet av objektiva prov dessutom främjar lärandet av repetitiva kunskaper som är lätta att memorera.

Den andre typen av prov; alternativa bedömningar, testar andra typer av kunskaper. Till skillnad från de objektiva proven är det här analysfrågor, olika typer av bedömningar från elevens håll samt reflektioner som testas (Frederiksen J & Allan, 1989, s. 28). Det finns normalt sett två typer av alternativa bedömningar. Dels är det *performance assessment* som betyder att något ska presteras. Även om det inte är helt fastslaget vad detta begrepp innebär kan det sägas att när eleven ska utföra någonting som ligger nära kriteriet som mäts så görs detta i ett performance assessment, exempelvis bilkörning eller målning (Palm, 2001, s. 4). Jönsson (2013, s. 56) definierar performance assesment som ”Sådana bedömningar, där eleven bedöms under tiden som hen utför en uppgift”. *Authentic assessment* å sin sida betyder att något ska vara äkta – genuint, men vad detta är, är inte helt fastslaget. En tolkning är att ett prov med öppna frågor faller under denna kategori (Palm, 2001, s. 11). Men det räcker inte med att endast ha öppna frågor för att provet ska vara autentiskt. Palm (ibid s.12) beskriver ett antal kriterier som behöver vara uppfyllda. För det första behöver provfrågorna vara knutna till livet utanför skolan. Frågorna behöver alltså bygga på kognitiva förmågor som är hämtade från exempelvis yrkeslivet. För det andra behöver frågorna vara kopplade till läroplanen och till den undervisning som föregått provet. För det tredje ska frågorna dessutom understödja vidare inläring. Snarare än att testa kunskap ska de alltså stimulera eleven att lära sig mer. Denna formativa del av provet är unikt för authentic assesment. Teorierna kring objektiva prov och alternativa bedömningar kan användas för att ge ytterligare information om vilken kunskap lärarna anser att de vill mäta med bokföringsprovet. Vad är det för typ av prov lärarna konstruerar och har provtyp valts beroende på vilka kunskaper lärarna vill mäta med provet? Har detta val gjorts för att lärarna ska kunna mäta elevernas kunskaper kopplat till kunskapskraven?

3.1.6 Alignment

Med alignment menas samstämmigheten mellan läroplansmålen och provet (Näsström, 2008, s. 66). Detta är tätt knutet till authentic assesment. Linn & Herman (1997, s. iii) menar att länder som har målrelaterade utbildningsystem och målrelaterade prov skapar en tätare koppling mellan prov och vad som har lärts ut. Vidare menar de att ett prov som bygger på multiple choice eller liknande, dvs objektiva prov inte kan visa elevernas fulla kunskaper kopplade till läroplanen. Detta innebär att det för provkonstruktören är viktigt att skapa en logisk kedja mellan läroplan, vad som lärs ut i klassrummet och vad som sedan testas på provet. Då är samstämmigheten, alignment hög. Denna teori kan användas för att addera kunskap om vilka kunskaper lärarna säger sig vilja mäta med sina bokföringsprov. Är dessa kunskaper samstämmiga med läroplanen och hur resonerar lärarna kring konstruktionen av sina bokföringsprov då det gäller att mäta elevernas olika kunskapsnivåer kopplat till kunskapskraven? Finns det en samstämmighet mellan kunskapskraven och den kunskap som valts ut?

3.1.7 Konstruktionsregler

Hur ska då ett prov konstrueras för att mäta den kunskap man önskar att mäta? Wikström (2013, s. 103) sammanfattar forskningsläget genom följande konkreta råd:

- Läraren behöver specificera vad vederbörande vill veta och ställa frågor som svarar mot dessa kunskaper.
- Läraren behöver använda diskret formulerade frågor. Detta innebär att inga frågor ska vara beroende av någon annan provfråga. Det innebär också att eleven inte ska behöva svara rätt på en fråga för att kunna svara rätt på en annan fråga.
- Läraren behöver konstruera okomplicerade frågor som ej ger anledning till tveksamheter och tvetydigheter. Frågorna ska vara förenklade.

När provet är konstruerat behöver konstruktören bestämma sig för hur det ska poängsättas. Eriksson & Henriksson (2012, refererad i Wikström 2013, sid 120) ställer upp två aspekter som behöver beaktas. För det första om svaret i sig är korrekt och sedan vilken tyngd frågan har i provet. Den första frågan kan i regel inte besvaras förrän eleven skrivit provet men vilken tyngd frågan har behöver läraren ta ställning till innan eleverna skrivit det. Här påpekas också att om provet innehåller öppna frågor av lite mer komplicerad art eller om alla frågor inte är lika viktiga så behöver en viktning göras. Vidare påtalar Wikström att det är enklare att särskilja elevernas prestationer åt om mer avancerade frågor vägs tyngre. Ett sätt att hantera detta som Wikström (ibid sid 122) lyfter upp är att dela upp frågorna efter betygskategorier och sedan se hur många poäng eleven har i respektive kategori. Men detta kan samtidigt innebära att om eleven lyckas med A-frågorna men inte E-frågorna så är troligtvis kopplingen till kunskapskraven för svag. Hur lärarna resonerar kring ovanstående konstruktionsregler ger ytterligare information till studiens samtliga forskningsfrågor. Huruvida en viktning sker adderar information om hur lärarna resonerar kring konstruktionen av sina bokföringsprov då det gäller att mäta elevernas olika kunskapsnivåer kopplat till kunskapskraven. Viktas provfrågorna för att mäta dessa kunskapsnivåer? Dessutom ger konstruktionsreglerna vägledning för att få mer information om vilka kunskaper lärarna säger sig vilja mäta samt hur de bestämmer vilka dessa kunskaper är.

3.2 Tidigare forskning

Här lyfts några av de studier som kan kopplas till den teoretiska ram som ovan har presenterats fram. En genomgång av tidigare forskning har visat att det inte finns några studier på hur lärare i företagsekonomi resonerar när de konstruerar prov. Däremot finns motsvarande forskning knuten till andra ämnesområden. Denna forskning presenteras nedan.

I en doktorsavhandling av Jesper Boesen (2006, s. 12) undersöks vilken typ av matematisk förmåga som behövs för att lösa frågor i skriftliga prov. Här konstateras att i merparten av de prov som lärare själva konstruerar så testas endast repetitiv kunskap och går inte in på de djupare förståelsenivåerna. Boesen (ibid, s.12) menar att detta framförallt beror på tre faktorer: för det första att lärare har svårt att bedöma skillnader i förmåga hos eleverna att resonera. För det andra att lärarna har låga

förväntningar på vad eleverna egentligen klarar av. För det tredje konstaterar Boesen (ibid, s.12) att lärarna vill att så många elever som möjligt ska klara proven och att detta blir enklare om frågor som handlar om resonemang utesluts.

Liknande slutsatser återkommer i andra studier. Christina Odenstad (2010, ss. 174-175) har i sin doktorsavhandling undersökt hur prov i samhällskunskap är konstruerade och vilken typ av kunskap som eleverna testas på. Slutsatserna som Odenstad drar är att detta är beroende av vilket arbetsområde provet testar. I vissa områden är det mest fokus på faktakunskaper medan andra arbetsområden ger mer utrymme för förståelsefrågor. Vidare konstaterar Odenstad att på studieförberedande program ligger tyngdpunkten på förståelsefrågor medan elever på yrkesprogram i större utsträckning får prov med större andel faktafrågor istället.

Men det finns också studier som visar på motsatsen. Tobias Jansson (2011, s. 150) har i sin licentiatuppsats *–Vad kommer på provet* tittat närmare på viken typ av kunskap som lärare testar i prov. Bland de prov som Tobias undersökt hittar han såväl provfrågor som testar faktakunskaper som mer kognitivt avancerade frågor även om tyngdpunkten på frågorna i de undersökta proven ligger på faktafrågor. Men även de lärare som främst använder sig av faktafrågor använder sig i någon mån av mer kognitivt avancerade frågor. Jansson menar samtidigt att proven främst är summativa till sin karaktär eftersom lärarna önskar använda provresultaten som underlag när betyg sätts. Slutligen noterar Jansson (ibid, s.152) att många prov lider av validitetsbrist men att detta beror på tidsbrist hos provkonstruktören.

I sin avhandling *Att hantera historia med ett öga stängt* undersöker David Rosenlund (2011, ss. 164-165) i vilken mån det finns en samstämmighet mellan kursplaner och lärarnas uppgifter och prov i kursen Historia A. Det ska här tilläggas att denna studie bedrivits gentemot de gamla läroplanerna. Undersökningen gav vid handen att samstämmigheten mellan kursplaner och prov/uppgifter ofta är låg. Rosenlund drar det så långt att han menar att samstämmigheten är på en oacceptabel nivå och att det inneburit att proven testar kunskaper på en lägre kognitiv nivå än vad kursplanen kräver. Rosenlund konstaterar att det föreligger en underrepresentation i många frågor och ger följande exempel:

Lärarna efterfrågar till exempel en jämförelse, men då jämförelsen redan är formulerad i boken är det egentligen elevernas förmåga att minnas som bedöms. Då dessa frågor har kategoriserats i den kognitiva processen minnas i stället för, till exempel, förstå, leder det till att graden av samstämmighet påverkats i negativ riktning (2011, s. 169).

Med hjälp av dessa tidigare studier kan ytterligare kunskap fås om denna studies forskningsfrågor. Denna tidigare forskning kommer att jämföras med resultaten från denna studie i arbetets diskussionsdel.

4. Metod

I detta kapitel beskrivs de metoder som valts och vilka överväganden som lett fram till valen. Här beskrivs också problem som uppstått och vägval som behövts göras. Först beskrivs metoden för undersökningen. Därefter beskrivs det urval som gjorts. Avsnittet fortsätter sedan med att förklara hur materialet samlats in och hur själva intervjuerna gått till. Efter det beskrivs hur analysen skett av det material som framkommit genom intervjuerna. Avslutningsvis framläggs de etiska överväganden som studien har fört med sig samt en metoddiskussion.

4.1 Val av metod för undersökningen

Intervjun är upplagd som en semi-strukturerad intervju. En semi-strukturerad intervju innebär att varje fråga inte är specificerad utan istället utgår från teman som sedan följs av följdfrågor (Johannessen & Tuft, 2002, s. 99). Lundahl & Skärvad (1999, s. 102) påpekar att vid själva intervjun är det viktigt att inte vara alltför styrd av på förhand formulerade frågor. Det innebär att intervjuaren behöver vara beredd på att komma med följdfrågor och att lyssna in intervjuobjektets utgångspunkter för att kunna föra samtalet framåt. Undersökningen sker genom kvalitativ metod. Enligt Lundahl & Skärvad (1999, s. 101) används denna metod när man vill undersöka människor och deras syn på någonting och få en bild av hur dessa resonerar och varför de gör som de gör. Vidare påtalas att denna metod inte används för att dra slutsatser om hur någonting förhåller sig utan om hur människor uppfattar sig själva eller uppfattar något fenomen. Detta faller väl inom undersökningen eftersom studien undersöker hur lärarna resonerar och vad de anser är viktigt vid provkonstruktion. Valet av metod har alltså utgått ifrån denna studies syfte och dess forskningsfrågor. Samtalsintervju valdes som metod. Om man vill veta hur människor uppfattar någonting är det enligt Esaiasson et al (2012, s. 253) lämpligt att använda sig av samtalsintervjuer. Denna studie går på djupet i frågeställningarna för att verkligen förstå hur lärarna resonerar. Vidare är samtalsintervju enligt Esaiasson et al. (2012, s. 250) lämpligt då man vill kunna ta in det icke förväntade och då kunna ställa följdfrågor. Denna typ av undersökning innebär att det inte går att göra en generalisering ur ett statistiskt perspektiv. Det går inte på grundval av de lärare som intervjuats påstå att detta skulle gälla för en viss andel av lärarkåren. Istället används då begreppet analytisk generalisering vilket innebär att resultaten kan användas för att utveckla teorier och ställa resultatet gentemot tidigare studier och den teori som undersökningen vilar på (Lundahl & Skärvad, 1999, s. 195).

4.2 Urval

Denna studie använder sig av ett bekvämlighetsurval. Esaiasson et al (2012, ss. 188-189) menar att detta ofta sker när det är svårt att hitta något sätt att göra ett urval ur den totala populationen. Den totala urvalsramen i detta fallet skulle vara alla behöriga företagsekonomilärare i Sverige som undervisar i Företagsekonomi 2 och som dessutom har prov i bokföring. Det faller på sin egna orimlighet att hitta denna population. Detta kan anses vara ett första-bästa-urval. Detta eftersom tiden för undersökningen är begränsad till 10 veckor och att det bedömdes som viktigt att träffa lärarna på plats för intervju. Bekvämlighetsurvalet syftade till att urvalet gjordes på ett begränsat område i södra Sverige. Urvalet var strategiskt så till vida att endast lärare som undervisar i företagsekonomi 2 och ger prov i bokföring kunde

väljas och vara en del av urvalet. Totalt intervjuas 4 lärare i studien. Följande lärare ingår i studien:

Eva är lärare på en gymnasieskola i en mellanstor kommun.

Eva har varit lärare i mer än 20 år och undervisar i företagsekonomi 1 och 2, redovisning 2 samt juridikämnen.

Roger undervisar på en gymnasieskola i en stor kommun och har undervisat i 8 år. Roger undervisar i företagsekonomi 1 och 2, företagsekonomi specialisering samt samhällskunskap.

Hans undervisar på en gymnasieskola i en annan stor kommun. Hans har varit lärare i 10 år och undervisar i företagsekonomi 1 och 2, företagsekonomi specialisering samt entreprenörskap och företagande.

Karin är lärare i en mellanstor kommun och har varit lärare i över 20 år. Karin undervisar i företagsekonomi 1 och 2, marknadsföring samt entreprenörskap.

4.3 Genomförandet

Av intervjuerna har tre genomförts fysiskt och en via telefon. Något enstaka svar har kompletterats via mail. Intervjuerna har genomförts på respektive lärares arbetsplats, detta för att få intervjuobjektet så bekvämt som möjligt i situationen. Intervjuerna genomfördes semi-strukturerat som ovan beskrivits.

Frågorna i intervjun är konstruerade utifrån syfte/frågeställningar och har med handledares hjälp bearbetats så att dessa inte går utanför syfte/frågeställningar. Frågorna är också konstruerade efter teorier och tidigare forskning. Detta för att få en tydlig teoretisk förankring.

Samtliga intervjuer har spelats in och transkriberats i ordbehandlare. Efter bearbetningen av texten skrevs alla transkriptioner ut och nyckelord, teman och kategorier markerades. Det bearbetade materialet presenteras i detta arbets resultatdel, sorterat efter frågeställningarnas tema. Svaren presenteras utifrån vad som särskiljer/förenar intervjuobjekten. Om något fall sticker ut eller något oväntat kommer fram belyses detta.

4.4 Analys av empiri

Analysen av materialet har huvudsakligen skett utifrån forskningsfrågorna och det teoretiska ramverket samt tidigare forskning. Respektive del av det teoretiska ramverket samt den tidigare forskningen har ställts upp var för sig och analys sker för varje del. Svaren från lärarna knyts till teoriavsnittet och analys sker utifrån detta. Här återknyts dessutom analysen till syfte/frågeställningar.

Huvudsakligen används kartläggningsmetoden som Esaiasson et al (2012, s.272) presenterar även om resultatet inte presenteras på det viset. Det innebär att man försöker kartlägga alla aspekter av ett fenomen. Varje del av fenomenet undersöks för sig och börjar exempelvis med om lärarna sätter betyg på provet för att sedan gå

vidare och titta på de lärare som sätter betyg och de lärare som inte sätter betyg på provet. I detta fall är fenomenet lärarnas resonemang vid provkonstruktion.

4.5 Etiska överväganden

Vetenskapsrådet (2002, ss. 5-12) betonar vikten av att ta hänsyn till de forskningsetiska principerna. Informationskravet uppfylldes genom att informationsbrev lämnades ut i samband med själva intervjun där syftet med uppsatsen förklarades samt vad det är för typ av undersökning. Då informerades lärarna dessutom om att deltagandet är helt frivilligt och att deltagarna så långt det går kommer att avidentifieras. I informationsbrevet upplystes dessutom om att rådata i form av ljudfil på diktafon samt transkriberat material kommer finnas inläst. Uppgifterna från intervjuerna kommer inte användas till något annat än den här uppsatsen så då har även nyttjandekravet uppfyllts. Möjlighet för lärarna har givits att få information om när uppsatsen beräknas vara klar för publicering så att de kan ta del av resultaten.

4.6 Metoddiskussion

Det är värt att ta i beaktande att deltagande i intervjun kan vara känsligt för lärarna. Det kan upplevas som att något lärarna konstruerat ska värderas och kritiskt granskas. Det påtalades för lärarna att någon sådan värdering inte sker, men risk för att lärarna ändå ger de svar som de tror att författaren förväntar sig föreligger och beaktades vid analysen av materialet.

Enligt Esaiasson et al (2012, s. 267) är intervjuareffekt någonting som måste tas i beaktande. Detta innebär att intervjuaren medvetet eller omedvetet påverkar intervjuobjektet. Medveten påverkan är av det mer allvarigare slaget och innebär att intervjuaren ställer ledande frågor eller försöker styra svaret i en viss riktning. Omedveten påverkan har minimerats genom att lyssna intresserat och undvika noteringar under mötet. Intervjuobjektet får också tala till punkt. Hur mimik och kroppspåk uppfattas, är svårt att kontrollera men dessa faktorer har åtminstone tagits i beaktande under intervjuerna.

Av ca 20 kontaktade lärare uttryckte tre sitt intresse av att delta i undersökningen. Med hjälp av handledare fångades sedermera ytterligare en lärare upp som via telefon ställde upp på intervju. Antalet kan uppfattas som litet men samtidigt ges då större möjligheter att gå på djupet i dessa intervjuer. Urvalet föll sig sådant att det där finns lärare som arbetat under både den gamla och nya läroplanen, kvinnor och män. Det blev även en bra spridning i ålder. Ingen tidigare relation finns till någon av de intervjuade lärarna. Lärare i Företagsekonomi i södra Sverige kontaktades för att tillfrågas om de ville vara en del av denna undersökning. Det visade sig sedermera vara svårt att få tag i lärare i Företagsekonomi som ville delta i undersökningen. Det finns ofta inte särskilt många lärare i företagsekonomi i varje kommun. Lärare från författaren av denna uppsats egen kommun valdes bort då detta är lärare som författaren känner sedan tidigare. Möjligheten att göra ett sannolikhetsurval fanns inte på grund av tidsramen för arbetet.

5.Resultat

Här presenteras det bearbetade materialet från intervjuerna. I avsnitt 5.1 görs en snabb överblick över lärarnas provkonstruktion genom att undersöka om lärarna använder provet som huvudsaklig bedömningsform samt om lärarna känner att de behöver komplettera med ytterligare information för att kunna sätta betyg. Denna del är intressant om provkonstruktionen påverkas av att vissa delar av området testas med andra bedömningsformer. Här undersöks även vad det är för typ utav prov. Avsnitt 5.2 redogör för lärarnas svar på frågor kopplade till studiens forskningsfråga om vilka kunskaper lärarna önskar mäta med sina bokföringsprov samt forskningsfrågan hur lärarna säger sig välja ut de kunskaper de vill mäta med bokföringsprov. Kapitel 5.3 används i huvudsak för att redogöra för lärarnas svar på frågor kopplade till studiens fråga om hur lärare resonerar kring konstruktion av bokföringsprov då det gäller att mäta elevernas olika kunskapsnivåer kopplat till kunskapskraven. Här delges även vad lärarna tycker att de får för vägledning av Skolverket för att kunna konstruera sina prov samt vilka svårigheter de ser med att konstruera prov för bedömning.

5.1 Lärarnas provkonstruktion

Samtliga lärare i studien säger sig använda prov i bokföring som den huvudsakliga bedömningsformen även om andra underlag kan ligga till grund för bedömningen av detta arbetsområde. Förutom provet används observationer i klassrummet samt elevernas förmåga att hantera bokföringsprogram. Eva förklarar detta:

Jag har alltid prov i bokföring i företagsekonomi 2. Men jag brukar även plocka in att de jobbar med något bokföringsprogram, alltså att jag ser att de klarar av att bokföra även i ett bokföringsprogram. Där har jag inte hitintills haft något prov [...] men alla har fått jobba på och jag har bedömt att de klarar av det. Jag tar in deras resultat och balansräkning för att se vad de gör rätt och fel. (Eva)

Eva säger att hon sätter betyg på arbetsområdet och frågan ställdes om Eva behöver komplettera med ytterligare underlag för att kunna sätta detta betyg:

Det är provet som styr [...] Vi har någon gång haft att de gjort intervjuer på företag och lämnat in en rapport om hur de gör i verkligheten, jämfört med i teorin. Då har de haft muntlig redovisning på det. Det kan alltså vara lite olika. Det är ett bra sätt för att man då kan få med mer resonemang och nyansering. Det är en bra komplettering. (Eva)

Karin berättar att observationer i det dagliga arbetet tillsammans med provet utgör det huvudsakliga underlaget för bedömning men att det ibland kan vara aktuellt med ytterligare underlag för bedömning. Karins förklarar att hennes prov skrivs på dator av eleverna men om hon tycker att hon behöver ytterligare underlag får de ibland göra uppgifter även på papper. Detta kan bli aktuellt om Karin misstänker att de fuskat på datorprovet.

Roger förklarar hur han resonerar kring provet som bedömningsunderlag:

Alltså, man jobbar ju mycket på lektionerna och jag tänker att man testas av förståelse, begrepp och sådana saker i det dagliga, även om det till viss del kommer på provet. Har man förstått liksom vad arbetsgivaravgift är för

begrepp, då vet man hur man ska bokföra det. Annars är det prov jag testat av med. (Roger)

Liknande resonemang har Hans som också han menar att det i huvudsak är provet som utgör bedömningsunderlaget men att han också tar hänsyn till hur de arbetar på lektionerna, hur de ställer frågor och hur de klarar av att arbeta med olika bokföringsprogram. Hans, Eva och Karin säger att de sätter ett betyg på provet medan Roger resonerar så här:

Jag ger ett omdöme. Betyg i formell mening sätts ju bara i slutet av terminen [...] men de får ett omdöme och jag säger med ”viss säkerhet”, ”säkerhet” och så vidare och då förstår de ju att exempelvis ”med säkerhet” är ett A. (Roger)

Karin berättar att hon lämnar ut de rätta svaren till eleverna där hon skriver individuellt till varje elev hur de lyckats i förhållande till kunskapskraven. Varje elev får en text med det de behöver öva mer på. Dessa utdrag ur kunskapskraven har eleverna sett redan innan provet genom att de får inlärningsinstruktioner:

Då så är det så här att jag har satt upp inför varje prov inlärningsinstruktioner som innefattar de här kunskapsmålen och hur man ska svara på det. Vad man ska kunna. Sen så får de ett prov som innehåller det. (Karin)

5.2 Innehåll på proven

Vad är det för typ av prov som de olika lärarna i undersökningen konstruerar, vilka kunskaper säger de att de vill mäta och hur säger de sig välja ut dessa kunskaper? Samtliga lärare uppger att provet består av verifikationer med tillhörande text som de ska bokföra. Karin särskiljer sig genom att berätta att hennes prov helt och hållet skrivs på dator i bokföringsprogram. Eleverna får varje uppgift som en verifikation som de sedan ska bokföra. Eva skiljer ut sig på annat sätt genom att berätta att hon också lägger in begreppskunskap och förståelsefrågor, vilket de andra lärarna inte gör:

Så mycket det går gör vi det på de olika proven. Sen är inte begreppen så jättemånga på något vis i bokföring. Det är inte så många begrepp. Förståelse [...], varför är inte moms en kostnad för företaget? Det är ju någon form av förståelse. Inte bara bokföra utan också att förstå. (Eva)

Eva uttrycker svårigheter med att testa hela betygsskalan på resonemangsuppgifterna där hon tolkar kunskapskraven som att eleverna ska kunna jämföra företeelser och sätta dessa i ett sammanhang. Det är enligt Eva mycket svårt att konstruera uppgifter för detta. Ofta väljer Eva att lyfta ur detta ur bokföringsprovet för att bedöma den delen med andra uppgifter istället då det är svårt att få till dessa uppgifter i en provsituation. Även Roger säger sig finna kunskaper som han skulle vilja mäta i provet men väljer att inte testa. För Rogers del handlar det om lagstiftning kring bokföring. Anledningen till att Roger lyfter ut detta ur sitt prov är enligt honom tidsskäl och att han vill ha så lite som möjligt att behöva bedöma eftersom han anser att han har en stor mängd elever. Hans tycker till skillnad från de andra att de testat just det han vill. Men Hans lyfter att han under tidigare läroplan kunde ta med kunskaper som han själv tycker är viktiga men att nuvarande läroplan hindrar honom från det. Han kan dock inte precisera vad detta

skulle kunna vara för typ av uppgifter. Men han uppger att han tycker att Skolverket styr för mycket över vilka kunskaper som är viktiga:

Vad det gäller vilka kunskaper de intervjuade lärarna väljer ut för att testa elevernas kunskaper på bokföring är de allihop styrda av läroböckerna. Eva använder *e3000*, Hans *e2000* samt eget konstruerat material och Karin använder *Take off*. Roger använder *Goodwill*. Lärarna säger sig testa eleverna på de bokföringsoperationer som anges i böckerna men Hans uppger att de för en diskussion i arbetslaget kring vilka av dessa begrepp som är viktigast att kunna. Karin betonar att även om hon använder läroböckerna som grund för vilka olika bokföringshändelser som eleverna ska klara så stämmer hon av dessa med en redovisningsbyrå för att se om också de tycker att kunskaperna är relevanta. Hon berättar att hon tidigare bedrivit eget företag och tycker sig därför själv ha en god uppfattning om relevant kunskap som ska testas. Inga av lärarna uppger att de underlåter att testa någon typ av affärshändelse som tas upp i läroböckerna. Eva berättar att de på hennes skola även köpt in andra böcker för att få variation och inspiration till uppgifter. Eva beskriver de kunskaper hon vill mäta så här:

För mig handlar det mycket om att kunna bokföra och vara säker på de olika momenten i bokföring. Sen styrs man nog tyvärr ganska många gånger av det som finns i läroboken [...] men i Företagsekonomi 2 bokför vi alltid med moms, momsredovisning, löner, lite svårigheter som till exempel kundförluster. (Eva)

Samtidigt menar Eva att det kan vara en fördel att använda läromedlet:

På något vis kan jag tycka att det är skönt att någon har bestämt var ligger nivån på detta för jag tänker att då kan vi hålla oss inom det häradet, nu kan vi i arbetslaget vara överens om innehållet även om det satts av läromedelsförfattarna. (Eva)

Hans berättar att även han till stor del använder läromedlet för att välja kunskaper som han säger sig vilja mäta men också utifrån vad han och hans kollegor uppfattar som viktigt på bokföringsområdet. Men det faller tillbaka på litteraturen även det:

Vi utgår alltid från någon slags litteratur ändå. Sen skulle jag vilja påstå att all litteratur påminner om varandra, så jag vet inte om de härmar varandra eller hur det fungerar. Men jag tycker att jag känner igen innehållet. (Hans)

Hans menar att det är till stor del litteraturen som styr vilka kunskaper som ska mätas. Roger menar att deras kunskaper som ska testas kommer ifrån vad de i arbetslaget tycker att eleverna ska ha med sig för att klara nästa kurs, redovisning. Men han konstaterar att han får god vägledning av litteraturen för detta. Hans, Roger och Eva uppger att deras prov konstrueras tillsammans i arbetslaget och är ingenting som görs av en enskild lärare. De uppger samtliga att syftet är att öka rättssäkerheten gentemot eleverna. Karin säger sig däremot konstruera sina prov helt själv men eftersom hon arbetar i ett arbetslag med flera ekonomilärare har de god kännedom om varandras prov ändå:

Vi har jobbat ihop ett tag, Vi tittar på varandras prov och ser ungefär vilken nivå vi ligger på. Vi tittar på varandras prov för att se hur jag ligger till jämfört med han eller henne. (Karin)

5.3 Konstruktion av prov för betygsättning och bedömning

Lärarna uppger att Skolverket inte ger någon vägledning i hur eleverna ska bedömas på prov. Eva menar att det hade varit bra om även ämnet företagsekonomi hade haft nationella prov. Detta för att hon på så vis hade fått en uppfattning om vad Skolverket tycker att eleverna ska kunna och vad som motsvarar vilken nivå. Alla lärare i undersökningen uppger att de har svårt att veta exakt vilka kunskaper som är hänförliga till respektive betygsnivå och hur gränsdragning ska ske. Detta ges följande uttryck:

Svårigheten är ju bedömningen. Att konstruera provet är ju bara att göra verifikationer på de moment som vi bestämmer att vi ska ha. Det svåra är i så fall att bedöma det här lite luddiga Skolverket säger. Var är gränsen för att göra det ”med säkerhet” eller inte? (Hans)

På frågan om det är enkelt eller svårt att översätta dessa kriterier från Skolverket till provfrågor svarar Hans vidare:

Det enda verktyget vi har för detta är sambedömning, samsyn. De beprövade erfarenheterna som sätter gränserna. (Hans)

Även Eva och Roger är inne på svårigheterna i detta:

Det är jättesvårt att veta. Pratar man med en lärare i ämnen som har nationella prov så kan de säga vad Skolverket vill. Det är detta de vill åt. Detta är deras krav för A-nivå och så vidare [...] Medan vi någonstans famlar, vi vet inte. Det är våra tolkningar som påverkar. Särskilt mycket strukturer från Skolverket får vi inte. [...] Vilka är de grundläggande begreppen? (Eva)

Man står helt ensam i provkonstruktionen. Jag är glad att jag har kollegor att resonera med. Verkligheten för ekonomilärare är nog att man ofta är ensam på sin skola[...] Det är oroväckande att Skolverket förlitar sig på förnuft och att man ska läsa mellan rader. (Roger)

Roger uppger att det inte finns någon katalog på området som beskriver vilka begrepp och kunskaper eleverna ska kunna. Han förlitar sig på läroböckerna. På frågan om han kopplat till kunskapskraven vet vilka kunskaper han ska testa på provet säger han:

Absolut inte. Men det gäller i andra ämnen också[...] Man förlitar sig i hög grad på litteraturen. (Roger)

Karin lyfter aspekten att det är upp till varje skola att göra en tolkning av kunskapskraven och vad som ska provas. I hennes fall har hon valt att stämma av dessa kunskaper med en redovisningsbyrå för att fånga deras uppfattning av vad en elev bör kunna för de olika nivåerna. Roger skiljer ut sig i provkonstruktionen på ett sätt. Han säger att han testat samtliga elever med ett E-prov med mycket grundläggande uppgifter som är konstruerade för att vara enkelt att klara av för eleverna. Senaste gången detta testades klarade 25 elever provet och en elev blev underkänd. De elever som siktar mot högre betyg än E får göra ytterligare ett prov för att visa sina kunskaper. Frågan ställdes då om samma typ av frågor används på bägge prov:

Jag skulle nog ända säga till stor del. Säkert 70 % eller något sådant. Samma prov men att det kommer till svårare saker för de högre nivåerna. (Roger)

På frågan om proven är indelat i A, C och E-frågor skiljer sig lärarnas svar något åt. De skiljer sig även åt avseende om de anser att detta är någonting önskvärt. Hans säger att han har delat in uppgifterna men skriver inte ut detta till eleverna. Hans bedömning är att elever inte försöker tillräckligt hårt om frågorna är indelade efter svårighetsgrad. Han vill att alla elever ska försöka så gott det går och att vissa elever tror att de inte klarar av exempelvis A-uppgifter eftersom de siktar på C i betyg. Roger berättar, som tidigare nämnts att han använder olika prov för E och för C-A. I provet för C-A ligger A-frågorna sist men eleverna kan inte se om en fråga är på C eller A-nivå. Bägge typer av prov bedömer Roger med poäng. Eva säger att hon har nivåindelade frågor men att hon inte hittills har skrivit ut detta gentemot eleverna. Hon använder övningsuppgifterna i böckerna som mall när de olika frågorna ska konstrueras. Eva berättar att dagen före intervjun har arbetslaget tagit beslut om att provfrågorna ska skrivas ut som E, C, A-frågor gentemot eleverna. Karin säger sig dela in provet efter om det är en E, C eller A-fråga och hennes elever får veta vilken nivå respektive fråga är på. Eleverna har före provet fått ut inlärningsinstruktioner där det framgår hur de olika nivåerna ser ut. Alla lärare i undersökningen säger sig i nuläget bedöma provet som en helhet och inte varje fråga för sig. Roger berättar att när han testar flera kunskapskrav på ett prov lägger han dock frågorna sorterade efter kunskapskrav och bedömningen gentemot elev är också utefter respektive kunskapskrav. Det kan resultera i att en elev får flera betyg på samma prov. Eva, Karin och Hans testar samtliga nivåer från E till A på samma prov. Hans säger att han har konstruerat provet på sådant vis att varje fråga ger ett rätt och att eleverna sedan får avdrag för varje fel. Men han säger att han bedömer varje fråga för sig när han sätter betyget sedan:

Det är inte så att 12 poäng är C och 13 A. [...] Det är mer min bedömning av det här provet. Då kan det , i alla fall i teorin bli så att någon som fått 12 poäng har A och en annan som har 13 poäng får C.[...] Jag försöker bedöma att om det är "med säkerhet". Kan de hantera bokföring på ett säkert sätt så har de ett A. (Hans)

Eva säger att hon också sätter betyg efter poängen. Hon menar att om eleven har många poäng så visar de också kunskaper som motsvarar respektive betygsnivå. Hon säger att hon är noga med att provet är konstruerat så att eleverna måste klara av även de svårare uppgifterna för att kunna få A. Det räcker inte med att klara av samtliga E-frågor:

Vi tycker att det är komplexiteten som gör att vi måste se till att det finns frågor i själva provet så att det motsvarar A. Gör vi bara enkla problem så kanske 28 av 30 ändå bara motsvarar E. (Eva)

Karin berättar att hennes prov inte är poängsatt men att provet är indelat i olika moment som hon sedan ger ett omdöme på. Det är hennes bedömning av helheten av detta som i slutänden resulterar i betyget/omdömet på provet.

Roger uppger vidare att det inte räcker med att enbart lyckas bra på de svårare frågorna för att få A i betyg:

Det kräver att de klarat de kriterierna för de underliggande nivåerna. Det går inte att klara de tre sista frågorna men inte de tidigare. Då är man inte säker. (Roger)

Roger gör en annorlunda tolkning av kunskapskraven jämfört med de andra lärarna. För att bedöma vilken betygsnivå eleverna når upp till berättar Roger att han även lagt in en tidsaspekt. Han menar att provet är konstruerat för att klaras av på ca 70 minuter. Om eleverna behöver ytterligare tid för att bli klara men ändå får full pott på provet är det inte möjligt att få A i betyg. Ingen av de övriga lärarna uppger att de lägger in en tidsaspekt i bedömningen

[...]För att i min bedömning menas ”med säkerhet” att även om man gör rätt men det tar en och en halv timme så bedömer jag dem inte vara säkra. (Roger)

Eva och Karin är de som tydligast säger sig vikta sina prov. Eva säger att hon sätter poäng efter om uppgiften är en E, C eller A-fråga. Karin berättar att även hon viktar på samma sätt.

Om du har en vanlig kontering till exempel: bokför en transaktion mellan kassa och bank, eller om du räknar ut varukostnaden och bokför varulager och rätt varukostnad. Det är klart att det är olika tyngd i dem. (Karin)

Hans berättar att han också har en typ av viktning men inte i poängsättningen. Ett rätt svar ger ett poäng men vid bedömningen av poängen kan ett poäng på en uppgift vara mer värd än ett poäng på en annan uppgift. Roger uppger att han viktar men inte så strukturerat. Om en elev gör fel på en enklare uppgift ser han det som allvarligare än fel svar på en svårare uppgift men att det ändå är mängden rätt som i slutänden avgör vad betyget blir. Han förväntar sig att en vad han kallar ”A-elev” ska klara 18 av 20 uppgifter.

Gällande mellanbetygen D och B resonerar lärarna på olika sätt. Roger säger att han inte sätter ut mellanbetyg på prov utan endast vid betygssamtal och när betyget sätts vid avslutad kurs. Hans säger att inte heller han sätter ut mellanbetyg på provet utan enbart E, C och A. Karin och Eva berättar att de sätter ut mellanbetyg. Men det sker inte utifrån att eleverna ska klara alla uppgifter på betygsnivån under och till övervägande del för betygsnivån över. Eva säger att det helt är poängen som styr.

Det är lite så att vi har en nivå. Så här många poäng behöver du för att få E, C, A, D och B blir något mellanting. [...]. Rättningsmallen styr var man hamnar. (Eva)

Karin säger att hon sätter D och B i betyg på bokföringsprovet. I inlärningsinstruktionerna till eleverna framgår bara kunskaperna för E, C och A men vid bedömning av provet används mellanbetygen. Karin berättar att det rör sig om att elever exempelvis inte klarat av en uppgift för C men lyckats med svårare uppgifter. En sådan elev kan få D i betyg.

Roger redogjorde för de för- och nackdelar han ser med sitt bedömningssystem. Eftersom han väljer att enbart testa ett område per prov, dvs bokföring som ett prov och bokslut som ett annat så menar han att han får högre betyg på eleverna. Rogers bedömning är att provresultaten och betygen hade varit lägre om han exempelvis

hade testat bokföring och bokslut i samma prov. Samtidigt säger han att han tror att eleverna inte får djupinläring när de högintensivt ska plugga till ett prov på ett enskilt område. Eva säger sig se en nackdel i att översätta ett visst antal rätt till en viss betygsnivå men hon konstaterar samtidigt att det är lättare för eleverna att förhålla sig till och att Skolverket inte ger något stöd i hur hon annars skulle konstruera provet. En fördel som hon ser med sitt bedömningssystem är att de konstruerar provet i arbetslaget och därmed alltid har diskussioner kring hur Skolverkets kunskapskrav ska tolkas. Eftersom de i Evas arbetslag är fyra lärare blir rättssäkerheten större för eleverna. Nackdelen som Hans ser med sitt system är att elever ibland hävdar att de har en dålig dag och att de hade vissa kunskaper på morgonen men inte fick fram detta i provet.

Jag kan ha upplevelsen att det här kan de, av min observation. Sen underpresterar de på proven. Anledning okänd för min del. (Hans)

Kan kunskaperna generaliseras och extrapoleras? Dvs. kan läraren dra slutsatsen att eleverna har en viss förmåga enligt kunskapskraven efter hur de presterar på bokföringsprovet? Hans menar att han kan dra den slutsatsen, åtminstone ur ett gymnasieperspektiv. Han menar att han testat dem i bokföringsprogram där de ska bokföra på riktigt och att de som lyckas väl på provet också lyckas där. Eva konstaterar att hon skulle vilja ha mer verklighetsnära uppgifter för att kunna dra den slutsatsen. Hon säger att hon skulle vilja se frågor där eleverna får se en riktig verifikation som de ska tolka och bokföra. Hon uppger att hon av tidsskäl inte har haft möjlighet att göra så idag. Roger säger att han tycker att det går att dra slutsatser kring hur bra eleverna är på metoden bokföring eftersom hans prov ser ut som vid kontering i verkligheten. Karin berättar att hon stämmer av med redovisningsbyrå för att se om provet säger något om relevant kunskap i bokföring och eftersom provet görs direkt i bokföringsprogram kan hon säga hur deras förmåga i bokföring ser ut men tillägger att det är ur ett gymnasieperspektiv.

Hans lyfte att han anser att provet inte ska testa saker som Skolverket inte kräver. Hans säger att han styrs av kunskapskraven när han konstruerar sitt prov. Detta för att inte testa för kursen irrelevanta kunskaper. Även Eva lyfter upp denna aspekt.

Det blir annars lite mer svenska och läsförståelse. Det är mycket information, ungefär som ett läsetal i matte. (Eva)

Eva och Hans berättar att de även använder sig av sambedömning när bokföringsproven rättas. Karin berättar att hon inte gör så eftersom hon menar att prov i just bokföring handlar om rätt eller fel. Hans säger att han lämnar prov som han känner sig osäker på till någon kollega för en andra bedömning. Eva berättar att hennes prov alltid rättas av fyra lärare som rättar några frågor var och sedan kommer fram till en gemensam bedömning. På det viset menar hon att rättssäkerhet ökar för eleverna eftersom samma prov bedöms av flera olika lärare.

6. Diskussion

I detta avsnitt analyseras studiens resultat mot frågeställningarna utifrån de begrepp som presenterades i det teoretiska ramverket. Dessa begrepp är validitet, Kanes validitetsmodell, reliabilitet, över- och underrepresentation, objektiva prov och alternativa bedömningar, alignment samt konstruktionsregler. I sista delen av avsnittet jämförs resultatet med tidigare forskning. I varje del av analysen framgår det dessutom till vilka forskningsfrågor ytterligare kunskaper har tillförts.

6.1 Validitet

Som Wikström (2013, s. 44) påtalar utgår begreppet validitet historiskt från att provet mäter det som det är avsett att mäta men att man numera istället fokuserar på validiteten av resultatet som eleverna får. Wikström menar (ibid, s.44) att validitet kan testas genom en direkt observation av elevens kunskaper, tex genom att låta eleven utföra någonting där förmågan visas. Detta skiljer sig ifrån indirekt mätning där en omväg behöver tas för att hitta indikatorer på den kunskap som ska mätas. Min tolkning är att om provet innehåller bokföringsoperationer så observeras förmågan att bokföra direkt. Jag menar vidare att om något mer abstrakt såsom begrepp, teorier och modeller ska testas så behöver lärarna hitta indikatorer för att mäta denna kunskap. Lärarnas beskrivning av deras prov visar att de samtliga, i varierande grad, prövar förmågan att bokföra genom att eleverna antingen konterar på papper eller bokför i bokföringsprogram. Närmast bedömer jag att Karin kommer, som ju låter eleverna göra sitt prov helt och hållet i ett bokföringsprogram. Hans och Roger använder sig helt av affärshändelser som ska bokföras men eleverna får bokföra i pappersform istället för på dator. Jag bedömer att lärarna i huvudsak väljer ut kunskaper de vill mäta genom direkt mätning. Eva säger att även hon har verifikationer på provet men det är endast en del av provet då hon även testar begreppskunskap och dessutom har med förståelsefrågor. Eva tolkar förmågorna i kunskapskraven som att eleverna inte bara ska kunna bokföra utan också få en djupare förståelse. Min bedömning är att detta hänger ihop med att Eva tolkar hela kunskapskravet medan de andra lärarna delar upp och endast testar den del av kunskapskravet som belyser företagsekonomiska metoder. Eva belyser ju också att de andra delarna är svåra att fånga i en provsituation. Tittar man på hela kunskapskravet så står det i sista delen att eleven också utvärderar metodens, alltså bokföringens användbarhet (Skolverket, 2011, ss. 9-10). Det är tydligt att det endast är Eva som i sina prov kan fånga denna aspekt av kunskapskraven eftersom inga av de andra intervjuade lärarna har några förståelsefrågor. Samtidigt berättar de andra lärarna att de testar exempelvis begreppskunskap under lektionen och att det lämpar sig bättre att observera dessa kunskaper i en klassrumssituation än i ett prov. Svaren lärarna här ger tillför kunskap om vilka kunskaper de önskar mäta på bokföringsproven

Lärarnas svar tillför dessutom information om hur de säger att de väljer ut de kunskaper de önskar mäta. Om jag tittar på den indirekta mätningen för att bedöma validitet, dvs att lärarna hittar indikatorer för den kunskap de vill mäta bedömer jag att lärarna använder läroböckerna som indikatorer. Dvs de hittar de begrepp som de tycker är relevanta för området och testar dessa på provet. Men samtidigt kan de inte veta att dessa begrepp är de som är relevanta utan får förlita sig på läroböckernas förmåga att hitta rätt kunskap för bokföringsområdet. Hans säger sig se att de olika

läroböckerna liknar varandra och därmed får han en kvalitetssäkring av att innehållet är relevant. Flera av lärarna känner sig osäkra på vad Skolverket anser. De är osäkra dels på vilka kunskaper området och därmed provet ska mäta och dels så är de osäkra på vilken nivå av kunskap som kan hänföras till respektive nivå i kunskapskraven. Eva och Rogers önskan om att ha en kanon, eller nationella prov på området ser jag som ett uttryck för denna osäkerhet. När jag frågar vilken kunskap det är de vill mäta och hur de väljer ut dessa kunskaper är svaret samstämmigt att det är utifrån läroböckerna. De kunskaper som läroböckerna väljer att ta upp är också de som lärarna säger att de vill mäta. Då handlar det om ett antal typer av bokföringsoperationer såsom kundförluster, försäljning och kundreskontra etc. Kriterierelaterad validitet innebär enligt Wikström (2013, s. 47) att provet mäter den kunskap som provkonstruktören vill mäta samt att läraren kan komma till samma slutsats om elevens förmåga med andra bedömningsinstrument. Om proven är konstruerade enligt definitionen av kriterierelaterad validitet ger det ytterligare kunskap för att förstå hur lärare resonerar kring konstruktionen av bokföringsprov då det gäller att mäta elevernas olika kunskapsnivåer kopplat till kunskapskraven. Frågan ställdes om man kan se att de elever som i hög utsträckning presterat bra på provet också presterar bra i en verklig bokföringssituation. Samtliga lärare svarade ja på den frågan och i Karins fall görs provet direkt i ett bokföringsprogram. Eva uttrycker också att det bästa vore att göra provet direkt i dator för att lättare kunna uttala sig om förmågan att bokföra. Samtidigt konstaterar hon att provet hade kunnat göras mer verkligt om riktiga bokföringsunderlag såsom fakturor etc. hade kunnat utgöra frågorna. Utifrån definitionen av begreppet provvaliditet som Cronbach & Meehl (1955, ss. 281-282) ger och utifrån lärarnas svar tycker jag att de i huvudsak har prov som innehållsmässigt är valida eftersom de utgår från läroböckernas innehåll. Jag kan också dra slutsatsen att den kriterierelaterade validiteten är god eftersom de samtliga vet vilken kunskap det är de vill fånga och eftersom andra bedömningsinstrument (i detta fall lärarens observation av elevens prestation i bokföringsprogram) leder till ungefär samma resultat enligt lärarna själva.

6.2 Kanes validitetsmodell

Kanes validitetsmodell är, som tidigare nämnts, en metod för att undersöka validiteten i ett bedömningsinstrument. Denna modell utgår från att validitetstestet görs i flera steg (Jönsson, 2013, s. 60). Med hjälp av denna modell kan ytterligare kunskap adderas kring hur lärarna resonerar kring konstruktionen av sina bokföringsprov då det gäller att mäta elevernas olika kunskapsnivåer kopplat till kunskapskraven, eftersom en analys kan göras av lärarens förmåga att översätta provresultatet till kunskapskravens betygsnivåer. I det första steget görs en bedömning av elevens resultat och detta skall översättas till kunskapskraven. Min bild är att lärarna tycker sig kunna bilda en uppfattning huruvida eleverna når upp till olika kunskapsnivåer endast utifrån provresultatet. Eva menar dock att det är bra med en komplettering i form av rapporter från eleverna efter att de besökt företag för att koppla det de teoretiskt har lärt sig till en verklig bokföringssituation. Men alla svarar att provet är den huvudsakliga bedömningsformen. Samtidigt medger lärarna att det inte är konstruktionen av provet som erbjuder mest motstånd för dem utan det är bedömningen utifrån kunskapskraven. Modellen hjälper dessutom till att skapa förståelse kring hur lärarna säger sig välja ut de kunskaper de vill mäta. Kanes validitetsmodell beskriver hur förmågan att dra slutsatser kring huruvida eleverna också klarar andra likartade uppgifter med samma resultat underlättas om provet innehåller flera liknande uppgifter. Förmågan som testas på provet är förmågan att

använda sig av företagsekonomiska metoder och i detta fall specifikt bokföring. Bokföring testas igenom hela provet hos alla lärarna och mest repetitivt är ju Karins, Hans och Rogers prov. Å andra sidan behöver eleverna kunna bokföra flera olika affärshändelser för att läraren ska kunna dra slutsatsen att eleverna behärskar metoden bokföring. Lärarna uppger att på deras prov är det enstaka uppgifter av varje typ. Exempelvis en varuförsäljning med moms, en kreditfaktura, en arbetsgivaravgift etc. Lärarna menar att deras prov inte testat samma typ av uppgift flera gånger.

Sedan skall läraren kunna extrapolera resultatet på provet (Jönsson, 2013, ss. 61-62). Det innebär att elevens resultat på provet ska kunna översättas till kunskapskraven genom att också säga någonting om elevens förmåga i situationer som inte har testats på provet. Frågan är i sig svår att svara på då svaret kommer först då eleverna bokför någon gång i framtiden i en verklig situation. Men lärarna uppger att när eleverna arbetar i bokföringsprogrammen så går det att säga att de elever som lyckas bra på proven också klarar nya uppgifter. Samtidigt är ett svar från Eva intressant. Hon menar att det inte bara handlar om att kunna bokföra utan också om att kunna förstå. Hon testat elevernas förståelseförmåga på provet genom att de ska få förklara varför moms inte är en kostnad för företaget. Om en elev exempelvis vet att moms inte är en kostnad eller vet hur skatten på lön beräknas så borde elevens förmåga att kunna bokföra liknande affärshändelser som inte har testats jämfört med en elev som bara präntat in kunskapen och gör konteringen mekaniskt vara högre. Både Karin och Hans lyfter fram att de visst tycker att eleverna är mycket duktiga på bokföring i allmänhet om de har A på provet men att det samtidigt är utifrån ett gymnasieperspektiv och att eleverna hade haft svårt att bokföra på exempelvis en redovisningsbyrå. Som jag tolkar teorin blir frågan om validitet till stor del en fråga om en subjektiv bedömning från varje lärare. På en direkt fråga om lärarna anser att eleverna också klarar bokföring i en verklig situation efter hur bra de presterat på provet så är ju svaret ja från lärarna. Jag tolkar det som att lärarna menar att validiteten är god i deras konstruerade prov utifrån Kanes validitetsmodell. Genom att använda Kanes validitetsmodell för att analysera lärarnas svar har kunskap tillförts om hur lärarna resonerar kring konstruktionen av sina bokföringsprov då det gäller att mäta elevernas olika kunskapsnivåer kopplat till kunskapskraven.

6.3 Reliabilitet

Ett annat centralt begrepp inom testteorin, som tidigare nämnts är reliabilitet, som säger oss hur tillförlitligt provet är. En hög reliabilitet innebär att eleven skulle få samma resultat på provet oavsett när detta görs, om antagandet görs att eleven innehar samma kunskaper vid de olika provtillfällena (Wikström, 2013, ss. 39-40). Endast Hans har reflekterat över vad det innebär för reliabiliteten om en elev har en dålig dag eller har sovit dåligt på natten. Han hanterar detta genom att låta dessa elever muntligen redogöra för delar av provet för honom eller att de får göra om provet. Det är ett sätt att hantera bristen på reliabilitet. En annan brist i reliabilitet kan enligt Wikström (ibid, s. 39–40) vara att det är för få frågor på provet. Detta kan då bli ett problem om för stor del av kursen testas i ett prov då varje delområde endast representeras av någon enstaka fråga. Med hjälp av reliabilitetsbegreppet kan kunskap nu adderas kring vilka kunskaper lärarna anser att de vill mäta med bokföringsproven samt hur lärarna säger sig välja ut denna kunskap. Eva, Roger och Hans säger att de endast testat bokföring i sina prov på bokföring. Karin berättar att

hon även testar bokslut vilket borde göra att reliabiliteten på varje delområde blir sämre. Roger berättar att det är en särskild strategi att testa endast små områden i taget då han bedömer att eleverna lyckas bättre då. Eva behöver å sin sida hantera att hon testat fler aspekter av bokföringen i form av begreppskunskap och djupare förståelse. Detta torde enligt ovan påverka reliabiliteten negativt.

Ett sätt att hantera reliabiliteten är enligt Skolverket (2011b, s. 35) att sambedöma proven, vilket Eva går längst i. Att flera lärare bedömer provet och att de sedan enas om ett betyg borde medföra ökad reliabilitet. Interbedömarreliabilitet beskriver enligt Lundahl (2016, s. 39) om också andra lärare skulle göra samma bedömning av elevernas kunskapsnivå som konstruerande lärare gör utifrån provet. Hans är den som tydligast testat interbedömarreliabiliteten genom att berätta att han låter andra lärare göra en andra bedömning av hans prov då han är osäker. Han berättar då oftast inte för den andre läraren vad han har kommit fram till för betyg och enligt honom själv kommer de oftast fram till samma resultat. Troligen påverkas interbedömarreliabiliteten av att både Eva, Roger och Hans säger att de konstruerar sina prov tillsammans med andra kollegor. De berättar att de är överens med sina kollegor om vad som ska ingå i provet och om vilken nivå frågorna för de olika nivåerna ska ligga på. Om en lärare från en annan skola skulle rätta proven är det däremot oklar om resultatet skulle bli detsamma. Det är inte orimligt att utgå från att lärare som arbetat länge tillsammans lär sig hur sina kollegor resonerar och bedömer olika elevsvar. Jag tycker också att det är värt att i sammanhanget notera att just bokföring till stora delar består av frågor som ger eleverna rätt eller fel. För Karin, Hans och Roger som endast testat affärshändelser borde det med andra ord inte vara särskilt svårt att enas om en bedömning av provens resultat kopplat till kunskapskraven. För Eva som har förståelsefrågor och begreppskunskap på provet kan det däremot ges större utrymme för tolkningar och därmed större grad av subjektivitet. Hans och Eva lyfter fram att en form av bedömning som diskuteras är om en elev inte nått hela vägen fram men ändå förstått principen eller räknat rätt men bokfört på fel sida debet eller kredit. I detta fall har ju eleven gjort uppgiften fel utifrån att de skulle bokföra en affärshändelse men läraren kan ändå bedöma att eleven når upp till D eller B i kunskapskraven. Interbedömarreliabiliteten blir enligt Lundahl (2016, s. 38) bättre med fler betygssteg. Det innebär att Karin och Evas prov torde ha högre interbedömarreliabilitet eftersom de använder fler betygssteg. Det är ju logiskt eftersom skillnaden för eleven ju är större mellan E och C än mellan D och C.

6.4 Under- och överrepresentation

Under- och överrepresentation beskriver enligt Wikström (2013, ss. 90-91) risken att provet antingen inte testat de kunskaper som läraren önskar eller testat för området icke relevanta kunskaper, vilket förklaras vidare i avsnitt 3.1.4. Teorin kring under- och överrepresentation är användbar för att dels studera vilka kunskaper lärarna anser att vill testa med bokföringsprovet och dels hur dessa kunskaper väljs ut av lärarna. Finns det risk att dessa kunskaper testat icke-relevanta kunskaper eller för lite kunskaper? När lärarna intervjuas framkommer det att de i stort sett följer den rekommendation som denna studies teoriavsnitt ger, dvs att definiera vilket innehåll som ska utgöra kunskaperna som ska mätas på provet. Alla lärarna säger sig använda läroböckerna för att definiera detta innehåll. Det är alltså därmed inte de själva som definierat innehållet - konstruktet. Jag bedömer att detta kan medföra problem. Hur kan lärarna vara säkra på att läromedelsförfattarna har

inkluderat alla relevanta delar av arbetsområdet? Däremot är det flera lärare som säger att de stämmer av kunskaperna som ska mätas med andra personer när proven konstrueras. Eva, Roger och Hans uppger ju att de konstruerar proven tillsammans med kollegor. Roger påtalar också att innehållet behöver vara en god grund för att eleverna ska klara av nästa kurs Redovisning 2. Karin säger sig stämma av innehållet med redovisningsbyrå och har haft eget företag och säger att hon därigenom vet vilka kunskaper som behövs i bokföring för att driva ett företag. Ur denna aspekt ser jag en risk att Karin och Rogers prov överrepresenterar innehållet, dvs de testar för mycket. Om Rogers prov tar upp kunskaper som först presenteras i kursen Redovisning 2 kan provet komma att testa för mycket som jag ser det. Karin borde riskera att testa kunskaper som ligger helt utanför skolvärlden om hon tar för mycket hänsyn till vad en redovisningsbyrå tycker att eleverna bör kunna. Samtidigt är det min bedömning att Karin minimerar den här risken genom att vara medveten om sitt sätt att resonera och genom att konstatera att kunskaperna som eleverna får i bokföring är relevanta för kunskapskravet om företagsekonomiska metoder men ur ett gymnasieperspektiv. Vidare uppger ju även Karin att litteraturen är en bra grund för att veta vilken nivå kunskaperna bör ligga på. Jag bedömer att Eva är den som mest utvecklat hanterar frågan om överrepresentation eller rättare sagt konstruktirrelevans. Hon säger att hon tänker mycket på att uppgifterna inte får innehålla för mycket information eftersom provet då riskerar att även testa läsförmågan vilket inte är en del av något kunskapskrav eller centralt innehåll i kursen Företagsekonomi 2. Samtidigt vill jag lyfta fram att även Hans för ett resonemang kring risken att testa irrelevanta kunskaper. Hans menar ju att han fick förändra sitt sätt att konstruera prov efter att läroplanen Gy11 introducerades. Tidigare kunde han testa kunskaper som han tyckte var intressanta men som han upptäckte inte svarade mot något av kunskapskraven.

Ingen av lärarna lyfter fram eller funderar över huruvida de däremot testar och mäter för lite kunskapsinnehåll. Samtliga lärare tycker att de håller sig inom ramarna eftersom de rättar sig efter vad läroböckerna tar upp i den kunskap de vill mäta. Men framförallt Eva och Hans resonerar kring att det finns en risk att proven testat för mycket. Karin ser det inte som en risk att det kan leda till att för mycket kunskap testas men stämmer av med redovisningsbyrå, vilket jag ser kan leda till en överrepresentation. Även Roger riskerar att testa för mycket eftersom han uppger att han till viss del bygger sina prov på vad som bör vara en god grund för nästa kurs. Eva, Roger och Hans säger att de inte vet exakt vad Skolverket vill att de ska ta upp i kursen och uttrycker viss oro för detta. Men ingenstans i intervjuerna med mig ger denna oro upphov till funderingar kring om de testat för lite.

6.5 Objektiva prov och alternativa (subjektiva) bedömningar

Frederiksen & Allan (1989, s. 28) menar att prov kan delas in i två kategorier. Objektiva prov där eleven ska fylla i enkla uppgifter och där rätt eller fel enkelt kan bedömas. Alternativa bedömningar handlar om prov som innehåller kognitivt svårare frågor där en noggrannare bedömning behöver göras av elevens prestation. Det är värt att upprepa det som skrevs i avsnittet om objektiva prov/alternativa bedömningar i teoriavsnittet, dvs att teorierna kring objektiva prov och alternativa bedömningar kan användas för att ge ytterligare information om vilken kunskap lärarna anser att de vill mäta med bokföringsprovet. Vad är det för typ av prov lärarna konstruerar och har provtyp valts ut beroende på vilka kunskaper lärarna säger sig vilja mäta med provet? Har val av provtyp gjorts för att lärarna ska kunna

mäta eleverna kunskaper kopplat till kunskapskraven? Karin, Hans och Rogers prov är enligt min bedömning att anse som objektiva prov utifrån de intervju svar de ger mig. Även Evas prov borde till stora delar vara objektivt utifrån hennes svar. Denna slutsats drar jag eftersom lärarna uppger att deras prov i bokföring till stora delar består av affärshändelser som ska bokföras och som flera av lärarna framhäver, så är det uppgifter av typen rätt eller fel. Förståelsefrågor är enligt Frederiksen & Allan (ibid, s.28) mer öppna i sin karaktär och faller snarare in under kategorin alternativ bedömning eftersom det är elevens bedömningar och reflektioner från eleverna som testas. Samtidig kan jag se att gränsen inte är helt knivskarp mellan vad som är en "authentic assesment" och ett objektivt prov. Min bedömning av Karins prov är att även hennes prov är att anse som ett objektivt prov eftersom siffror ska fyllas i på antingen debet eller kredit. Samtidigt är ju detta exakt den situation som den som bokför i skarpt läge befinner sig i. Å andra sidan menar Palm (2001, s. 12) att fler kriterier behöver vara uppfyllda för att provet ska kunna betecknas som "auhtentic". Provet behöver vara knutet till livet utanför skolan, vilket det ju är eftersom eleverna enligt lärarna ska bokföra affärshändelser men Eva tycker inte själv att hennes prov är tillräckligt verklighetsnära. Sedan behöver provet vara kopplat till läroplanen och till undervisningen som föregått provet. Min bedömning är att detta uppfylls genom att bokföring just kan knytas till det centrala innehållet och till kunskapskraven. Eftersom Eva säger att hon följer läroboken både i undervisningen och sedan använder begreppen från denna till provet uppfylls det kriteriet. Slutligen behöver frågorna stimulera till vidare inläring. Den delen bedömer jag att hennes prov har svagare koppling till eftersom det är ett summativt prov där elevernas befintliga kunskap ska testas av.

För objektiva prov finns det enligt Frederiksen & Allan (1989, s. 28) inte särskilt mycket utrymme för tolkningar och kan användas för att testa mindre djupa kunskaper. Jag tror att Eva, Rogers och Hans poängprov är ett uttryck för detta. Eftersom åtminstone Roger och Hans säger att de endast testat affärshändelser är det enkelt att sätta poäng efter rätt eller fel och sedan med hjälp av detta mäta elevernas olika kunskapsnivåer kopplat till kunskapskraven. Dessa tre lärare menar också att antalet poäng motsvarar en specifik betygsnivå. Samtidigt ger de svar som till viss del motsäger sig själva. Dels tycker de att poängen styr betyget men å andra sidan tycker de att det är svårt att bedöma proven eftersom Skolverket inte ger klara instruktioner för var gränsen mellan "säkerhet" och "viss säkerhet" går. Roger framför att han tycker att det är skönt med ett prov som ger rätt eller fel eftersom det är lätt att rätta och sätta betyg på. Men tidsaspekten som Roger säger att han sätter in i provet tillför en subjektiv bedömning.

Som jag tolkar det använder Karin sig av objektiva prov men använder sig inte av poäng vid rättandet utan gör en bedömning utifrån de olika nivåerna i kunskapskravet om eleverna kan sägas kan bokföra med "säkerhet" eller med "viss säkerhet". Jag bedömer att lärarna har olika sätt att hantera just konstruktionen av sina prov avseende bedömningen av elevernas kunskaper. Alla lärarna i studien säger att de viktar proven. Roger, Eva och Hans berättar att de tar hänsyn till detta i sina poängprov medan Karin tar hänsyn till detta när hon bedömer elevens kunskap i förhållande till kunskapskraven. Karin menar ju att en elev som klarar av en mer komplex kontering visar större säkerhet. Roger och Eva berättar att de hanterat detta genom att de ger fler poäng på svårare frågor. Roger reflekterar samtidigt över att den elev som siktar på A i betyg inte kan lyckas enbart med de svåra uppgifterna och inte på de mer grundläggande uppgifterna.

Min bedömning är att de ingående lärarna konstruerar sina prov som objektiva prov för att lättare kunna bedöma och sätta betyg. Detta ger en svagare koppling till kunskapskraven eftersom det är svårare för läraren att motivera var gränsen går mellan ”säkerhet” och ”viss säkerhet” när antalet poäng styr. I intervjuerna med lärarna framgår det ju att detta är ett sätt att hantera en svår uppgift, dvs att konstruera ett prov där elevens prestation ska kunna översättas till en nivå i kunskapskraven. Vad som är rätt eller fel kan jag dock inte uttala mig om eftersom Skolverket ju ger läraren frihet att själv utforma bedömningsformer för de olika kunskapskraven.

6.6 Alignment

Denna del är tätt kopplad till förra delen och belyser problematiken i att kunna koppla prov till kunskapskraven eftersom alignment enligt (Linn & Herman, 1997, s. iii) berättar hur samstämmigt ett prov är med läroplanen. I intervju med lärarna framgår det att de brottas med den här problematiken när de väljer vilka kunskaper de vill mäta med sina bokföringsprov. Eftersom lärarna beskriver att Skolverket inte ger någon vägledning kring vad de förväntas ta upp i området och hur de ska tolka de olika betygsnivåerna är de beroende av att läromedlen är utformade efter läroplanen. Svårigheten ligger enligt flera av lärarna i att konstruera proven på ett sådant vis att de kan mäta elevernas olika kunskapsnivåer kopplat till kunskapskraven för kursen. Karin påtalar dock att hur man tolkar kunskapskravens olika nivåer är upp till respektive skola och lärare. De andra lärarna är mer frustrerade över bristen på vägledning. Men om läromedlen är utformade efter läroplanen, undervisningen utgår från denna och eleverna sedan testas på den kunskapen de lärt sig i undervisningen så borde samstämmigheten vara hög. När lärarna väljer ut de kunskaper som de vill mäta så uppger de att det helt är efter läroböckernas stoff. Men läroböckerna graderar inte uppgifterna efter E, C, A-nivå så det är upp till respektive lärare att göra den bedömningen själv. Det försvagar kopplingen mellan lärarnas bedömning av elevernas kunskapsnivåer och läroplanen.

6.7 Konstruktionsregler

Här ställs resultatet mot de tre kriterier som Wikström (2013, s. 104) anser behöver vara uppfyllda för att ett prov ska mäta de kunskaper som läraren vill mäta. Dessa kriterier ger dessutom mer kunskap kring hur lärarna resonerar kring konstruktionen av sina bokföringsprov då det gäller att mäta elevernas olika kunskapsnivåer kopplat till kunskapskraven? Viktas provfrågorna för att mäta dessa kunskapsnivåer? För det första behöver lärarna specificera vilka kunskaper de vill mäta och ställa frågor som motsvarar dessa kunskaper. Jag har tidigare varit inne på att lärarna menar att den kunskap de vill mäta till största delen består av vad som tas upp i läroböckerna och att de väljer ut frågorna utifrån de olika delar som tas upp där. Lärarna berättar att frågorna består av affärshändelser som liknar de övningsuppgifter som finns i läroboken. Evas förståelsefrågor är däremot egenkonstruerade men utgår enligt henne från de olika begrepp som tas upp i läroboken. Jag tycker att samtliga lärare är tydliga med vad det är för kunskaper de vill mäta. De är däremot mindre tydliga med hur denna kunskap ska mäta elevernas olika kunskapsnivåer kopplat till kunskapskraven för kursen. Dvs hur ska eleven svara för att visa på kunskaper ”med säkerhet” eller med ”viss säkerhet”? Där uppfattar jag att lärarna inte riktigt kan ge tillfredsställande svar. De är också medvetna om detta och tar upp det som ett

problem eftersom Skolverket inte ger dem tydliga instruktioner. De uppger att de därför väljer att bedöma efter antal poäng. Eva och Hans lyfter fram att nationella prov hade varit ett bra sätt att veta vilken nivå respektive fråga bör vara på. För det andra behöver lärarna använda diskret formulerade frågor. Detta innebär att inga frågor ska vara beroende av någon annan fråga. Det innebär också att eleven inte ska behöva svara rätt på en fråga för att kunna besvara en annan fråga. Lärarna uppger att de väljer ut kunskaper och ställer frågorna diskret, dvs varje fråga står för sig. För det tredje behöver lärarna använda sig av okomplicerade frågor som ej ger anledning till tveksamheter och tvetydigheter. Frågorna ska vara förenklade. Detta resonerar de intervjuade lärarna mycket kring. Eva säger att hon är noga med att frågorna inte ska ha karaktären av läsetal och är dessutom noga med att eleverna ska förstå begreppen. Hon berättar att hon har konstruerat ett bokföringslexikon som eleverna har tillgång till just för att eleven inte ska misslyckas på en fråga bara för att de inte förstår ett ord eller begrepp. Lärarna lyfter också att frågor på just bokföringsområdet går rakt på sak och att affärshändelserna som beskrivs, beskrivs med en enkel text. Alltså bedömer jag att lärarnas prov inte ger upphov till vare sig tveksamheter eller tvetydigheter samt att de tar hänsyn till detta när de konstruerar prov. Det är uppenbart att lärarna i intervjuerna med mig har ett väl underbyggt resonemang kring detta.

Vidare lyfter Wikström (2013, s. 120) vikten av att fundera över hur tung respektive fråga ska vara i förhållande till andra frågor eftersom det då blir lättare att bedöma elevernas olika kunskapsnivåer. Detta har lärarna ett väl utvecklat resonemang kring. Hans säger att han endast ger rätt eller fel och därmed max 1 poäng per fråga. Men han berättar att när han sedan bedömer proven ger han betyg efter vilka frågor eleverna svarat rätt på. Hans och Roger uppger att de har tänkt ut vilka som är E, C och A-frågor och bedömningen grundar sig sedan på elevens förmåga att svara rätt på dessa frågor. Eva berättar att hon fr.o.m. nu kommer sätta ut dessa nivåer i provet så att eleverna vet hur tungt respektive fråga väger. Karin uppger att hon redan idag sätter ut detta på provet och att eleverna klart och tydligt ser vad det är för typ av fråga gällande betygsnivå. Sammanfattningsvis konstaterar jag att de intervjuade lärarna tar hänsyn till dessa konstruktionsregler när de konstruerar prov för att mäta de olika nivåerna i kunskapskraven.

6.8 Analys av resultat mot tidigare forskning

I likhet med Boesens (2006, s. 12) slutsatser i sin avhandling konstaterar jag att den kunskap som lärarna i min studie säger sig vilja mäta till allra största del är repetitiv kunskap som inte går djupare i förståelsenivåerna. Undantaget är Eva som ju säger att hon även lägger in förståelsefrågor och begreppskunskap på sina prov. Samtidigt konstaterar Eva att hon tycker att förståelsefrågor och begreppskunskap är svårare att konstruera provfrågor på så att de mäter elevernas olika kunskapsnivåer kopplat till kunskapskraven för kursen. Odenstad (2010, ss. 174-175) drog ju slutsatsen att det till viss del är beroende av vilket arbetsområde det är som testas. Vissa arbetsområdens prov består till större del av faktakunskap och endast till mindre del av förståelse. De intervjuade lärarna framför att bokföring handlar mycket om ett hantverk som behöver bemästras för att förstå områdena som kommer efter bokföring, dvs bokslut och nyckeltalsanalys. Det är också min uppfattning. Om eleverna ska kunna förstå senare moment i kursen bygger det på att de förstår

grunderna. Det finns egentligen ingen värdering i just bokföring utan det är rakt av en kontering av olika poster. I nästa steg när bokslutet ska upprättas, vilket bygger på bokföringen är det mer relevant att ställa sig frågor av typen: är detta ett bra resultat?

Tobias Jansson (2011, s. 152) konstaterade att många prov lider av validitetsbrist men att detta beror på tidsbrist. Det är inte min bild efter intervjuerna. Lärarna uppger att de mäter de kunskaper de vill mäta. Samtliga lärare uppger ju dessutom att den bedömning de gör av elevernas olika kunskapsnivåer kopplat till kunskapskraven för kursen sammanfaller med bedömningen från andra bedömningsinstrument än prov. Inga av de ingående lärarna pekar på att provet hade konstruerats bättre om eleverna hade fått mer tid på sig att göra provet eller om lärarna hade fått mer tid för att konstruera provet. Eva lyfter förvisso problemet med att provet skulle kunna innehålla mer verklighetsanknutna frågor om hon fått mer tid med att hitta fakturor och liknande från verkligheten.

David Rosenlunds studie är intressant eftersom han drar slutsatsen att lärarens prov har en oacceptabelt låg nivå av samstämmighet med läroplanerna eftersom de ofta testar eleverna på lägre kognitiv nivå än vad kursplanerna kräver (Rosenlund, 2011, ss. 164-165). Det kan säkert vara så vad gäller begrepp och förståelsefrågor. Eva lyfter ju att det framförallt är svårt att testa de högre betygsnivåerna för begreppen. Men vad gäller kunskapskravet om företagsekonomiska metoder kan jag inte se att de olika betygsnivåerna rör sig på olika kognitiva nivåer. Här handlar det om att eleverna ska välja, använda och förklara oavsett betygsnivå. Valet av metod är ju gjort av läraren i och med att provet handlar om just den dubbla italienska bokföringen. Men eleverna testas på förmågan att använda denna metod i och med att de enligt lärarna förväntas kontera affärshändelserna. Det som lärarnas prov missar i samtliga fall, förutom hos Eva är att eleverna inte får möjlighet att förklara bokföringen. Eva menar ju att hon lägger in frågor om varför moms inte är en kostnad eller vad en nettolön och bruttolön är. De övriga lärarna missar detta i sina resonemang. Eva resonerar mycket kring detta när hon väljer ut de kunskaper som ska mätas på provet och framförallt hur elevernas olika kunskapsnivåer kopplat till kunskapskraven för kursen ska mätas, även om hon medger att bedömningen inte är lätt, kopplat till elevernas förmåga att förklara. De andra lärarna har inte några direkta resonemang kring denna problematik.

7. Slutsatser

Den här uppsatsen har adderat kunskap kring hur lärare resonerar när de konstruerar sina bokföringsprov i kursen företagsekonomi 2. Lärarna uppger att de vill att bokföringsprovet ska vara så likt en verklig bokföringssituation som möjligt. I vissa fall handlar det om att provet görs direkt i bokföringsprogram medan eleverna i andra prov får kontera på papper. För att välja ut de kunskaper som lärarna ville mäta använder de sig i hög grad av sina läroböcker i företagsekonomi men en av lärarna gör även en avstämning av detta innehåll med en redovisningsbyrå. Genom sitt sätt att välja ut kunskap och genom att resonera kring vad litteraturen tar upp vill lärarna säkerställa att proven inte lider av validitetsbrist. Denna slutsats förstärks av att lärarna anser att prestation på provet också speglas i förmågan att använda bokföringsprogram. Men, konsekvenserna av att hämta de kunskaper som ska mätas från läroböckerna blir att lärarna inte kan vara säkra på att de får med relevanta kunskaper eftersom de då blir beroende läromedelsförfattarnas urval.

Lärarna medger att det är svårt att göra bedömningen av elevernas kunskap kopplat till kunskapskraven i proven. De hade önskat ytterligare vägledning från Skolverket för att veta vad som ska testas för respektive nivå. Detta hanterar tre av lärarna genom att ha poäng på provet och att poängresultatet styr vilket betyg eleven får på provet. Tre av fyra lärare använder sambedömning och samma tre lärare använder sig av samkonstruktion. På detta vis vill de försäkra sig om att de säkerställer rättssäkerheten gentemot eleven, vilket också ökar reliabiliteten i provet. Två av lärarna i studien resonerar kring risken att testa för mycket på provet. En av lärarna har dessutom fått ändra vad som ska mätas eftersom han anser att den nya läroplanen Gy 11 inte ger honom utrymme att testa det innehållet.

Lärarna bedöms i huvudsak konstruera proven som objektiva prov. Detta gör att reliabiliteten ökar eftersom den typen av prov är lättare att rätta. Lärarna kan då avgöra om eleven har rätt eller fel på respektive uppgift. Däremot säger lärarna att det är svårare att i ett sådant typ av prov bedöma var skillnaden mellan olika betygsnivåer ligger. Detta bekräftar lärarna också genom att berätta att det är svårt att förstå var Skolverket tycker att nivåerna bör ligga. De har dock själva gjort en bedömning om vad som är lättare respektive svårare nivå men kan samtidigt inte säga om detta är en bedömning som delas av Skolverket. De intervjuade lärarna uppger att de viktar provet så att svårare uppgifter väger tyngre vid rättning. En del av dem visar denna viktning mot eleverna medan två av dem inte gör det och en av dessa tycker dessutom inte att det är önskvärt att visa detta för eleverna. Lärarna tycker att viktningen är konstruerad på ett sådant vis att eleverna måste klara av även de svårare uppgifterna för att kunna erhålla exempelvis betyg A på provet. Lärarnas konstruktion av sina prov bedöms ha en tydlig samstämmighet med läroplanen vad gäller innehållet genom att de uppger att utgår från innehållet i läroböckerna som ju är utformade efter läroplanen. Däremot är det sämre med samstämmigheten avseende betygsdifferentieringen. Återigen påtalar lärarna att detta beror på att Skolverket inte ger dem någon vägledning. En av lärarna finner sig i detta och menar att det faktiskt är upp till varje skola att göra den bedömningen.

Det är tydligt att lärarna använder sig av Wikströms (2013, s. 103) rekommendationer för att konstruera ett rättvisande prov. Även om lärarna inte känner till forskningen kring detta resonerar de kring de frågeställningar som

Wikström hänvisar till. Det kan konstateras att de i denna studie ingående lärarna, vid provkonstruktion till vissa delar bekräftar den bild som framkommer i den tidigare forskningen på området. Samtidigt finns det delar som den här studien motsäger. I en tidigare studie visades att lärarnas prov led av validitetsbrist och att detta beror på tidsbrist. Inga av de lärare som intervjuades klagade på tidsbrist. En annan studie visade att samstämmigheten mellan prov och läroplaner ofta är oacceptabelt låg och att detta beror på att lärarna konstruerar sina prov så att fel kognitiva nivåer testas. Denna bild bekräftas inte i denna studie.

8. Förslag till vidare forskning

Den här studien har inriktat sig på hur lärare resonerar när de konstruerar prov i bokföring och har en kvalitativ ansats. Det vore intressant att med en kvantitativ inriktning undersöka hur prov i bokföring ser ut när de är konstruerade. Vilken typ av kunskap visade det sig sedan att dessa prov mätte? Finns frågor upptagna på proven som mäter de olika metakognitiva nivåerna enligt kunskapskraven? I det sammanhanget kan det vara intressant att mäta i vilken utsträckning detta sker. För att göra en sådan undersökning behöver troligen andra metoder användas än vad som använts i denna studie. Vidare skulle det kunna vara intressant att studera hur lärare som inte använder sig av bokföringsprov resonerar när de ska bedöma elevernas färdigheter på detta område. Vilken typ av bedömningsunderlag använder de istället? Varför tycker dessa lärare att bokföringsprov inte lämpar sig som bedömningsform för området?

Referenser

- Boesen, J. (2006). *Assesing mathematical creativity*. Umeå universitet: Department of Mathematics and Mathematical Statistics.
- Bolagsverket. (2017). *Verksamt.se*. Retrieved 12 01, 2017, from <https://www.verksamt.se/starta/forbered-din-start/bokforing>
- Crocker, L., & Algina, J. (1986). *Introduction to classical & modern test theory*. Orlando, Florida, USA: Holt, Rinehart and Winston.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), pp. 281-302.
- Esaiasson, P., Gilljam, M., Oscarsson, H., & Wägnerud, L. (2012). *Metodpraktikan - Konsten att studera samhälle, individ och marknad* (4:e upplagan ed.). Stockholm: Norstedts Juridik.
- Frederiksen J, R., & Allan, C. (1989). A Systems Approach to Educational Testing. *Educational Researcher*, 18(9), pp. 27-32.
- Goldstein, H. (2015). Validity, science and educational measurement. *Assesment in Education, Policy & Practice*, 22(2), pp. 193-201.
- Jönsson, A. (2013). *Lärande bedömning* (3:e upplagan ed.). Malmö: Gleerups.
- Jansson, T. (2011). *Vad kommer på provet - Gymnasielärares provpraxis i samhällskunskap*. Karlstads Universitet, Fakulteten för samhälls- och livsvetenskaper Statsvetenskap. Karlstad: Karlstads Universitet.
- Johannessen, A., & Tufte, P. A. (2002). *Introduktion till samhällsvetenskaplig metod* (1:a upplagan ed.). Malmö: Liber.
- Kane, M., Crooks, T., & Cohen, A. (1999). Validating Measures of Performance. *Educational Measurement: Issues and Practice*, 18(2), pp. 5-17.
- Linn, R. L., & Herman, J. L. (1997). *Standards-led assessment: Technical and policy issues in measuring school and student progress*. University of California. Los Angeles: University of California.
- Lundahl, C. (2016). *Bedömning för lärande* (2:3 ed.). Lund: Studentlitteratur.
- Lundahl, U., & Skärvad, P.-H. (1999). *Utredningsmetodik för samhällsvetare och ekonomer* (3:e upplagan ed.). Lund: Studentlitteratur.
- Messick, S. (1989). Meaning and Values in Test Validation: The Science and Ethics of Assesment. *Educational Researcher*, 18(2), pp. 5-11.
- Näsström, G. (2008). *Measurements of alignment between standards and assesment*. Doktorsavhandling, Umeå Universitet, Department of Educational Measurement , Umeå.
- Odenstad, C. (2010). *Prov och bedömning i samhällskunskap - En analys av gymnasielärares skriftliga prov*. Karlstads Universitet, Fakulteten för samhälls- och livsvetenskaper Statsvetenskap. Karlstad: Karlstads Universitet.
- Palm, T. (2001). *Performance and Authentic Assessment, Realistic and Real Life Tasks: A Conceptual Analysis of the Literature*. Umeå Universitet, Department of Educational Measurement. Umeå: Em No 39.
- Riksdagsförvaltningen. (2011). *Riksdagen*. Retrieved 12 05, 2017, from Svensk författningssamling: https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/forordning-2011686-om-kompletterande_sfs-2011-686

- Rosenlund, D. (2011). *Att hantera historia med ett öga stängt - Samstämmighet mellan Historia A och lärares prov och uppgifter*. Forskarskolan i historia och historiadidaktik. Lund: Lunds Universitet.
- Skolverket. (2011). *Läroplan för gymnasieskolan - kursplan Företagsekonomi*. Retrieved 12 15, 2017, from <https://www.skolverket.se/laroplaner-amnen-och-kurser/gymnasieutbildning/gymnasieskola>
- Skolverket. (2016). *Statens skolverk*. Retrieved 12 08, 2017, from Handfasta råd vid provkonstruktion: <https://www.skolverket.se/skolutveckling/forskning/bedomning/betygsattning/handfasta-rad-om-provkonstruktion-1.214797>
- Skolverket. (2017). *Nationella prov*. Retrieved 11 21, 2017, from Statens skolverk: <https://www.skolverket.se/bedomning/nationella-prov>
- Skolverket. (2017b). *Statens skolverk*. Retrieved december 17, 2017, from Betygsättning: <https://www.skolverket.se/bedomning/betyg>
- Statens skolverk . (2011b). *Kunskapsbedömning i skolan - Praxis, begrepp, problem och möjligheter*. Stockholm: Statens skolverk.
- Tengberg, M. (2016b). Rätt mätt? *Kapet*, 12(1), pp. 21-37.
- Wells, C. S., & Wollack, J. (2003). *An instructor's guide to understanding test reliability*. . university of Wisconsin, Testing & Evaluation Services. Madison: University of Wisconsin.
- Vetenskapsrådet. (2002). *Forskningsetiska principer inom humanistisk-samhällsvetenskaplig forskning* . Retrieved 12 17, 2017, from https://www.gu.se/digitalAssets/1268/1268494_forskningsetiska_principer_2002.pdf
- Wikström, C. (2013). *Konsten att göra ett bra prov - Vad lärare behöver veta om kunskapsmätning* (1:a ed.). Stockholm: Natur och Kultur.

Bilagor

Bilaga 1. Informationsbrev

Information om deltagande i studie om lärarens resonemang kring provkonstruktion

Du tillfrågas härmed om deltagande i denna undersökning.

Syftet med detta arbete är att få mera kunskap om hur lärare resonerar kring sitt konstruerande av prov i bokföring i ämnet företagsekonomi. Jag vill huvudsakligen undersöka hur lärarna tänker kring konstruktion av prov och bedömning av elevers kunskaper. Studien är en del av mitt examensarbete.

För att genomföra undersökningen intervjuar jag lärare i företagsekonomi 2 som konstruerar sina egna prov om bokföring. Intervjun tar ca 60 minuter. Intervjun kommer att spelas in och transkriberas. Fram tills dess att arbetet är klart kommer ljudfiler samt transkriberad text att finnas på min lösenordskyddade dator samt på diktafon som finns inlåst i min bostad. Efter att arbetat är bedömt och klart kommer jag att makulera rådata, såsom ljudfil och transkription. När uppsatsen är färdig är det möjligt för er att ta del av densamma. Jag kommer i uppsatsen aidentifiera era namn och andra uppgifter för att göra er medverkan så anonym som möjligt. Endast jag själv, min handledare samt övrig behörig personal vid Högskolan i Dalarna kommer att ha tillgång till rådata.

Ditt deltagande i undersökningen är helt frivilligt. Du kan när som helst avbryta ditt deltagande utan närmare motivering. Undersökningen kommer att presenteras i form av en uppsats vid Högskolan Dalarna.

Om du behöver ytterligare upplysningar så lämnas de av nedanstående ansvariga

Ronneby den 26/11 2017

Tim Svanberg
Lärarstudent
0733-91 77 53
Tisv04@hotmail.com
Fribergs väg 14. 372 31 Ronneby

Steven Hunter Lindqvist
Handledare
073-941 08 12
shl@du.se
Falun

Bilaga 2. Frågeformulär – Semistrukturerad intervju

Frågeformulär: Tim Svanberg

Inledande (basinformation)

4. Namn på lärare
5. Namn på lärarens skola
6. Undervisar du i företagsekonomi 2?
7. Hur länge har du arbetat som lärare?
8. Är du en behörig ekonomilärare?
9. Vilka ämnen undervisar du i det här läsåret?
10. Är bokföring ett område som bedöms enbart med prov eller med andra bedömningsformer?
11. Sätter du betyg på provet? Varför?

Huvudfrågor:

1. Vad önskar du mäta för kunskaper kring bokföring i provet?
 - Är det frågor som du skulle vilja inkludera på ett bokföringsprov som du inte inkluderar? Vilka typ frågor?
 - Varför (tid, storlek på provet, eleverna har inte tillräckliga kunskaper)
 - Mäter provet de olika kunskapsnivåer du vill mäta?
 - Hur hittar du en modell för att mäta denna kunskap i prov?
 - Har du olika typer av frågor på bokföringsprovet som du har valt för att få olika kunskapsnivåer?
 - Kan samma typ av frågor användas för olika betygsnivåer på bokföringsprovet?
2. Hur resonerar du kring provkonstruktion för att mäta olika betygsnivåer kopplat till kriterierna i kursen för just bokföring?
 - Har du något sätt att rätta provet så att varje fråga på bokföringsprovet resulterar i E,C,A?
 - Får alla frågor samma tyngd eller viktar du frågorna på bokföringsprovet?
 - E, C, A-frågor?
 - Hur resonerar du kring D och B på bokföringsprovet
 - Vilka för/nackdelar kan du hitta med ditt bedömningssystem
 - Behöver provresultatet kompletteras med ytterligare underlag för att kunna sätta betyg på bokföringsprovet?
 - Hur kan elevernas resultat på bokföringsprovet översättas till en förmåga i kunskapskraven?
 - Använder du dig av sambedömning på bokföringsprovet?
3. Vilka för/nackdelar ser du med prov för att mäta olika kunskapsnivåer på bokföring?
 - Vari ligger utmaningarna/svårigheterna?
 - Kan alla betygsnivåer i bokföringsprov för just bokföringsmoment?
4. -Om eleven presterar på detta moment men inte i andra moment. Hur sker betygsättning då?

Kan jag ta kontakt med dig igen om jag behöver ytterligare upplysningar?