

Den vetenskapliga otryggheten

En studie i systematisk utvärdering av examensarbeten

Pedher Johansson och Marie Nordström

SAMMANFATTNING

Vi presenterar en metodik där SOLO används till att klassificera grad av vetenskaplighet i studentuppsatser. Vår studie har kunnat påvisa att genomförda förändringar i en utbildning haft effekt och dessutom har den gett oss värdefulla kunskaper om systematiska kvalitetsbrister. Målet är att vår metodik ska vara ett effektivt verktyg i ett systematiskt kvalitetsarbete.

Vi resonerar vidare kring betydelsen av operationalisering i en uppsats och vikten av en kollegial diskussion när man vill åstadkomma kvalitetshöjande åtgärder på en utbildning.

Introduktion	2
Val av Taxonomi	2
Metod	4
Anpassning av SOLO	4
Datainsamling	5
Resultat	6
Diskussion	8
Val av aspekter	8
SOLO som metod	8
Koppling mot frågeställningar	8
Reflektioner	9
Referenser	10
Appendix A	11

Introduktion

Studenter från Datavetenskap ska vara erkänt duktiga i sin profession när de lämnar våra utbildningar. Detta mål har länge präglat våra utbildningar och gör så fortfarande. Våra studenter är också duktiga på det mjukvarumässiga hantverket. Däremot, efter i olika roller varit inblandad när studenterna gör sina examensarbeten på kandidatnivå har vi över åren sett tydliga brister i uppsatserna. Speciellt när det kommer till vetenskapligt angreppssätt och resonering. Studenterna uppvisar en osäkerhet när det kommer till utvärdering, diskussion och analys.

Villkoren för kandidatarbetena har genomgått flera förändringar de senaste åren. Den första förändringen gjordes 2012. Då infördes ett nytt upplägg, med en ny pedagogik och en mer reglerad examinationsform. Året efter kom sedan utvärderingen av kandidatprogrammet utförd av UKÄ som gav programmet kritik. I första hand var det bristande vetenskapligt arbetssätt som konstaterades. Detta inledde ett förändringsarbete av programmet vilket innebar ett ökat teoridjup och en kurs i vetenskaplig metod. Exjobbskursen fick också tydligare examenskriterier. De första studenterna på det reviderade programmet skrev sitt examensarbete 2017.

Med de genomförda förändringarna borde studenterna vara bättre förberedda. Vår magkänslan sa också att uppsatserna blivit något bättre. Frågan var om denna magkänsla enbart berodde på att vi ville att så skulle vara fallet, eller om vi också kunde se tydliga förändringar i studenternas angreppssätt i uppsatsskrivandet. Vår hypotes var att brister i diskussionsdelen beror på en oklar frågeställning och en bristfällig uppfattning om vad det innebär att sätta in problemet i en kontext och använda befintliga arbeten som utgångspunkt för det egna arbetet. Därmed blir också operationaliseringen av frågeställningen, dvs metodavsnittet bristfälligt.

Med detta som utgångspunkt initierades ett projekt med två syften. Dels att utveckla ett verktyg för att kvantifiera graden av vetenskapligt angreppssätt med avseende på operationalisering, litteraturbehandling och diskussion. Dels att se om insatserna som gjorts på utbildningen haft effekt. Projektet finansierades av Umeå Universitet via PUNKTUM. I förlängningen är vårt mål att systematiskt utvärdera och utveckla alla våra utbildningars innehåll utifrån karaktären av eventuella brister vi kan identifiera med hjälp av resultat från denna studie.

Val av Taxonomi

Det förekommer flera etablerade taxonomier i utbildningsvetenskapliga sammanhang som har olika utgångspunkter. De flesta utgår dock från arbeten av Bloom, Piaget och Perry.

Den mest kända är nog Blooms taxonomi (Bloom 1954) som är ett sätt att med olika verb kategorisera en hierarki av abstraktionsnivåer i lärandesammanhang. I det här sammanhanget är det del 1, beskrivningen av kognitiva förmågor, som skulle kunna vara intressant eftersom det rör på vilken abstraktionsnivå individen kan förhålla sig till kunskap (se tabell 1). Taxonomin har även använts i snarlika studier, exempelvis för att avgöra den epistemologiska nivån på diskussionen (Granello 2001).

Piagets läror kring personens utvecklingsstadium ligger till grund för flera taxonomier såsom Model of Hierarchical Complexity (MHC) och SOLO , Structure of Observed Learning Outcome (Biggs and Collis 1982)(Commons et al. 1998; Stålné et al. 2016). MHC tittar på vilka uppgifter som kan utföras i något av de 14 stadier som anges. Stadierna 9 till 12 är de som kan sägas beröra högre utbildning (se tabell 2).

Tabell 1 Blooms originaltaxonomi

Knowledge	Faktakunskaper	Repetera någon annans definition av en princip
Comprehension	Förståelse	Förklara principen genom exempel på dess användning inom andra områden
Application	Tillämpning	Personligen använda principen på procedurer i verkliga situationer
Analysis	Analys	Kan separera fakta från antaganden i principer
Synthesis	Syntes	Kombinerar ett antal principer till en ny fungerande strategi
Evaluation	Värdering	Bedöma användningen av den nya strategin

Tabell 2 Stadier i MHC

9	Abstract	Stereotyper, kvantifiering (alla, ingen, alltid), abstrakta variabler
10	Formal	Två abstrakta variabler koordineras (typiskt orsakssamband: om x så y), logisk deduktion, referens till empiriska bevis, enkla samband formuleras (lagar i fysik, funktioner i en variabel $f(x)$)
11	Systematic	Minst två formella relationer koordineras för att skapa ett sammanhängande system ($f(x,y)$) typiska ord: kontext, ideologi, struktur, kultur, multivariat, system
12	Metasystematic	Kapacitet att diskutera egenskaper och karakteristika hos system (sammanhängande, begränsade). Jämföra och koordinera olika system (kombinera, kontrastera, visa på interaktion)

Teorier kring epistemologisk utveckling utgår från Perry's ursprungliga studie (Perry 1970). Med undervisning och lärande som utgångspunkt definieras en hierarki för kognitiv utveckling med olika stadier. Den ursprungliga studien refereras ofta, men kan bl.a. kritiseras för det snäva urvalet av respondenter. Underlaget består av vita medelklassmän i åldern 18–25 på Harvards universitet. Teorierna har därför såväl förfinats och utökats där bland annat Belenky (Belenky et al. 1986) varit framträdande.

Den taxonomi vi i slutänden valde var SOLO. SOLO är en beskrivning av hur information koordineras, och vilken komplexitet den har. I den här modellen är det studentens uppvisade resultat som värderas och inte studentens faktiska förståelse eller utvecklingsstadium som t.ex. Piaget gör (Brabrand and Dahl 2009; Chan et al. 2002; Stålné et al. 2016). Detta strävar alltså efter att fastlägga den komplexitet som finns i ett specifikt resultat av utförandet av en uppgift, vid ett specifikt tillfälle. Denna taxonomi har använts för en mängd olika områden och olika slag av information. T.ex. har Campbell använt SOLO för att karakterisera den kognitiva strukturen på uppsatser i sin helhet (Campbell et al. 1998). Brabandt och Dahl beskriver användningen av SOLO för att analysera progression i kursplaner (Brabandt). För en gedigen genomgång av bakgrundteori, styrkor, begränsningar och förslag till modifikationer av SOLO hänvisas till Stålné et.al (Stålné et al. 2016).

I SOLO indelas komplexiteter in i fem nivåer (se tabell 3). Transitionen mellan stadierna *Multistructural* och *Relational* är ett kvalitativt skift, eftersom det i det senare stadiet sker en syntes till en större helhet (Stålné et al. 2016).

Tabell 3 SOLO-taxonomin fem nivåer.

Nivå 1	Prestructural	Studenten missuppfattar problemet eller använder tautologier i resonemanget.
Nivå 2	Unistructural	En aspekt eller orsak specificeras.
Nivå 3	Multistructural	Flera aspekter/orsaker anges, men koordineras ej till en sammanhängande helhet.
Nivå 4	Relational	Ett kvalitativt skift från tidigare nivåer, nu demonstreras hur olika aspekter interagerar i ett system.
Nivå 5	Extended abstract	Förståelsen generaliseras till andra och större kontexter på en högre nivå av komplexitet.

För att sammanfatta så tittar Blooms taxonomi alltså på kunskapspresentation, dvs. hur en person kan beskriva och använda sin kunskap om ett visst fenomen. MHC fokuserar på tillämpning av kunskaper i olika stadier av kunskap. Vår studie inriktar sig i första hand på användningen av olika vetenskapliga redskap. SOLO som tittar på den strukturella komplexiteten i en färdig produkt passar därför bra in på de komponenter som ska observeras.

Metod

I vår studie bryter vi alltså ner uppsatsinnehållet i tre specifika komponenter, operationalisering, litteraturanvändning och diskussion. Att se på vilken komplexitetsnivå en författare lyckas formulera sin frågeställning är erkänt svårt (Rienecker 2003) varför vi valde att avstå den aspekten. De övriga valdes i huvudsak baserat på att det var komponenter vi tyckte oss se brister i.

Anpassning av SOLO

Baserat på tidigare forskningsresultat (Chan et al. 2002; Trigwell and Prosser 1991) beslutade vi oss för att använda undernivåer. Användningen av SOLO får på det sättet en bättre *interrater reliability*, dvs överensstämmelse mellan bedömare. Vi valde att dela in de två mittersta (och mest förekommande) nivåerna, *Multistructural* och *Relational* i två, med noteringen *low* och *high*. Chan et.al. (Chan et al. 2002) använde tre undernivåer, vilket vi ansåg skulle försvåra klassificeringen och inte tillföra något till våra frågeställningar.

För att validera utvärderingsprotokollet utfördes två pre-tester. I första pre-testet valdes tre uppsatser från Uppsala Universitet ut som bedömdes av oss båda. Formuleringar och exempel i protokollet justerades för att bli mer entydigt.

Vi såg också ett behov av att göra avgränsningar i vad som skulle bedömas. Eftersom det enbart var operationalisering, litteraturanvändning och diskussion vi skulle bedöma var det också nödvändigt att bortse från andra kvalitetsaspekter. Vi ville bedöma studentens uppvisade förmåga att använda vetenskapliga verktyg, inte vad de använde dem till. Beaktade och icke beaktade aspekter redovisas i tabell 4.

I det andra pre-testet hämtades tre uppsatser från Mälardalens högskola. Dessa bedömdes dels av oss två och dels en tredje lärare med erfarenhet av examination av kandidatuppsatser. Resultatet av detta test blev mindre justering, t.ex. ordval vid beskrivningen av nivåerna.

Generellt visade resultaten av pre-testerna på en god samsyn mellan bedömarna. Svårigheterna var som tidigare nämnts tydligheten i avgränsningarna. Det fullständiga verktyget återfinns i appendix A.

Tabell 4 Beaktade faktorer ur respektive aspekt samt vilka begränsningar som gjorts.

<i>Operationalisering:</i> Studentens beskrivning av hur problem ska besvaras	
Beaktar	Beaktar inte
- hur olika faktorer i problemlösningen motiveras, problematiseras, och värderas.	- om operationaliseringen matchar frågeställningar - om operationaliseringen är vedertagen eller ej - implicit formulering i frågeställningar eller resultat - form (dvs var i rapporten beskrivningen finns)
<i>Litteraturhantering:</i> Studentens behandling av refererad litteratur.	
Beaktar	Beaktar inte
- hur olika delar av studien värderas i relation till känd litteratur	- litteraturens relevans - litteraturens kvalitet - om studenten verkar kunnig eller ej - form (dvs var i rapporten litteraturen används)
<i>Diskussion:</i> Studentens diskussion av resultat	
Beaktar	Beaktar inte
- hur resultaten och metod generaliseras och värderas	- om diskussionen besvarar frågeställningar - diskussionens ämnesmässiga kvalitet - resultatens kvalitet - form (dvs var i rapporten beskrivningen finns)

Datinsamling

Materialet till studien är de kandidatuppsatser som vid tiden för studien var godkända och publicerade i DiVA. Kandidatutbildningen startade 2007 och kursplanen för de självständiga arbetena var 5DV096. Genomförandet av de självständiga arbetena gjordes om 2012 med en ny kursplan, 5DV129. De främsta förändringarna var tydligare deadlines och en förändrad handledning. De tidigare beskrivna förändringarna av utbildningen skedde 2014. Detta innebar även nya rutiner för examination med tydliga examinationskriterier. De första studenter som följde detta upplägg gjorde sina självständiga arbetena 2017 på kursplanen 5DV165. Detta blev därför intressanta uppsatsgrupper att studera. Uppsatserna valdes ut baserat på enligt vilken av de tre kursplaner de genomförts under. 2017 blev 15 personer blev färdiga med examensarbetet på 5DV165. Samma antal uppsatser slumpades därför ut ur de tidigare två grupperna, 5DV096 och 5DV129. Detta redovisas i tabell 5.

Tabell 5 Det använda antalet och det totala antalet uppsatser i varje kategori

Grupp	# totalt	# i urval	Andel av befintliga uppsatser
5DV096	20	15	75%
5DV129	79	15	19%
5DV165	15	15	100%
	114	45	39%

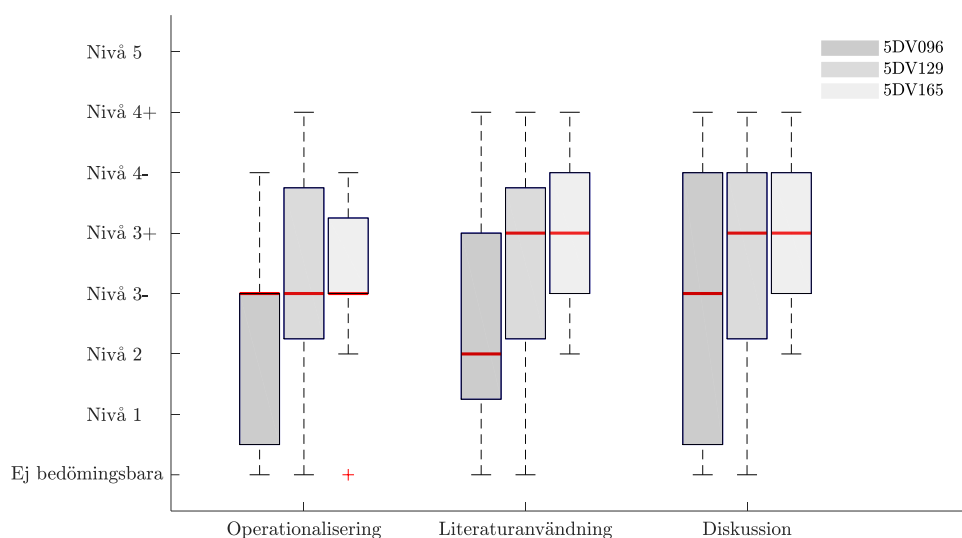
Det finns en liten eftersläpning mellan första och andra gruppen i att två studenter genomförde sin uppsats enligt den gamla kursplanen, trots att den nya startat. Fördelningen över vilket år en student skrev uppsatsen ses i tabell 6.

Tabell 6 Antalet uppsatser i varje kategori, separerat över år uppsatsen skrevs.

År	5DV096	5DV129	5DV165
2010	7		
2011	6		
2012	1	1	
2013		3	
2014	1	4	
2015		4	
2016		3	
2017			15

Resultat

Under studiens gång blev det tydligt att vissa uppsatser inte gick att bedöma enligt de kriterier vi utarbetat i de två pre-testerna. Den lägsta nivån i SOLO som innebär att en student missuppfattar uppgiften, var helt enkelt inte tillämplig i de fall studenterna inte hade försökt att tillämpa ett vetenskapligt angreppssätt. Dessa uppsatser var typiskt resultatet av en produktutveckling, där förvisso kriterier för funktionalitet kunde vara uppställda, men där det saknades en problematisering och en strategi för att uppnå dem. Vi bestämde därför att dessa uppsatser skulle klassificeras som "ej bedömningsbara". Samtliga resultat av vår studie redovisas i figur 1.



Figur 1 Spridning av uppnådd nivå enligt SOLO sett till respektive grupp med redovisat medianvärde (rött). Notera även en outlier för 5DV165 angående Operationalisering.

Populationerna är små och vi får därför nöja oss med att titta på eventuella trender. Vi tycker oss se indikationer på att det skett en förskjutning mellan grupperna. Ingångspunkten i förändringsarbetet av utbildningen var att höja nivån på de självständiga arbetena där det viktigaste var att se till att studenterna är så pass förberedda inför examensarbetet att de har förmågan att skriva uppsatser som håller en acceptabel vetenskaplig nivå. Av särskilt intresse är därför att studera andelen studenter som inte når upp till multistrukturall nivå vilket vi särredovisar i tabell 7. Andelen ser ut att ha minskat i alla tre aspekterna. Av tabell 8 kan man dock se att gruppen 5DV165 även signifikant har

fler som når upp till en multistrukturall nivå jämfört med 5DV096 när det gäller litteraturanvändning och diskussion.

Tabell 7 Andel under nivå 3 och andelen ej bedömningsbara i respektive grupp.

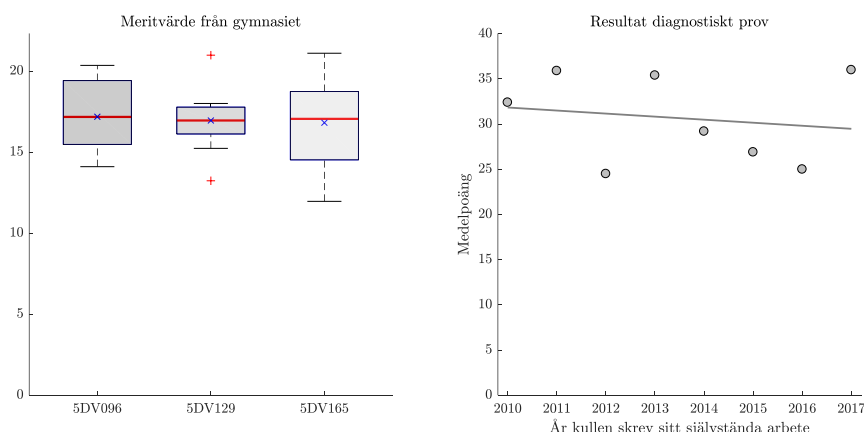
Andel under nivå 3 (inklusive ej bedömningsbara)			
Grupp	Operationalisering	Litteratur	Diskussion
5DV096	40,0%	60,0%	46,7%
5DV128	26,7%	26,7%	26,7%
5DV165	13,3%	13,3%	6,7%

Andel ej bedömningsbara			
Grupp	Operationalisering	Litteratur	Diskussion
5DV096	26,7%	20,0%	26,7%
5DV128	20,0%	6,7%	13,3%
5DV165	6,7%	-	-

Tabell 8 Korsjämförelse med Mann-Whitney U-test mellan urvalsgrupperna med avseende på hur många som når upp till minst nivå 3. * $p < 0.05$

	5DV129	5DV165
Operationalisering	p	p
5DV096	0.461	0.167
5DV129		0.496
Litteraturanvändning		
5DV096	0.074	0.019*
5DV129		0.496
Diskussion		
5DV096	0.275	0.027*
5DV129		0.212

En viktig fråga är hur representativ de olika populationerna är för respektive studentgrupp och hur jämförbara de är. För att se om studentgruppen förändrats över tid har vi tittat på meritvärden från gymnasiet i de studerade grupperna. Meritvärdena har inte ökat. Vi kan heller inte se någon ökande trend i de diagnostiska prov som Inst. för matematik och matematisk statistik genomför för att alla nya programstudenter. Dessa data redovisas i figur 2. Sammantaget kan vi inte se att studentgrupperna generellt har bättre förutsättningar inför utbildningen.



Figur 2 Meritvärden från gymnasiet i respektive studerad grupp, samt resultat på det diagnostiska matematikprovet för respektive kull.

Diskussion

Val av aspekter

Vi valde att titta på operationalisering, litteratur och diskussion, vilket visade sig vara både genomförbart och utslagsgivande. I uppsatsskrivande betonas gärna frågeställningens betydelse för såväl metodval, som diskussion och det är förvisso viktigt, men vår bedömning är att en frågeställning i sig kan vara svår att klassificera men att effekten av frågeställningen i såväl operationalisering och diskussion inte är det. Angående litteraturanvändning i argumentation kan givetvis en uppsats mycket väl använda annat underlag för ett fullvärdig vetenskaplig argumentation, men likväl blir litteraturanvändningen en viktig markör i hur mycket engagemang som lagts ned på att sätta studien i ett sammanhang. Detta i sin tur borde rimligen ge en förutsättning för att resultaten får en högre vetenskaplig kvalitet, även om vi inte tittat specifikt på det sambandet.

SOLO som metod

Vår utgångspunkt i det här arbetet var utbildningens kvalitet. Att då titta på den yttersta produkten av utbildningen med SOLO anser vi vara intressantare än att göra en epistemologisk analys av en students utveckling som är ansatsen i exempelvis Bloom. Med fokus på produkten kan vi värdera utfallet av den undervisning, instruktioner och betygsättning som är gjord. Att vara noga med att varje nivå måste styrkas med ett exempel i texten gör dessutom att en eventuell avsaknad av det mest basala kan bli pålägsamt uppenbart. Vad studenterna skulle ha varit kapabla till med en annan undervisning kan vi inte veta. På samma sätt var det viktigt att lyfta ut den inomvetenskapliga kvaliteten från bedömningen. Vad som avhandlas i en uppsats och med vilken metod blir oundvikligt en studie i handledar- och examinatorskompetens vilket är en separat fråga och den borde rimligtvis utvärderas annorlunda. Med det förhållandevis instrumentella angreppssättet med fokus på produkten och inte vare sig förmåga eller kvalitet kunde vi se i pre-testerna att bedömningen blev tämligen konsistent mellan bedömare.

De kriterier vi använt (Appendix A) kan givetvis diskuteras. Vi kan inte bortse från att det mycket väl kan finnas kulturella skillnader mellan såväl ämnesdiscipliner som lärosäten. Vi kan heller inte bortse att vårt förslag är baserat utifrån våra erfarenheter. En del tycker kanske att kraven är allt för låga, medan andra tycker att det är orimligt i en viss typ av uppsats. Målet har dock varit att göra kriterierna oberoende av ämne. Eftersom vårt fokus dessutom var de relativa förändringarna är heller inte den exakta skrivningen eller kvalificeringen central så länge klassificeringen görs konsekvent och där finns det sannerligen fallgropar. En svårighet som direkt uppstår är om det räcker med ett exempel eller om där bör vara ett genomgående drag. Ett annat är om det finns flera faktorer som är uppfyllda för en högre nivå men det saknas ett konkret exempel för en lägre nivå. Vi vill dock betona att det handlar om struktur i konkreta exempel i den bedömda texten. Det är nivån i struktur som ska vara avgörande när ord som exempelvis ”analys” ska klassificeras. Vi rekommenderar därför att man så långt det är möjligt är strikt och undviker bedömningslägen.

Koppling mot frågeställningar

Vad avser studiens första syfte, att identifiera karaktären hos brister samt även kvantifiera graden av vetenskaplighet, visade sig metoden vara effektiv. Den tydligaste insikten var att många av våra tidigare uppsatser inte hade ett vetenskapligt angreppssätt. Här har nog undervisningstradition och ämneskultur en stor betydelse. Som verksam på ett universitet är bristande vetenskaplig metod naturligtvis en smärtsam insikt.

Ett tydligt mönster i svaga uppsatser är en otrygghet kring den vetenskapliga begreppsapparaten och metodiken. Uppsatsen får karaktären av en avrapportering av ett utfört experiment eller produktutveckling. Inledningen beskriver snarare beställarens krav; operationaliseringen reduceras till kriterier för funktionalitet eller en

experimentuppställning, och diskussionen till en sammanfattning av vilka krav som kan anses vara uppfyllda, dock utan reflektion eller analys. Därtill kommer gärna inslag av meta-texter där ett avsnitt inleds med att beskriva vad som ska ingå i avsnittet. Något starkare uppsatser lyckas reda ut några operationaliserande faktorer och kan åtminstone tematisera nödvändig begrepp, dock utan någon ingående litteraturgenomgång eller beskrivning av orsakssamband. Uttalade litteraturstudier stannar inte sällan i en tematisering av ett antal källor. Däremot saknas en beskrivning av metod. Även en syntes av materialet saknas. Det sker sällan en problematisering i diskussionen.

En annan insikt är att vi på Datavetenskap har varit för lite stödjande vad gäller struktur i uppsatser. En studie av Helen Sword (Sword 2012) indikerar att artiklar i datavetenskap ofta har en hybrid- eller egen struktur. Orsaken till detta kan diskuteras, men inte sällan lånas metodik från närliggande discipliner inom nya tvärvetenskapliga grenar av ämnet. Undervisningen har också betonat att metoden ska bestämma strukturen och vi har därför varit motvilliga i att ge någon form av mall. I många fall lämnar det dock studenterna i ovisshet hur presentationen av arbetet bör ske. Teorier och resultat (Biggs 2011; Rust et al. 2003) kring constructive alignment inverkar möjligen även här, men även studier kring att diskutera och tydliggöra uppsatsinslagen kontinuerligt under kursen (Kapborg and Berterö 2002). En given struktur visar vad vi förväntar oss och vad studenten kommer att bli bedömd på samtidigt som en mall naturligt gör en återkommande diskussion nödvändig.

Vår hypotes att den bristande diskussionen har sitt ursprung i en bristande frågeställning har stärkts. Trots att vi inte tittat direkt på frågeställningen är avsaknaden av tydliga faktorer i att lösa ett problem en tydlig markör som påvisar en vag frågeställning. Avsaknaden av faktorer gör i sin tur att kunskapsluckor och därmed behov av inläsning i litteraturen inte blir uppenbar för studenten. I diskussionen finns sedan mycket lite att diskutera annat än kravuppfyllnad. Den slutsats vi drar är att större vikt bör läggas på hur man utformar en operationalisering utifrån en målbild.

Studiens andra syfte, att kunna avgöra om förändringar på utbildningen haft effekt, gick över förväntan. Urvalet är litet, men studien visar ändå på tydliga trender och i några fall även signifikanta skillnader. Utbildningen har haft som mål att höja lägstanivån, och våra resultat tyder på att så också har skett. I den tidigaste gruppen hade vi så pass många som en fjärdedel som inte hade en vetenskaplig ansats. Denna grupp har näst intill försvunnit. Den tydligaste förändringen har skett inom litteraturhanteringen. Att det stora hoppet skedde redan till den andra gruppen, dvs innan vi på allvar arbetat med att höja den vetenskapliga nivån, kan nog förklaras att det var den lägst hängande frukten. Tittar vi på operationaliseringen har vi förvisso fått markant färre i gruppen med en nivå under 3, men medianen är densamma. Tittar vi på diskussionen, dvs upprinnelsen till denna studie, kan vi se samma trend. Volymen på lägst nivå har minskat, men medianen har bara sakta rört sig uppåt. Likväl är resultaten mycket glädjande. Därtill har studien gett oss värdefulla insikter som gör att vi förhoppningsvis även på sikt kan lyfta kommande uppsatsarbetena som helhet.

Reflektioner

Vi ska ställa höga krav på våra utbildningar. En högkvalitativ utbildning ska förena praktisk färdighetsträning, med verktyg för ett livslångt lärande och ett kritiskt förhållningssätt. Tonvikten varierar mellan utbildningar och vi inom datavetenskap kommer alltid att prioritera ett stort mått av teknikkunnande och praktisk färdighet. Det innebär dock inte att vi kan dra ner på ambitionen att studenterna, i sina examensarbeten, måste kunna förklara varför lösningen ser ut som den gör och problematisera sitt arbete och sina resultat.

Att utvärdera en utbildning innebär ett stort mått av självrannsakan och att förändra en institutions kultur tar tid. Viktigt är därför att aldrig skuldbelägga vare sig tidigare handledare, examinatorer eller studenter. En betygsättning är gjord utifrån de instruktioner och den undervisning som tidigare bedrivits. Liksom i all undervisning måste även det

självständiga arbetet förhålla sig till studenternas ingångskunskaper. Det egentliga förändringsarbetet börjar med en kollegial diskussion, gärna baserad på konkreta exempel.

Vi har i vissa avseenden varit framgångsrika i att höja kvalitet på våra uppsatser, men det beror inte i huvudsak på förändringar som gjorts i undervisningen eller handledningen under uppsatskursen utan för att det finns en samsyn kring målbilden, kanske inte i hela kollegiet, men åtminstone bland lärare på ett antal nyckelkurser och bland examinatorerna på examensarbetet. Med resultaten från denna studie är nu vår målsättning att bredda diskussionen att omfatta ytterligare utbildningsprogram och att involvera fler lärare i diskussionen där även konkreta exempel kan användas.

Referenser

- Belenky, M. F., Clinchy, B. M., Goldberger, N. R., and Tarule, J. M. (1986). *Women's ways of knowing: The development of self, voice, and mind*: Basic books New York.
- Biggs, J. B. (2011). *Teaching for quality learning at university: What the student does*: McGraw-hill education (UK).
- Biggs, J. B., and Collis, K. F. (1982). *Evaluation the quality of learning: the SOLO taxonomy (structure of the observed learning outcome)*: Academic Press.
- Bloom, B. S. (1954). "Taxonomy of educational objectives. Handbook 1, Cognitive domain."
- Brabrand, C., and Dahl, B. (2009). "Using the SOLO taxonomy to analyze competence progression of university science curricula." *Higher Education*, 58(4), 531-549.
- Campbell, J., Smith, D., and Brooker, R. (1998). "From conception to performance: How undergraduate students conceptualise and construct essays." *Higher Education*, 36(4), 449-469.
- Chan, C. C., Tsui, M., Chan, M. Y., and Hong, J. H. (2002). "Applying the structure of the observed learning outcomes (SOLO) taxonomy on student's learning outcomes: An empirical study." *Assessment & Evaluation in Higher Education*, 27(6), 511-527.
- Commons, M. L., Trudeau, E. J., Stein, S. A., Richards, F. A., and Krause, S. R. (1998). "Hierarchical complexity of tasks shows the existence of developmental stages." *Developmental Review*, 18(3), 237-278.
- Granello, D. H. (2001). "Promoting cognitive complexity in graduate written work: Using Bloom's taxonomy as a pedagogical tool to improve literature reviews." *Counselor Education and Supervision*, 40(4), 292-307.
- Kapborg, I., and Berterö, C. (2002). "Critiquing Bachelor candidates' theses: are the criteria useful?" *International Nursing Review*, 49(2), 122-128.
- Perry, W. G. (1970). "Forms of Intellectual and Ethical Development in the College Years: A Scheme. Holt, Reinhart and Winston." *Inc., New York*.
- Rienecker, L. (2003). "Problemformulering". City: Liber, Malmö.
- Rust, C., Price, M., and O'Donovan, B. (2003). "Improving students' learning by developing their understanding of assessment criteria and processes." *Assessment & Evaluation in Higher Education*, 28(2), 147-164.
- Stålne, K., Kjellström, S., and Utriainen, J. (2016). "Assessing complexity in learning outcomes—a comparison between the SOLO taxonomy and the model of hierarchical complexity." *Assessment & Evaluation in Higher Education*, 41(7), 1033-1048.
- Sword, H. (2012). *Stylish Academic Writing*: Harvard University Press.
- Trigwell, K., and Prosser, M. (1991). "Relating approaches to study and quality of learning outcomes at the course level." *British Journal of Educational Psychology*, 61(3), 265-275.

Appendix A

	Operationalisering	Litteratur	Diskussion
Avgränsning	Studentens beskrivning av hur problem ska besvaras Beaktar inte: - om operationaliseringen matchar frågeställningar - om operationaliseringen är vedertagen eller ej - implicit formulering i frågeställningar eller resultat - form (dvs var i rapporten beskrivningen finns)	Studentens behandling av refererad litteratur Beaktar inte: - litteraturens relevans - litteraturens kvalitet - om studenten verkar kunnig eller ej - form (dvs var i rapporten litteraturen används)	Studentens diskussion av resultat Beaktar inte: - om diskussionen besvarar frågeställningar - diskussionens ämnesmässiga kvalitet - resultatens kvalitet - form (dvs var i rapporten beskrivningen finns)
Prestructural	--	Upprepar information utan att visa en förståelse. Exempel: Mycket citat. "räknar upp" fakta från olika källor.	--
Unistructural	En aspekt eller orsak specificeras. Exempel: "vilket är det bästa sättet att ..."	Visar enkel förståelse. Kan sammanfatta litteratur med egna ord.	Ett enkelt orsakssamband. Exempel: "A är större än B, alltså är A bättre".
Multistructural - low	Några faktorer anges till hur problemet ska besvaras	Viss tematisering inom problemområden.	Kan formulera resultaten i egna ord. Något orsakssamband med argumentation/kommentar.
Multistructural - high	En förklaring ges till hur faktorerna ska användas (gärna baserad på teoribildning).	Klarar av att se huvudidéer och resultat som direkt relaterar till det egna arbetet och gör en explicit koppling till frågeställningar.	Redogör för flera orsakssamband och komplexiteten i resultaten. En slutsats presenteras.
Relational - low	Mer operationaliserade faktorer specificeras där de tydligt/medvetet bidrar med olika komponenter i utredningen och hur de relaterar till varandra.	Kan identifiera komponenter, och dra slutsatser baserat på en analys av forskning eller idéer som presenteras. Bedömer ref. arbetets metod utifrån den egna metoden.	Resultaten kopplas ihop. Slutsatsen baseras på viss analys.
Relational - high	Förhållandet mellan komponenter görs explicit. Viss problematisering baserad på teoribildning.	Idéer och resultat från olika källor kan sammanföras tematiskt inom ett problemområde. Viss värdering av relevans för det egna problemet.	Kan sammanföra resultaten till en helhet. Kan ge exempel på tillämpningar av metoden/resultaten utöver den genomförda studien/resultaten. Kommentarer styrkor och svagheter med den genomförda studien i relation till andra resultat.
Extended abstract	En generell teoribildning bryts ned till faktorer som täcker det aktuella problemet. Faktorerna problematiseras och värderas innan de slutligen specificeras	Värderar studiens olika delar i relation till känd litteratur	Resonemang och generalisering kring tillämpligheten i en mer allmän kontext. Argumenterar för begränsningar, metodproblem, styrkor och svagheter utifrån teorimodellen.