

Egil Nodland, Kjemisk institutt, Universitetet i Bergen

Planlegging, utføring og rapportering av tiltak for utvikling av undervisning og læringsmiljø i
KJEM225/PTEK226/BIO301

Kursene som rapporten omtaler blir undervist ved UiB og Høgskolen på Vestlandet. Studentene ved UiB er som regel på masterprogram eller doktorgradsstipendiater i kjemi eller prosessteknologi. Studentene ved HVL deltar på etter- og videreutdanningsprogram for bioingeniører, i medisinske laboratorieanalyser.

Felles for alle kursene er en omfattende del om metoder til analyse av store datasett med mange, kovarianse variabler. I UiB-kursene anbefales det at studentene har de nødvendige forkunnskapene i grunnleggende lineær algebra. Begge kurs undervises av samme lærere og er bygget opp på klassisk vis. Undervisningen foregår som forelesninger, regneøvelser og obligatoriske data-øvinger. Etter at nødvendig teori er forklart i forelesninger, tar data-øvingene for seg prinsipalkomponentanalyse (PCA) som danner grunnlaget for visualisering av innholdet i store datasett. Partial Least Squares (PLS) brukes til å modellere sammenhengen mellom en ekstern respons og instrumentelle profiler. Klassifikasjonsteknikker som SIMCA og diskriminantanalyse tas opp. Data-øvingene foregår intenst over to uker med korte innleveringsfrister. All annen undervisning i fagene utgår. Alle oppgavene blir gjennomgått i plenum etter at alle studenter har levert sine første løsningsforslag. Studentene får skriftlig tilbakemelding på samtlige arbeider og leverer reviderte løsningsforslag inntil de faglige kravene er oppnådd. Avslutningsvis får studentene utlevert lærernes løsningsforslag.

I BIO301 ved HVL gis en grunnleggende innføring i lineær algebra. Vektor- og matriseregning presenteres, og det gis en innføring i egenverdier, egenvektorer og multippel lineær regresjon. Dette er en av tre hoveddeler i kurset. Del 2 tar for seg statistisk forsøksplanlegging og i den siste delen presenteres metodene for de fler-variable dataanalysene nevnt over. Kurset har tre lærere. Undervisningen er lagt opp med ukesamlinger der forelesninger, dataøvinger, artikkelanalyser og gruppearbeid er vektlagt. Mellom samlingene skal studentene arbeide med oppgaver og forberedelser til neste seminar, i tillegg til selvstudier. Det er 5 uker mellom del 1 og del 3.

Bakgrunn for tiltakene.

Å hente ut underliggende informasjon fra store datasett ved hjelp av flervariabel dataanalyse er som å be studentene om å gjennomgå en transformasjon som beskrevet av Meyer og Land [1], er vår erfaring. Flervariabel dataanalyse kan beskrives som et såkalt terskelkonsept som vil være med på å transformere studentene når de forstår tankegangen rundt latente variabler. Ved å ta spranget fra én- variabel-om-gangen til å håndtere mange variabler samtidig kan gi det Meyer og Land kaller en irreversibel oppfatning og forståelse av faget. Første gang BIO301 ble undervist i 2017 var en stor del av del 3 å lese og forstå vitenskapelige artikler om multivariat dataanalyse relevant for det medisinske laboratorium. Kurset ble avsluttet med en oppgave som involverte litteraturstudier og presentering av resultater i form av foredrag. Studentene var alt fra nyutdannede bioingeniører til dem med mer enn 20 års erfaring fra medisinske laboratorier. Det gikk mer tid med til artikkelanalyser og vurderinger om litteraturen var egnet og relevant nok for deres eget arbeidsfelt, enn til å anvende de aktuelle metodene til å analysere store datasett.

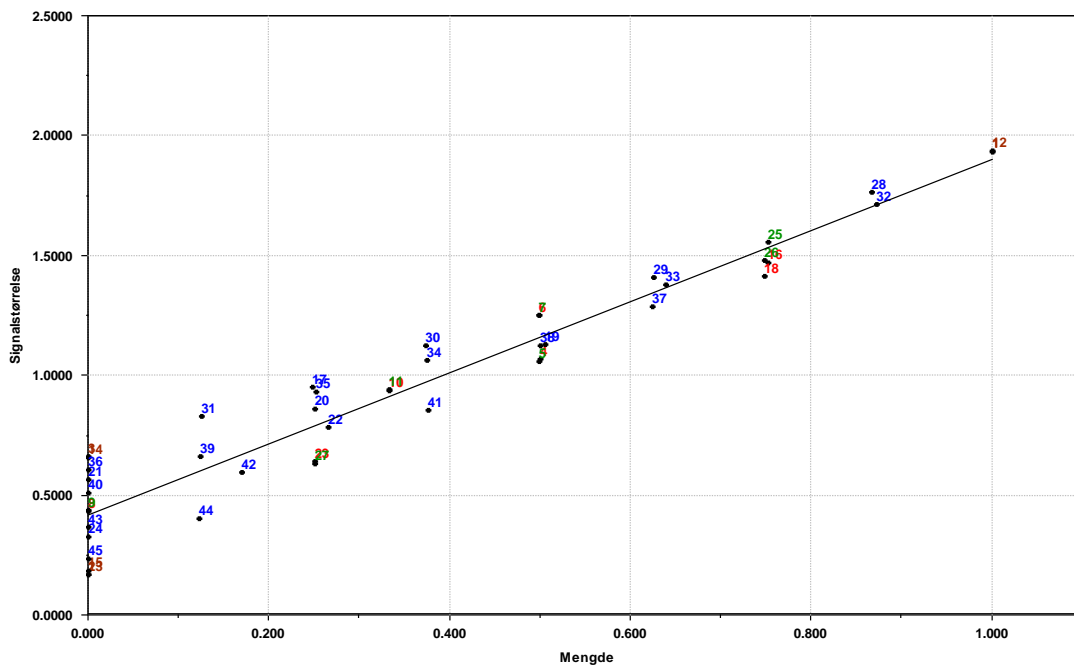
Studentene på UiB-kursene har ulike faglige bakgrunner. For PTEK226 er anbefalte forkunnskaper to emner i matematikk, herunder lineær algebra, samt et kurs i elementær statistikk. Studentene på KJEM225 har ofte kun brukerkurset i matematikk i bagasjen. De har til gjengjeld ofte bakgrunn i kurs i analytisk kjemi og instrumentell analyse som produserer de såkalte flervariabel datasettene som skal undersøkes. Selv om de ulike analyseapparatene de har benyttet i kjemikursene genererer flervariabel data, er de som regel opplært i å bruke kun en av variablene til såkalt univariat analyse. Likheter og ulikheter mellom forskjellige prøver blir da beskrevet ved bruk av middelveier og standardavvik.

Det matematisk-naturvitenskapelige fakultetet ved UiB har som et uttalt mål å modernisere undervisningen ved å legge om til former mer studentaktiv undervisning [2]. Det betyr bruk av siste generasjons klasserom og lærings- og undervisningsformer som ligner mer på laboratoriearbeid enn forelesninger. Laboratoriearbeid er velkjent for studentene på BIO301 og KJEM225. De bakenforliggende pedagogiske prinsippene hviler på samarbeid mellom student og lærer som veileder og ikke som den klassiske foreleseren. Såkalt aktivitetsbasert læring er en metode som legger stor vekt på å aktivisere studentane i mindre grupper «– Med denne metoden lærer studenten gjennom å gjere noko, i staden for berre å ta imot»[3]. Dee Fink beskriver at gode kurs blant annet innebærer at studenter blir utfordret til såkalte signifikante måter for å lære ved å bruke aktive læringsformer, har lærere som bryr seg om

faget, studentene, det å lære bort og det å lære. Gode kurs har lærere som vekselvirker godt med studentene [4,5]. Kjell Raaheim sier at «- Et sentralt kjennetegn på en god læringssituasjon er at du har en lærer som bryr seg og tror på deg» [6]. Tror en på disse uttalelsene er det lett å møte seg selv i døren, gjerne lettere enn å møte studenter i samme dør kl. 0815 til forelesing.

Undervisningen i flervariabel dataanalyse forsøker å aktivere såkalt inert kunnskap [1]. Dataanalysen er beskrevet av en av nestorene i faget til å være anvendt, kognitivt og filosofisk av natur for å motivere studentene til å benytte seg av dette i senere studier eller arbeidslivet [7]. Mønstergjenkjenning og visualisering av flervariable data er sentrale tema i faget. Spranget fra å lete etter mønstre basert på en variabel og flere objekter, til å se mønstre basert på et meget stort antall variabler og relativt få objekter, er som å passere Meyer og Lands terskel [1].

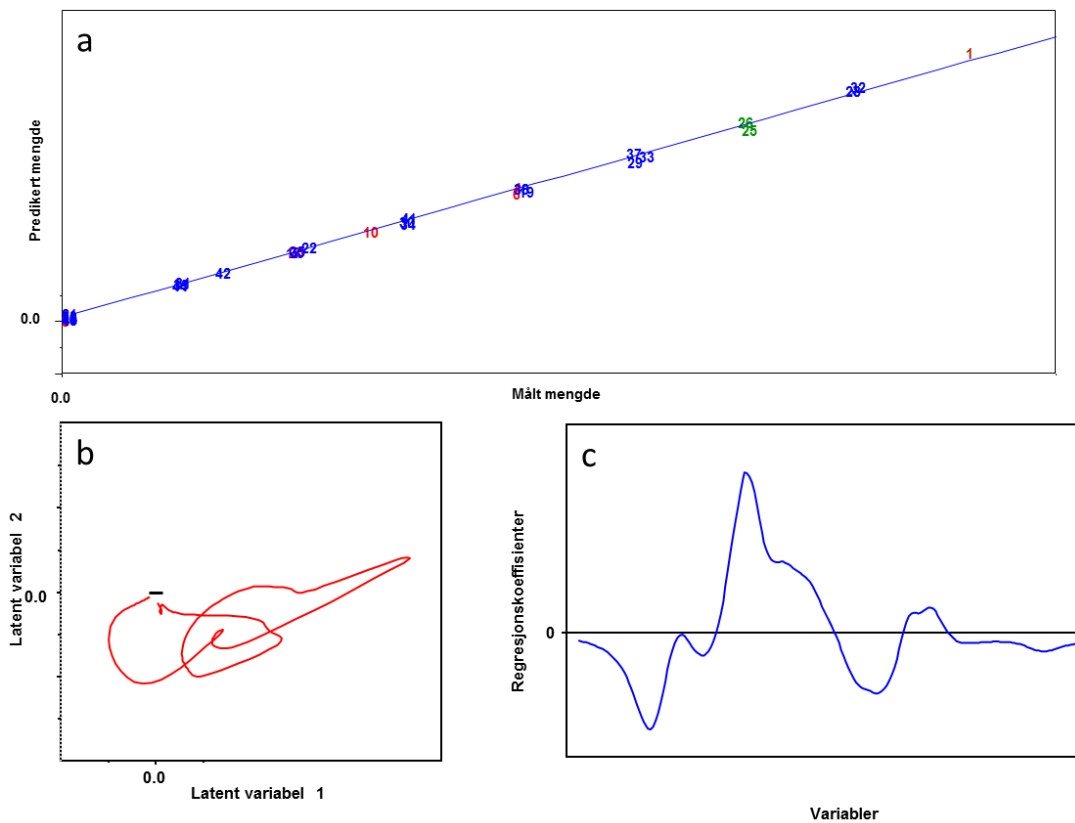
Figur 1. Mønster fra en variabel og flere objekter



Mønsteret i figur 1 er velkjent for mange studenter. Det viser at signalstørrelsen (y-aksen) for en variabel øker noenlunde lineært med mengde av en forbindelse i mange objekter og brukes

i kvantitative analyser. Når studentene benytter seg av alle variabler fra måleinstrumentene fremkommer ofte mønstre som vist i figur 2.

Figur 2



Fra figur 2a skal studentene lese at den kvantitative modellen blir bedre ved å bruke et lite antall latente variable som er lineærkombinasjoner av de originale variablene. I figur 2b skal de se hvilke av de originale variablene som bidrar mest til de latente variablene. Fra figur 2c skal de lese korrelasjonsmønsteret mellom de originale variablene og målt mengde av en kjent sambinding. Dette baserer seg på antagelsen om at studentene benytter seg av tidligere tilegnede kunnskaper og ferdigheter. Antagelser er roten til mye vondt. Målet er at studentene skal beherske en samtidig kvantitativ og kvalitativ analyse. Kolomitro og Gee skriver at i følge Blooms hierarkiske taksonomi må studenter først beherske de lavere nivåene før de kan nå de øvre nivåene [8].

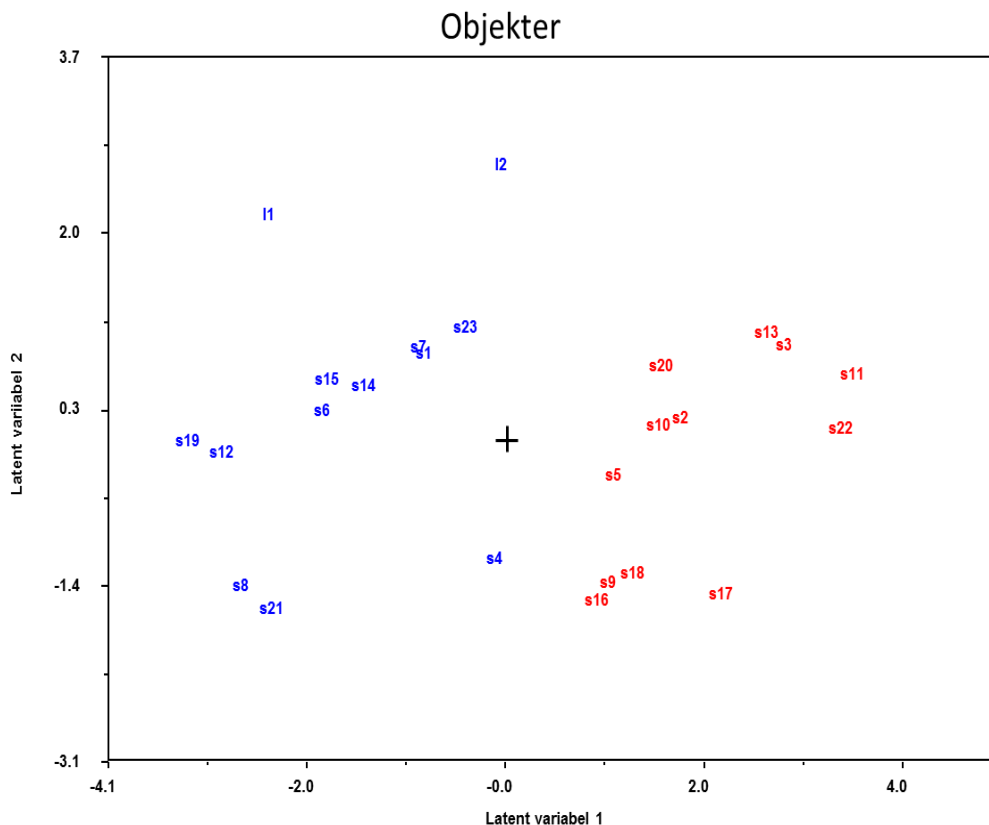
Tiltakene

Tiltakene som ble utført var derfor å gå tilbake i studentenes faglig spor til vi kom til «kjente steder», og ved å benytte oss av mye studentaktiv undervisning.

Et døme på det første tiltaket var å beskrive begrepet «lineær uavhengighet», som tidligere var antatt å være kjent. Det avstedkom kommentaren «Hvorfor har vi ikke lært dette før?» Deler av eksempelet på lineær uavhengighet var som følger; *Vektorene \mathbf{x} , \mathbf{y} , og \mathbf{z} er lineært uavhengige dersom det er umulig å skrive t.d. $\mathbf{x} = b_1\mathbf{y} + b_2\mathbf{z}$. Si at $\mathbf{x} = [1\ 2\ 3]$, $\mathbf{y} = [4\ 7\ 3]$ and $\mathbf{z} = [5\ 9\ 6]$. De er ikke lineært uavhengige fordi vi kan skrive $\mathbf{z} = \mathbf{x} + \mathbf{y}$. Er $\mathbf{z} = [-7\ -12\ -3]$ er de heller ikke lineært uavhengige fordi $\mathbf{z} = \mathbf{x} - 2\mathbf{y}$. Settes $\mathbf{z} = [0\ 13\ 1]$ er det umulig å uttrykke \mathbf{z} som en sum av de to andre vektorene, De tre vektorene er derfor lineært uavhengige. Lineær uavhengighet er gyldig for et hvert antall vektorer med et hvert antall elementer vektorene inneholder [9].*

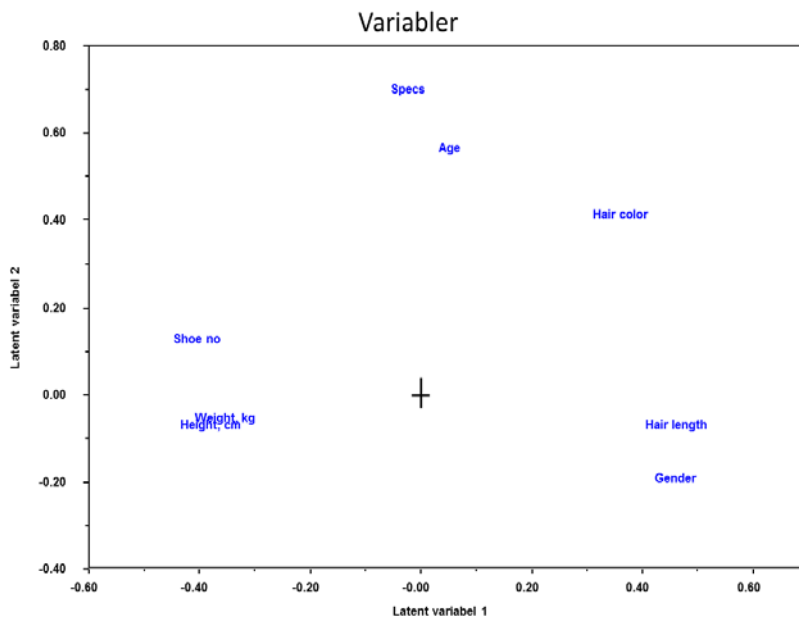
Tiltak to var å finne et datasett som kunne tolkes i plenum av studentene, uavhengig av deres alder, kjønn, faglige og kulturelle bakgrunn. Det ble samlet inn et persondatasett under første forelesning der studentene ble redusert til objekter, og åtte egenskaper fra hver av dem ble valgt som variabler. Variablene var deltagerens alder, skonummer, høyde, vekt, kjønn, hårlengde, hårfarge og om de brukte briller eller ikke. En såkalt prinsipalkomponentanalyse ble utført og tre figurer skulle tolkes i fellesskap.

Figur 3



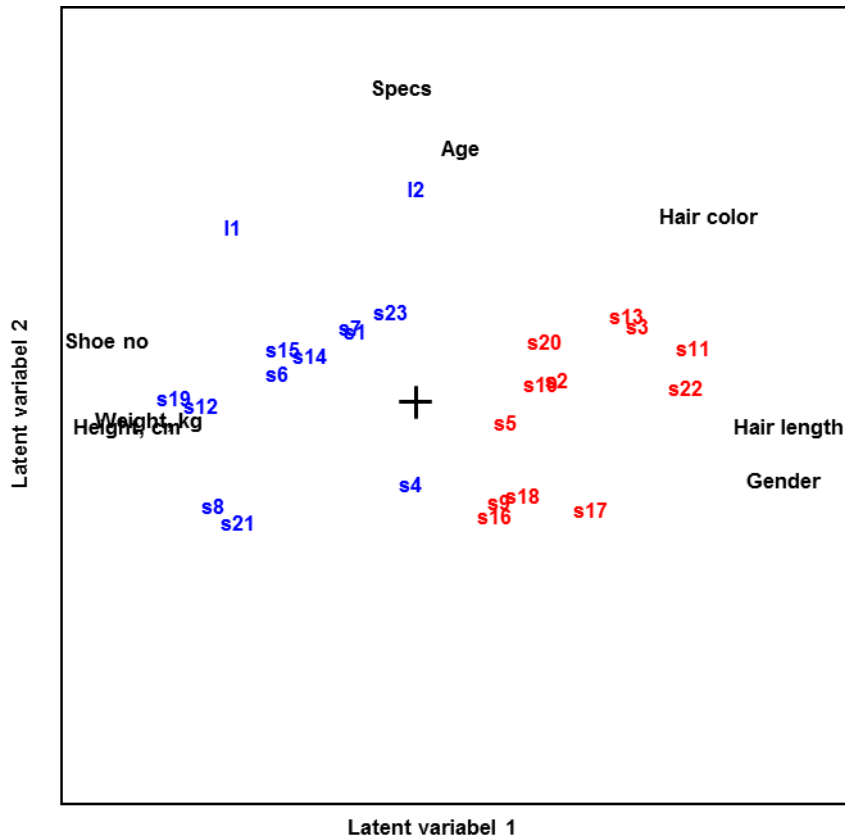
Figur 3 viser studenter (s1-s23) og to lærere (l1 og l2) i et såkalt skåreplott. Mannlige objekter er markert med blå farge. Kvinner og menn er separert langs den horisontale akse, representert av latent variabel 1.

Figur 4



Figur 4 er et såkalt ladningsplott for de åtte originale variablene. Det viser samvariasjon mellom variablene, og hvilke variabler som er viktige for de to latente variablene. De latente variablene er som nevnt lineærkombinasjoner av de originale. Kombineres disse figurene i et såkalt bi-plott ser man årsak-virkningssammenhenger i datasettet.

Figur 5



Figuren viser bi-plottet av persondatasettet som ble laget for et av kursene. Tolkningen av plottet var forståelig for alle studentene. Korrelasjonsmønsteret mellom variablene gav studentene mening og gruppering mellom objektene likeså. I grove trekk var tolkningen som følger; Samtlige kvinnelige studenter hadde langt hår fordi lang hårlengde og alle kvinner finnes på høyre side i figuren. Menn er generelt større enn kvinner fordi menn veier mer, bruker større sko og er høyere. Mannlige objekter og disse tre variablene finnes til venstre i figuren. Lærer 2 skiller seg ut på grunn av kombinasjonen høy alder og brillebruk, og finnes øverst i figuren.

Tiltak 3 gjaldt kun for kurset ved HVL. Som nevnt ble kurset i 2017 avsluttet med en oppgave som involverte litteraturstudier og presentering av resultater i form av foredrag. Studentene kommenterte og vurderte innhold i medstudentenes presentasjoner. Den studentaktive arbeidsformen i del 3 var i hovedsak å forberede og fremlegge presentasjoner under veiledning. Student- og lærerevalueringen tilsa at dette ikke gav særskilt gode læringssituasjoner. Del 3 i kurset ble i år utført på en måte jeg tror ligner på det såkalte omvendte klasserommet. Det ble utført fem korte bolker med teoriundervisning og

demonstrasjoner etterfulgt av lange og mange bolker med fire praktiske øvelser. Studentene jobbet som regel parvis, men tidvis i større grupper, under vei- og håndledning av læreren. Kurset ble avsluttet med en innlevering av en skriftlig rapport av en selvstendig data-analyse, for karaktersetting. Den eller de flervariabel metodene som studentene benyttet seg av i oppgaven var valgfri. Datasettene som ble analysert i innleveringsoppgavene ble hentet fra nettstedet kaggle.com. Det var satt av nok tid i siste del av seminaret til veiledning av studentenes selvvalgte datasett og metoder.

Effekter av tiltakene.

Studentkommentaren «Hvorfor har vi ikke lært dette før?» viser at antagelser om overføringsverdien fra tidligere kurs ofte er overdrevet. Det gjelder særskilt når emner fra t.d. matematikk kobles mot emner i kjemi. Heldigvis falt også kommentaren «Dette gav mening til lineær algebraen i MAT121» Å få studentene til å aktivisere forhånds-, *a priori*, inert-kunnskap fra persondatasettet gjorde overgangen til analyser av instrumentelle profiler noe mer håndterbart for de fleste. Datasettene som ble funnet på kaggle.com lignet i noen tilfeller mye på analyseresultater studentene hadde erfaring med fra egne arbeidsplasser. Det mykgjorde starten av den selvstendige oppgaven. Underveis gjorde studentene oppdagelser som at flere av datasettene bestod av «håndplukkede» ukorrelererte variabler. Latent-variabel metoder er ikke de mest velegnede for den typen variabler. Ved neste kurs vil fritt valg av datasett begrense seg til valg blant forhåndsvurderte tabeller.

Studentaktiv undervisning kan i dette tilfellet nesten oversettes til mindre lærersnakk og mer studenthandling. Korte bolker med forelesninger direkte etterfulgt av praktisk øvinger var en behagelig undervisningsform, og krever mye forberedelser. Motivasjonen hos studentene ved HVL var nok høyere enn i de regulære kursene ved UiB. Ved HVL-kurset er det krav om deltagelse i all undervisning, det er et frivillig etterutdanningskurs og studentene har fått mulighet av arbeidsgiver til å delta. Arbeidsviljen var stor hos de fleste alle dager, hele uken. Å få samme arbeidsmengde fra studenter som ikke har krav om oppmøte er erfaringsmessig usikkert. I referansene til Kolomitro og Gee [8] står det at det viktigste er hva eller om studentene lærer noe, ikke hvor eller når de lærer det. Kursene ved UiB gir studentene mulighetene til selv å velge tid og sted for de obligatoriske oppgavene. De som ikke velger å følge undervisningstilbudet er dem som oftest må pures på, om innleveringsfristene.

Om vi oppnår at de fleste studentene gjennomgår irreversible kunnskapstransformasjoner som beskrevet av Meyer og Land [1], er usikkert. De fleste studentene møter vi aldri igjen. Resultatmessig var det liten endring å spore i eksamenskarakterene for UiB-kursene. De var som vanlig meget gode. Det kan skyldes at eksamensoppgavetyperne er velkjente eller for enkle. De fleste studentene utfører eksperimentelt arbeid som en del av masteroppgaven hvor det er naturlig å koble inn metoder fra kursene PTEK226/ KJEM225. Om vi har lykket med å endre tankegangen blant studentene våre fra én til flere variable om gangen, gjenstår å se til kommende høstsemester. Om så er, vil etterspørselen etter den programvaren som de benyttet i kurset øke. Det som er sikkert, er at Kjell Raaheim sin antydning om at « utbyttet av slike forelesninger vil være avhengig av at man som student på forhånd iallfall i grove trekk har gjort seg fortrolig med innholdet i pensum» stadig er gyldig [10].

Referanser

[1] Meyer J H F and Land R 2003 'Threshold Concepts and Troublesome Knowledge 1 – Linkages to Ways of Thinking and Practising' in Improving Student Learning – Ten Years On. C.Rust (Ed), OCSLD, Oxford.

Threshold_Concepts__and_Troublesome_Knowledge_by_Professor_Ray_Land.pdf
(UPED620-litteratur)

[2] <https://www.uib.no/matnat/116128/undervisningsrom-aktivitetsbasert-laering>

[3] <http://pahoyden.no/2018/09/merittering-kritikken-liknar-klimafornekting>

[4] Creating significant learning experiences: an integrated approach to designing college courses. L. Dee Fink, Jossey-Bass San Francisco, 2013, 2nd ed. Kap. 2.

[5] <https://ebookcentral.proquest.com/lib/bergen-ebooks/reader.action?docID=1394307&ppg=52>

[6] <https://forskning.no/skole-og-utdanning-universitetet-i-bergen-partner/loser-i-laering/1080555> 22.5.2019

[7] <https://www.ntnu.edu/employees/harald.martens>

[8] Developing effective learning outcomes; A practical guide. K. Kolomito, K. Gee, Queens University 2015, s. 6.

[9] Kjemometri innen forskning og industri, R. Nortvedt red., Tidsskriftforlaget Kjemi AS (1996) s. 124.

[10] HVA ER EN FORELESNING? Kjell Raaheim, UiB. s.2. UPED620-litteratur H2017.