

La kvar student velje sitt eige tempo: Meistringslæring i intro programmering

Gabrielle Hansen, Guttorm Sindre, Guri Sivertsen Korpås, Joakim Pettersen Vassbakk, *NTNU*

SAMANDRAG: Meistringslæring går ut på å dele eit emne opp i mange små modular, som kvar for seg skal vere overkomelege. Studenten må så vise meistring av den første modulen før hen går vidare til neste. Med eit slikt opplegg kan kvar student få velje sitt eige tempo og ambisjonsnivå gjennom pensum, heller enn at alle studentar blir forventta å halde same tempo slik det gjerne blir i emne bygde opp rundt ein plenums serie av førelesingar eller liknande opplegg med ein fast plan veke for veke. Hausten 2023 blei eit emne med design basert på meistringslæring køyrt i intro programmering for studentar på lektorutdanning i realfag ved NTNU. Kvar student kunne velje sitt eige tempo og ambisjonsnivå i ein meistringsstige som inneheldt ein serie med testar i parallell med eit individuelt prosjekt. Testane hadde høg transparens, med treningstestar som matcha dei teljande testane. Opplegget resulterte i høg gjennomføring, men karaktersnittet vart lågt. Studentane var for det meste nøgde med emnet, sjølv om det også var somme aspekt ved det som var mindre tilfredsstillande. Artikkelen presenterer resultat frå ei spørjeundersøking, og diskuterer korleis emnet kan gjerast betre neste gong.

Mot slutten presenterer vi også nokre refleksjonar for andre som kunne vere interesserte i å prøve ut meistringslæring, anten det skulle vere i programmering eller heilt andre fag.

1 INNLEIING

Dei fleste emne legg opp til eitt sams tempo for alle. På universitetsnivå viser dette seg til dømes gjennom ein serie førelesingar, obligatoriske innleveringar og semesterprøver på spesifikke datoar, og til slutt ein eksamen på ein gitt dato. På engelsk blir dette gjerne kalla *instructor-paced* eller *teacher-paced learning* – i motsetnad til *student-paced* / *self-paced* der studenten styrer tempoet (Tullis and Benjamin 2011). Rimelegvis kan det ha mange føremoner at læraren styrer tempoet – elles ville neppe slike undervisningsopplegg ha vore så dominante. Det blir lettare å skalere undervisning og vurdering til store klassar dersom alle får same innhald på same tid, og det blir også lettare å leggje opp til samarbeid mellom studentane dersom alle held på med same tema i pensum. Samstundes er det også ulemper. Studentar har ulike førehandskunnskapar, eller lærer av andre grunnar raskare, somme lærer langsamare eller kan vere indisponerte i periodar t.d. pga. sjukdom. Studentar som opplever at undervisninga går vidare til eit nytt og meir avansert tema før dei har rekt å lære seg det førre meir grunnleggjande, vil gjerne få enda større problem med å lære dette nye fordi dei manglar basisen – og vil dermed ha lita nytte av undervisninga i det nye temaet (Robins 2010). Legg ein derimot opp undervisninga i så lågt tempo at alle skal hengje med heile tida, blir faget lett keisamt og understimulerande for studentar som kunne ha lært mykje raskare – og det blir lite undervisning på dei avanserte delene av pensum. Dette er dei sentrale argumenta for *meistringslæring* (Bloom 1971; Keller 1967): Studentar bør ikkje gå vidare med neste tema, før dei har meistra det førre temaet – og dersom studentar lærer i ulikt tempo, bør ein leggje til rette for at dei kan gjere nettopp det. Det typiske opplegget for meistringslæring er å dele opp eit emne i modular, der studentar kan vise meistring av ein modul ved å stå ein modultest. Studentar som står modultesten kan gå vidare med meir avansert stoff medan dei som ikkje har stått, kan øve meir og ta testen på nytt.

Hausten 2023 blei eit opplegg basert på meistringslæring nytta i eit emne i programmering for 48 førsteårsstudentar ved lektorutdanning i realfag ved NTNU. I ein artikkel i Læring om læring i fjor, forklarte vi ideane bak emnedesignet og ein prosess med studentinvolvering for å lage ein plan for emnet (Sindre et al. 2023). I denne nye artikkelen vil vi presentere erfaringane frå første gjennomføring av emnet, med svar på følgjande spørsmål:

- Korleis presterte studentane, kva resultat oppnådde dei, og korleis vart dette samanlikna med tilsvarande studentgruppe året før (som tok eit emne med veke-for-veke plenumsførelesingar, øvingsinnleveringar og ein avsluttande eksamen)?
- Kva studieåttferd førte opplegget til, og i kva grad mogeleggjorde opplegget at kvar student kunne velje sitt eige tempo? Eit spesifikt problem som mange andre har opplevd med meistringslæring er *prokrastinering* (Fox 2004; Francis, Figl, and Savage 2009; Puroo et al. 2017; Ott, McCane, and Meek 2021) – når studentar sjølv kan velje tempo, vil mange utsetje alt til siste liten, kanskje også lenger enn siste liten slik at meistringslæring gir auka fråfall (Reiser and Sullivan 1977).
- Kva tykte studentane om emnet? Og meir spesifikt: Kva tykte dei om opplegget med hyppige modultestar heller enn avsluttande eksamen? Fekk dei ei kjensle av meistring (som er eit viktig poeng med meistringslæring)? Kva aspekt ved emnet bør bli betre til neste år?

Resten av artikkelen er strukturert som følger: Seksjon 2 presenterer bakgrunns litteratur om meistringslæring både generelt og innan programmeringsundervisning på universitetsnivå. Seksjon 3 presenterer deretter undervisningsopplegget vårt og fortel om erfaringar frå førebuing og gjennomføring av emnet. Seksjon 4 forklarar forskingsmetoden for følgjeforskinga, og seksjon 5 presenterer resultatata frå denne følgjeforskinga, som konsentrerer seg om det siste kulepunktet: kva studentane tykte om emnet. Diskusjon kjem i seksjon 6, både med korleis resultatata kan tolkast og tankar om i kva grad eit liknande opplegg kunne vere mogeleg å bruke i heilt andre fag. Til slutt gir seksjon 7 nokre konkluderande tankar om korleis vi tenkjer å forbetre og vidareutvikle emnet med tanke på neste gjennomføring.

2 BAKGRUNN OM MEISTRINGSLÆRING

Bakgrunns litteratur om meistringslæring blei også gjennomgått i ein artikkel som forklarte planen og motivasjonen for emnet (Sindre et al. 2023) men vi gir også ein kort gjennomgang her for at denne artikkelen skal vere komplett i seg sjølv. Ein sentral idé i meistringslæring er at studenten skal meistre første modul før hen byrjar på andre modul, og så vidare. Systematiske metodikkar for meistringslæring blei foreslått på slutten av 1960-talet, slik som Bloom sin *Learning for Mastery* (LFM) (Bloom 1968) – seinare forenkla til *mastery learning* (Bloom 1971) og Keller sitt Personalized System of Instruction (PSI) (Keller 1967, 1968).

Både meistringslæring generelt (Kulik, Kulik, and Bangert-Drowns 1990) og PSI spesielt (Kulik, Kulik, and Cohen 1979) har vist god læringseffekt i empiriske studiar (jamvel om det også finst døme på negative effektar, t.d. (Reiser and Sullivan 1977)) og hadde ei blomstringstid på 1970-talet – om enn aldri i slik grad at det var i ferd med å overta for meir dominante undervisningsopplegg med eit felles lærarstyrt tempo. Utover 80- og 90-talet minka interessa for meistringslæring. I artikkelen (Buskist, Cush, and DeGrandpre 1991) blir det presentert ulike årsaker til at PSI gradvis mista popularitet, mellom anna at opplegg der studentane skulle bruke ulik tid var vanskeleg å passe inn i ein vanleg semesterstruktur, og at det var svært arbeidskrevjande for fagstaben å utvikle og gjennomføre alle dei testane som trongst. Artikkelen med den slående tittelen “*Keller's Personalized System of Instruction: Was it a Fleeting Fancy or is there a Revival on the Horizon?*” (Eyre 2007) spekulerer om digitalisering kan gi ein ny vår for PSI, sidan moderne læringsstøttesystem bør kunne gjere det lettare både å lage og gjennomføre testar. Snart 20 år har gått sidan dette, og det er nok framleis berre i horisonten – dei fleste faglærarar har knapt høyrte om PSI, langt mindre vurdert å bruke eit slikt opplegg i sitt eige fag. Samstundes finst det døme på vellukka bruk av meistringslæring (Pérez and Verdín 2023) og PSI (Mannion et al. 2023) i mange ulike fag på universitetsnivå også i nyare tid.

I ein review publisert i 2019 (Garner, Denny, and Luxton-Reilly 2019) blei det funne 12 døme på bruk av meistringslæring i programmering for førsteårsstudentar, og fleire andre har blitt publisert seinare (Aggarwal, Puthanveetil, and Gardner-McCune 2023; Alvarez, Samary, and Wise 2023; Campbell, Petersen, and Smith 2019; de Pontes, Guerrero, and de Figueiredo 2019; Izu, Ng, and Weerasinghe 2023; Ott, McCane, and Meek 2021; Toti, Chen, and Gonzalez 2023). Eit tidlegare publisert opplegg med meistringslæring / PSI i intro programmering i Noreg var ved UiA (Puroo et al. 2017). Programmering er typisk eit fag der det er viktig å ha forstått basiskonsept før ein går vidare til meir avanserte konsept (Robins 2010). I neste seksjon vil vi forklare vårt eige undervisningsopplegg og

korleis dette skiljer seg frå andre meistringslæringsopplegg i intro programmering som har vore publisert tidlegare.

3 EMNET VÅRT: IT1001 – DESIGN OG GJENNOMFØRING

Emnet vårt, *IT1001 Informasjonsteknologi grunnkurs*, vart køyrd første gong hausten 2023 og var eit emne berre for førsteårsstudentar ved lektorutdanning i realfag (LUR) ved NTNU (som hadde det som obligatorisk. Føremålet med å opprette eit eige emne akkurat for 1.årskurs LUR og tankane bak emnedesignet er forklarte i meir detalj i (Sindre et al. 2023) men vi gjentar kjapt dei viktigaste punkta så ein kan skjønne denne artikkelen utan å måtte lese den førre:

- *Betre kjensle av å høyre til i ein klasse.* 1. årskurs LUR mangla tidlegare noko emne som var spesifikt for dei.
- *Høgare relevans.* I tillegg til å lære programmering, reflektere over korleis programmering kan brukast pedagogisk.
- *Prøve ut meistringslæring.* Vi var nyfikne på dette, og for studentar som skal bli lærarar, kan det også vere spesielt relevant å ha vore eksponert for ulike undervisningsopplegg.

Opplegget vårt har følgjande eigenskapar:

1. *Karakter definert berre ut frå kor mange modular studenten har passert.* Dvs., ingen avsluttande eksamen. Emnet var delt i 9 modular, kalla I, H, G, F, E, D, C, B, A – der bokstav indikerer karakter. Dvs., ein student måtte stå 5 modular for å ta E i emnet, alle 9 modulane for å ta A.
2. *Inkrementelt, individuelt programmeringsprosjekt i parallell med testane.*
3. *Ingen plenumsførelesingar av fagstoff.* Faglærer hadde typisk ein 5-10 minutt motivasjonsprat i starten av kvart seminar, resten var til studentaktiv læring. Programmeringskonsept var forklarte i Jupyter Notebooks og videoar som faglærer hadde laga på førehand.
4. *Obligatorisk oppmøte* på desse seminarane (2 timar per veke, 80% oppmøtekrav).
5. *Høg transparens i vurderinga.* Modultestar blei brukt både formativt og summativt, formative testar som trekte oppgåver frå dei same spørsmålsbankane som dei summative.

Dette opplegget skil seg frå andre forsøk på meistringslæring i intro programmering. Somme har hatt avsluttande eksamen i tillegg til modulane (Purao et al. 2017; Offutt et al. 2017; Campbell, Petersen, and Smith 2019), eller tidfesta plenumsaktivitetar som inngjekk i karakteren (Ott, McCane et al. 2021). Opplegget til (Toti, Chen, and Gonzalez 2023) likna på vårt med at det stort sett var passerte modular som definerte karakteren, men dei hadde eit kortare tidsvindaug for kvar test (tre veker) så studentane sto mindre fritt til å velje eige tempo enn i vårt emne. Det finst nokre andre som også har hatt prosjekt i emne med meistringslæring (Jazayeri 2015; Toti, Chen, and Gonzalez 2023) men da som aller siste modul så berre dei høgst presterande studentane fekk prosjekt – hos oss fekk alle gjere prosjekt. Det er også uvanleg å ha så høg transparens for testane som vi hadde, der studentane kunne trene formativt med tilfeldig trekking frå same spørsmålsbankar som dei summative testane brukte. Dei fleste andre opplegg med meistringslæring har hatt separate oppgåvesett for første forsøk på ein test og gjentak for å sikre at studenten ikkje skal ha sett oppgåvene før, t.d. (Campbell, Petersen, and Smith 2019; Toti, Chen, and Gonzalez 2023). Somme har opplegg som liknar meir på vårt ved at første forsøk og gjentak har trekt spørsmål tilfeldig frå same store spørsmålsbank, t.d. (Ott, McCane, and Meek 2021; Offutt et al. 2017) men det verkar ikkje som dei har tillate studentane å ta formative treningstestar på dei same spørsmålsbankane. Målet vårt med denne auka transparensen var å redusere testangst (Wiggins et al. 2023) og sørge for kjapp tilbakemelding til studentar om kva dei hadde forstått og kva dei trengde å jobbe meir med (Epstein et al. 2002; Treviño and Cavazos 2018; Izu, Ng, and Weerasinghe 2023).

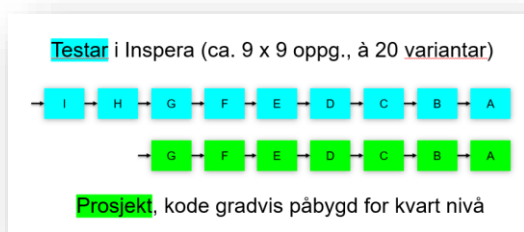


Fig. 1. Meistringsstigen, testar blå, prosjekt grøn

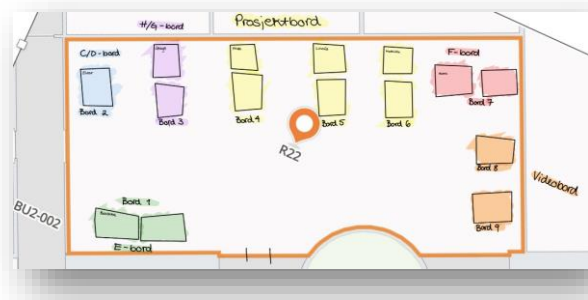


Fig. 2. Seminarrommet, gruppedeling 6. okt.

Figur 1 illustrerer meistringsstigen i emnet, med testar som byrja på nivå I, medan første prosjektlevering var på nivå G. I den opprinnelege planen publisert i (Sindre et al. 2023) tenkte vi å starte både testar og prosjekt på I, men vi gjorde om på dette da vi tenkte at mange studentar ikkje ville ha føresetnader for å setje i gong med eit prosjekt allereie på nivå I. Frå starten av var kravet for å bestå kvar einiskild test at ein klarte ein score på 90% når testen blei tatt under tilsyn (sjanse kvar fredag). Etter kvart fann vi ut at F- og E-testen var litt for vanskelege for mange av studentane ut frå kravet på 90%, så dei fem siste vekene av semesteret fekk kvar student 3 stk. 80%-sjansar som dei kunne bruke ved behov, anten dei hadde kome langt eller kort i testserien. Dvs., ein student som t.d. på det tidspunktet skulle ta F-testen, ville kunne passere F, E, D på 80% - eller ved å greie minst ein av desse på 90%, framleis ha 80% tilgjengeleg for B-testen. Prosjektet hadde tydelege rubrikkar for kva som trongst for å få godkjend kvart nivå, i hovudsak ei sjekkliste på om koden hadde gjort tenleg bruk av dei ulike programmeringskonseptane som inngjekk i den aktuelle modulen. Frå starten av var det dessutan eit krav til eit visst tal på kodelinjer per nivå, men dette blei fjerna samstundes med innføring av 80%-sjansane da vi såg at det førte til kvantitet snarare enn kvalitet i studentane sin programkode.

Figur 2 viser korleis rommet blei brukt under dei obligatoriske seminarane, med konkret døme 6. oktober, altså om lag midtvegs i semesteret. Kvar student sto fritt til å velje kva ein ville jobbe med i seminaret, typisk var det anten å bu seg til den neste testen eller å jobbe med prosjektet. Oransje bord er for studentar som ville sjå på faglærer sine videoar med forklaring av ulike programmeringskonsept. Vi fann det tenleg å separere ut dette ved eigne bord sidan ein student som sit med øyretelofonar og ser video ved eit anna bord lett ender opp med å hemme gruppediskusjonen dersom hen blir sitjande mellom to studentar som elles kunne hatt ein diskusjon. Dei tre gule borda er for studentar som ville jobbe med prosjektet, gruppert litt etter kor langt dei hadde kome med det. Dei andre borda hadde grupper av studentar som øvde på ulike testar: C/D-test lyseblå oppe venstre, G/H-test lilla, F-test raud oppe høgre, E-test grøn nede venstre. Typisk arbeidsmåte ved testborda var at ein av studentane tok opp testen i det digitale eksamenssystemet Inspera Assessment på den sams skjermen som kvart bord hadde, og så prøvde studentane å løyse testen samstundes som dei diskuterte kva som var rett og korfor på kvar oppgåve.

Det var stor arbeidsinnsats både med å førebu og gjennomføre emnet. Som forklart i (Sindre et al. 2023) nytta førebuingssfasen seg av fleire tilsette vit.assar som køyrde ein prosess med studentinvolvering – dette særleg fordi opplegget var så ulikt alt anna som var prøvd tidlegare. Assistentane hjelpte også til med å kvalitetssikre testoppgåver og lage modell-løysingar på prosjekt. Faglærer brukte fleire hundre timar på førebuing, kanskje så mykje som 500 timar berre på å lage dei store spørsmålsbankane med ulike variantar av testoppgåver – som var naudsynt for å kunne ha det transparente opplegget der studentane kunne trene på same spørsmål som fantest i dei summative testane – samt å lage Jupyter notebooks og læringsvideoar. Mykje av dette er likevel éingongsinvesteringar som ein ikkje vil måtte gjenta seinare, og ein del av læringsressursane kan også vere moglege å bruke på andre emnekodar enn berre IT1001.

Gjennomføringa brukte også mykje ressursar. Sidan det var første gong med ein pilot av eit svært annleis opplegg som vi på førehand ikkje visste korleis studentane ville takle, hadde vi langt fleire læringsassistentar enn det som ville vere normalt for eit emne med berre 50 studentar, om lag 700 assistenttimar totalt. Neste år vil vi prøve med halvparten av dette, og vidare reduksjonar kan vere

mogeleg ved å ha betre tilbakemelding direkte i modultestane og betre stillas for prosjektet. Erfaringane frå 2023 viste også at vi fekk til eit opplegg der studentane samarbeidde godt med kvarandre i seminara, som i nokon grad kan redusere behovet for læringsassistentar. Ein del tidsbruk for faglærer og administrativt personell hausten 2023 skuldast låg funksjonalitet i Inspera – særleg at systemet ikkje taklar at ein student kan ta same test mange gonger, som vi trong både for formative og summative testar. Dette gjorde at det måtte setjast opp nye testar kvar veke, der vi måtte dele ut eingongsbrukarar. Dersom eit nytt digitalt eksamenssystem (eller ein betre ny versjon av Inspera) meir direkte støttar gjentak av testar og organisering av ein serie testar i ein stige, vil slik tidsbruk hos faglærer og administrasjon bli vesentleg redusert.

Figur 3 viser karakterfordeling for vår klasse (lyse søyler) versus tilsvarande klasse året før, som hadde eit tradisjonelt opplegg med avsluttande eksamen (mørke søyler). Opplegget resulterte i høg gjennomføringsgrad, da alle som gjennomførte emnet sto, medan det var 11% stryk for '22-kullet av LUR-studentar. Også positivt er ein liten auke i prosent som greidde A og B, jamvel om desse i begge tilfelle er låge. Det som derimot vart monaleg dårlegare med karakterane i 2023 samanlikna med 2022 er at det var færre på C og D, og i staden ein stor topp på karakteren E. Det vil vere spekulativt å gjere for mykje ut av ei slik samanlikning sidan det er to ulike studentkull og heilt ulike undervisningsopplegg og vurderingsformer, men det kan verke som opplegget i positiv forstand har greidd å løfte studentar som elles kunne ha enda på stryk, opp til ein E, men samstundes også ført til at ein del studentar som kunne hatt potensial for C eller D, også har enda på E – meir om moglege årsaker til dette i diskusjonen mot slutten. Bokstavnippet for begge klassane er D, men 2022 låg litt betre enn D, 2023 litt svakare, så skilnaden i snitt er ein tredjedels karakter i negativ retning for 2023-kullet.

Fire veker ut i semesteret var faglærer sitt inntrykk at klassen sin snittfart oppover teststigen låg nærmare C enn D, som ville ha vore ei monaleg betring samanlikna med 2022. Men deretter vart framdrifta noko tregare, av fleire grunnar. Veke 2-4 hadde det vore høgt oppmøte på dei summative testane, med nesten heile klassen til stades – dette gjekk etter kvart noko ned, og det same gjorde prosenten som sto den testen dei forsøkte på. Det såg ut til å vere fleire grunnar til dette. Som det også går fram av emnerapporten, var F-testen monaleg vanskelegare enn G-testen (og mykje meir enn skilnaden frå H til G, t.d.), og E-testen var også vanskeleg. Medan dei fleste hadde greidd testane I-G på første forsøk og ingen trong meir enn 2 forsøk, vart F og E meir krevjande for ein del studentar. Sidan det berre var mogeleg å ta teljande testar éin gong i veka, gjorde dette at somme fekk tregare framdrift enn dei ideelt sett ville ha ønskt. Somme såg nok også frå resultat på treningstestar at dei enno ikkje hadde nivået inne for neste summative test, som kan forklare det gradvis lågare oppmøtet.

Faglærer konkluderte etter kvart at F- og E-testane var for vanskelege med eit bastant krav om 90% score for å stå. Som respons på dette blei det innført eit opplegg med at kvar student fekk 3 stk. «80% sjansar», dvs. at ein kunne passere 3 av testane med 80% score. Dette tiltaket kom nok litt seint (veke 10) men bidrog da til at fleire greidde seg gjennom dei testane som elles hadde vore litt for vanskelege. Motivasjonen for opplegget med 80%-sjansar heller enn å berre redusere terskelen på spesifikke testar (t.d. F, E) var å finne noko som blei rettferdig for alle uansett kor langt dei hadde kome. Dei som allereie hadde stått F og E på 90% kunne da i staden bruke 80%-sjansane sine på testar lenger opp i stigen.

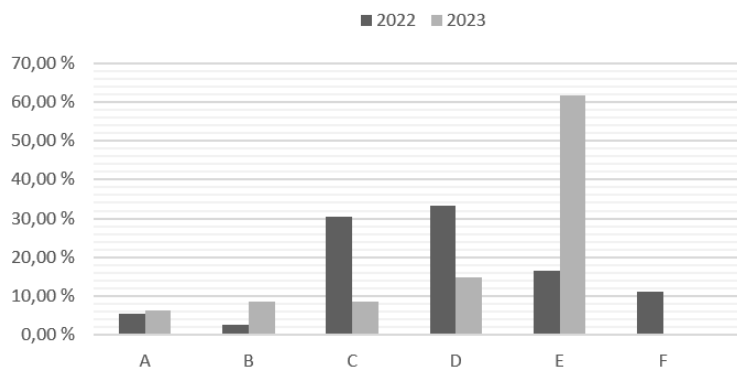


Fig. 3. Karakterar, '22 mørk, '23 lys

Ein annan årsak til tregare tempo var at studentar sleit med å kome i gong med prosjekta sine, der lite stillas gjorde at dei måtte finne ut av mange ting sjølve med korleis dei skulle byggje opp koden sin. Som vist i figur 1 starta prosjektet med innlevering på nivå G fordi vi tenkte at dei på nivå I og H enno ikkje hadde føresetnader for å starte med å skrive eit program. Dette gjorde at prosjekt G blei ein litt stor bøyg for mange, og såleis ein naturleg årsak til at progresjonen gjekk tregare etter testnivå G (sidan dei fleste tok testen før dei leverte prosjektet). Innleiingsvis hadde vi også eit krav om 50 kodeliner tillegg for kvart nivå for storleiken på prosjektet totalt (dvs. G 50 linjer, F 100, ...), dette var nok også for mykje, og førte til lang men dårleg kode for mange. Samstundes med innføringa av 80%-sjansane valde vi å kutte dette kravet, noko studentane var glade for (men også meinte burde ha vore gjort lenge før).

4 FORSKINGSMETODE

Sidan opplegget i emnet er såpass ulikt mykje anna som har vore forsøkt ved NTNU, bestemte vi oss for å gjere mykje følgjeforskning undervegs og etter semesterslutt for å finne ut korleis emnet og opplegget fungerte for studentane. Overordna blir det kanskje mest rett å sjå på dette som eit case-studium der caset er gjennomføringa av emnet hausten 2023 (Johannesen et al., 2016). Fleire ulike metodar for datainnsamling er blitt nytta for å få brei og djup innsikt i caset:

- Observasjon av seminara (14 dobbelttimar, kvar torsdag gjennom semesteret)
- Observasjon av testane (14 timar, kvar fredag gjennom semesteret)
- Spørjeundersøking til studentane (44/48 svarte, dvs. responsrate 92%)
- Fokusgruppeintervju med studentar som melde seg friviljug til dette etter spørjeundersøking. Fem grupper med om lag tre til fire studentar i kvar. Kvart intervju varte i om lag ein og ein halv til to timar.

I denne artikkelen fokuserer vi berre på funn fra spørjeskjemaundersøkinga sidan data frå observasjonar og intervju er under analyse og vil bli presenterte ved eit seinare høve.

Føremålet med spørjeundersøkinga var å få innsikt i korleis studentane hadde opplevd og erfart eit semester med meistringslæring. Resultata fra spørjeundersøkinga er òg tenkt brukt inn i vidareutvikling av emnet. Skjemaet var inndelt i ulike aspekt: generelle inntrykk, fagleg tilhøyrgheit, individuell framdriftsplan, sjølvbestemt tempo, seminar, læringsressursar, teststige og prosjekt. Undersøkinga hadde 42 spørsmål, der 30 var lukka spørsmål (dei fleste av desse på ein 5-trinns Likert-skala, somme hadde berre 3 trinn), medan 12 var opne tekstspørsmål. Undersøkinga blei gitt i seminar nummer 11 for å maksimere responsraten da ein måtte ha møtt på minst 11 seminar for å få emnet godkjent. Dette høvet blei også brukt for å invitere til fokusgruppeintervju, og over halvparten av dei frammøtte melde interesse for dette.

Spørjeundersøkinga, som blei utført i Nettskjema, var anonym. Det blei heller ikkje samla inn personopplysningar, derfor var det ikkje naudsynt å melde ho inn til Sikt. I denne artikkelen har vi berre deskriptiv statistikk for ein del av svara, ikkje noko hypotesetesting.

5 RESULTAT FRÅ FØLGJEFORSKINGA

Det var eit par veker igjen av semesteret da spørjeundersøkinga blei utført, men verka allereie klart at det kom til å gå mot eit lågt karaktersnitt. Såleis var vi ganske nervøse for kva undersøkinga ville vise, og vart positivt overraska når vi såg resultatata. Figur 4 viser fordelinga av svar på spørsmål om kor nøgde studentane var med emnet – med ei klar overvekt av nøgde eller nøytrale, få misnøgde. Figur 5 viser fordelinga av svar på spørsmål om korleis dei ville samanlikne dette emnet med andre emne dei tok samstundes. Her var det berre tre alternativ: «betre» / «dårlegare» / «kjem for det same», vist som grønt, raudt, gult, og med klar overvekt for betre.

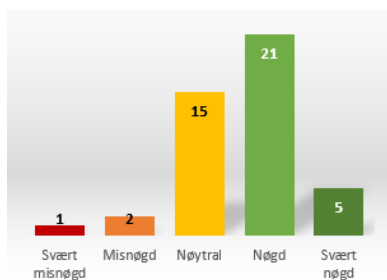


Fig. 4. Kor nøgde var studentane med emnet?



Fig. 5. ...og samanlikna med andre emne?

I fritekstsvara knytt til dei innleiande spørsmåla utdjuar studentane korfor dei likte emnet. Studentane var letta over å ha eit emne utan avsluttande eksamen. Dei tykte at dette emnet, med teststigen og prosjektet, gav dei som studentar meir kontroll og ansvar, de kjende at dei blei motivert til å jobbe meir kontinuerleg, noko dei opplevde som positivt for eiga læring. Eksamen derimot, assosierte dei med stress, pugging og dårleg læringsutbytte. Somme studentar skildra det slik:

Det å slippe å ha en dags eksamen i vente og heller jobbe kontinuerlig med faget, samt å hele tiden vite hvilket karakternivå man ligger på, elsker jeg. Det gjør at jeg har én eksamen mindre å bekymre meg for/grue meg til, og jeg føler at jeg har mye bedre kontroll. Det er godt å vite så spesifikt hva man kan jobbe med. I tillegg, med en slik mestringsstige, "tvinges" man litt til å jobbe litt med faget hver uke, og det synes jeg har vært veldig greit, og jeg føler et stort læringsutbytte av det. Større enn jeg ville følt ved å pugge til en stor eksamen.

Emnet gir mindre stress med tanke på at vi ikke har en avsluttende eksamen. Jevnlige tester gjør at jeg får lært meg faget bedre enn disse ukentlige obligatoriske øvingene som er i andre fag.

Jeg har mer kontroll på hvilken karakter jeg har og føler jeg kan ta ting i mitt eget tempo.

I tillegg til dei overordna spørsmåla over, blei det stilt diverse meir spesifikke spørsmål om ulike aspekt ved emnet. Vi kan ikkje gå gjennom alle desse her, men vil trekkje fram det som synest mest sentralt. Eitt motiv for å opprette eit eige programmeringsemne for LUR i staden for å ha dei i ein større klasse saman med andre studieprogram, var å bidra til fagleg-sosialt fellesskap for første årskurs LUR. Figur 6 viser prosentvis fordelt resultat for tre spørsmål relatert til dette, med svært lite / svært usamd til venstre (mørk raud), svært mykje / svært samd til høgre (mørk grøn), og nøytral som grå. Dei tre spørsmåla var om emnet bidrog til å gi ei kjensle av å høyre til på LUR-programmet (øvtst), om det gjorde det lettare å bli kjend med medstudentar i klassen (midt) og om det var viktig med samarbeid med medstudentar i dette emnet. Alle desse spørsmåla viser klare positive utslag, med nesten ingen studentar som er usamde. Ein av studentane skildrar det slik (fritekstsvaer i spørjeundersøking):

Det er koselig å møte resten av klassen, og samarbeide med medelever og læringsassistenter hvis man bygger på prosjektet og trenger innspill, eller for å gjennomgå treningstester sammen. Er glad vi har obligatoriske seminarer, hvor det faller veldig naturlig å samarbeide med andre og lære av hverandre, i stedet for å få en obligatorisk øvingsoppgave hver uke som ender opp med at man kanskje heller velger å gjøre den alene hjemme.

Generell motivasjon for meistringslæring er at alle skal kunne oppnå meistring, samt at inndelinga gi modular gjer krava meir tydelege for studentar og gir dei sjansen til å ha betre kontroll på korleis dei ligg an i emnet. Figur 7 viser resultat for fire spørsmål relatert til dette, igjen med mest negative svar som mørkt raudt til venstre, mest positive mørkt grønt til høgre og nøytralt som grått. Øvste påstand var «Eg har opplevd meistring» der klart fleire er samd enn usamd, og liknande svarfordeling kom for andre spørsmål om «tydelege krav for kva som skal oppnåast for kvar modul». «Kjensle av kontroll på karakter eg ønskjer å oppnå» er meir blanda. Ein del hadde på dette tidspunktet enno ikkje stått E-testen og kan ha vore urolege for om dei ville greie det, eller om dei hadde stått E-testen, kan ha vore urolege for om dei ville klare ytterlegare testar dei hadde ambisjon om å ta. Det nedste spørsmålet, om dei mange testane blei opplevd som stressande, har mest negative svar av desse fire, med 12% som svarer «Svært stressande».

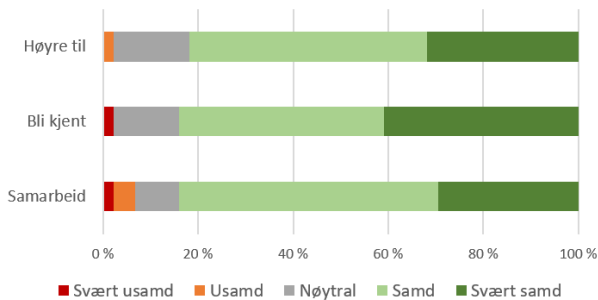


Fig. 6. Fagleg-sosialt fellesskap

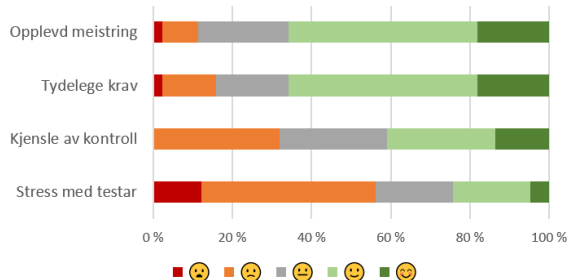


Fig. 7. Krav, kontroll, meistring, stress

Det studentane var minst nøgde med, var ein del aspekt ved prosjektdelen av emnet, som at det var for lite læringsressursar spesifikt for prosjektet og såleis vanskeleg å kome i gong – særleg sidan mange av dei ikkje hadde noko erfaring med programmering frå vidaregåande skule. Nedste rad i Figur 8 viser fordelinga av svar på «Korleis var det å kome i gong med prosjektet?», frå «svært utfordrande» (raudt), via utfordrande, nøytral, lett, til svært lett (mørk grøn). Som vi ser var det ei klar overvekt av studentar som tykte dette var vanskeleg – og dei som er på raudt sikkert *for* vanskeleg. Nokre sitat frå spørjeundersøkinga kan også kaste lys over dette:

Veldig overveldende i starten. Jeg har aldri vært borti koder og programmering før, så er mye derfor. Skjønte liksom ikke helt hva programmering var. Når jeg kom på et visst nivå og skjønnte the basics, var det lettere å begynne å diskutere med andre.

Til å begynne med var det litt mye for jeg hadde aldri programmert før og kunne ikke programmere ordentlig før jeg kom til f testen, men var heldigvis relativt raskt på det nivået så fikk kommet i gang greit. Etter det har kodingen gått bedre og bedre og har fått ganske grei kontroll på hva jeg gjør.

Referansegruppa gav også uttrykk for at det hadde vore for lite hjelp og læringsressursar til å kome i gong med prosjektet i starten, medan testane var mykje lettare å kome i gong med sidan det var svært konkret kva ein skulle gjere der, ikkje minst på grunn av høg transparens med treningstestar over same spørsmålsbankar. Dette tyder likevel ikkje at studentane i det store og heile var misnøgde med prosjektet. Dei to midtre radene i Figur 8 viser svar på spørsmål om vanskegrad på testar og prosjekt og prosjektet er berre opplevd som marginalt vanskelegare enn testane. Når det gjeld opplevd nytteverdi (øvre rad) ser vi at det var ein god del studentar som oppfatta prosjektet som nyttigare enn testane, færre som oppfatta testane som det mest nyttige. Spørsmålet for øvre rad hadde berre tre svaralternativ: testar mest nyttig / likt / prosjekt mest nyttig.

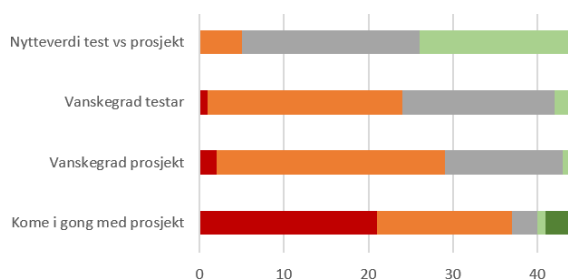


Fig. 8. Opplevd nytteverdi og vanskegrad

6 DISKUSJON

6.1 Tolking av funn

I forhold til studentane sine prestasjonar var det blanda resultat. Det var positivt med redusert stryk, men snittet gjekk ned med ein tredels karakter på grunn av den høge toppen på E. Det kan vere fleire årsaker til dette. Ein grunn kan vere gradvis mindre søking til lektorstudium dei siste åra. Ifølgje Samordna opptak hadde LUR-studiet opptaksgrense 47,3 (ORDF) i 2022, medan 2023 står med grense 0, som vil bety at alle søkjarar som tilfredsstilte dei generelle opptakskrava for programmet, kom inn.

Ein annan faktor kan vere at ein del studentar bestemte seg for å stoppe på E når dei hadde sikra seg ståkarakter i emnet, for å konsentrere seg meir om andre emne der dei framleis var usikre på å stå. IT1001 var berre 7,5 studiepoeng så studentane hadde 22,5 fleire studiepoeng å ta, som kan kjennest hardt for førsteårsstudentar. Alle dei andre emna klassen tok i same semester, hadde avsluttande i siste veke av november og første veke av desember – og fleire av studentane gav uttrykk for å vere nervøse for desse eksamenane. Ein del studentar sleit også mykje med å greie testane F og E og måtte prøve fleire gonger om att. Desse kan ha tenkt at det ville bli minst like vanskeleg å greie D-testen, og at dette var ein innsats dei ikkje hadde råd til av omsyn til andre emne – særleg sidan innsatsen ville kjennest bortkasta dersom dei likevel ikkje skulle greie D. For studentar som hadde greidd E-testen allereie tidleg i semesteret (veke 10 eller før) kan dette likevel neppe vere grunnen – desse har truleg vore godt i stand til å ta fleire testar. Somme av desse låg likevel eit stykke etter på prosjektet, og innan dei hadde sikra seg E på prosjektet, var det kanskje blitt såpass seint i semesteret at freistinga for å konsentrere seg om eksamensfaga vart større. Kombinasjonen av testar og prosjekt – der ein måtte opp eit nivå på begge deler for at det skulle få effekt på karakteren – kan også ha bidratt til at mange har vegra seg for å satse på kome seg frå E til D dei siste vekene. Dersom ein jobbar prosjektet opp til D, men så ikkje greier D-testen likevel, vil det kjennest litt bortkasta – og dersom ein skal vente med prosjektet til ein har klart testen, kjem prosjektjobbinga for tett opp i leseperioden for andre eksamenar.

Likevel verka studentane generelt sett nøgde med emnet, og dei fleste såg det som betre enn andre emne dei tok parallelt. Ein kan likevel ikkje trekke for bastante konklusjonar av dette. Spørjeundersøkinga blei gjennomført i veke 44, det vil seie tre veker før slutten av undervisningsperioden. På dette tidspunktet hadde mange studentar allereie sikra seg ståkarakter i vårt emne, som dei ikkje hadde gjort i dei andre emna. Derfor kan somme ha vore meir urolege for og dermed mindre nøgde med andre kurs av den grunnen – men om undersøkinga hadde vore gjort etter at dei hadde sensurresultata frå alle emna, ville kanskje resultatet ha vore eit anna.

Noko ein kunne vere redd for på førehand med dette undervisningsopplegget, var at einskildstudenten skulle ende opp med å jobbe mykje aleine – sidan kvar kunne halde sitt eige tempo, og prosjektet også var individuelt. Tvert om viste det seg at studentane samarbeidde mykje i grupper, som både blei observert på seminara og sett på svara i spørjeundersøkinga, jfr. resultat om samarbeid i fig. 6. Seminara med obligatorisk oppmøte viste seg å bli ein vellukka arena for samarbeid, og sjølv om studentane var på ulike nivå, var det stort sett alltid mange nok som var på same nivå til at dei kunne danne ei gruppe av passeleg storleik rundt eit bord. Kanskje var det slik at studentar med ulike tempo nettopp gjorde det lettare å danne effektive grupper der alle hadde ulike utfordringar. Det gjorde også at dei blei kjent med ulike personar i klassen sidan grupperingane ikkje var konstante for kvar veke, men endra seg etter kven som hadde greidd testar og ikkje, kven som ville jobbe med test og kven med prosjekt i seminaret.

6.2 Pedagogiske refleksjonar

Mange andre opplegg for som har brukt meistringslæring generelt eller PSI (Personalized System of Instruction) spesielt, har opplevd prokrastinering som eit stort problem (Reiser 1984; Ott, McCane, and Meek 2021; Purao et al. 2017; Offutt et al. 2017). I vårt opplegg verka det tvert om som ein del studentar gjekk i motsett retning og prøvde å gjere seg ferdig tidleg – men at ein del deretter slo seg til ro med lågaste ståkarakter. Ein av årsakene kan vere at vårt opplegg eksplisitt gjorde det mogeleg å sikre ein ståkarakter tidleg, medan mange andre opplegg har hatt ein kombinasjon av innleveringar og avsluttande eksamen som uansett har gjort at ein må jobbe intensivt på slutten av semesteret. Såleis kan det tenkjast at det ikkje er sjølvvald tempo i og for seg som skaper prokrastinering, men sjølvvald tempo i kombinasjon med avsluttande eksamen eller tilsvarande, som gjer at det er umogeleg å gjere seg ferdig med emnet tidleg, eller sjølvvald tempo i kombinasjon med at studentane er usikre på korleis dei skal kome seg vidare (slik det litt meir vart med prosjektet i vårt emne).

Spesielt sidan målgruppa var lektorstudentar, kan det vere ein ekstra verdi også i at dei er blitt eksponerte for fleire ulike lærings- og vurderingsmetodar i studiet, utan at det eine treng vere meir vellukka enn det andre. Meistringslæring har lange røter tilbake til 1960-talet, (Bloom 1968; Keller 1968), men har trass i positive resultat (Kulik, Kulik, and Bangert-Drowns 1990) aldri tatt heilt av (Eyre 2007). Studentane hadde på dette tidspunktet enno ikkje lært noko pedagogikk – det kjem først

frå andre år av studiet og utover – men vonaleg kan dei når dei seinare i studiet lærer meir om dette, også vere i stand til å trekkje trådar tilbake til det dei opplevde i dette undervisningsopplegget, både med tanke på kva dei tykte fungerte bra og det dei ikkje var like nøgde med, og såleis vurdere meistringslæring som ein del av sitt pedagogiske repertoar i eigen praksis i yrkeslivet. Jamvel om studentane som framtidige lærarar på 8.-13.trinn ikkje får så full kontroll over opplegget i emne dei underviser at dei plutselig kan leggje om til å bruke meistringslæring, kan element av det dei har opplevd vere verdfullt i mindre radikale tilpassingar også. Eitt døme på dette kan vere korleis seminara blei køyrte: Vi unngjekk plenumsundervisning og lét i staden studentane samarbeide i grupper på nivået der dei var – medan faglærar og assistentar fasiliterte. Eit anna døme er korleis digitale læringsressursar kan gjere det mogeleg for ulike studentgrupper å halde ulike tempo, og korleis modultestane i seg sjølve også blei ein viktig læringsressurs gjennom den kombinerte formativ-summative bruken, som gjorde det svært konkret for studentane kva som måtte til for å kome seg opp til neste nivå, og korleis testar kan brukast både formativt og summativt. Gjennom det individuelle prosjektet har dei også fått erfaring med prosjektbasert læring, sjølv om dette nok var den delen av emnet som har mest potensial for betring.

I lys av det ovanståande kunne det vere interessant å gjennomføre ein studie nokre år fram i tid, når studentar som tok emnet har kome ut i pedagogisk praksis i skulen, samt når dei er ferdige med studiet og ute i lektorjobb på fulltid, for å undersøkje om opplevinga av meistringslæring i eit emne på universitetet på nokon måte har vore med på å forme deira eigen pedagogiske praksis eller forståinga av denne.

7 KONKLUSJON

Som tittelen seier: «La studentane velje sitt eige tempo...» - lukkast emnedesignet med å tilby dette? Svaret er i stor grad JA – studentar heldt ei rekkje ulike tempo, på det meste var det seks ulike testnivå som var aktive samstundes. Ein kan også seie at studentane kunne velje sin eigen karakter. Sjølv sagt var det i utgangspunktet mange som kunne ha tenkt seg ein *betre* karakter – som t.d. hadde C som slutt mål i framdriftsplanen dei sette opp i starten av semesteret men enda opp med å ta E. Å velje karakter i positiv retning inneber at ein må klare mange nok testar, ha nok arbeidskapasitet til å lære seg materialet for desse testane, pluss utvide prosjektet tilsvarende – men iallfall er det svært konkret kva ein må lære seg for å gå opp eit nivå til, det er berre å sjå på den formative treningstesten kva slags oppgåver som inngår. Studenten vil alltid kunne velje karakter i negativ retning – straks ein har nådd E, kan ein om ein vil, leggje inn årene og til dømes konsentrere seg om andre emne fram mot eksamen. For faglærar kan det vere trist med ein såpass høg topp på E, men for ein student som har sikra ståkarakter i eitt emne men er nervøs for tre ventande eksamenar, kan det også vere god optimalisering å velje slik – for dei fleste formål vil det vere betre å stå i alle emna i eit semester enn å ha betre karakter i somme men stryke i andre. Dessutan vil dei fleste studentar tenkje at det er viktigast med gode karakterar i emne som dei skal spesialisere seg vidare i. Det var berre om lag 10% av studentane på LUR som gjekk retninga matematikk + informatikk, og som dermed skulle ha eit vidaregåande programmeringsemne det påfølgjande semesteret. Dei resterande 90% skulle spesialisere seg i fag som matematikk, fysikk, kjemi og biologi – og da vil det gjerne kjennest viktigare å prioritere desse emna høgare. Det treng ikkje vere ei dårleg side ved eit emnedesign at det faktisk gir studenten mogelegheit til eit slikt val. Om emnet i staden hadde avsluttande eksamen, ville ein del studentar som var nøgde med E, sikkert ende på D eller C ut frå uvisse om kva som trengst for å stå på eksamen og behov for ein tryggleiksmargin – som på sett og vis ville vere å «overprestere» dersom ein da enda opp med å stryke på ein annan eksamen som kanskje eigentleg var viktigare.

Det var fleire aspekt ved den første gjennomføringa som ikkje fungerte normalt. Når det gjeld testane, tenkjer vi derfor å endre litt på eit par av oppgåvene på F- og E-testen, som no anten var unauddsynt vanskelege, eller der det var for stor variasjon i vanskegrad mellom variantar av same oppgåve, slik at studentar kunne ha flaks eller uflaks med kva oppgåver dei trekte på den summative testen. Dessutan vurderer vi å ha ordninga med 3 stk. 80%-sjansar på plass allereie frå starten av semesteret, heller enn som ein improvisasjon midt i oktober, men i så fall slik at det likevel ikkje er lov å bruke desse på testane I, H, G, som var enkle nok til at 90% var eit rimeleg passeringskrav. Når det gjeld prosjektet, tenkjer vi å lage vesentleg meir læringsressursar for startfasen, i form av malar som studentane kan bruke for å få til dei første nivåa. Sannsynlegvis var det ein tabbe å ha første prosjektinnlevering på G medan testane starta på I – dette gjorde at alle kom i gong kjapt med testane medan prosjektet blei sett på vent. Neste år vil vi sannsynlegvis prøve å ha prosjektinnlevering også på nivå I og H, men da

svært enkle innleveringar, t.d. at prosjekt I ikkje skal krevje meir enn éin times innsats, der studenten berre treng å gjere nokre endringar i ein utdelt mal heller enn å skulle starte med å skrive kode i eit heilt blankt skjermbilete.

For andre som måtte vere nyfikne på å prøve ut meistringslæring i sine emne, er det viktig å merke seg at jamvel om denne artikkelen handlar spesifikt om eit emne i intro programmering, er det ingenting ved meistringslæring som avgrensar bruken til programmering eller andre tekniske fag. PSI vart i utgangspunktet utvikla innan psykologiundervisning (Fred Keller) og Bloom sine idear om meistringslæring galdt for fag i skulen generelt. Einaste krav er at ei må kunne klare å dele emnet opp i ein del mindre modular som det finst ei naturleg rekkjefølgje for, men slik rekkjefølgje har dei fleste emne i alle fall, til dømes gjennom ein førelesingsplan eller kapittel i ei pensumbok. Det store timetalet som er brukt på førebuing og gjennomføring, kan også verke avskrekkande, men det skal vere mogeleg å gjennomføre meistringslæring med monaleg mindre tidsbruk. I førebuinga vår inngjekk eit stort tal på assistenttimar mellom anna for å ha ein prosess med studentinvolvering (Sindre et al. 2023). På grunn av den formativ-summativ bruken av testar vart det dessutan naudsynt å lage mange variantar av kvar oppgåve – ein vil kunne sleppe billegare dersom testane er noko mindre transparente slik at ein kan ha færre variantar (sjølv om vi opplevde transparensen som ein positiv verdi i gjennomføringa av emnet). Likeeins vil ein kunne gjere det enklare for seg sjølv ved til dømes å berre ha modultestar i staden for testar + prosjekt (både i førebuing og gjennomføring). Ein annan faktor som kanskje kan synest skremmande er det svært radikale opplegget med modular som er kopla direkte til karakter, og der studentar kan stogge straks dei har stått, som nok bidrog til den høge toppen vi fekk på E. Her er det også mogeleg å gå mindre radikalt til verks dersom ein ønskjer det, til dømes ved at alle må ta alle modulane, med noko lågare ståkrav, t.d. 70% som hos (Offutt et al. 2017), men at studentar får betre karakter ved å score betre enn ståkravet, i staden for å ha berre stått/ikkje for kvar modul slik vi hadde, med 90% grense. Det finst mange mogelege måtar å implementere meistringslæring på og vår treng slett ikkje vere den beste eller rette for alle fag.

Keller sin klassiske artikkel om PSI frå 1968 hadde tittelen «Good-bye teacher...» (Keller 1968) – ut frå ideen om at læraren ikkje lenger skulle halde plenums førelesingar, som blir meningslause dersom studentar held mange ulike tempo og såleis er på ulike stader i pensum. Keller sine hovudprinsipp var at læraren sin kommunikasjon med studenten primært gjekk via skriftleg materiale (study guides) og at direkte kommunikasjon med studentane andlet-til-andlet mest var ved læringsassistentar, som også typisk hadde ansvar for å rette modultestane (som den gongen måtte gjerast manuelt) – og desse modultestane skulle ha høge ståkrav (t.d. 90%, som vi også hadde) fordi ein student skulle ha meistra ein modul før ein gjekk vidare til neste. Vi har behalde liknande høge ståkrav, og skriftleg materiale (vel å merke Jupyter Notebooks med kodeeksempel, pluss tilhøyrande videoar, men det kan sjåast som ei naturleg utvikling – Keller ville sannsynlegvis også ha brukt digitale læringsressursar dersom han skulle ha funne opp PSI no). Faglærer i IT1001 fekk likevel ikkje noka «Good-bye teacher»-kjensle av undervisningsopplegget som vi køyrtte, men opplevde tvert imot å kome mykje nærmare studentane enn han nokon gong har gjort i førelesingsbaserte emne med om lag same studenttal. Dette skuldast nok dels at vi gjorde noko som ikkje var vanleg i tradisjonell PSI, nemleg ha seminar med obligatorisk oppmøte, der studentane plasserte seg ved ulike bord etter ønskje og behov og ofte samarbeidde i grupper. Ved å berre gi ein kjapp motivasjonsprat dei første fem minutta av eit seminar, og deretter gå rundt mellom borda og sjå på kva studentane dreiv med å svare på spørsmål, fekk faglærer mykje betre inntrykk av korleis studentane jobba og kva dei sleit med. Ved å sjå over testresultata frå kvar veke visste dessutan faglærer heile tida korleis klassen låg an: Kven har allereie stått, kven held eit høgt tempo og ser ut til å sikte mot A, kven har kome seint i gong og kan vere i fare for å stryke, kven kan trengje litt ekstra merksemd og oppmuntring?

REFERANSER

Aggarwal, Ashish, Neelima Puthanveetil, and Christina Gardner-Mccune. 2023. "Who Attempts Optional Practice Problems in a CS1 Course? Exploring Learner Agency to Foster Mastery Learning." Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1. https://dl.acm.org/doi/abs/10.1145/3545945.3569854?casa_token=6kY8SDMIJ-gAAAAA:fKIQRU29pMrVAd-npTvhsPZanzG3QG5qrSvv1P0_dFX3DO2YZR0kqBFmF60HOMbNjP6h-OnpRipC

- Alvarez, Claudio, Maira Marques Samary, and Alyssa Friend Wise. 2023. "Modularization for mastery learning in CS1: a 4-year action research study." *Journal of Computing in Higher Education*:1-44. <https://link.springer.com/article/10.1007/s12528-023-09366-1>
- Bloom, Benjamin S. 1968. "Learning for Mastery. Instruction and Curriculum. Regional Education Laboratory for the Carolinas and Virginia, Topical Papers and Reprints, Number 1." *Evaluation comment* 1 (2):n2. <https://eric.ed.gov/?id=eD053419>
- Bloom, Benjamin S. 1971. "Mastery learning." In *Mastery learning: Theory and practice*, edited by J. H. Block, 47-63. New York: Holt, Rinehart & Winston.
- Buskist, William, David Cush, and Richard J DeGrandpre. 1991. "The life and times of PSI." *Journal of Behavioral Education* 1 (2):215-234. <https://link.springer.com/article/10.1007/BF00957005>
- Campbell, Jennifer, Andrew Petersen, and Jacqueline Smith. 2019. "Self-paced mastery learning cs1." Proceedings of the 50th acm technical symposium on computer science education. https://dl.acm.org/doi/abs/10.1145/3287324.3287481?casa_token=8C2c90X9O4IAAAAA:cX_HaAGAj4f3ujawjzS28rk_37WV3fwWE7pCXWimDfmvaXI87EDT_znmLGHmJvR6Kd4jTiOfNPa
- de Pontes, Rafael G, Dalton DS Guerrero, and Jorge CA de Figueiredo. 2019. "Analyzing gamification impact on a mastery learning introductory programming course." Proceedings of the 50th ACM technical symposium on computer science education. https://dl.acm.org/doi/abs/10.1145/3287324.3287367?casa_token=8LGUe_wSnzYAAAAA:HB2TvXEt6F9qhOGJs-CjB_xyAR-DkFWR-f3JYM5W_MqejaJCSr_eded4KSfL5lc0EG4ftROQIZQ
- Epstein, Michael L, Amber D Lazarus, Tammy B Calvano, Kelly A Matthews, Rachel A Hendel, Beth B Epstein, and Gary M Brosvic. 2002. "Immediate feedback assessment technique promotes learning and corrects inaccurate first responses." *The Psychological Record* 52:187-201. <https://link.springer.com/article/10.1007/BF03395423>
- Eyre, Heidi L. 2007. "Keller's Personalized System of Instruction: Was it a Fleeting Fancy or is there a Revival on the Horizon?" *The Behavior Analyst Today* 8 (3):317. <https://psycnet.apa.org/fulltext/2008-05986-007.html>
- Fox, Eric J. 2004. "The personalized system of instruction: A flexible and effective approach to mastery learning." In *Evidence-based educational methods*, 201-221. Elsevier. <https://www.sciencedirect.com/science/article/abs/pii/B9780125060417500139>
- Francis, Paul, Cristina Figl, and Craig Savage. 2009. "Mastery learning in a large first year physics class." Proceedings of The Australian Conference on Science and Mathematics Education. <https://openjournals.library.sydney.edu.au/IISME/article/view/6218>
- Garner, James, Paul Denny, and Andrew Luxton-Reilly. 2019. "Mastery learning in computer science education." Proceedings of the Twenty-First Australasian Computing Education Conference. https://dl.acm.org/doi/abs/10.1145/3286960.3286965?casa_token=XQNxr6m81i8AAAAA:Gc0ZJNIFIAXzIY6NJ71QzJ7ErXALRUbb08v4zLhKwt-avDr4z9VCH9NT-Ap-KW5fD0z0ToNoa_C-
- Izu, M, D Ng, and H Weerasinghe. 2023. "Mastery Learning and Productive Failure: Examining Constructivist Approaches to teach CS1." <https://www.ppig.org/files/2022-PPIG-33rd-izu.pdf>
- Jazayeri, Mehdi. 2015. "Combining mastery learning with project-based learning in a first programming course: An experience report." 2015 IEEE/ACM 37th IEEE International Conference on Software Engineering. https://ieeexplore.ieee.org/abstract/document/7202980?casa_token=7in5P-zYibkAAAAA:GLPC1VRpmRYcVDC5zh8E677Dti4XIY86zrjydBcNpb2nv9y-ofSCyoFTfAZRku-wg4tL-3
- Keller, Fred S. 1967. "Engineering personalized instruction in the classroom." *Revista Interamericana de Psicología/Interamerican Journal of Psychology* 1 (3). <https://journal.sipsych.org/index.php/IJP/article/view/445>
- Keller, Fred S. 1968. "Good-bye, teacher...." *Journal of applied behavior analysis* 1 (1):79-89. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1310979/>
- Kulik, Chen-Lin C, James A Kulik, and Robert L Bangert-Drowns. 1990. "Effectiveness of mastery learning programs: A meta-analysis." *Review of educational research* 60 (2):265-299. <https://journals.sagepub.com/doi/abs/10.3102/00346543060002265>
- Kulik, James A, Chen-Lin C Kulik, and Peter A Cohen. 1979. "A meta-analysis of outcome studies of Keller's personalized system of instruction." *American Psychologist* 34 (4):307. <https://psycnet.apa.org/record/1979-25160-001>
- Mannion, Arlene, Rory Coyne, Chiara Ferrari, Melike Neşeli, Ciara McGee, Sumeyye Mollaoglu, and Geraldine Leader. 2023. "Personalized System of Instruction in Higher Education: A Systematic Review." *Journal of Behavioral Education*:1-41. <https://link.springer.com/article/10.1007/s10864-023-09530-8>
- Offutt, Jeff, Paul Ammann, Kinga Dobolyi, Chris Kauffmann, Jaime Lester, Upsorn Praphamontriping, Huzefa Rangwala, Sanjeev Setia, Pearl Wang, and Liz White. 2017. "A novel self-paced model for teaching programming." Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale. <https://dl.acm.org/doi/abs/10.1145/3051457.3053978>
- Ott, Claudia, Brendan McCane, and Nick Meek. 2021. "Mastery learning in cs1-an invitation to procrastinate?: Reflecting on six years of mastery learning." Proceedings of the 26th ACM Conference on Innovation

- and Technology in Computer Science Education V. 1.
https://dl.acm.org/doi/abs/10.1145/3430665.3456321?casa_token=iPt7MYcPuPcAAAAA:HWkbGq7xrM03aEFzUDQETd6TW7GyJovev0CJeq8y8cf5jbsHWPBZgfDJJHDJx78MPBCsZwnrpKFi
- Pérez, Carlos L, and Dina Verdín. 2023. "A Systematic Literature Review for Mastery Learning in Undergraduate Engineering Courses." *International Journal of Engineering Education* 39 (6):1358-1385. <https://peer.asee.org/mastery-learning-in-undergraduate-engineering-courses-a-systematic-review>
- Purao, Sandeep, Maung Sein, Hallgeir Nilsen, and Even Åby Larsen. 2017. "Setting the Pace: Experiments With Keller's PSI." *IEEE Transactions on Education* 60 (2):97-104. https://ieeexplore.ieee.org/abstract/document/7553543?casa_token=d1UGBGi7snIAAAAA:aVtxqjt0kMeneIYRET5JPZXCJnJMDrBDLnAEr1vrBSnZ7URGvM7MIXDJ5rdvy8YZPNPFBBIO
- Reiser, Robert A. 1984. "Reducing student procrastination in a personalized system of instruction course." *ECTJ* 32 (1):41-49. <https://link.springer.com/article/10.1007/bf02768768>
- Reiser, Robert A, and Howard J Sullivan. 1977. "Effects of self-pacing and instructor-pacing in a PSI course." *The Journal of Educational Research* 71 (1):8-12. <https://www.tandfonline.com/doi/abs/10.1080/00220671.1977.10885024>
- Robins, Anthony. 2010. "Learning edge momentum: A new account of outcomes in CS1." *Computer Science Education* 20 (1):37-71. <https://www.tandfonline.com/doi/abs/10.1080/08993401003612167>
- Sindre, Guttorm, Gabrielle Hansen, Guri Sivertsen Korpås, Aksel Kirknes, Johanna Xiaoli Skjøien, and Juni L Magnussen. 2023. "Meistringslæring i intro programmering: plan for eit nytt emne." *Læring om læring* 10 (1). <https://www.ntnu.no/ojs/index.php/lo/article/view/5087>
- Toti, Giulia, Guoning Chen, and Sebastian Gonzalez. 2023. "Teaching CS1 with a Mastery Learning Framework: Impact on Students' Learning and Engagement." Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1. https://dl.acm.org/doi/abs/10.1145/3587102.3588844?casa_token=whbfzzN1zlwAAAAA:X_48oQ1cog u7XkZtWGp6le2N7zbl6hbX890VQ3_VF0vf5JpBnQu6luHaY0KOiTYQyTwPI5EHVDs
- Treviño, Yolanda Martínez, and María Raquel Landa Cavazos. 2018. "Effects of immediate feedback using ICT in a CS1 course that implements Mastery Learning." 2018 IEEE Frontiers in Education Conference (FIE). https://ieeexplore.ieee.org/abstract/document/8658845?casa_token=IeYWEJk-jVcAAAAA:37XKvg0sirZwslcq45MOe37EFyHi2Kn6gveSCMQSPN3tG06INGLSWOWDOPyI7Bjvo322Z5Qa
- Tullis, Jonathan G, and Aaron S Benjamin. 2011. "On the effectiveness of self-paced learning." *Journal of memory and language* 64 (2):109-118. https://www.sciencedirect.com/science/article/pii/S0749596X10000999?casa_token=ELViPnPfPhcAAA:NRRLi8tJGdr_T9ZK9c_USwPIxb2MnwpSzFzqQ-NFYOuYpobweEW6OdmH2CV_YMLUQZdZMcCizQ
- Wiggins, Benjamin L, Leah S Lily, Carly A Busch, Méta M Landys, J Gwen Shlichta, Tianhong Shi, and Tandi R Ngwenyama. 2023. "Public exams may decrease anxiety and facilitate deeper conceptual thinking." *Journal of STEM Education: Innovations and Research* 24 (2). <https://www.jstem.org/jstem/index.php/JSTEM/article/view/2624>