

Automatisk rettet eksamen i generell kjemi

M. T. P. Beerepoot & J. A. Kosonen

Institutt for kjemi, Fakultet for naturvitenskap og teknologi, UiT Norges arktiske universitet

SAMMENDRAG: En god avsluttende eksamen vurderer i hvilken grad intendert læringsutbytte har blitt oppnådd på en nøye, objektiv og reproducerbar måte. Det er imidlertid lite tilfredsstillende hvis en faglærer bruker mye tid på vurdering *av* læring framfor vurdering *for* læring underveis i semesteret. Derfor kan det være attraktivt å utforske muligheter med en digital automatisk rettet eksamen, spesielt i emner med mange studenter. I dette bidraget beskriver vi hvorfor og hvordan vi har lagt om en eksamen i et emne i generell kjemi fra en sekstimers bruk-og-kast-eksamen på papir til en tretimers gjenbrukbar automatisk rettet digital eksamen. Vi drøfter også fordeler og ulemper med å bruke en slik eksamen som skoleeksamen eller hjemmeeksamen, systematisk kvalitetssikring av eksamensoppgavene, og spørsmål knyttet til sensorens rolle i sensur av en automatisk rettet eksamen.

NØKKEORD: Digital eksamen, kjemi, hjemmeeksamen, skoleeksamen, kvalitetssikring

1 INTRODUKSJON

En god avsluttende eksamen vurderer i hvilken grad intendert læringsutbytte har blitt oppnådd på en nøye, objektiv og reproducerbar måte. Et ønske om å bruke mest mulig av faglærerens tid på vurdering *for* læring kan gjøre det attraktivt å utforske muligheter med en digital automatisk rettet eksamen, spesielt i emner med mange studenter. En slik eksamen stiller ekstra høye krav til kvalitet på spørsmålene og kan gi utfordringer med for eksempel fusk og overfladisk læring blant studentene (Denyer & Hancock, 2002). Disse utfordringene er overkommelige og det finnes gode hjelpemidler til å lage gode spørsmål (Towns, 2014) og for å unngå ureglementert samarbeid og bruk av hjelpemidler (Nguyen et al., 2020).

Eksamenen i KJE-1001 Introduksjon til kjemi og kjemisk biologi har lenge vært en sekstimers skriftlig skoleeksamen på papir. Eksamen besto av 20 deloppgaver der hver deloppgave for hver av de cirka 200 studentene ble manuelt rettet av en faglærer og skåret på en skala fra 0 til 5 poeng. Prosessen med skriving og kvalitetssikring av eksamensoppgavene startet på nytt hvert år. Flere personer bidro vanligvis med oppgaver og én faglærer forsøkte å sette sammen en eksamen som var balansert, feilfri og med cirka samme vanskelighetsgrad som i tidligere år. Studentene fikk ikke lov å bruke boka eller andre hjelpemidler under eksamen, med unntak av et ark med egne notater, en kalkulator og periodesystemet. Eksamenssensur ble utført av vanligvis tre faglærere rett før, i og rett etter juleferien slik at studentene først fikk eksamensresultat tre til fire uker etter eksamen. Et mindretall av studentene ba om begrunnelse av eksamensresultat, men de aller fleste fikk i praksis bare sluttkarakter som tilbakemelding. Tidligere eksamensoppgaver med løsningsforslag ble gjort tilgjengelig til studentene for øvelse.

Høsten 2020 ble eksamensformen lagt om til en digital automatisk rettet eksamen på tre timer. Følgende prinsipper var førende i utvikling av den nye eksamenen, som kan brukes som (digital) skoleeksamen eller som hjemmeeksamen:

1. Faglæreren skal kunne prioritere vurdering *for* læring på bekostning av utvikling av en bruk-og-kast-eksamen og vurdering *av* læring etter gjennomføring av eksamen. Faglærere og sensorene skal ikke lenger bruke juleferien for å sensurere eksamenen.
2. Eksamen skal være forutsigbar, rett på læringsmålene i emnet og samstemt med øvelsetester og obligatoriske tester i løpet av semesteret slik at eksamensformen oppfordrer til øvelse og læring. Det skal være lov å bruke hjelpemidler på eksamen og i testene i løpet av semesteret, akkurat som i anvendelse av den oppnådde kunnskapen etterpå.
3. Studentene må vise at de mestrer *alle* intenderte læringsmål for å få adgang til eksamen. Studentene som har bestått arbeidskravene uten fusk skal kunne bestå eksamen.

Målet med dette bidraget er å beskrive den nye eksamensformen, samt å drøfte fordeler og ulemper med å bruke en slik eksamen som skoleeksamen eller hjemmeeksamen, systematisk kvalitetssikring av eksamensoppgavene, og spørsmål knyttet til sensorens rolle i sensur av en automatisk rettet eksamen.

2 METODE

2.1 Beskrivelse av emnet

KJE-1001 Introduksjon til kjemi og kjemisk biologi (10 studiepoeng) er et obligatorisk emne i over ti studieprogram ved tre ulike fakulteter. Emnet gis på høsten og de fleste studenter tar emnet i første semester. Emneinnholdet er fordelt over 14 moduler/uker med fire til seks læringsmål per modul. Hver uke har studentene to til tre timer med pollbasert fellesundervisning (Coucheron & Beerepoot, 2023) og to timer med seminarundervisning med gruppediskusjon basert på begrepsoppgaver (Netland et al., 2018). For å få adgang til eksamen må studentene ha bestått tre arbeidskrav: 1) laboratoriekurs, 2) oppmøte på minst 10 av 14 seminarer og 3) minst 10 av 12 poeng på begge innleveringene og alle fem obligatoriske automatisk rettede tester. Innleveringene blir rettet manuelt og tester læringsmål som ikke kan bli vurdert gjennom automatisk rettede tester (f.eks. å tegne, å forklare). Studenter som tar emnet på nytt må kun bestå det tredje arbeidskravet på nytt. Emnet legger opp til bruk av dokumentert effektive læringsstrategier og studentene har blant annet tilgang til en digital øvelsestest for hver modul, der studentene får ulike spørsmål hver gang de tar testen (Beerepoot, 2022).

2.2 Beskrivelse av eksamenen

Den digitale gjenbrukbare eksamenen på tre timer er utformet i programmet WISEflow og består av tre oppgavetyper som alle kan rettes automatisk:

1. **Flervalgsoppgaver** med vanligvis tre eller fire svaralternativer (Townes, 2014). Eksempel: «Bestem hvilket av følgende molekyler har høyest kokepunkt ut fra type og styrke av intermolekylære krefter» med svaralternativene for én av variantene: H_2O , H_2S og H_2Te . Oppgaven gir to poeng.
2. **Matriseoppgaver** med vanligvis tre til fem rader og to til tre kolonner. Eksempel: «Bestem om følgende redoksreaksjoner er spontane prosesser» med fem redoksreaksjoner (rader) der studentene må velge enten «Spontan» eller «IKKE spontan». Oppgaven gir ett poeng per rad.
3. **Rangeringsoppgaver** med vanligvis tre grunnstoff, ioner, molekyler, salter eller bindinger som må rangeres på en bestemt måte. Eksempel: «Ranger følgende grunnstoff etter atomstørrelse fra det minste (venstre) til det største (høyre)» med svaralternativene for én av variantene: aluminium, silisium, svovel og natrium. Oppgaven gir to poeng.

I 2021 ble eksamen gitt som hjemmeeksamen med ti ulike versjoner som var forskjellige på alle 51 oppgaver. Noen av disse oppgavene hadde i tillegg ulike varianter innenfor samme versjon, slik at ingen student fikk en helt identisk eksamen. Høsten 2022 ble eksamen gitt som skoleeksamen med én versjon med 54 oppgaver hvorav 33 flervalgsoppgaver, 13 matriseoppgaver og 8 rangeringsoppgaver. Antall oppgaver med ulike varianter innenfor én versjon hadde da blitt utvidet til cirka halvparten, som er en avgjørende faktor for at eksamenen skal være gjenbrukbar. De ulike variantene på en oppgave har samme formulering, men for eksempel ulike molekyler som svaralternativer, ulike reaksjoner som rader i en matriseoppgave eller ulike grunnstoff som må rangeres. Vanskelighetsgraden er ideelt sett likt mellom de ulike variantene og for noen oppgaver har dette blitt undersøkt og justert med hjelp av resultater fra øvelsestester. I tillegg til disse strategiene for å redusere *faktisk* likhet mellom eksamenssettene ble to strategier brukt for å redusere *opplevd* likhet: svaralternativene i flervalgsoppgaver og rader i matriseoppgaver ble stokket om, og hver student fikk oppgavene i en tilfeldig rekkefølge. Selv om rekkefølge av svarene kan påvirke vanskelighetsgraden, vil denne effekten jevne seg ut over eksamenen som helhet (Tellinghuisen & Sulikowski, 2008). På denne måten vil to studenter som sitter ved siden av hverandre nesten aldri få opp to like oppgaver. Det var imidlertid mulig for studentene å navigere fram og tilbake mellom oppgavene og studentene ble anbefalt å bruke den muligheten. Selv om begrensning av denne navigasjonen kan motvirke ulovlig samarbeid, bør man være forsiktig med å bruke dette virkemiddelet (Sindre, 2021). I tillegg til nevnte strategier for å unngå samarbeid fikk studentene også en antifuskerklæring (Nguyen et al., 2020) tilpasset denne eksamenen, som de bekreftet ved å levere eksamen. Det er all grunn til å tro at det er nødvendig å ta eksamensfusk seriøst i utforming av eksamener (Nguyen et al., 2020; Sindre, 2021) og for denne eksamenen var fokus

på *prevensjon* framfor *deteksjon* av fusk. Bruk av hjelpemidler som lærebok, egne notater og ressurser på læringsplattformen er tillatt for denne eksamenen.

En generisk sensorveiledning ble utviklet, som inneholder blant annet en tabell for konvertering fra poengsum til bokstavkarakter (A-F). En ekstern sensor har deltatt i utforming av eksamen og sensorveiledning. Den interne sensoren har hatt muligheten til å se på studentenes svar før poengsum ble konvertert til karakter.

2.3 Spørreundersøkelse: studentenes perspektiver på eksamensform

Studentenes perspektiver på valget mellom skole- og hjemmeeksamen ble undersøkt i en digital avsluttende kursevaluering i november 2022. Spørreundersøkelsen inneholdt spørsmål om ulike aspekter av emnet og ble gjennomført i siste undervisningstime og i tillegg sendt ut til studenter som ikke var til stede. I en innledende tekst til delen om eksamen ble det tydelig spesifisert at spørreundersøkelsen ikke hadde noen innvirkning på eksamensform høsten 2022, men at eksamensformen for høsten 2023 ikke var avklart ennå. Ett lukket spørsmål med mulighet for utdypning handlet om eksamen: «Ville du foretrekke hjemmeeksamen eller skoleeksamen om du kunne velge?» med tre svarmuligheter «jeg ville foretrekke skoleeksamen», «jeg ville foretrekke hjemmeeksamen» og «vet ikke / nøytral». For å samle en begrunnelse for valget fikk studentene som valgte skoleeksamen (hjemmeeksamen) mulighet til å utdype hva de ser som fordeler med skoleeksamen (hjemmeeksamen) og hva de ser som ulemper med hjemmeeksamen (skoleeksamen). Studentene som svarte «vet ikke / nøytral» fikk også mulighet til å utdype sitt svar. Utsagnene i utdypningene ble kategorisert ut fra argumentene som ble brukt med mulighet for at utsagn med flere argumenter kunne tilhøre flere kategorier. Begge forfatterne telte antall utsagn i hver kategori og diskuterte avvik fram til enighet.

3 RESULTATER: SKOLEEKSAMEN ELLER HJEMMEEKSAMEN?

Studentenes perspektiver på valget mellom skole- og hjemmeeksamen ble undersøkt *før* selve eksamenen. De fleste studentene hadde på dette tidspunktet ingen eksamenserfaring fra høyere utdanning. Resultatene representerer dermed i større grad studentenes forventninger enn deres erfaringer. Av de 104 studentene som svarte på kursevalueringen antydte nesten halvparten (N=50; 48%) at de ville foretrekke hjemmeeksamen og 30% (N=31) at de ville foretrekke skoleeksamen. Resten (N=23; 22%) svarte «vet ikke / nøytral». Til sammen var det 115 utdypninger. Utdypningene med argumenter for eller mot én av alternativene har blitt kategorisert og er presentert i Tabell 1.

Tabell 1. Argumenter for skoleeksamen og mot hjemmeeksamen fra studenter som foretrekker skoleeksamen, og argumenter for hjemmeeksamen og mot skoleeksamen fra studenter som foretrekker hjemmeeksamen. Kategorier med minst fire utsagn er gjengitt i tabellen.

Utsagn fra studenter som foretrekker skoleeksamen	# utsagn
Det er lettere å konsentrere seg på en skoleeksamen.	22
Det er lett å jukse på en hjemmeeksamen.	9
En skoleeksamen er lettere.	5
En skoleeksamen fører til mer øvelse og mer læring.	5
Utsagn fra studenter som foretrekker hjemmeeksamen	
Det er lettere å konsentrere seg på en hjemmeeksamen.	22
En hjemmeeksamen er mindre stressende / mer behagelig enn en skoleeksamen.	48

Interessant nok er det like mange utsagn som tyder på at det er lettere å konsentrere seg på en skoleeksamen som utsagn som tyder på det motsatte. Eksempler fra studenter som tror det er lettere å konsentrere seg på en skoleeksamen er at «det er lettere å konsentrere seg i et arbeidsmiljø» og at det er «lettere å bli distraheret om man sitter hjemme». Eksempler fra studenter som tror det er lettere å konsentrere seg på en hjemmeeksamen peker imidlertid på det motsatte, som for eksempel «mener selv det er mer ro når jeg sitter alene og kan konsentrere lettere» og «av egne erfaringer vet jeg at jeg er en som lett blir påvirket av lyder eller bevegelser som skjer i rommet. Det kan lett utgjøre en forskjell på resultatet jeg vil ende opp med på eksamen». Det er ikke sikkert at studentenes forventninger om en

skoleeksamen stemmer overens med den faktiske situasjonen, gitt at de fleste aldri hadde hatt en eksamen i høyere utdanning da de svarte på spørreundersøkelsen.

Det er derimot mer enighet at idéen med en skoleeksamen oppleves som mer stressende og at en hjemmeeksamen er «litt mer chill». Mange utsagn tyder på at studentene synes det er mer komfortabel hjemme, for eksempel fordi «man kan sitte der man synes [det] er behagelig og gjøre situasjonen mindre stressende». Noen skriver at stress og nervøsitet fører til at man glemmer ting. Én student argumenterer mot skoleeksamen fordi vedkommende «ser på det som problematisk at man skal bli tatt ut av sitt vanlige miljø når man skal vise hva man kan».

Andre argumenter for en skoleeksamen som kommer tilbake i minst fire utsagn er knyttet til vanskelighetsgrad på eksamen, juks, effekt på øvelse/læring eller kombinasjoner av disse. Utsagn som forklarer disse argumentene er for eksempel: «det er mye lettere å jukse på hjemmeeksamen og dermed lærer man ikke stoffet like bra, tror jeg»; på en skoleeksamen er det «[...] ikke noe kommunikasjon. Det gjør eksamen mer fair, da de med svakere nettverk ikke vil kunne få like mye hjelp [på en hjemmeeksamen] som de med sterkt nettverk rundt seg»; «i tillegg vil kanskje ikke eksamensoppgavene bli mer avanserte [på en skoleeksamen] fordi de vet at folk ikke kommunisere og kan hjelpe hverandre»; «Man øver ikke like mye da man tenker man googler og kan få hjelp» på en hjemmeeksamen og det er «lett å jukse og kommunisere med andre. Blir ikke en real karakter vis man skal kunne få svar av andre».

Én av studentene som svarte «vet ikke / nøytral» oppsummerer: «Tenker at hvor man sitter ikke er så viktig så lenge det er en eksamensform som tester hvordan en klarer å vise og bruke den kunnskapen man har. Men det vil være mye lettere å kontrollere at folk ikke samarbeider med en skoleeksamen».

4 DISKUSJON

Utvikling av en automatisk rettet eksamen med mange spørsmålsvarianter krever en stor tidsinvestering og stiller høye krav til kvalitet på oppgavene. Når man er ferdig med utviklingen, er imidlertid fordelene mange. Faglæreren bruker lite eller ingen tid for forberedelse og sensur av eksamenen. Studentene får raskt tilbakemelding på eksamensresultat. Det er mulig å vurdere mange ulike læringsmål på kort tid slik at det ikke er nødvendig å ta et utvalg av læringsmål som vurderes på eksamen. Eksamensnivået er garantert likt fra år til år og fra ordinær eksamen til kontinuasjonseksamen. Sensur er helt objektiv og studentene vet dette, som fører til færre eller ingen klager. Til slutt vil ikke eksamenen endre seg fra år til år, som muliggjør systematisk kvalitetsarbeid med eksamen over årene.

Studentenes perspektiver fra kursevalueringen tyder på at det er noen studenter som mener de kan konsentrere seg best hjemme, mens andre mener de kan konsentrere seg best på en skoleeksamen. Økt stress på en skoleeksamen er imidlertid noe som taler for en hjemmeeksamen, i hvert fall ifølge studentene. Det er relativt få studenter som argumenterer for en skoleeksamen fordi det fører til mer læring og fordi det er lett å jukse på en hjemmeeksamen, men begge er viktige argumenter for en faglærer. Særlig rettferdighetsprinsippet og prevensjon av fusk er viktig å tenke på. Sindre (2021) deler inn fusk i tre kategorier: alenefusk (for eksempel bruk av ulovlige hjelpemidler), samarbeid (mellom kandidater som tar eksamenen) og hjelp fra tredjeperson (som ikke tar eksamenen samtidig). Alenefusk er ikke et stort problem for denne eksamenen fordi hjelpemidler er tillatt. Både samarbeid og hjelp fra tredjeperson er betydelig lettere å motvirke – men ikke helt mulig å utelukke – på en skoleeksamen der studentene er under observasjon. Selv om hjelp fra en tredjeperson absolutt ikke kan utelukkes verken i arbeidskrav eller på eksamen, så er dette trolig et mindre problem i dette emnet siden en student må bestå både eksamen og ulike tester, innleveringer og andre arbeidskrav gjennom semesteret. Samarbeid er i praksis umulig å unngå på en hjemmeeksamen og er sannsynligvis det største problemet for denne eksamenen. Selv om antifuskerklæringen og den tilfeldige rekkefølgen av spørsmål og svaralternativer vil redusere mengde samarbeid til en viss grad, vil de ikke kunne forhindre det. Økt tidspress kan gjøre at færre har mulighet å delta i samarbeid, men kan også føre til mer eksamensangst og dermed mer fusk (Sindre, 2021). Den mest effektive og rettferdige måten å begrense samarbeid på er derfor å gi eksamenen som skoleeksamen.

Uansett om eksamenen blir brukt som skole- eller hjemmeeksamen, stiller en automatisk rettet eksamen enda høyere krav til kvalitet på oppgavene enn en manuelt rettet eksamen. Studentene har ingen mulighet til å begrunne sine svar og dermed er det vanskelig for en sensor å oppdage uheldige formuleringer i oppgaveteksten. Studentene har som regel ingen innsikt i sin besvarelse etter eksamen, og dermed kan

de heller ikke argumentere for sine svar i etterkant av sensur. Dette legger et stort ansvar på faglærerne som utvikler eksamenen, og det stilles høye krav til kvalitetssikring. Ulike rutiner for kvalitetssikring har blitt brukt i utviklingen av den aktuelle eksamenen. Flervalgsoppgavene har blitt brukt og kvalitetssikret i *no-stake* øvelsestester og *low-stake* obligatoriske tester før de ble brukt i eksamenen. Gjennom analyse av testresultatene har eventuelle feil blitt rettet opp og dårlige spørsmål fjernet. Tre faglærere har testet hele eksamen flere ganger, diskutert eventuelle gale svar og forbedret oppgavene i prosessen. Under sensuren av første bruk av eksamenen har kun én feil blitt funnet og rettet opp for alle studenter som fikk den aktuelle spørsmålsvarianten. Kvalitetssikringsrutinen som ble brukt for å finne denne feilen var å gå gjennom besvarelsene fra studentene som hadde kun én eller to feil i eksamen med en hypotese at det var en feil i eksamenssettet. Hypotesen viste seg å være riktig kun denne ene gangen. Den store fordelene med denne eksamensformen er at slike rutiner for kvalitetssikring fører til en systematisk forbedring av kvalitet på eksamenen gjennom årene, i motsetning til kvalitetssikring av en bruk-og-kast-eksamen der man starter på nytt hver eneste gang.

Bruk av en automatisk rettet eksamen med mange spørsmålsvarianter stiller også andre krav til rollen til en ekstern sensor enn en mer tradisjonell eksamen. En ekstern sensor kan ha en rolle i utforming av en eksamen. Det er imidlertid uvanlig både at en eksamen har mye variasjon fra student til student innenfor samme år og at den har lite eller ingen variasjon fra år til år. Burde en ekstern sensor se på hele eksamenssettet på nytt hvis ett spørsmål blir lagt til eller byttet ut? Burde en ekstern sensor gå gjennom alle mulige varianter av et spørsmål og – hvis ja – hvordan? En ekstern sensor kan også ha en rolle i vurdering av studentenes besvarelser. Denne vurderingen gjennomføres imidlertid i utgangspunkt automatisk. Hvordan kan en ekstern sensor bidra til kvalitet av sensur? Hva gjør en eventuell klagesensor, som blir bedt om å vurdere besvarelsen på nytt ut fra samme sensorveiledning, som inneholder konverteringstabellen fra poengsum til karakter? Hvordan kan vi ivareta studentenes rettigheter på klage hvis verken den første sensoren eller klagesensoren egentlig foretar en vurdering av studentenes besvarelse? Disse er noen av spørsmålene som burde diskuteres når man bruker en automatisk rettet eksamen.

REFERANSER

- Beerepoot, M. T. P. (2022). Effektive læringsstrategier – Innsikter, implementering og tverrfaglig erfaringsutveksling. *Læring om læring*, 9(2), 9. <https://www.ntnu.no/ojs/index.php/lo/article/view/5020>
- Coucheron, D. A. & Beerepoot, M. T. P. (2023). Studentaktiv læring med store studentgrupper – pollbasert undervisning på nett. *Konferansebidrag til MNT-konferansen 2023*.
- Denyer, G., & Hancock, D. (2002). Graded multiple choice questions: Rewarding understanding and preventing plagiarism. *Journal of Chemical Education*, 79(8), 961. <https://doi.org/10.1021/ed079p961>
- Netland, K. Ø., Sivertsen, A. & Olufsen, M. (2018). Innføring av studentaktive arbeidsformer i seminarundervisningen. Hvilken betydning har dette på læringsutbyttet og klassemiljøet? *Nordic Journal of STEM Education*, 2(1), 15. <https://www.ntnu.no/ojs/index.php/njse/article/view/2346>
- Nguyen, J. G., Keuseman, K. J., & Humston, J. J. (2020). Minimize Online Cheating for Online Assessments During COVID-19 Pandemic. *Journal of Chemical Education*, 97(9), 3429-3435. <https://doi.org/10.1021/acs.jchemed.0c00790>
- Sindre, G. (2021). Kan fusk på hjemmeeksamen forhindres? *Nordic Journal of STEM Education*, 5(1), 5. <https://doi.org/10.5324/njsteme.v5i1.3918>
- Tellinghuisen, J., & Sulikowski, M. M. (2008). Does the answer order matter on multiple-choice exams? *Journal of Chemical Education*, 85(4), 572. <https://doi.org/10.1021/ed085p572>
- Towns, M. H. (2014). Guide to developing high-quality, reliable, and valid multiple-choice assessments. *Journal of Chemical Education*, 91(9), 1426-1431. <https://doi.org/10.1021/ed500076x>