

# Evaluering af fællesfaglig problemløsning – med afsæt i en netbaseret SSI-case



Lars Brian Krogh,  
lektor, VIA  
University College,  
læreruddannelsen  
i Aarhus



Peer Daugbjerg,  
VIA University  
College,  
læreruddannelsen  
i Nørre Nissum

**Abstract:** Vi har designet og pilottestet et netbaseret evalueringsformat for elevernes evne til fællesfaglig problemløsning. Formatet er bygget op omkring en fællesfaglig SSI-case, og elevernes arbejdsproces stilladseres gennem samme trin som den fælles prøve i naturfagene. Der indsamles procesnære elevartefakter fra processen som kulminerer med at eleverne laver en argumenteret anbefaling. I artiklen deklarerer prøveformatet og vores analysetilgange til elevsvar (Toulmin- og SOLO-kriterier). Formatet har vist sig særdeles brugbart i et pilotstudie med elever (34 "prøvehold") fra fire skoler. Fra pilotstudiet rapporteres resultater mht. elevernes evne til selvstændigt og uden forberedelse at indsamle, validere og syntetisere viden mhp. fællesfaglig problemløsning.

## Baggrund

I 2017 indførte Undervisningsministeriet en fællesfaglig prøve for naturfagene i grundskolens overbygning. Samtidig med prøvens indførelse startede en følgeforskning af prøven og det fællesfaglige arbejde hen imod denne. Følgeforskningen, som vi er en del af, viser at prøveformatet langt hen ad vejen er velfungerende, og at de fleste grundskolelærere grundlæggende mener at den kompetencerettede og problembaserede prøve er en god måde at efterprøve centrale mål for naturfagsundervisningen (Rambøll et al., 2018). Følgeforskningen påpeger dog samtidig at den problembaserede prøveform også har sine problemer. Eleverne har her typisk fire uger hvor de ved hvilket overordnet fokusområde de skal til prøve i. I denne periode og inden for området skal de så formulere deres egen konkrete problemstilling og faglige arbejdsspørgsmål som de arbejder med at besvare i tiden frem mod prøven. Den lange prøveforberedelsestid betyder at mange elever møder op med artefakter og vejledningspræparation (samtalepapirer, modeller, forhåndsopstillede øvelser) som de i prøvesituationen har svært ved at frigøre sig fra. I vores følgeforskning ser vi også at pigerne karaktermæs-

sigt klarer sig markant bedre end drengene i den fælles prøve, tilmed ganske meget bedre end i den foregående praktisk orienterede fysik/kemiprøve (Krogh et al., 2021). Mange lærere mener at denne bias opstår fordi drengene har svært ved at strukturere deres arbejde i den lange prøveforberedende periode (Rambøll et al., 2020; Rambøll et al., 2018). Endelig mener lærerne at det er svært at forberede eleverne til *begge* fagernes meget forskellige prøveformer, den mundtlige fællesfaglige over for de skriftlige selvrettende prøver i enkeltfagene.

I forlængelse heraf springer en række centrale spørgsmål i øjnene: hvad vil eleverne kunne på egen hånd? Hvad vil de kunne uden den lange prøveforberedende periode? Vil bias til fordel for pigerne forsvinde hvis man fjerner denne periode? Svarene på disse spørgsmål burde kunne interessere både udskolingslærere og kolleger i feltet som arbejder med evaluering.

I et forsøg på at tilvejebringe svar på spørgsmålene har vi udviklet en alternativ prøve som afvikles uden forberedelse over to lektioner i en almindelig undervisningskontekst. Eleverne arbejder her problembaseret og i par med afsæt i et såkaldt socio-scientific issue (SSI), og deres delbesvarelser indsamles i skriftlig form via internettet. Inden for rammerne af det ændrede format arbejder eleverne sig igennem samme procestrin som til den fælles prøve. Det alternative prøvedesign er inspireret af forskningslitteraturen om socio-scientific issues (SSI) og nature of science (NOS), som ikke tidligere har været anvendt i en dansk evalueringskontekst. Den udviklede prøve og selve afprøvningen af formatet skal således ses som et bidrag til forskningen i evaluering i dansk naturfag, herunder evaluering af (aspekter af) naturfaglig dannelse. Hvilket berettiger at de metodiske aspekter får deres eget forskningsspørgsmål (RQ1 nedenfor), som vi lidt utypisk søger at besvare i artiklens afsnit om metode.

Empirisk er der tale om et pilotstudie med data fra fire 9.-klasser fra forskellige skoler. Dataene belyser såvel evalueringsformatets duelighed som elevernes evne til selvstændigt og uden forberedelse at udvikle et bud på løsning af en fællesfaglig problemstilling. Eller mere mundret og kompakt: deres evne til "fællesfaglig problemløsning". Selvom begge dele på forskellig vis kaster lys over den eksisterende fællesfaglige prøve, så er det vigtigt for os at pointere at det aldrig har været et mål at vores alternative prøveformat skulle erstatte denne.

Med denne pointe på plads overgår vi til at besvare artiklens konkrete forskningsspørgsmål:

RQ1: Hvordan evaluere elevens evne til at belyse en fællesfaglig problemstilling vha. en SSI-baseret case og et netbaseret, undervisningsnært evalueringsformat?

RQ2: Hvilket billede tegner der sig af elevernes evne til situeret og selvstændigt at belyse en fællesfaglig problemstilling (herunder informationssøgning, arbejdsspørgsmål, syntese og argumentation)?

## SSI-evalueringsformatet – det teoretiske grundlag

I udgangspunktet har vi ønsket at det alternative prøveformat skulle have størst mulig sammenlignelighed med den fælles prøve i naturfagene. Denne er rammesat i *Bekendtgørelse om folkeskolens prøver* (Bekendtgørelse om folkeskolens prøver, 2019) og detaljebeskrevet i *prøvevejledningen for den fælles prøve* (STUK, 2019). Her tages afsæt i et af elevernes mere eller mindre selvstændige arbejder med problemstillinger inden for fællesfaglige fokusområder som kobler samfundsmæssige udfordringer med naturfaglige aspekter. I bedømmelsen vægtes det højt at eleverne formår at belyse en selvvalgt problemstilling med anvendelse af naturfaglige kompetencer, herunder bl.a. at de “kan argumentere for naturfaglige forhold” og “kan anvise og begrunde relevante handlemuligheder”.

Der er en rig undervisningsagenda og forskningslitteratur omkring socio-scientific issues (SSI) hvor udgangspunktet helt tilsvarende er fagoverskridende problemstillinger (konkrete SSI), og sigtet er at eleverne lærer at argumentere, tage kritisk stilling og handle (“transformative education”, Zeidler, D.L.; Applebaum, S.M.; Sadler, 2011). Det nære slægtskab mellem SSI og hjemlige fællesfaglige problemstillinger fremgår bl.a. af deklARATIONEN på SSI:

“SSI er kontroversielle, samfundsmæssigt relevante og autentiske (“real-world”) problemstillinger, som informeres af naturvidenskab og som ofte inkluderer moralske aspekter.”  
(Sadler et al., 2007).

I SSI-undervisningen formuleres SSI-opgaven typisk som et kontroversielt spørgsmål som eleverne i grupper skaffer sig indsigt i og diskuterer. I SSI-forskningen har der været stort fokus på hvordan eleverne argumenterer og tager stilling til det kontroversielle spørgsmål (fx Garrecht, Eckhardt, Hoffler, Harms, 2020; Kolsto, 2006). En række studier har fokuseret på andre typer af elevudbytte, først og fremmest elevernes vækst i socio-scientific reasoning (Sadler et al., 2007) og øget forståelse af naturvidenskabelige arbejdsmåder (“Nature of Science”) ud fra SSI-drøftelser (se fx Karisan, Zeidler, 2017). Rationalet er her at elever som arbejder kritisk med at belyse en SSI-case, samtidig får erfaringer med science-in-the-making, fx at denne viden er tentativ, kulturelt indlejret og socialt forhandlet m.m. – hvilket er centrale NOS-pointer. Dette er umiddelbart interessant i en dansk kontekst, hvor den fællesfaglige naturfagsprøve har SSI-inspirerede elementer, og de bagvedliggende kompetencemål har NOS-inspirerede elementer. Vigtigere i sammenhæng med nærværende artikel er det imidlertid at der er et overlap mellem NOS-undervisning med afsæt i en *aktuel* kontekst og SSI-undervisning med samme fokus hvorfor det har været muligt at hente inspiration til vores SSI-baserede prøve i et casebaseret NOS-evalueringsformat udviklet af Douglas Allchin (Allchin, 2011). Det er en pointe hos Allchin at NOS-viden må

forstås og evalueres kontekstnært. Vigtigheden af en kriteriebundet og dybdegående analyse af elevrespons er endnu en metodisk pointe som vi tager med fra dette studie. I metodeafsnittet vil disse aspekter blive yderligere udfoldet. Her vil vi også omtale analytiske tilgange som kan bruges til at afdække kvaliteten af elevernes fællesfaglige problemløsning, i særdeleshed deres evne til at syntetisere vidensbaserede svar og argumentere for handlemuligheder i tilknytning til SSI-casens problemstilling.

## SSI-evalueringsformatet – metodiske aspekter

I udgangspunktet har vi ønsket at evalueringsformatet blev håndterbart for lærere og for eleverne, som i 9. klasse må forventes at have visse erfaringer med fællesfaglig tænkning og arbejdsmåder. Vi har derfor tilstræbt at evalueringsformatet ligner autentisk fællesfaglig undervisning mest muligt. Konkret betyder det at evalueringsaktiviteten kan afvikles over to lektioner og som en del af almindelig undervisning. Dette bidrager i høj grad til evalueringens validitet ud fra et socio-kulturelt perspektiv (se fx Gipps, 1999).

Sammen med den internationale SSI-inspiration og ønsket om at inkorporere flest mulige træk af den eksisterende fællesfaglige prøve har det resulteret i følgende specifikke designkriterier for det alternative evalueringsformat:

- Evalueringen skal tage afsæt i en problemstilling der rammesættes af en *SSI-case*.
- Problemstillingen og SSI-casen skal relatere til et af de kanoniserede fællesfaglige fokusområder, være fællesfaglig af natur og omfatte alle tre udskolingsnaturfag.
- Casen fastlægger problemstillingen, men eleverne skal selv formulere faglige arbejdsopgaver og syntetisere argumenterede svar ligesom de skal ekspliciteres hvordan de tre fag har bidraget til svaret.
- Alle sædvanlige hjælpemidler er tilladte – eleverne må arbejde i par hvis de ønsker det, ligesom alle håndte lærebøger og internetkilder må anvendes. Af praktiske årsager har eleverne dog ikke mulighed for selv at gennemføre undersøgelser i laboratoriet eller felten.
- De naturfaglige kompetencer søges inkluderet – så langt som muligt. Da eleverne ikke får mulighed for selv at lave undersøgelser, så vil *praktiske* aspekter af undersøgelseskompetence ikke blive indfanget.
- Der gives en ensartet og meget begrænset stilladsering af elevernes problemløsende proces. Med prompts holdes eleverne således fast på at overvejelser, kilder og argumentation fastholdes i bestemte trin af processen. Fx at deres arbejdsopgaver bliver fastholdt, at deres litteratursøgning ekspliciteres, og at deres konklusion udfoldes og begrundes.
- Elevernes overvejelser, kilder og argumentation undervejs indsamles på en passende platform (konkret Google Forms da dette muliggør at eleverne også kan

uploade kilder som de har fundet værdifulde). Denne artefaktindsamling dokumenterer både elevernes proces og formulerede udbytte og udgør det empiriske grundlag for evalueringen og for forskningsmæssig publikation.

I det følgende vil vi uddybe den anvendte SSI-case, stilladseringen og i øvrigt redegøre for de væsentligste analytiske tilgange ift. det empiriske materiale.

### Evalueringens SSI-case

Ved evalueringens start får eleverne et link til en hjemmeside: ([https://docs.google.com/forms/d/e/1FAIpQLSc\\_-y75iH6AdACkLyPitFOWzjnLRCLTxgNm1fCsA3f5y08SQQ/viewform?usp=sf\\_link](https://docs.google.com/forms/d/e/1FAIpQLSc_-y75iH6AdACkLyPitFOWzjnLRCLTxgNm1fCsA3f5y08SQQ/viewform?usp=sf_link)) hvor de straks præsenteres for casen, her blot omtalt som en “fællesfaglig problemstilling” for at linke til termer og en praksis som eleverne burde være bekendte med i 9. klasse. Indbygget i casebeskrivelsen finder man den reelle problemstilling: er det en god idé at få foretaget en mammografi eller ikke? (Når en række kontekstfaktorer medtænkes).



Sektion 1 ud af 7

## Mammografi - skal, skal ikke?

En fællesfaglig problemstilling, som I skal belyse ved hjælp af Jeres naturfaglige viden og kompetencer.

**Problemstillingen:**

"I har besøg derhjemme af en af din mors veninder, som du kender godt. Hun er lige blevet 40 år og fortæller om sin bekymring for at få brystkræft. Bekymringen startede da nogen på hendes vej fandt ud af, at der var meget Radon i deres hus. Der er vist noget med, at der kommer en stråling fra Radon, som kan give kræft. Nu overvejer din mors veninde, om hun skal få lavet en mammografi-undersøgelse inden for de næste måneder. Men: hun mener, at have hørt, at undersøgelsen i sig selv er noget med stråling, som vel også kan give kræft. Hun har også læst et sted, at der er en anbefaling om, at kun kvinder over 50 får bør få lavet mammografi. I kender alle tre en anden kvinde, der døde i 43-års-alderen af en brystkræft, som blev opdaget for sent. Din mors veninde er derfor usikker på, om hun skal få foretaget undersøgelsen - og spørger Jer til råds.

Du lover at prøve at undersøge tingene nærmere, vel vidende at hun skal gå om 2 timer. Du har internettet til rådighed - og selvfølgelig de sædvanlige naturfaglige lærebøger."

Figur 1. Casen som den præsenteres for eleverne.

Casen relaterer sig til det fællesfaglige fokusområde *Strålings indvirkning på levende organismers levevilkår*, som normalt henlægges til undervisningen i 9. klasse da eleverne først må interagere med radioaktive kilder på dette tidspunkt. Grundidéen er hentet fra D. Allchin (se ovenfor), men da dennes evalueringsscase ikke indlysende gav mulighed for at inddrage geografifaget, har vi selv indføjet radonaspekterne i formuleringen. Dermed skønnede vi at eleverne havde rimelig mulighed for at knytte an til alle tre udskolingsnaturfag i deres besvarelser. Ifm. følgeforskningen omkring

den fælles prøve har vi både observeret klare 12-talspræstationer og velvilligt bedømte O2-præstationer med nært beslægtede problemstillinger som afsæt hvorfor casen burde kunne udspænde og differentiere elevernes formåen på dækkende vis.

### *Processtilladseringen via skriveprompts*

Elevernes selvstændige arbejdsproces guides af nedenstående prompts som er tænkt at fastholde eleverne på at bestemte procestrin skal gennemløbes, men uden at levere *indholdsmæssig* stilladsering. Det er virkelig elevernes selvstændige formåen som indfanges.

#### **Delopgave 1: Formulér arbejdsspørgsmål**

##### **Hvilke naturfaglige spørgsmål er det vigtigt at finde svar på?**

Læs problemstillingen godt igennem. Overvej hvad du behøver af naturfaglig viden for at komme med en anbefaling til din mors veninde. Find de tre-fem NATURFAGLIGE arbejdsspørgsmål som det er vigtigst at finde svar på for at give et godt svar.

SKRIV DE TRE-FEM ARBEJDSSPØRGSMÅL IND I SPØRGSMÅLSFELTET NEDENFOR.

#### **Delopgave 2: Find relevant information**

Du/I skal nu prøve at finde svar på arbejdsspørgsmålene. Dertil må du gerne bruge både lærebøger og internettet.

Lav de mest relevante søgninger, og find materiale som I mener er både *brugbart* og  *troværdigt*.

- Angiv evt. lærebogssider og i hvert fald tre internetmaterialer i boksene nedenfor.
- Hvorfor er disse sider relevante og vigtige?
- Internetsøgningen – hvilke søgeord har I brugt?
- Upload de mest relevante hjemmesider eller internetmaterialer I fandt – hvorfor er de gode?

#### **Delopgave 3: Eksperimenter og undersøgelser som kunne hjælpe dig/er til at give en bedre anbefaling**

Hvis I mener at eksperimenter eller praktiske undersøgelser kunne hjælpe jer til at give bedre svar til veninden, så skriv lidt om disse i nedenstående svarfelt.

- Skriv ud for hver undersøgelse hvilket spørgsmål det hjælper med at belyse.

#### **Delopgave 4: Din/jeres anbefaling og begrundelserne for den**

Overvej på basis af den indsamlede information om du/I vil anbefale veninden at få foretaget en mammografi. Giver den indsamlede viden anledning til andre anbefalinger til hende, så skriv også gerne disse.

- Anbefalinger til veninden.
- Hvorfor hun skal følge anbefalingen? Hvordan vil I argumentere
- over for hende?

#### **Delopgave 5: De forskellige fags bidrag til din/jeres anbefaling**

Du/I skal nu vurdere hvordan de forskellige naturfag har bidraget til din/jeres anbefaling og argumentation: Hvilke af de forskellige naturfag har været i spil? Hvordan har de hver især bidraget til at belyse problemstillingen?

- Vigtigste bidrag fra fysik/kemi (evt. også eksperimenter og modeller) (tilsvarende for biologi og geografi).

**Tabel 1.** *Processtilladsering og procesnær artefaktindsamling.*

Forskningsspørgsmålene i denne artikel knytter især an til delopgaverne 1, 2 og 4 – med delopgave 4 som den centrale problembaserede, SSI-orienterede og handlekompetence-rettede syntese. Som det vil fremgå af analysebeskrivelsen, så er det afgørende at processtilladseringen holder eleverne fast på både at komme med en anbefaling (syntese) og tillige levere en udfoldet argumentation (begrundelse). I resultatafsnittet vil der drypvist blive inddraget empiri om elevernes oplevelse af det alternative evalueringsformat.

## Analytiske tilgange

### A. Analysen af elevernes arbejdsspørgsmål m.m.

I deres review over elevens spørgsmål i science fremhæver Chin & Osborne (2008) at elevernes spørgsmål har flere funktioner som er væsentlige ifm. SSI og fællesfagligt arbejde: De giver indblik i elevernes aktuelle formåen og muliggør dermed evaluering mens de for eleverne både fokuserer og driver deres arbejde.

Vi har hurtigt indset at det er vanskeligt at vurdere elevernes formåen ud fra deres arbejdsspørgsmål alene. Uden mulighed for opfølgende dialog med eleverne har vi her indskrænket os til at vurdere elevernes spørgsmål ud fra prøvevejledningens kategorier (STUK, 2019): videns- og dataspørgsmål, forklarings- og forståelsesspørgsmål, holdnings- og vurderingsspørgsmål og handlingsspørgsmål. Eksempler fra denne analyse er vist i tabel 2:

Elevformulerede arbejdsspørgsmål	Undervisningsfag	Spørgsmålstype
<i>Hvor stor er chancen for at få kræft af radonstrålerne?</i>	Biologi Fysik/kemi	Holdning og vurdering
<i>Og hvad ville konsekvenserne være af ikke at få taget en?</i>	Biologi	Forståelse og forklaring
<i>Hvor i landet kan man få en mammografi?</i>	Geografi	Viden og data
<i>Hvad er radon?</i>	Fysik/kemi	Viden og data

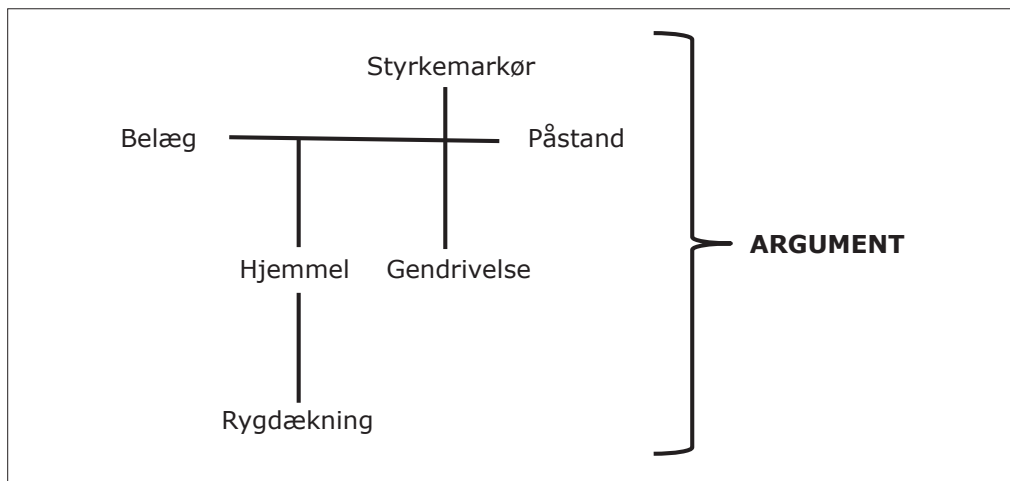
**Tabel 2.** Eksempler på kategoriserede arbejdsspørgsmål.

Det er selvsagt interessant hvorledes arbejdsspørgsmål gennem deres fokuserende funktion indvirker på elevernes formåen. I reviewet af Chin & Osborne er der indikationer (“suggesting ...”) af at der er en sammenhæng mellem kvaliteten af elevernes spørgsmål og deres tilgang til læring (“deep learning”) hhv. deres endelige begrebsforståelse (Chin & Osborne, 2008, p. 19). Dette aspekt har vi også forsøgt at forfølge i vores analyse idet vi forsøgsvist har undersøgt korrelationen mellem elevernes højeste taksonomiske niveau i spørgsmål ift. deres niveau af fællesfaglig argumentation (se nedenfor).



## B. Analysen af elevernes anbefalinger og begrundelserne for disse

**Toulmin-analyse af elevernes argumentation:** Som det ultimative udtryk for elevernes evne til at belyse en fællesfaglig problemstilling har vi analyseret deres anbefalinger til den bekymrede veninde i SSI-casen og den tilhørende argumentation. Andre har tidligere analyseret elevers argumentation, herunder også socio-scientific argumentation (fx Osborne et al., 2004). Mange analyser har fokuseret på i hvilken udstrækning eleverne har været i stand til at generere en argumentation som i struktur og kompleksitet lever op til S. Toulmins argumentmodel (Toulmin, 1958). Modellens hovedkomponenter fremgår af figur 2:



Figur 2. Toulmins argumentmodel.

Den strukturelle kvalitet af et elevargument er typisk gjort op iht. følgende taksonomi (Osborne et al., 2004):

- Niveau 1 (uden belæg): Argumentationen består udelukkende af en eller flere påstande, men ikke belæg af nogen art.
- Niveau 2 (belæg med hjemmel): Argumentation på dette niveau indeholder en eller flere påstande med belæg og forskellige grader af hjemmel m.m. Men: ikke noget forsøg på gendrivelse.
- Niveau 3 (svag gendrivelse af hjemmel): Argumentation på dette niveau indeholder niveau 2-elementer, men tillige svage tilløb til gendrivelse.
- Niveau 4 (tydelig gendrivelse af hjemler): Argumentation på dette niveau svarer til niveau 3, blot er der nu tale om en klart identificerbar gendrivelse. Argumentationen kan godt indeholde flere påstande og belæg m.m., men det er ikke afgørende.
- Niveau 5 (flere gendrivelser af hjemler): Her er tale om en udvidet argumentation med flere påstande, belæg m.m. og dertilhørende flere gendrivelser.



Fra forskellig side (fx Christenson, 2015) er det blevet påpeget at elevernes argumentation godt kan være strukturelt i orden uden at den fagligt i øvrigt er gyldig. Med henvisning til en række andre studier fastslår Christenson således (p. 27): “Det er vigtigt med analytiske rammesætninger som gør det muligt på samme tid at sikre at eleverne laver reelle argumenter (strukturaspektet), og at disse giver mening fagligt (indholdsaspektet)” (vores oversættelse). Da vi deler dette synspunkt, har vi forsøgt på to forskellige måder at tilføje analysen efter Osborne-taksonomien et indholdsaspekt:

Inden for rammerne af Toulmin-analysen har vi helt enkelt tilføjet et par indholdsrettede krav så det ikke længere er nok at eleverne anfører et belæg eller en gendrivelse. Disse skal være både fagligt relevante, der skal være fagligt acceptabel hjemmel, og den faglige viden skal bruges på gyldig vis for at de tæller med. På denne måde sikrer vi at eleverne faktisk bruger faglig viden i deres argumentation, at denne faglighed er relevant, og at den bruges meningsfuldt. Med denne væsentlige indholdssikring finder vi i øvrigt Osborne et al.s taksonomi anvendelig omend den måske lægger for stor vægt på gendrivelse og differentierer for dårligt i det interval hvor flest elever præsterer. Her har vi fundet det relevant at underinddele alt efter om der gives flere belæg, evt. er tale om flere påstande med relevante og gyldige belæg. En argumentation hvor der faktisk er ét belæg med gyldig hjemmel, placeres således på niveau 2.1, mens ét argument med flere gyldige belæg berettiger til en indplacering på niveau 2.2.

Et eksempel på elevrespons og tilhørende Toulmin-analyse er vist nedenfor:

<p><b>P8: Anbefaling til veninden:</b> Hun skal tage mammografiundersøgelsen.</p>	<p><b>P8: Argumentation:</b> Skaden er langt større hvis hun har kræft og ikke får taget undersøgelsen, i forhold til hvis man tog prøven, og man ikke havde kræft. Man kan også med mammografi se om man har kræft, uden at det er muligt at man kan mærke det med de blotte fingre. Der er også den chance at der er radon i dit hus, og det kan betyde at der er en langt større chance for du har kræft, end hvis du ikke har været i kontakt med en for koncentreret mængde radongas, og eftersom man ikke kan smage/lugte det, er det for stor en chance for radon at det vil være for dumt ikke at få taget undersøgelsen. Det er også vigtigt at du kommer afsted og får taget den prøve så hurtigt som muligt, for jo længere man venter, jo sværere bliver det at kurere.</p>
---------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Reformuleret analysestruktur som gør det muligt at overskue argumentationen:

Påstand: Du er tjent med at få en mammografi!

Belæg: fordi ...

1. Konsekvenserne af en udiagnosticeret kræft er store  
Og de bliver kun større jo længere du venter – sværere at kurere.
2. Sundhedsrisikoen ved mammografi er i hvert fald mindre.

3. Det er en fremmelig diagnostisk metode (kan se kræft bedre end at mærke efter).
4. Du (grundet geografisk område) har en forhøjet risiko for kræft pga. mulig radon i undergrunden til dit hus.

Her er altså tale om en argumentation med hele fire belæg som alle er fagligt relevante og gyldige. Her godskriver vi eleverne for deres inddragelse af radon selvom en evt. radonforekomst snarere vil påvirke lungevævet end være en kilde til brystkræft.

Selvom argumentationen virker solid, så er der ikke tilløb til gendrivelse, så responsen holder sig på niveau 2. Alt i alt analyseres argumentationen til at være på niveau 2.4.

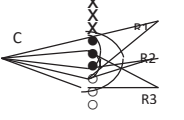
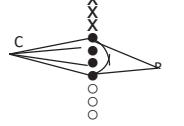
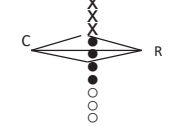
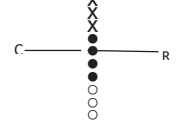
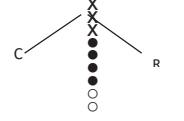
**SOLO-analysen af den indholdsmæssige kompleksitet i elevernes begrundede anbefalinger:** Vi har også analyseret elevernes svar ud fra den oprindelige SOLO-taksonomi (Biggs, 1982). I denne afgøres et svars kvalitet af hvor mange relevante videnselementer der indgår ("kapacitet"), om disse bringes relationelt i spil ift. hinanden ("Relationer"), og om dette sker konsistent og evt. overskridende. Nedenstående figur 3 illustrerer de forskellige SOLO-niveauer og den underliggende tænkning om kompleksitet i elevernes besvarelser.

Pointen er, at en SOLO-analyse i oprindelig forstand både er håndterbar og kan laves pålideligt hvis man forlods opregner de videnskomponenter som eleverne i sammenhængen (her altså 9. kl) burde kunne bringe i spil. I den aktuelle case fremgår disse "forventede vidensbrikker" i højrekolonnen af figur 3. Undervejs føjede vi ét nyt element til listen da et elevsvar fik os til at indse at det manglede. Afhængigt af om ét eller flere af disse dukker op i elevsvaret på rimelig vis, så er svaret på uni- hhv. multistrukturelt niveau. Hvis flere komponenter dukker op, tjekker man hvorvidt nogen af disse forbindes og/eller diskuteres op mod hinanden. I bekræftende fald er svaret relationelt. Hvis endelig der dukker relevante og udvidende svar op, så berettiger det til at svaret karakteriseres som udvidet & abstrakt ("extended abstract").

En SOLO-analyse af det tidligere anvendte eksempel på et elevsvar ("P8") identificerer dele af følgende vidensbrikker:

- 2: "i forhold til hvis man tog prøven og man ikke havde kræft."
- 4: Henvisning til radon som noget man bør medtænke/måle/ ...
- 7: "skaden er langt større ..."

Samtidig kobler den første svarsætning brikkerne 2 og 7, så svaret er ikke blot multistrukturelt, men faktisk *relationelt*.

<b>SOLO-NIVEAU</b> ● Relevant og forventet vidensbrik ○ Relevant vidensbrik som udvider, perspektiverer, generaliserer X Irrelevant vidensbrik	<b>SVARSTRUKTUR</b> Visualisering af spørgsmål (C=cue) og svar (R=response)	<b>RELEVANTE OG FORVENTEDE VIDENSBRICKER I DEN KONKRETE SSI-SAMMENHÆNG</b>
<b>ABSTRAKT</b> Overskridende det forventede, perspektivering og generalisation		1. Inddragelse af, at skadevirkningen af forskellige strålekilder kan estimeres vha. den ækvivalente strålingsdosis 2. relevant henvisning til typisk strålingsdosis og -type og den tilhørende risiko for strålingskader (herunder andre kræftformer) ifm. mammografi
<b>RELATIONELT</b> Et antal relevante elementer bringes i diskussion ift. hinanden og integreres evt.		3. relevant henvisning til strålingsdosis for Radon (enten typisk vha. radonkort eller aktuelt i form af undersøgelse) 4. at Radon-aktivitet med fordel kan måles & evt. reduceres 5. at Radon-eksponering først og fremmest udgør en risiko for lungerne
<b>MULTISTRUKTURELT</b> Flere relevante videnselementer, men ikke bragt i relation til hinanden.		6. at mammografi er en nyttig men ingenlunde sikker diagnostisk metode inddrages, dvs. viden om falsk positive og falsk negative undersøgelsesvar
<b>UNISTRUKTURELT</b> En enkelt relevant vidensbrik bringes på gyldig vis ind i svaret.		7. medtænkning af risikoen og de typiske konsekvenser af at gå rundt med en udiagnosticeret brystkræft, fx. risiko for spredning 8. Relevant inddragelse af vurdering af symptomer på en evt. spirende brystkræft.
<b>PRÆSTRUKTURELT</b> Kun irrelevante vidensbrikker bringes i spil		9. Relevant inddragelse af viden om kræftirisikoen for særlige brugergrupper (fx kvinder i den relevante aldersgruppe, rygere, arveligt belastede....) 10. overvejelser over sundhedsrisici ved at få foretaget en mammografi (uden henvisning til strålingsdosis)

Figur 3. SOLO-taksonomien efter Biggs (1982) og de vidensbrikker som har været udgangspunktet for artiklens SOLO-analyse.

### C. Analytisk kvalitetssikring

Alle analyser af elevresponsene er gennemført af artiklens forfattere uafhængigt af hinanden. Interrater-reliabiliteten af de to analyser af elevernes begrundede anbefalinger blev således i *udgangspunktet* estimeret via Cohens kappa til hhv. 0.26 (Toulmin-analysen) hhv. 0.56 (SOLO-analysen). Mens værdien for SOLO-analysen er akkurat acceptabel efter gængse standarder (Cohen, 1960), så var pålideligheden af Toulmin-analysen klart under det acceptable. I forlængelse heraf blev begge sæt af scoringer diskuteret igennem indtil der var opnået enighed. Dette er dog en indikation af at Toulmin-analyse ikke er ukompliceret og kræver både eksplicitering af procedurer, træning og synkroniserende diskussion. Omvendt vil fx interesserede lærere rimelig hurtigt kunne lære sig at vurdere elevernes niveau på en SOLO-taksonomisk skala.

Hermed har vi ekspliciteret teoretiske forankringspunkter, designprincipper, et SSI-case-eksempel, en tænkning om "hård" stilladsering af elevernes arbejde samt forskellige analytiske tilgange til at kvalitetsbedømme elevsvar fra vores alternative

SSI-format. Tilsammen udgør det et foreløbigt svar på artiklens forskningsspørgsmål 1 om udformningen af et fællesfagligt SSI-prøveformat. I den afsluttende diskussion vil vi tilføje yderligere metodiske refleksioner.

## Resultater (RQ2)

Hvilket billede tegner der sig af elevernes evne til situeret og selvstændigt at belyse en fællesfaglig problemstilling?

I det følgende vil vi lave tre nedslag i analysen af elevernes formåen i forskellige trin af deres arbejde med SSI-problemstillingen: deres evne til at formulere arbejdsspørgsmål, deres evne til at finde relevant og troværdig information samt mest udtømmende deres evne til at syntetisere argumenterede forslag til handling med afsæt i SSI-casen.

### A Elevernes arbejdsspørgsmål

Analysen af elevernes arbejdsspørgsmål viser at eleverne kun undtagelsesvist får stillet spørgsmål som vi (med begrænsninger som biologi- hhv. fysik/kemi-undervisere) kan knytte til geografifaget. Til trods for at casen åbner for at spørge til både jordbundsforhold og demografi. Om det primært skyldes casen, eleverne og/eller den foregående undervisning har vi svært ved at afgøre. Evt. er geografispørgsmålene "gemt" i den nye "fællesfaglige" kategori, som vi måtte indføre for at opfange spørgsmål som kunne relateres til *flere* fag.

Eleverne stiller en overvægt af viden- og dataspørgsmål som kan besvares med ja, nej, et tal eller reference til leksikal viden. Selv når man betænker at en stor del af eleverne ikke forudgående var blevet undervist i det mest relevante indhold, så er det overraskende at det taksonomiske niveau fremstår så lavt. I følgeforskningen om den fælles prøve vurderes det at "det kognitivt-taksonomiske niveau i problemstillinger og arbejdsspørgsmål (...) er uangribeligt, tenderende til højt." (Rambøll et al., 2018, p. 33). Det modstridende indtryk her indikerer enten at eleverne lærer at formulere bedre spørgsmål i løbet af 9. klasse, eller at arbejdsspørgsmålene til prøven er kraftigt influeret af lærerne.

Vi har forsøgt at afdække om der er en sammenhæng mellem det taksonomiske niveau af elevernes arbejdsspørgsmål (viden og data: 1, forklaring og forståelse: 2 og vurdering: 3) og kvaliteten af de begrundede anbefalinger som eleverne når frem til (se nedenfor). Her finder vi en signifikant korrelation på 0.27. Så der er *en vis* sammenhæng, men kausaliteten er uklar: Handler det om at de gode arbejdsspørgsmål fokuserer arbejdet og dermed leder til mere argumenterede anbefalinger? Eller er det blot sådan at gode elever formulerer gode arbejdsspørgsmål og uafhængigt heraf argumenterer godt?

Spørgsmålstype	Antal spørgsmål i kategorien (total N=158)	Fag	Eksempel på arbejdsspørgsmål
Viden og data	24	Biologi	Har du børn i alderen under 20?
	50	Fysik/kemi	Hvor meget stråling udleder mammografimaskinen?
	9	Geografi	Er der meget radon i undergrunden?
	18	Fællesfaglig	Hvor kommer radon fra?
Forklaring og forståelse	18	Biologi	Hvad er fordele og ulemper ved mammografi?
	5	Fysik/kemi	Kan man fjerne radon i hjemmet eller på en måde formindske strålingen fra det?
	0	Geografi	
	24	Fællesfaglig	Hvad er farligt ved radon?
Vurdering og holdning	1	Biologi	Hvor stor chance er der for at man dør af brystkræft og lungekræft?
	1	Fysik/kemi	hvor farligt er radon?
	0	Geografi	
	8	Fællesfaglig	Hvad er det dårlige og gode ved mammografi?

**Tabel 3.** Kategoriseringen af elevernes arbejdsspørgsmål.

## B Elevernes evne til at finde relevant og troværdig information

Selvom flertallet af eleverne ikke forlods har haft detailkendskab til det underliggende stofområde, så har en stor del af dem alligevel fundet relevant information som de med vekslende kvalitet tilegner sig og bruger i deres anbefaling og udfoldede begrundelser. Kilderne til denne viden er naturligt nok diverse lærebøger og internettet. Her er det interessant at alle elever anførte netsider som kilder, mens under halvdelen anførte lærebogskilder. Evnen til at lave en dækkende og kritisk netsøgning er således ultimativt vigtig for kvaliteten af elevernes stillingtagen og anbefaling. Vi har bedt eleverne redegøre for deres søgning (fx søgeord/streng). Af svarene fremgår det at de fleste anvender meget simple søgestrategier. Typisk søges direkte på enkelttermer,

fx “brystkræft”, “mammografi” eller “radon”. Alternativt søges overdrevent restriktivt på lange søgestrengene a la “chancen for at få kræft af stråling fra mammografi?”. Den første strategi lægger op til at man finder løsrevne vidensbrikker, den anden er i princippet relationel, men med stor risiko for at søgningen ikke giver nogen hits. På trods af disse noget ubehjælpssomme tilgange til søgning, så anfører praktisk taget alle en/flere internetkilder af både relevans og stor lødighed, fx cancer.dk, videnskab.dk, sundhedsstyrelsen osv. Når man læser deres uddybende svar på: *”Hvorfor er disse internetmaterialer gode? Hvad kigger I efter for at afgøre det?”*, så indser man at det langt hen ad vejen må være Googles fortjeneste idet ca. to tredjedele af eleverne blot henviser til at der står det de skal bruge. Et typisk elevsvar er således: *”Ja, det er de fordi de fortæller om hvad man har søgt på, og viser ikke bare en hel side reklamer”*. Et mindre antal elever betegner deres kilder som “ troværdige”, nogle tilmed med henvisning til relevante kriterier (fx at de har lavet mange undersøgelser), mens andre blot henviser til institutionen som (indlysende?) garant for den downloadede viden. Med fare for at generalisere ud over hvad et pilotstudie kan bære, så ligner det en tanke at elevernes fællesfaglige arbejde vil kunne styrkes hvis de lærer systematiske strategier til informationssøgning og kildekritik. For indeværende viser analysen af elevernes vidensgrundlag at den typiske elev i situationen og med Googles hjælp har formået at finde relevante vidensbrikker af rimelig kvalitet, men sikkert løsrevne fra hinanden og uden vished for at de dækker hele casens problemfelt.

### Elevernes SSI-anbefalinger og tilhørende argumentation

**Resultaterne fra Toulmin-analysen af elevernes begrundede anbefalinger** går som tidligere omtalt både på at der strukturelt er tale om et argument (med belæg, hjemmel, rygdækning, gendrivelse), og på at indholdet fagligt er i orden (belæg, hjemmel m.m. er fagligt relevante og gyldige).

De fleste danske naturfagsundervisere vil nok være betydeligt mere opmærksomme på det indholdsmæssige end på det strukturelle i elevernes argumentation. Vanligvis vil de forvente at eleverne er i stand til fagligt at begrunde at noget sker (“kausal forklaring”). Argumentation er ikke det samme som forklaring, men på beslægtet vis gives også her en begrundelse i form af belæg, hjemmel m.m. Overføres forventningsniveauet for elevforklaringer til elevs argumentation, så vil de fleste lærere nok stille sig tilfredse med at elever kan håndtere argumentation på Toulmin-niveau 2-3, dvs. med visse faglige belæg, gyldig hjemmel og spæde tilløb til gendrivelse (gendrivelse fylder normalt ikke meget i den traditionelle forklaringsorienterede naturfagsundervisning).

Tabel 4 nedenfor viser resultaterne af vores Toulmin-analyse.

Toulmin-argumentniveau	Hyppighed
Niveau 1 (uden belæg)	8/34
Niveau 2 (belæg med hjemmel)	16/34
Niveau 3 (svag gendrivelse af hjemmel)	5/34
Niveau 4 (tydelig gendrivelse af hjemmel)	4/34
Niveau 5 (flere gendrivelser af hjemler)	1/34

**Tabel 4.** Fordeling af elevresponser på forskellige Toulmin-argumentniveauer.

Det fremgår at ca. halvdelen af eleverne præsterer argumentation på niveau 2.

Vores tal er ikke umiddelbart sammenlignelige med internationale undersøgelser af elevs SSI-argumentation da disse normalt stiller sig tilfredse med *at* der argumenteres med belæg osv., uden at forholde sig til den faglige kvalitet af belæg m.m. Med vores krav om *fagligt holdbar* argumentation er det selvfølgelig sværere at nå de højeste niveauer. Under en femtedel af eleverne får argumenteret på måder som inkorporerer en eller flere tydelige gendrivelser. Vanskeligheden ved denne (strukturelle) del af argumentation går igen i internationale undersøgelser både med yngre og alderssvarende samples (se fx Erduran et al., 2004). Mere opbyggeligt viser samme forskning også at dette aspekt kan trænes og læres.

## Resultater fra SOLO-analyse af elevernes SSI-anbefalinger

Af tabel 5 nedenfor fremgår det at ca. to tredjedele af elevsvarene er klassificeret som værende uni- hhv. multistrukturelle. Dvs. at disse elever har fundet en eller flere relevante vidensbrikker som de uafhængigt af hinanden bringer i spil i deres udfoldede anbefalinger.

SOLO-taksonomisk niveau	Hyppighed
Niveau 1: Præstrukturelt	1/34
Niveau 2: Unistrukturelt	10/34
Niveau 3: Multistrukturelt	11/34
Niveau 4: Relationelt	10/34
Niveau 5: Udvidet abstrakt	2/34

**Tabel 5.** Fordeling af elevsvar på SOLO-taksonomiske niveauer.



De hyppigste vidensbrikker som indgår i elevernes svar, er 2, 7 og 9, dvs. inddragelse af risikoen for strålingskader ved mammografi, medtænkning af risikoen og de typiske konsekvenser af at gå rundt med en udiagnosticeret brystkræft, inddragelse af symptomer på brystkræft samt inddragelse af viden om kræfttrisikoen for særlige brugergrupper. Det er tydeligt at mange elever har taget grundlæggende pointer til sig, men det tilhørende fagligt-tekniske sprog (aktivitet, strålingsdosis, intensitet) bruges reelt ikke i elevernes svar. Således har eleverne fået godskrevet brik 2 når de har omtalt "risikoen", uden at de i øvrigt har nævnt noget om "strålingstype" og/eller "strålingsdosis". Fraværet af egentligt fagsprog afspejler rimeligvis at eleverne ikke forudgående var undervist i dette faglige indhold. Den ekstra pointe her må være at elevens kortvarige omgang med internetsider ikke er tilstrækkeligt til at etablere et aktivt fagsprog i relation til et fænomen.

Et par eksempler som illustrerer en vis spændvidde, både hvad angår den konkrete SSI-anbefaling og det SOLO-taksonomiske svarniveau:

*Eksempel: unistrukturelt elevsvar:*

*Anbefaling:* "Vi anbefaler at veninden skal til mammografiscreening."

*Begrundelse:* "Fordi jo tidligere man opdager sygdommen, jo tidligere overlever man den."

Her genfinder vi i en lidt sprogligt uskarp formulering vidensbrik 7 i SOLO-analysegrundlaget, som handler om "medtænkning af risikoen og de typiske konsekvenser af at gå rundt med en udiagnosticeret brystkræft".

*Eksempel: relationelt svar*

*Anbefaling:* "Vi syntes ikke du skal tage en mammografiundersøgelse da radon kun øger risiko for lungekræft og ikke brystkræft. Du er også kun 40 år gammel, så sandsynligheden for du har brystkræft, er lille."

*Begrundelse:* "Det er ikke godt for din krop at få stråling, så når sandsynligheden for du har brystkræft, er så lille, så synes vi ikke du skal tage den."

I dette tilfælde genfinder vi vidensbrikkerne 2 (risiko v. mammografi – implicit gemt i "det er ikke godt for din krop at få stråling"), 7 samt 9 (risiko for særlige grupper (50+). I begrundelsen kobles brikkerne 2 og 9 relationelt til hinanden.

Interessant, men ikke egentlig overraskende er der stor overensstemmelse mellem kvalitetsbedømmelsen af et elevsvar efter SOLO-taksonomien hhv. Toulmin-analysens niveauer. Korrelationen mellem scoringen på hver af de to skalaer viser sig at være 0.78. SOLO-analysen giver således et nogenlunde sammenfaldende indtryk af kvalitet i elevernes respons som den noget sværere Toulmin-analyse.

Alt i alt tegner der sig et billede af at mange elever faktisk formår at ekstrahere relevante vidensbrikker fra deres kildetekster og drage dem ind i deres argumentation på relevant vis. Det hører dog med til billedet at kun ca. 30 % af eleverne i vores sample formår at skabe sammenhængende relationel viden, og kun ca. 15 % udfolder en argumentation med antydning af for og imod.

## Diskussion

Vores pilotstudie godtgør at det er muligt at udprøve elevernes evne til fællesfaglig problemløsning vha. en fællesfaglig SSI-case og en netbaseret indsamling af elevartefakter fra deres problemløsende arbejdsproces. Grundlæggende har både den elaborerede SSI-case, den procesnære artefaktindsamling og de anvendte analytiske tilgange vist sig at fungere. Hvad angår SSI-casen har studiet dog antydning at detaljer i formuleringen kan være afgørende for om alle tre udskolingsfag trækkes ind i opgaveløsningen på dækkende vis. Derfor vil det være relevant at udarbejde yderligere SSI-cases for på denne måde at afklare om et andet afsæt vil få geografi til at fremstå stærkere i elevernes besvarelser. Mere grundlæggende vil man også kunne evaluere samme elever på flere *forskellige* cases og derved afklare hvor robuste resultaterne er over for den anvendte case. Dette ville være et vigtigt bidrag til at belyse reliabiliteten af SSI-evalueringsformatet.

Det har aldrig været vores ambition at SSI-prøveformatet skulle erstatte den nuværende fællesfaglige prøve. Først og fremmest fordi vi anser denne for at matche de centrale mål for naturfagsundervisningen, men også fordi lokalt forankrede og dialogiske prøveformer generelt har høj validitet. Ved at spørge ind vil en lærer fx kunne afklare om et mangelfuldt elevsvar er udtryk for sproglige og/eller forståelsesmæssige problemer, på en måde, som vores ikkedialogiske SSI-evaluering ikke giver mulighed for. Alligevel har vi bestræbt os på at processen i vores SSI-format indeholder samme komponenter som processen hen mod den fælles afsluttende prøve – blot tidligt kondenseret. Dette åbner for at evalueringsformatet meningsfuldt vil kunne indgå i praksis, samtidig med at det har skaffet indsigter som fremadrettet kan kvalificere den prøverettede undervisningspraksis.

Hvad angår det første har underviserne i vores pilotafprøvningsklasser efterspurgt at få SSI-prøven til rådighed for at træne senere klasser mhp. den fælles mundtlige prøve fordi den nemt kan passes ind i undervisningen, træner relevante kompetencer og kan bruges formativt. Begrænsningerne er her dog at SSI-evalueringen kun indfanger en mindre del af undersøgelseskompetencen og i usikkert omfang involverer modellerings-kompetence. Til gengæld vil evalueringsformatet automatisk betone literatursøgning og kildekritik på nye måder, naturfaglige dannelsesaspekter, som efter alt at dømmes forsømmes i gængs naturfagsundervisning. Derudover kræver formativ

anvendelse af prøven at underviseren trænes i at bedømme typiske elevsvar på åbne SSI-spørgsmål og give eleverne forbedrende feedback. Her indikerer vores pilotstudie at en evt. efteruddannelsesindsats nok bør fokusere på at naturfagsunderviserne lærer at bruge SOLO-taksonomien som redskab for bedømmelsen fordi den efter alt at dømme er nemmere at håndtere end bedømmelse efter Toulmin-kriterier, og fordi de to typer af bedømmelser alligevel korrelerer i betragtelig grad.

En mere ekstrem praksisanvendelse af SSI-formatet ville være at lade det erstatte den nuværende *skriftlige udtræksprøve*, som kritiseres massivt af lærerne for at være for forskellig fra den fælles mundtlige prøve. Med SSI-formatet ville man få et sammenhængende og fint samstemt evalueringssystem hvor de to delprøver vil kunne kompensere hinandens relative styrker og svagheder, fx mht. undersøgelseskompetence og forberedelses-impact. Omkostningen (set fra officielt hold) vil være at man dropper forestillingen om at meningsfulde læringsudbytter kan udmåles med selvrettende prøver.

Vores pilotstudie med SSI-evalueringssystemet har givet nye indsigter i hvorledes eleverne håndterer arbejdet med fællesfaglig problemløsning. Ved at indsamle og analysere empiri fra de forskellige trin i elevernes arbejdsproces giver evalueringssystemet indblik i hvad eleverne kan og gør på hvert enkelt trin, men det giver samtidig også mulighed for at kvalificere arbejdet med dette aspekt.

I det omfang man kan generalisere ud fra vores afgrænsede pilotstudie, så ligger der en vigtig praksisrettet indsats i at lære eleverne at søge og validere information på nettet. Meget tyder også på at man med fordel kan lære elever at skrive taksonomisk bevidste arbejdsspørgsmål *på et tidlige* tidspunkt. Det er påfaldende så få vurderingsspørgsmål vi ser på testtidspunktet i 9. klasse. Det er også påfaldende så svært det øjensynligt har været for elever at identificere og beskrive undersøgelser som kan være med til at belyse SSI-casen. Denne slags overvejelser vil kunne trænes som et relevant bidrag til undersøgelseskompetence, endda uden at eleverne rent faktisk gennemfører de konkrete undersøgelser. Endelig indikerer studiet at eleverne med fordel trænes i at diskutere for og imod for på denne måde at opbygge relationel viden og lægge grunden for argumenteret stillingtagen med øje for gendrivelse.

Samtidig med at vores studie således udpeger bud på forbedring af praksis, så er den opløftende erkendelse i nærværende pilotstudie at mange elever – med Googles hjælp – faktisk formår at finde og ekstrahere relevante vidensbrikker og drage dem ind ifm. en argumenteret stillingtagen. Tilmed uden nævneværdig fagfaglig ballast fra forudgående undervisning. Dette er i overensstemmelse med Lewis & Leach (2006), som fandt at elever kunne indgå i “reasoned discussion” ifm. en SSI-case på et moderat grundlag af viden. Målforestillingen om at elever kan opnå naturfaglig dannelse nok til at tage informeret stilling i SSI-tilfælde, forekommer pludselig ikke så urealistisk endda.

Konklusionen er herefter at vi har vist at man kan designe en meningsfuld og vel-fungerende SSI-baseret evaluering som er rimelig pålidelig og undervisningsmæssigt valid. Som noget hidtil uset har vi anvendt og vurderet flere analysemetoder – en metodetriangulering som øger pålideligheden af vores resultater. Med evaluerings-konceptets procesnære empiriindsamling har vi tillige identificeret dele af elevernes arbejdsproces som med fordel kan forbedres. Samtidig må vi erkende at vores afgræn-sede pilotstudie ikke formår at besvare samtlige spørgsmål som vi stillede i optakten. Fx har vi for få data til at lave en signifikant kønsopdelt analyse som afdækker om SSI-formatet – uden forberedelsesperiode – faktisk fjerner prøvebias til fordel for pigerne. Et næste forskningsskridt kunne således være at udvide samplet og forfølge dette spørgsmål. Det ville også være interessant at forfølge spørgsmålene: Hvor langt giver SSI-evalueringsformatet resultater som matcher elevernes afsluttende prøvere-sultater? Hvor robuste er “resultaterne” over for de SSI-cases som anvendes?

SSI-evalueringsformatet åbner således for både yderligere forskning og praksisper-spektiver.

## Referencer

- Allchin, D. (2011). Evaluating knowledge of the nature of (whole) science. *Science Education for Everyday Life – Evidence-Based Practice*, 95(3), 518-542.
- Bekendtgørelse om folkeskolens prøver (2019). BEK nr 1128 af 14/11/2019. <https://www.retsinfo-formation.dk/eli/lta/2019/1128>.
- Biggs, J. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy*. Academic Press.
- Chin, C. & Osborne, J. (2008). Students' questions: A potential resource for teaching and learning science. *Studies in Science Education*, 44(1), 1-39. <https://doi.org/10.1080/03057260701828101>.
- Christenson, N. (2015). *Socio-scientific argumentation. Aspects of content and structure*. Dis-ertation. Faculty of Health, Science and Technology. Karlstad University Studies. 2015:26.
- Cohen J. (1960). A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 20:37-46.
- Erduran, S., Simon, S. & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Educa-tion*, 88(6), 915-933.
- Garrecht, C.; Eckhardt, M.; Hoffler, T.N.; Harms, U. (2020). Fostering students' socioscientific decision-making: exploring the effectiveness of an environmental science competition. *Disciplinary and Interdisciplinary Science Education Research*, 2, article 5.
- Gipps, C. (1999): *Review of Research in Education*, 24, p. 355-392.
- Karisan, D.; Zeidler, D.L. (2017). Contextualization of Nature of Science within the Socioscientific Issues Framework: A Review of Research. *International Journal of Education in Mathematics, Science and Technology*, 2, 139-152.

- Kolsto, S.D. (2006). Patterns in Students' Argumentation Confronted with a Risk-Focused Socio-Scientific Issue. *International Journal of Science Education*, 28(14), 1689-1716.
- Krogh, L.B.; Elgaard, J.F.; Secher, A.; Daugbjerg, P. (2021). Pigerne stikker af fra drengene i karakterer til den fællesfaglige prøve. *MONA (Matematik Og Naturfagsdidaktik)*, 1, 6-26.
- Lewis, J. & Leach, J. (2006). Discussion of socio-scientific issues: The role of science knowledge 1. *International Journal of Science Education*, 28(11), 1267-1287.
- Osborne, J., Erduran, S. & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching*, 41(10), 994-1020.
- Rambøll; VIA; KP. (2020). *Indførelse af den fælles prøve i fysik/kemi, biologi og geografi – prøvens betydning for elevernes læring og motivation*. <https://www.uvm.dk/-/media/filer/uvm/publikationer/2020/juni/200616-statusrapport-endelig-juni2020.pdf>.
- Rambøll, V. U. og U. (2018). *Statusnotat – evaluering og følgeforskning: Indførelsen af den ny fælles prøve i fysik/kemi, biologi og geografi – prøvens betydning form og indhold undervisningens form*. <https://uvm.dk/-/media/filer/uvm/aktuelt/pdf18/180319-statusrapport-faelles-naturfagsproeve.pdf>.
- Sadler, T.D.; Barab, S.A. & Scott, B. (2007). What do students gain by engaging in socio-scientific inquiry. *Research in Science Education*, 37, 371-391.
- STUK (2019). *Fælles prøve i fysik-kemi, biologi og geografi – prøvevejledning*. <https://www.uvm.dk/-/media/filer/uvm/udd/folke/pdf19/nov/191121-faelles-proeve-i-fysik-kemi-biologi-og-geografi-proevevejledning-ua.pdf>.
- Toulmin, S. (1958). *The Uses of Argument*. Cambridge University Press.
- Zeidler, D.L.; Applebaum, S.M.; Sadler, T.D. (2011). Enacting a socio-scientific issues classroom – transformative transformations. In T.D. Sadler (Ed.), *Socio-scientific Issues in the Classroom*. Springer.

## English abstract

*We have designed and pilot-tested an internet-based assessment of students' competence for making informed decisions regarding socio-scientific issues. Students are guided through steps resembling, but condensing, an interdisciplinary oral science-exam in Denmark. Student artefacts are collected from each step to get insight into how they manage to search and validate information, consider inquiries, and in the end are able to argue for their decisions. We present the assessment format along with our analytical approaches (SOLO and Toulmin). We also present results from a pilot-study with students from four schools, demonstrating comforting as well as problematic sides of students' performance.*