

Der blæser nye vinde i testverdenen, hvor der løbende kommer nye typer af test til. Vi tester fx elevers demokratiske dannelse, deres samarbejdsevner, deres empati og deres evner til at få idéer og løse problemer. Det er positivt, at vi tester andet end de traditionelle fag, men det kalder på nye og skærpede krav til kvaliteten af testene, mener to DPU-forskere.

TEST PÅ TVÆRS



FOTO: SCANPIX

* Flere af de nye test, bl.a. *PISA Collaborative Problemsolving* og *Game Based Learning in the 21st Century*, tester eleverne i samarbejde og problemløsning – herunder deres empati og evne til at sætte sig i andres sted.

”Du er direktør i en ny forlystelsespark, som vil fejre åbningen med en indvielsesfest. Der er inviteret 2000 gæster, men nu viser det sig, at der kun er plads til 1000. Der bliver indkaldt til krisemøde i ledelsen, og du skal som leder argumentere for to forskellige løsninger.”

Sådan en opgave stiller testen *Game Based Learning in the 21st Century* (GBL21) elever i 5. og 7. klasse. Men hvad er det, man tester? Det handler blandt andet om elevernes empati, forklarer professor på DPU, Aarhus Universitet, Jeppe Bundsgaard, der har været med til at formulere spørgsmålene i testen.

”Empati er en forudsætning for at kunne forstå og argumentere for begge løsninger. Hvordan kan eleven sætte sin indlevelse i spil i konflikt- og problemløsning? Eleverne kan lide disse test. De får udfordringer, der giver mening, og de har ikke oplevelsen af, at de fejler. De oplever, at de har et godt argument, uanset hvad de kommer frem til,” siger han og tilføjer, at forskerne vurderer elevernes svar ud fra en kodevejledning, der i detaljer beskriver, hvad der skal til for at få point i de enkelte delaspekter af opgaverne.

Formålet med testen er at teste elevernes designkompetencer med henblik på at blive klogere på, hvordan man kan bruge spilbaseret læring i undervisningen. Designkompetencer skal forstås bredt, nemlig som evnen til at sætte sig ind i problemer, udtænke løsningsstrategier og skabe løsninger. Hermed tester man nogle af de kompetencer, der af OECD betegnes som ’21st Century Skills’. Det drejer sig blandt andet om kritisk tænkning og problemløsning, kreativitet og innovation samt kommunikation og samarbejde. GBL21-projektet relaterer sig til forskellige fag; i matematik skal eleverne fx udvikle spil, og i dansk skal de være journalister på en avisredaktion.

Test som termometer

Den enkelte elev får aldrig at vide, hvordan han eller hun klarede sig i testen. Hverken eleven, læreren eller skolelederen får resultaterne: De bliver bearbejdet af forskere, blandt andre Jeppe Bundsgaard, med henblik på at sammenligne, hvordan hhv. elever, der har deltaget i de undervisningsinterventioner, forskerne har igangsat, og de elever, der ikke har, klarer sig i testen. Denne type test kaldes med et fagudtryk for ’Large Scale Assessments’.

De internationale test fra OECD og IEA er eksempler på internationale Large Scale Assessments. De skal dels tage temperaturen på de enkelte landes skolesystemer, og dels kan landene bruge resultaterne som afsæt for at udvikle deres skolepolitikker ved blandt andet at lade sig inspirere af de lande, der klarer sig godt i testen. Særligt OECD's PISA-test har fået tildelt en betydelig politisk funktion, idet den bliver brugt til at legitimere politiske budskaber i debatter om skolen.

Jepp Bundsgaard deltog i arbejdet med at rapportere fra den seneste PISA-test, og han er dansk forskningsleder for IEA-testen ICILS (*International Computer and Information Literacy Study*). Flere af de test, han arbejder med, har det til fælles, at de tester innovative områder, der ikke er direkte relateret til et specifikt fag. Og så alligevel, for ifølge Jepp Bundsgaard er der tale om kompetencer, der netop er relevante for de fleste fag i skolen. Han er derfor glad for, at OECD's PISA-konsortium har fået øjnene op for vigtigheden af disse tværgående kompetencer.

”Vi skal blive bedre til at teste andet end de traditionelle fagområder. OECD har på det seneste forsøgt at sige, ’skoler er meget mere end det’, og de har derfor udviklet test, der tester andre og mere innovative kompetencer end de direkte fagrelaterede,” siger han.

Kan man teste samarbejde?

PISA Global Competence, PISA Collaborative Problem Solving og PISA Digital Problem Solving er eksempler på den udvikling. Hvert år tester man på skift ét af hovedområderne (enten læsning, matematik eller naturfag) plus et af de nye innovative områder, der tester tværgående kompetencer.

PISA Collaborative Problem Solving (2015), som blev rapporteret i 2017, tester elevernes samarbejdsevner ved at undersøge deres digitale interaktioner med selve testen, blandt andet via en chat-funktion. Spørgsmålene lægger typisk op til, at eleverne skal gennemtænke nogle scenarier. De bliver bedt om at forestille sig, at de står i en specifik situation, hvor der opstår et problem, de så skal reflektere over og komme med løsningsforslag til. Men er der ikke noget principielt selvmodsigende i at teste samarbejde i en individuel test? Ifølge Jepp Bundsgaard er det ikke uladsigørligt, men han påpeger nogle udfordringer:

”Den slags test er lidt for meget hængt op på kognitiv forestillingsevne. Det er for let at regne ud, hvad det rigtige svar eller argument er. Og der er ingen garanti for, at eleven rent

»Vi skal blive bedre til at teste andet end de traditionelle fagområder. OECD har på det seneste forsøgt at sige, ’skoler er meget mere end det’, og de har derfor udviklet test, der tester andre og mere innovative kompetencer end de direkte fagrelaterede.«

Jepp Bundsgaard

INTERNATIONALE TEST

IEA (International Association for the Evaluation of Educational Achievement) er et vigtigt forum i den internationale uddannelsesverden. Ud over læsning (*Pirls*) samt matematik og naturfag (*Timss*) tester IEA elevernes politiske dannelse samt deres computer- og informationskompetencer.

ICCS (International Civic and Citizenship Education Study): Tester 8.-klasse-elevens viden om demokrati, samfund og politik med henblik på at undersøge deres politiske dannelse i det 21. århundrede i et internationalt komparativt perspektiv.

ICILS (International Computer and Information Literacy Study): Tester computer- og informationskompetencer og datalogisk tænkning. Gennemføres på 8. klassetrin.

faktisk ville handle sådan i virkeligheden. Typisk er det de elever, der generelt klarer sig godt i de traditionelle PISA-test, der scorer højt her, fordi de kan ’regne den ud,’ fortæller han.

Danske elever har growth mindset

Jepp Bundsgaard fremhæver også det til-lægsspørgsmål, den seneste PISA-test fra 2018 stillede eleverne: ’I hvilken grad føler du, at din intelligens er noget ved dig, som du ikke kan ændre meget ved?’ Bagtanken var at få et indblik i, hvorvidt eleverne ser

deres intelligens som noget fastlåst eller noget, de kan rykke ved. Det teoretiske afsæt er teorien om ’growth mindset’, der er udviklet af den amerikanske professor i psykologi Carol Dweck. Hvis man har et growth mindset, betragter man ikke sine fejl som en stopklods for sin udvikling. Tværtimod ser man, at fejl kan fundere ens videre udvikling. Man tror på, at man kan rykke sig og klør derfor på. Har man derimod et ’fixed mindset’, ser man sin intelligens som noget, man ikke kan ændre, og som derfor hindrer ens udvikling. Begår man fejl, ser man det som et nederlag. Derfor giver man også nemmere op over for udfordringer og svære opgaver. Mange elever vil selvfølgelig have lidt fra begge typer mindset, men som oftest vil de hælde mere i den ene end den anden retning.

I PISA-resultaterne tegner der sig et interessant og tydeligt mønster på tværs af landegrænser, pointerer Jepp Bundsgaard: De dygtigste elever svarer overvejende ja til, at de tror på, at de kan rykke ved deres intelligens.

”Kigger man på sammenhængen mellem det at have et growth mindset og elevernes læsescore, kan man se, at danske elever, der tilslutter sig et growth mindset, i gennemsnit opnår en læsescore, der er 54 point højere end elever, der har en mere fastlåst forståelse af intelligens. Sammenhængen falder til 46 point, når der justeres for elevernes socio-økonomiske baggrund, men der er stadig en stærk og signifikant sammenhæng,” siger Jepp Bundsgaard og peger på, at netop den variabels sammenhæng med PISA-resultatet er større end alle andre variabler – fx køn og kulturel baggrund.

”Det siger noget om vigtigheden af det, vi med et engelsk ord kalder ’self efficacy’ – altså troen på egne ressourcer og på, at man kan mestre noget.”

Jepp Bundsgaard ser således en sammenhæng mellem det growth mindset, som testes i PISA Læsning, og graden af self efficacy. Og her, pointerer han, er det bemærkelsesværdigt, at 75 procent af de danske elever har et udpræget growth mindset. På tværs af lande er gennemsnittet 63 procent. Også i forhold til de andre nordiske lande ligger danske elever højt.

”Det er selvfølgelig positivt, at så mange danske elever figurerer højt her, for troen på at man kan udvikle sig gennem uddannelse og læring, er en vigtig egenskab at besidde i det 21. århundrede, der blandt andet stiller krav om livslang læring. Der er de danske elever altså over en bred kam ret godt stillet,” siger han, men understreger samtidig, at vi ikke må miste blikket for de 25 procent, der

ikke har samme store tiltro til deres udviklingsmuligheder.

”Det fortæller os, at vi skal være opmærksomme på elever, der føler sig fastlåste og tror, de ikke kan blive bedre, for det spænder ben for deres læring. Det er vigtig information i forhold til, hvordan vi skal arbejde med elever som mennesker.”

Testens begrænsninger

En anden test, som rækker ud over det fagrelaterede, er IEA-testen *ICCS (International Civic and Citizenship Education Study)*, som tester elevers demokratiske dannelse. Jonas Lieberkind, der er lektor på DPU, Aarhus Universitet, er en af de to danske forskere, der står for at udarbejde den danske del af testen samt indsamle og behandle data. Ligesom Jeppe Bundsgaard mener han, at det kan være vigtigt og interessant at teste færdigheder og kundskaber, der rækker ud over – eller går på tværs af – fagene. Men man skal være opmærksom på testens begrænsninger, understreger han.

”Det særlige ved test er, at de er meget lette at omsætte politisk og ideologisk. De kræver ikke umiddelbart en kompliceret analytisk fortolkning, som kun forskeren og fagkyndige inden for et givent forskningsfelt har adgang til. Testens binære koder, om man klarer sig godt eller dårligt, lægger op til hurtige og gerne politiske tolkninger,” siger Jonas Lieberkind og peger på, at netop kombinationen af at give enkle svar på komplicerede spørgsmål og lave internationale sammenligninger gør testen populær. Men ofte også utilstrækkelig.

ICCS bliver gennemført i 8. klasse, som er det klassetrin, hvor danske elever får samfundsfag på skemaet. Det betyder imidlertid ikke, at den tester elevernes viden relateret til samfundsfag. Men, forklarer Jonas Lieberkind, i Danmark går testen ikke på selve indholdet i samfundsfag, men mere på, hvordan skolen forvalter folkeskolens formålsparagraf om at danne eleverne til demokratiske borgere. *ICCS* tester således hele skolens virke. Og her er der noget at være stolt af for os danskere – i hvert fald umiddelbart. For danske 8.-klasse-elever har begge de gange, testen er gennemført (2009 og 2016), placeret sig blandt de allerdygtigste. Men konklusionen kræver nuancering, understreger Jonas Lieberkind. Da testresultaterne blev kombineret med en spørgeskemaundersøgelse, fik forskerne øje på et interessant paradoks: Danske unge ved en masse om politik og demokrati, men de omsætter ikke den viden til handling.

»Det særlige ved test er, at de er meget lette at omsætte politisk og ideologisk.«

Jonas Lieberkind

FREMtidENS TEST

PISA 2021 fokuserer på matematik med en til-lægstest i kreativ tænkning.

PISA 2024 skal undersøge elevernes evner til at 'lære i en digital verden' ved at teste deres evner til at regulere deres egen læring ved brug af digitale værktøjer. De skal måle deres egen fremgang, bruge feedback og holde fokus, alt imens de bygger modeller og udvikler løsninger med de digitale værktøjer. Disse kompetencer er, ifølge OECD, væsentlige for elevernes selvstændige læring i en stigende kompleks og digital verden.

Kilde: oecd.org/p

Gamebased Learning in the 21st Century (GBL21): dansk undersøgelse, som er støttet af Innovationsfonden og ledet af Thorkild Hanghøj fra Aalborg Universitet. Undersøgelsen er i gang og projektet afsluttes i 2022.

Læs mere: gbl21.aau.dk

”Når de danske unge klarer sig godt i testen, vidner det selvfølgelig om et vist engagement i samfundet. Men de er ikke aktive og involverede i demokratiet i den forstand, at de aktivt går ud og demonstrerer for en sag. Men den viden kom ikke frem i testresultaterne, og det blev kun afdækket i spørgeskemaundersøgelsen – ikke forklaret,” siger han.

Det har han forsøgt at råde bod på ved selv at sætte en kvalitativ undersøgelse i værk, hvor han har interviewet repræsentanter for den generation, der blev testet med henblik på at forstå, hvorfor de unge er umiddelbart passive og forbeholdne over for at deltage i politiske aktiviteter, men samtidig politisk engagerede.

”De har en enorm viden om demokrati og samfundsforhold. Men de fravælger bevidst det politiske engagement med den begrun-

delse, at fordi de er unge, skal de fokusere på at leve deres ungdomsliv og på at få en uddannelse. De begrundet det også med, at de grundlæggende er glade for det danske samfund,” siger Jonas Lieberkind, der generelt taler for at kombinere test med andre typer af undersøgelser.

Intimiderende test fejler

Når Jonas Lieberkind formulerer spørgsmål til sine kvalitative såvel som kvantitative undersøgelser, anstrenger han sig for at sætte etikken højt.

”Folk bliver hurtigt intimiderede. Og bliver de det, så er der risiko for, at de ikke svarer oprigtigt på spørgsmålene. Og det gælder i særdeleshed også, når man formulerer spørgsmål til en test. For testen kan i selve sin udformning overskride nogle grænser hos dem, der testes,” forklarer han.

Netop det var en af årsagerne til, at Jonas Lieberkind sammen med sin DPU-kollega Jens Bruun og en række andre forskere rundt om i verden gav dumpekarakter til en ny PISA-undersøgelse, der skulle teste elevers globale kompetencer. *PISA Global Competence 2018* overskred ifølge ham en etisk grænse for, hvordan man kan teste. Det var særligt de spørgsmål, der gik på, hvordan man forholder sig til mennesker, som er udsatte eller tilhører en minoritetsgruppe, der var problematiske.

”Rundt omkring i klasserne vil der jo sidde elever fra sådanne minoritetsgrupper, som også skal svare på disse spørgsmål. Det vil tvinge majoritetskulturen til at forholde sig til minoritetsgrupperne, og det er intimiderende. Elever fra minoritetsgrupperne vil føle sig forkerte og intimiderede på en anden måde. Begge dele er problematiske – ikke mindst for de data, man indsamler. Bliver det for intimiderende, så svarer eleverne ikke det, de mener, men hvad de tror, de bør mene. Pointen er, at hvis man udfordrer dem for meget, risikerer man at få urigtige svar,” siger han.

Når man vurderer en tests kvalitet, handler det blandt andet om at tage stilling til, om den lever op til forskningsmæssige standarder og metoder, forklarer Jonas Lieberkind. I den vurdering indgår de etiske overvejelser, men man ser også på, om resultaterne fra en given test risikerer at forsimple emnet, om testemnet lader sig omsætte til enkle binære svar, om den snarere måler kulturelle og økonomiske forskelle end selve emnet, samt hvilke politiske konsekvenser, den viden, man genererer, eventuelt kan få.

”*PISA Globale Competence* faldt på samtlige niveauer. Selve konstruktionen af testen var

»I mit perspektiv går grænsen ikke nødvendigvis ved, hvorvidt emnerne der testes, er bløde eller abstrakte. Det handler mere om, hvorvidt det er forskningsmæssigt interessant at omsætte en given problematik i denne reducerede og enkle form, som testen lægger op til, men som også definerer dens begrænsninger.«

Jonas Lieberkind

problemet – metodisk såvel som etisk. Hertil kom, at den var ekstremt forudindtaget af vestlige forestillinger om demokrati og samfund. Det fortæller os, at forskningsetik, bias etc. bliver desto mere udfordrende og afgørende, når det handler om emner, der er normative og politiske,” siger Jonas Lieberkind.

En ommer

Betyder det, at *PISA Global Competence* overskred grænsen for, hvad man reelt kan teste? Var emnet for blødt, for bredt, for abstrakt eller uhåndgribeligt? Jonas Lieberkind understreger, at der selvfølgelig er grænser for, hvad man kan teste – og i særdeleshed også hvor mange test, man kan udsætte eleverne for. Men når det er slået fast, så er han mere optaget af, hvordan man tester eleverne i de internationale komparative test, og om testen overhovedet kan indfange det, man søger svar på.

”I mit perspektiv går grænsen ikke nødvendigvis ved, hvorvidt emnerne der testes, er bløde eller abstrakte. Det handler mere om, hvorvidt det er forskningsmæssigt interessant at omsætte en given problematik i denne reducerede og enkle form, som testen lægger op til, men som også definerer dens begrænsninger,” siger Jonas Lieberkind og fortæller, at han faktisk er begejstret for selve idéen og ambitionen om at teste elevernes globale kompetencer.

”*PISA Global Competence* vil bryde med vores snævre forestilling om, hvad test er. Det fantastiske i mit perspektiv er, at den har som ambition at måle noget, der går ud over den nationale opfattelse af demokrati. En national undersøgelse ville ikke kunne løse den opgave, netop fordi det ikke er den snævre nationale opfattelse af demokrati, der tæller her,” siger Jonas Lieberkind.

Så selvom hans dom lige nu lyder på en ’ommer’, så håber han, at folkene bag testen tager kritikken til sig og kommer igen med et nyt og bedre forslag, der også er bedre afprøvet end den første version.

Nuancer skal formidles

De forskere, der udvikler PISA-testene, er heldigvis ret lydhøre, fortæller Jeppe Bundsgaard. De er modtagelige for feedback – også den kritiske. Derfor er testene i konstant udvikling, for de kan hele tiden blive bedre, understreger Jeppe Bundsgaard, der følger nøje med i, hvilken vej vinden blæser, når det kommer til udviklingen af nye typer test. For ham er det væsentligt, at test tester det, der er hensigten, og ikke alt muligt andet. Det er især vigtigt at være opmærksom på i disse år, hvor testsfæren konstant udvider sig, og nye testtemaer bringes i spil.

”I forskningen er vi interesserede i, hvordan skolen fungerer på dens egne præmisser. Derfor har vi brug for valide test, der afdækker, om det, vi sætter i værk, virker. Dér bliver det interessant at teste andre ting – fx om man er god til at få ideer, og om man kan samarbejde. Det, mener jeg, er vigtigt og helt i orden, at forskere tester. Men til forskningsbrug og til at udvikle skolen, men ikke nødvendigvis som en national test,” siger Jeppe Bundsgaard.

Selvom sådanne testresultater i princippet kun er relevante for forskere og skoleudviklere, bliver de ofte kommunikeret bredt til offentligheden, fx via universiteternes kommunikationsafdelinger og derpå via medierne. Her peger Jonas Lieberkind på, at der i det meste forskning, men særligt i forbindelse med test, ligger en svær formidlingsopgave, som han mener, man som forsker skal gå til med omhu:

”Det er en trend, at forskning skal være let omsættelig i politik og medier. Men det bliver hurtigt unuanceret, særligt når vi spørger til holdninger og følelser. Det er fx ikke entydigt, hvad der er god eller dårlig trivsel. Det betyder ikke, at man ikke skal teste trivsel i skolen, men formidlingen er vanskelig. Her er vi som forskere forpligtede på at formidle konjunkturer og nuancer og balancere den efterfølgende videnspolitiske debat.” ■

TVERGÅENDE PISA-TEST

Organisationen OECD (Organisation for Economic Cooperation and Development) står bag PISA-test, som i Danmark gennemføres på 9. klassetrin. Ud over de traditionelle PISA-test, der tester eleverne i matematik, naturfag og læsning, er der gennem de senere år kommet en del nye test til:

PISA Global Competences (2018): Tester elevernes evne til at undersøge lokale, globale og interkulturelle forhold samt deres evne til på tværs af kulturelle forskelle at kunne forstå, anerkende og interagere med mennesker, der har andre perspektiver og verdensanskuelser end deres egne. Testen er gennemført i 2018, og resultaterne præsenteres sommer 2020. Danmark deltager ikke i denne test.

PISA Collaborative Problemsolving (2015): Tester elevernes evne til at interagere og samarbejde og på den baggrund løse komplekse problemstillinger.

PISA Digital Problemsolving (2012): Tester elevernes evne til at løse problemer i teknologirige kontekster

PISA Electronic Reading (2009): Tester elevernes evne til at læse og forstå elektroniske tekster

PISA Financial Literacy (2012 og 2021): Tester elevernes viden om og erfaring med at administrere penge. Danmark deltager i 2021.

Hvert tredje år tester man et af de enkelte hovedområder plus et af de nye områder. Dvs., hvert år kommer der en PISA-test (enten læsning, matematik eller naturfag) plus et tværkompetenceområde.

Kilde: oecd.org/pisa og Jeppe Bundsgaard

Læs mere:

Jonas Lieberkind og Jens Bruun: Unge, skole og demokrati. Hovedresultater af ICCS 2016. Aarhus Universitetsforlag, 2018.

Jeppe Bundsgaard: Pædagogisk brug af test. Sakprosa, årgang 10, nr. 2, 2018.



JEPPE BUNDSGAARD er professor i danskfagets didaktik. Han forsker bl.a. i test, dansk didaktik og it i undervisningen. Han underviser på didaktik-uddannelsen.



JONAS LIEBERKIND er lektor i pædagogisk sociologi. Han forsker i unges politiske engagement og politiske dannelse. Han underviser på Kandidatuddannelsen i pædagogisk sociologi.