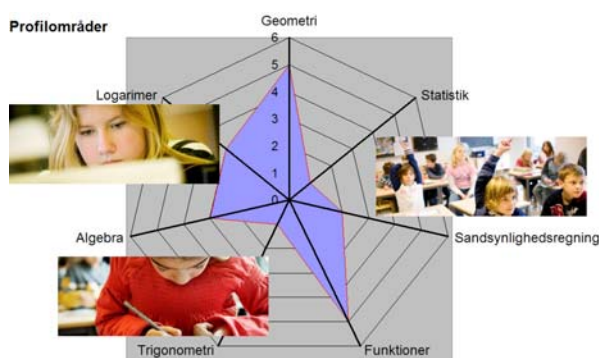


# Masterspeciale: IKT-baserede test i gymnasiet



"Learning is intellectual. Learning is social. Learning is emotional. It is ordered, and it is erratic. It happens by design and by chance. We all do it, and we take it for granted, even though we do not have a clear understanding of what it means or how to make the most of it."

Lorna M. Earl

*Skrevet af:*

MIL gruppe 14:  
Chris Østergaard  
Hans Olivur Joensen

*Vejleder:*

Jørgen Lerche Nielsen

Afleveringsdato: 29.05.2007



## Indholdsfortegnelse

<b>ABSTRACT .....</b>	<b>4</b>
<b>1. INDLEDNING .....</b>	<b>6</b>
1.1 INDLEDNING.....	6
1.2 PROBLEMFELT.....	6
1.3 PROBLEMFOMULERING .....	7
1.4 AFGRÆNSNING.....	8
1.5 LÆSEVEJLEDNING .....	8
<b>2. EVALUERING I LÆRINGSFELTET .....</b>	<b>9</b>
2.1 EVALUERINGSKULTUR (CØ).....	9
2.2 EVALUERING I GYMNASIET (CØ) .....	12
2.3 EVALUERINGSFORMER I UNDERVISNINGEN (HJ).....	14
2.4 OPSAMLING .....	18
<b>3. IKT-BASEREDE TEST I EVALUERINGSFELTET.....</b>	<b>19</b>
3.1 TESTNING I HISTORISK PERSPEKTIV, PLACERING OG TYPOLOGISERING (HJ).....	19
3.2 FOLKESKOLENS NATIONALE TEST (CØ) .....	22
3.3 TESTVÆRKTØJET QUIA (CØ) .....	27
3.4 PSYKOMETRISKE PRINCIPPER (HJ) .....	30
3.5 OPSAMLING .....	34
<b>4. METODE - TEORI .....</b>	<b>35</b>
4.1. MULTIPLE INTELLIGENSER OG LÆRINGSSTILE (CØ: 4.1.1, 4.1.4 OG HJ:4.1.2, 4.1.3).....	35
4.2 FAGLIGE TEST (CØ 4.2.1, 4.2.3 OG HJ: 4.2.2, 4.2.4) .....	49
4.3 HOLDNINGSUNDERSØGELSE – ELEVSURVEY (CØ) .....	63
4.4 FOKUSGRUPPEINTERVIEW - LÆRERE (HJ) .....	66
4.5 OPSAMLING .....	69
<b>5. EMPIRI – ANALYSE .....</b>	<b>70</b>
5.1 FOKUSGRUPPEINTERVIEW ANALYSE (CØ:5.1.2 OG HJ:5.1.3).....	70
5.2 FAGLIGE TEST – ANALYSE (CØ:5.2.1 OG HJ:5.2.2) .....	82
5.3 LÆRINGSSTILE – ANALYSE (HJ).....	88
5.4 MULTIPLE INTELLIGENSER – ANALYSE (CØ).....	90
5.5 HOLDNINGSUNDERSØGELSE ANALYSE (HJ).....	92
5.6 KORRELATIONSANALYSER (CØ).....	96
<b>6. DISKUSSION OG PERSPEKTIV .....</b>	<b>104</b>
6.1 DIDAKTISK TESTDESIGN, MI OG LÆRINGSSTILE (CØ) .....	104
6.2 FAGLIGE TEST OG TAKSONOMISKE NIVEAUER (HJ) .....	109
6.3 VYGOTSKY OG ADAPTIVE TEST (CØ) .....	111
6.4 FORMATIV EVALUERING, FEEDBACK OG LÆRING (HJ) .....	114
<b>7. KONKLUSION.....</b>	<b>119</b>
<b>8. BILAGSOVERSIGT .....</b>	<b>123</b>
<b>9. LITTERATURLISTE .....</b>	<b>124</b>

## Abstract

This thesis holds the title 'ICT-based Tests in Secondary School'.

We use the hypothetical/deductive approach, where our starting point is the problem formulation: How can we take advantage of ICT-based testing as a tool for enhancing learning? This formulation is specified in five hypotheses that we intend to confirm or disprove:

- The students performance/proficiency in testing academic subjects is determined by their Learning Styles and Multiple Intelligence preferences
- ICT-based testing can be used as an evaluation tool for higher taxonomical level than first-order knowledge
- Feedback has significance for the utility value of the test in a learning context
- The ICT-based approach will have a positive effect on the students experience carrying out the tests
- The concepts testing and learning are in fundamental conflict with each other, because they are based on different paradigms

In chapter two we look at the evaluation field in the Danish School system from an overall perspective, where the presentation of the subject is based on central figures in the evaluating/learning literature. We demonstrate the influence from international evaluation institutes on the decision-making and on the dialectic-based school culture. The term evaluation-culture is a central term in the debate, and is also discussed in this chapter.

In chapter three the focus is shifting towards testing, with a short historical review and an account of some of the central psychometric principals and concepts. We will also take a look at the Mandatory ICT-based Danish National Tests in the Primary School, which are in progress from the spring of 2007. The chapter also holds a short description of the testsoftware that we use in our empirical test.

Chapter four contains the development of the theoretical foundation for the empirical studies. The empirical research includes two surveys for determining the students Multiple Intelligence and Learning Style preferences. The research also includes two subject-based tests in biology and music, respectively. Afterwards we conducted an opinion survey to the same pupils asking them about

their attitude towards these tests and testing in general. In addition to this we carried out two focus-group interviews with teachers from the two schools involved in the project.

Data from the 72 students answering the tests from two different schools are the foundation of the quantitative empirical material. Two focusgroups consisting of 3 biology-teachers in the first group, and 2 math/phys, 2 history/sports, 3 language teachers in the other group, are the foundation of the qualitative empirical material.

The theoretical background in the studentoriented investigations is Howard Gardner, Dunn & Dunn and Benjamin Bloom.

Chapter five holds the analysis of the empirical data, and chapter six contains four selected areas from the research subjects, which have evolved in the process, that we want to bring into focus for a deeper discussion. In this chapter Vygotsky comes into play as a widening perspective.

The main conclusions we can extract from the research is, that a shift from summative to formative assessment and from behaviouristic to socioconstructivistic views of learning is significant to enhance learning of higher order, but both positions exist as a continuum and the theoretical starting point depends heavily on the purpose of the academic intentions. Computermediated testing holds some potential that can improve contextualization of testing.

# 1. Indledning

## 1.1 Indledning

De nationale test og de digitale afgangsprøver i folkeskolen skriver sig ind i en evalueringsdiskurs, som er opstået på baggrund af flere internationale evalueringer af det danske uddannelsessystem. Som det vil være mange bekendt, er disse evalueringer endt op med en del kritiske bemærkninger om det faglige niveau hos især folkeskoleeleverne. Indførelsen af testene har fået en del medieomtale og på det seneste har der jo også været en betydelig omtale af de problemer, der har været med afviklingen af prøverne i dette forår. For at gøre ondt være, har der så også været problemer med afviklingen af folkeskolens digitale afgangsprøver i biologi og geografi.

Evalueringsdiskursen er ikke blot begrænset til folkeskolen, den har slået rødder mange andre steder i samfundet herunder også i gymnasiet, hvor vi projektforfattere har vores daglige virke. Vi er gymnasielærere i henholdsvis biologi og musik og har herigennem en faglig interesse, hvor evaluering gerne skulle være en vigtig del af den undervisningsmæssige sammenhæng. Desuden er der en øjensynlig fælles interesse i IKT og læring, ekspliciteret gennem valg af MIL-uddannelsen, således at vi har en faglig motivation for at gå ind på afhandlingens område, og med den styrke, at vi har forskellig uddannelsesteoretisk baggrund, henholdsvis en naturvidenskabelig og en humanistisk.

## 1.2 Problemfelt

Undersøgelsesfeltet for nærværende afhandling er IKT-baserede test, som en del af evalueringskulturen i gymnasiet. Der er i gymnasiet ikke tale om ”nationale test” som i folkeskolen, men derimod lokalt forankrede test som en del af den formative evaluering. Som det er antydnet i indledningen, er det ikke uproblematisk at indføre IKT-baserede test. Det er det heller ikke med test i det hele taget. Test har ikke været en fremtrædende del af dansk skole- og evalueringskultur, men kommer i disse år, påtvunget eller frivilligt, ind i undervisningssektoren, i grundskolen som en bestemmelse oppefra og i andre uddannelsesinstitutioner på mere frivillig basis. Det er næppe usandsynligt, at der også kommer internationale og nationale evalueringstiltag til de højere skoletrin i den nærmeste fremtid.

Vi har valgt at undersøge om IKT-baserede test kan være et meningsfuldt lokalt forankret redskab i den gymnasiale undervisningspraksis. Der er stor forskel på om test eller evaluering påtvinges en sektor eller en institution udefra med standardiserede krav til indhold, form og dokumentation, eller om det sker som et lokalt initiativ med uddelegering af beslutningskompetence til den enkelte lærer. Det er ikke frivilligt for gymnasierne at evaluere, undervisningsministeriet har fastsat krav om evalueringen af gymnasiets undervisning, men det enkelte gymnasium kan selv vælge hvordan evalueringen skal foregå.

### ***1.3 Problemformulering***

#### **Hvordan kan IKT-baserede test inddrages i gymnasiet som et redskab der fremmer læring?**

Problemformuleringen operationaliseres gennem nedenstående hypoteser, der danner baggrund for den empiri vi tilvejebringer:

- **Elevers præstation i faglige test er betinget af deres læringsstile og multiple intelligensformer**

Faglige test skal her forstås i betydningen evaluering af elevernes læring inden for afgrænsede fagområder i de to fag biologi og musik. De læringsstile vi tager udgangspunkt i er Dunn & Dunns. Ved multiple intelligensformer forstår vi de intelligensformer Howard Gardner har introduceret.

- **IKT-baserede faglige test kan anvendes til at evaluere kundskaber på højere taksonomiske niveauer end paratviden**

De højere taksonomiske niveauer vi henviser til, er de niveauer, der fremgår af Blooms taksonomi.

- **Feedback har betydning for testens nytteværdi i læringssammenhæng**

Med feedback hentyder vi til den respons der gives, enten medieret via testsystemet eller via en lærers tilbagemelding.

- **IKT-basering påvirker elevernes oplevelse af at gennemføre test i positiv retning**

Alternativet til en IKT-basering, vil være en "papir og blyant test". Dvs. antagelsen går på, at eleverne foretrækker en IKT-baseret test frem for en papirtest.

- **Testning og læring er i grundlæggende modstrid med hinanden, idet de baserer sig på forskellige paradigmer**

Antagelsen kan også formuleres på følgende måde: Læring opfattet som en konstruktivistisk videnskonsstruktion er uforenelig med test, idet test anses for at bygge på et behavioristisk grundlag.

Ud over disse hypoteser benytter vi os af følgende teorifelter i vore undersøgelser af problemstillingen: læringsteori, kognitionsteori, læringsstilteori, evalueringsteori og testteori.

### ***1.4 Afgrænsning***

Som det fremgår ovenfor, er der i afhandlingen fokus på, hvordan faglige test kan indgå i evalueringsspraksis i gymnasiet. Vi vil derfor kun perifert behandle de andre evalueringsformer, der kan bringes i anvendelse. Vi ser desuden bort fra den evaluering, der vedrører selve undervisningen og lærerens indsats. Endelig vil vi koncentrere os om det operationelle niveau og kun som led i kontekstualiseringen beskæftige os med det strategiske og politiske niveau.

### ***1.5 Læsevejledning***

I indledningen i **kap. 1** gives en kort introduktion til afhandlingens genstandsfelt. Derefter skitseres problemfeltet efterfulgt af problemformuleringen og de fem heraf udledte hypoteser. Som afslutning på kapitlet afgrænser vi afhandlingen. I **kap. 2** beskriver og diskuterer vi evalueringssfeltet i lyset af den aktuelle kontekst, affødt af indførelsen af de nationale test i folkeskolen og gymnasireformen. Vi diskuterer herunder evalueringsskulturbegrebet. I **kap. 3** redegør vi for det teoretiske grundlag bag test herunder psykometriske principper. Som konkret og aktuelt eksempel på en test beskrives og diskuteres de nationale test, der er indført i folkeskolen. I **kap. 4** redegøres der for den teori og de metoder, der danner grundlag for vores empiri. I **Kap. 5** fremlægges en forholdsvis struktureret analyse af den empiri afhandlingen har tilvejebragt. I **Kap. 6** fremlægger vi en tematiseret diskussion og perspektivering på grundlag af den fremlagte teori og empiri. Endelig afrunder vi med en konklusion, i **Kap. 7**, på afhandlingens resultater og teori i relation til den problemstilling og de hypoteser vi indledningsvis har præsenteret som ramme for fremstillingen.

Bilagsdelen foreligger på CD-rom. Her ligger bilagene i et pdf-dokument og følger bilagsoversigten fra **Kap. 8** kronologisk. **Kap. 9** er litteraturlisten. En del henvisninger til kilder er i afhandlingen angivet som links, idet kilderne er hentet fra internettet. Opgaven er på 110 normalsider. Ansvarsfordelingen fremgår af indholdsfortegnelsen med deltagernes initialer CØ og HJ.



## 2. Evaluering i læringsfeltet

### 2.1 *Evalueringens kultur*

Der er indenfor den offentlige sektor og uddannelsessektoren i særdeleshed opstået et øget fokus på kvalitetssikring og evaluering. Det er et fokus som, i hvert fald delvis, er politisk betinget. Myndighederne har i stigende grad fokus på "value for money" - får samfundet tilstrækkelig værdi for pengene (Dolin 2006:125)<sup>1</sup>. Det er en tænkning som formentlig er inspireret af den allestedsnærværende og meget omtalte globalisering. Evaluering er blevet et styringsinstrument og konkret har det udmøntet sig i ny lovgivning inden for mange områder herunder også uddannelsesområdet.

I 2004 udkom en OECD rapport om den danske folkeskole, hvor en af anbefalingerne lyder:

Da en stor del af det danske samfund ser ud til at have taget skolevæsenets succes for givet, og vurderingen af dets funktionsduelighed i forhold til andre systemer kun har været mulig i sjældne tilfælde, er det klart, at det vil være vanskeligt at etablere en ny evalueringens kultur. Men det er sikkert den enkeltforandring, det er mest vigtigt at opnå, hvis andre initiativer skal kunne indføres, så de får virkning, og standarderne kan hæves. (Mortimore et al 2004: 129)

Peter Dahler-Larsen<sup>2</sup> mener at disse anbefalinger til den danske skole er mere et produkt af tidstypiske begreber end af eksakte data om den danske skoles situation (Dahler-Larsen 2005:2), hvor han antyder at disse internationale eksperter ikke har nok grundlæggende kendskab til det danske skolesystem.

Dahler-Larsen diskuterer og problematiserer evalueringens kulturbegrebets betydning og når frem til, at der i sammenstillingen af de to begreber evaluering og kultur er et indbygget paradoks, (Dahler-Larsen 2006:38ff). Paradokset anskueliggøres af Dahler-Larsen ved at opstille de to begreber som modsætninger: F.eks. er evaluering et eksplicit begreb, hvorimod kultur er et implicit begreb. Ideelt set skal evalueringens metoder og evalueringens resultater være så eksplicite som muligt og eksplicit begrundet. I modsætning hertil er kultur og kulturfænomener ofte implicite. Dahler-Larsen illustrerer det på følgende vis: " Et evalueringens resultat skal i princippet være til at forstå for ethvert menneskeligt fornuftsvæsen, hvorimod et kulturelt fænomen ikke behøver at være forståeligt for andre end

---

<sup>1</sup> Jens Dolin er lektor ph.d. og forskningsleder ved Institut for Filosofi, Pædagogik og Religionsstudier, Syddansk Universitet.

<sup>2</sup> Peter Dahler-Larsen er ph.d. og professor på Institut for Statskundskab, Syddansk Universitet.

medlemmerne af den pågældende kultur” (ibid:38). Man kan måske sige, at kultur og kulturfænomener ofte beror på tavs viden – altså viden, der ikke er ekspliciteret. Et andet eksempel som han nævner, er viden som forudsætning for evaluering, hvor kultur er baseret på tro: ”Evaluering er baseret på viden, kultur på tro – forstået som to fundamentalt forskellige principper for, hvordan det, som tages for gældende, i en given social situation produceres, vedligeholdes og sættes igennem” (ibid:38ff). Dette paradoks er der næppe enighed om, eftersom der findes forskellige opfattelser af, hvordan begrebet kultur kan eller skal defineres. Den oprindelige latinske betydning af ordet kan betyde ”at dyrke”<sup>3</sup> dvs. noget, der er menneskeskabt. Ud fra sådan en forståelse, giver det i vores optik ikke mening at definere kultur som noget der bygger på tro. Efter vores opfattelse, er der er i lige så høj grad tale om viden, både i betydningen eksplicit og tavs viden. Ud fra en sådan forståelse kan der gives mange eksempler på artefakter og værdier, der er frembragt af en kultur, der baserer sig på viden.

Paradoks eller ej så fremhæver Dahler-Larsen, at evalueringskultur indebærer, at man ikke er bange for at bruge data som sådan, og samtidig har et kyndigt og afmålt forhold til, hvad de forskellige datakilder kan. Endvidere påpeger han at åbne kommunikationskanaler og vidensdeling er forudsætninger for en evalueringskultur (Dahler-Larsen 2005:3). Dokumentation og systematik er også blevet et voksende udefra kommende krav til uddannelsesinstitutioner og nu til underviseren. Institutionerne må synliggøre og legitimere deres praksis og ”gøre dem til genstand for egne og andres evalueringer” (Bendixen 2005:8)<sup>4</sup>. Det eksterne dokumentationskrav bliver i mindre grad pædagogisk begrundet (Bendixen 2005:23). Niels Egelund<sup>5</sup> nævner, at danske lærere utvivlsomt hele tiden har evalueret deres undervisning, men at det bør ske på en langt mere systematisk basis. Han påpeger det som værende vigtigt, at der sker en løbende intern og individuel evaluering af elevernes progression, og det er vi for dårlige til i Danmark, og peger på at årsagerne ligger i den svage evalueringskultur (Egelund 2005:3ff).

Vi kan konstatere, at der er voksende fokus på evaluering og at Undervisningsministeriet har taget begrebet evalueringskultur til sig og indskrevet det i lovgivningen for folkeskolen ”Udviklingen af en evalueringskultur i folkeskolen er både vigtig og nødvendig. Det er baggrunden for den store fokus på evalueringskultur i den nye folkeskolelov” (Undervisningsministeriet 2007).

---

<sup>3</sup> Se f.eks. <http://gyldendalsleksikon.dk/leksikon/leksikon.htm> og Schultz 2004:11ff

<sup>4</sup> Carsten Bendixen er mag. art. I psykologi lektor Videncenter for Didaktik og Professionsudvikling CVU Storkøbenhavn

<sup>5</sup> Niels Egelund Professor Institut for Pædagogisk Psykologi ved Danmarks Pædagogiske Universitet

Som et delelement i denne evalueringskultur foreslås indførelsen af de nationale test: ”Et vigtigt pædagogisk værktøj til styrkelse af evalueringskulturen er test. Der er derfor udviklet en række it-baserede test, som bliver en del af den fremtidige evaluering” (Undervisningsministeriet 2007).

Ejendommeligt er det, som Egelund peger på, at PISA hverken peger på, at standardiserede test øger eller mindsker de faglige kompetencer, men at en øget løbende evaluering af undervisningen synes at nytte (Egelund 2005:46). Fagligheden i folkeskolen har i debatten siden sidst i 90’erne fået et skud for boven, hvor undervisningen kritiseres for en forskydning mod pædagogisk arbejde mere end undervisning efter de rent faglige mål. Mange af de seneste initiativer fra Undervisningsministeriet har gået på at øge de faglige kvalifikationer hos eleverne. Bendixen påpeger, at i mange af de kvalitetsudviklingsprojekter, der er sat i gang på folkeskoleområdet, er der usikkerhed og uklarhed om evalueringens ”metodik, erkendelsesgrundlag og ikke mindst sammenhæng med didaktisk forståelse” (Bendixen 2005:9).

Obligatoriske IT baserede test introduceres i foråret 2007, som en metode til at måle om skoleelever opnår de fastsatte mål i den danske folkeskole. Andre uddannelsesinstitutioner (f.eks. gymnasierne) er ligeledes blevet pålagt at evaluere undervisningen og dokumentere denne. Der tegner sig således et billede af, at elevernes læring og opnåelse af de fastsatte mål kan ske ved hjælp af test og evaluering. På den anden side, er det stærkt omdiskuteret om test vil kunne føre frem til de mål, der ønskes såvel politisk som fagligt.

For så vidt angår gymnasiet er det bemærkelsesværdigt, at begrebet evalueringskultur ikke anvendes i beskrivelsen af gymnasiereformen, der trådte i kraft i august 2005. Her tales der i stedet for om kvalitetssikring og resultatvurdering. Af gymnasieloven fremgår det således af § 37: ”Rektor fastlægger et system til kvalitetsudvikling og resultatvurdering i forbindelse med uddannelsen og undervisningen. Stk. 2. Undervisningsministeren fastsætter nærmere regler herom” (Undervisningsministeriet 2004).

Grunden til, at der ikke anvendes den samme terminologi som i fastsættelsen af målene for folkeskolen er sikkert den, at evalueringskultur begrebet ikke endnu ikke har manifesteret sig eftersom loven er skrevet før OECD-rapporten slår igennem med begrebet. Men der er næppe nogen tvivl om, at intentionerne med gymnasiereformen svarer til målene i folkeskolen. Det skal dog bemærkes, at der ikke er noget krav i gymnasiereformen om, at eleverne skal testes som i folkeskolen i dag.

Dahler-Larsen påviser en latent risiko ved det store fokus på evaluering, at evaluering kan blive institutionaliseret og objektiviseret, hvor evaluering bliver til for evalueringens skyld. Scenariet er, at hvis evaluering får en højere status end pædagogikkens teori og praksis og løsrives, snarere end at integreres med praksis, kan det føre til en forsimpning og reduktion af undervisningens væsen (Dahler-Larsen 2005:8).

## 2.2 Evaluering i gymnasiet

Som redegjort ovenfor er evaluering og evalueringskultur begrebet kommet på dagsordenen i den uddannelsespolitiske debat og objektet for nærværende projekt er evaluering i de gymnasiale uddannelser. Vi vil derfor give en fremstilling af baggrunden for den evalueringsdiskurs, der er i spil i gymnasiet i dag.

Af den gældende Bekendtgørelse om uddannelsen til studentereksamen (stx-bekendtgørelsen)<sup>6</sup> fremgår det af § 120:

at skolen skal udarbejde en evalueringsplan og af stk. 2 fremgår det at "Planen skal sikre, at der sker en løbende evaluering af den enkelte elev, så der sker en vurdering af elevens faglige, almene og personlige kompetencer og sikres udvikling af disse.

Og videre fremgår det af § 122, at

den løbende evaluering skal sikre, at den enkelte elev får et klart billede af egne styrker, svagheder og fremskridt og give grundlag for at justere undervisningen.

Stk. 2. Evalueringen skal danne grundlag for vejledning af den enkelte elev med hensyn til både faglig udvikling og udvikling af arbejdsmetoder.

Stk. 3. Resultatet af evalueringen drøftes med klassen eller holdet og med den enkelte elev og anvendes til at justere undervisningens progression og niveau i forhold til elevernes faglige formåen.

Selvom det ikke nævnes direkte i teksten, kan der ikke være tvivl om, at det evalueringsbegreb der her introduceres er *formativ evaluering*. Men det betyder ikke at den summative evaluering er afskaffet. Eleverne skal stadigvæk evalueres med standpunktskarakterer, prøver og eksaminer. Ellen Krogh<sup>7</sup> udtrykker det på denne måde: "I gymnasireformen 2005 er evaluering indskrevet som en

---

<sup>6</sup> [http://www.uvm.dk/nyheder/gymnasireform/udd\\_bekendtgørelser.htm?menuid=152510](http://www.uvm.dk/nyheder/gymnasireform/udd_bekendtgørelser.htm?menuid=152510)

<sup>7</sup> Ellen Krogh er ph.d. i danskfagets didaktik og lektor ved institut for Institut for Filosofi, Pædagogik og Religionsstudier, Syddansk Universitet.

nøglefunktion i uddannelserne gennem en dobbelt strategi af styring, kontrol og mere formative evalueringsstrategier” (Krogh 2006:359). Når vi yderligere lægger det tidligere omtalte krav om, at rektor skal fastlægge et system til kvalitetsudvikling og resultatvurdering står det klart, at evalueringen skal køre ad to spor i form af såvel formativ som summativ evaluering. Der er dog ikke noget nyt i at undervisningen skal evalueres. Det blev den også under den forhenværende gymnasielov. Det nye er som det fremgår ovenfor, at evalueringen (den formative) skal dokumenteres.

Torben Spanget Christensen<sup>8</sup> taler om evaluering på makro- og mikroniveau, hvor makroniveauet er samfunds niveauet eller forvaltningsniveauet. Hans pointe er, at styringsproblemer på dette niveau har afstedkommet et krav om evaluering – et krav som har været gældende i flere år – med henblik på at løse dette problem f.eks. via øget effektivisering og ressourceudnyttelse. Problemet er så, at dette krav på makroniveauet slår igennem på mikroniveauet (undervisningen lokalt). Kvaliteten af undervisningen betegner han som et mikroproblem, der kvalitativt og kvantitativt er forskelligt fra makroproblemet. Det kan således være meget svært meningsfuldt at overføre evalueringskravet fra makroniveauet til mikroniveauet (Christensen 2004:269). Vi aner således konturerne af en konflikt her. Torben Spanget Christensen konstaterer da også følgende: ”På mikroplanet i det enkelte klasseværelse, medfører evalueringskravet en risiko for, at den tradition og den professionalisme, der er opbygget gennem årene, bliver fejet af bordet af makroniveauets logik. Mikroniveauet mangler et effektivt modtræk, der både kan forsvare denne professionalisme og indarbejde makroniveauets evalueringskrav meningsfuldt” (Christensen 2004:269ff). Vi ser her en konflikt mellem den formative evaluering, der er forankret internt i gymnasiet og den summative evaluering der kommer til udefra som en kontrollerende evaluering. Det er næppe så meget det, at evalueringen er summativ som det, at den bruges i en kontrollerende funktion der vækker anstød. Peter Dahler-Larsen forklarer det med, at

i et kontrolsigtet ønsker man typisk at beskrive relativt komplekse aktiviteter og udfald heraf ved hjælp af relativt simple kriterier. Dette er forudsætningen for, at der er god styringsøkonomi i et givet kontroltiltag, og at kontrollen netop understøtter en beslutning om en sanktion, hverken mere eller mindre. .... Ikke overraskende har kontrol-evalueringer derfor et ret dårligt image hos de kontrollerede. Det er ikke rart at få en kompleks indsats belyst ved hjælp af simple kriterier, især når det kan føre til sanktioner (Dahler-Larsen 2006:124ff).

---

<sup>8</sup> Torben Spanget Christensen er ph.d. i undervisningsevaluering. Konsulent ved Center for Undervisningsmidler.

Et andet konflikt punkt, som allerede er antydnet i forrige afsnit, kan tænkes at bestå i, at gymnasierne og gymnasielærernes selvforståelse, som professionelle udøvere af lærergerningen, trues af formaliserede eksterne evalueringskrav (såvel formative som summative). Man kan forestille sig, at professionaliseringen har medført, at lærerne betragter gerningen som uantastelig. Men det er så netop pointen ved evaluering, at man antaster det som nærmest fremstår som naturgivet (eller uantasteligt). Evalueringsgenstanden skal bevise sit værd, ellers kan der rejses tvivl om dens eksistensberettigelse eller ved dens nuværende måde at fungere på (Dahler-Larsen 2006:31).

Noget tyder så alligevel på, at gymnasielærerne er ved at flytte sig i hvert fald konstaterer Christensen:

Med den stærke markering af evalueringskravet i den gymnasiereform, der træder i kraft i 2005, er det som om, at en mere imødekommende holdning til evaluering breder sig i de gymnasiale uddannelser. Alt i alt kan man karakterisere situationen som en blanding af modvilje (evalueringer er noget vi er nødt til), rådvildhed (vi ved ikke hvordan vi skal gribe det an), nysgerrighed (vi vil meget gerne høre om, hvordan man kan gøre det og hvilke virkninger det kan have, vi vil også meget gerne prøve forskellige metoder) og uafklarethed (vi ved ikke rigtigt, hvilke veje vi kan gå) (Christensen 2004:271).

### **2.3 *Evalueringsformer i undervisningen***

Carsten Bendixen påviser det komplekse didaktiske forhold mellem evaluering af undervisning og læreprocesser. Han opsummerer at i nutidens skole "eksisterer traditionel måling og pædagogisk prøvetagning side om side med nye pædagogiske selvevalueringskomponenter som logbøger, portfolioer og kvalitetsvurderinger, og det er ikke underligt, at mange pædagoger og lærere metodisk og begrebsmæssigt har vanskeligt ved at navigere i dette felt (Bendixen 2005:10-11). Der ses ovenfor en skelnen mellem evalueringsformerne, hvor logbøger og portfolioer beskrives som selvevalueringskomponenter. Logbøger og portfolioer er blevet udbredte redskaber i procesorienteret undervisning, både som dokumentation og evalueringsredskaber. Sammen med it-integrationen og særligt asynkrone konferencesystemer, har disse redskaber lagt op til en øget skriftlig baseret kommunikation. Dette har, sammen med et stigende fokus på betydningen af den lærendes refleksion, stillet nye didaktiske muligheder til rådighed for undervisningen. Refleksion fremmes i kommunikationen med omgivelserne og er en forudsætning for læring på læringsniveau 2, hvor man kommer ud over klassisk betingning forstået som udenadslære eller instrumentel indlæring (jf. Hermansen 2005:23).

Vi vil i et selvstændigt afsnit i den perspektiverende del komme ind på feedback og 'scaffolding' som forudsætning for læring. Ellen Krogh skriver om gensidig feedback i undervisningen, at det "opøver elevernes refleksionskompetence, dvs. deres kompetence i at vurdere, begrunde og dokumentere kvalitet i produkter og processer" (Krogh 2006:405). Hun skriver om portfolioevalueringens tre nøgleaktiviteter: indsamling, udvælgelse og refleksion (Krogh 2006:401). Elevinddragelse kendetegner en portfolio og kan både bruges som et fagdidaktisk læringsværktøj, hvor vægten ligger på faglige kompetencer, og til mere alment udviklende kompetencer. Desuden kan den være både proces- og produktevaluerende. Portfolio kan betragtes som byggende på et konstruktivistisk læringssyn, graden afhænger lidt af, hvilke læreprocesser der arbejdes med, om der er tale om private porfolier, eller de er en del af en større social undervisningsmæssig sammenhæng. Vi har valgt at nævne portfolier i denne sammenhæng fordi det virkelig sætter test, som værktøj i evalueringen, i relief. Graaff (Graaff 2000:8ff) taler om et rationelt paradigme på test eller vurdering (assessment). Assessment karakteriseres som en videnskabelig måling, eller rationalistisk: "The corner stone of the rational, or scientific measurement paradigm is static. In order to measure something it has to remain constant, at least for a reasonable period of time" og det antydes endvidere, at der kunne være et processuelt paradigme:

Consequently, criteria like reliability and validity must be re-defined. Instead of repetition, reliability could for instance be defined in terms of trustworthiness, emphasizing that the students' own judgement should be an integral part of assessment. Validity could be re-defined in terms of the degree to which assessment contributes to improvement. Applied to assessment itself, that would mean that assessment is good as long as it contributes to the general objectives of education" (ibid:10).

Her ses elevinddragelse som en del af et nyt evalueringsparadigme.

Carsten Bendixen ser begreberne i historisk perspektiv og sammenfatter det perspektivskifte, der er sket på evalueringsområdet gennem de seneste årtier med 1970'ernes fokusering på den interne evaluering som et fælles pædagogisk projekt til 1980'ernes og 1990'ernes udvikling af individualiseret og forhandlet selvevaluering og den aktuelle situation med nationale Fælles Mål og indførelsen af obligatoriske test fra 2006 sideløbende med udstrakt anvendelse af selvevalueringsmetoder som logbøger, portfolioer og samtalerark under samlebetegnelsen *Evalueringskultur* (Bendixen 2005:15).

Bendixen taler også om evalueringsparadigmer, hvor han, efter Guba og Lincoln, opdeler i 4 generationer af evaluering. Den første evalueringsopfattelse, som psykometriken bygger på, er den traditionelle objektive testning som blev udviklet i forbindelse med intelligencetest (se også kapitel 3). Han kalder det også det naturvidenskabelige evalueringsparadigme forankret i en behavioristisk teoriramme (Bendixen 2005:28-29). I 2. generationsforskningen taler han om en curriculumpåvirket forskning, hvor der opstod den såkaldte undervisningsteknologi, med en indsnævring af det pædagogiske projekt, med koordination af undervisningsmidler, undervisningsmetoder og undervisningsmål (Bendixen 2005:32). Undervisningsformen kan kritiseres for at danne basis for instrumentel indlæring, hvor der læres udenfor en meningsfuld kontekst. Spørgsmålet er om ikke den internationale påvirkning som vi mærker i høj grad har sin grobund i dette paradigme? En del af kritikken af det danske system har været, at målbeskrivelserne har været for ”bløde” og mere hensigtserklæringer end konkrete mål.

Med den sociologiske orientering i 70’erne kom der modspil til målopfyldelsesevalueringen.

Bendixen beskriver 3. generation på den måde, at den holdnings- og værdibaserede interne evaluering var en modreaktion mod den rationelle instrumentelle evaluering (Bendixen 2005:42).

Den 4. generation af evalueringsparadigmer beskrives som et socialkonstruktivistisk perspektiv, hvor forhandling er et nøgleord (Bendixen 2005:43). Det konstruktivistiske perspektiv er først og fremmest orienteret omkring processuelle forhold (Bendixen 2005:52).

Situationen i dag er at alle disse perspektiver eksisterer side om side, og at der umiddelbart er store modsætninger mellem retningerne, både ideologiske og læringsteoretiske. Vores teoretiske ”dilemma” har været at finde en faglig evalueringsform som både fremmer personligt udviklende læreprocesser og samtidig kan honorere kravene til de faglige mål som bekendtgørelserne foreskriver. Bendixen belyser samme problemstilling hvor ”det grundlæggende dilemma i 1990’ernes evalueringsforståelse i den danske folkeskole, at resultatorienterede ’objektive’ målinger af elevfærdigheder gennem standardiseret testning..kom i et metodologisk modsætningsforhold til den procesorienterede forhandlingsevaluering” (Bendixen 2005:59ff). Han nævner at i den aktuelle evalueringsdebat synes begrebet om selvevaluering at være en midlertidig platform for konsensus (ibid:90).

Det er et udbredt krav i bekendtgørelsen, at der foretages løbende evaluering i fagene. Der skelnes ofte mellem formativ og summativ evaluering. I engelsk litteratur tales ofte om ’high stakes assessment’ og ’low-stakes assessment’, med en betydning som ligger tæt på ekstern og intern evaluering. Low-stake assessment bruges også i betydningen formativ, hvor det knyttes til lærings-



og feedbackprocesser. Krogh nævner, at de gymnasiale uddannelser har traditionelt opfattet evaluering som summativ, hvor den har været rette mod produkt og afslutning af forløb (Krogh 2006:356). Der er kommet mere fokus på den formative evaluering og der er tendenser til at opdelingen imellem summative og formative metoder er flydende. Torben Spanget Christensen skriver, at de summative metoder fokuserer på effekter og virkninger, mens de formative fokuserer på lære- og udviklingsprocesser (Christensen 2006:361). Han påpeger også evalueringskrav på forskellige niveauer: skoleeksterne, skoleinterne og undervisningsinterne krav. Det sidste krav svarer til vores fokus i nærverende projekt, hvor det er evaluering i forbindelse med elevernes læring der er temaet. Han argumenterer for en systematisk dataindsamling, og nævner den intuitive vurdering, som er legitim, men medfører en risiko for skred i fokus, pga. manglende forankring i klare evalueringskriterier og metoder (ibid:362).

Sammenhæng mellem mål og metoder er ikke altid helt klart i den pædagogiske praksis, hvilket vi som praktikere kan nikke genkendende til, hvor der i store klasser ofte er problematisk at få faglige mål og pædagogisk-didaktiske mål til at gå op i en højere enhed.

Han opstiller to grundlæggende kriterier for en formativ evalueringsmetode. Den skal indeholde en, måske overraskende, summativ kerne med systematik omkring måloppstilling og dataindsamling og være indrettet på at stimulere lære- og udviklingsprocesser (ibid:366). I figuren nedenfor vises Christensens systematisering af den formative evalueringsproces:

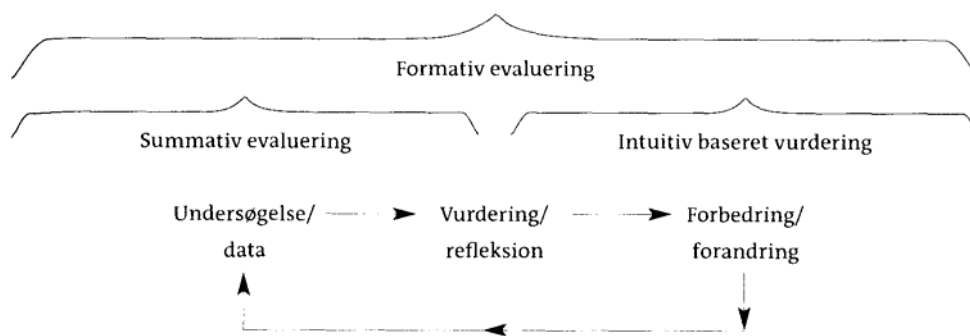


Fig. 1: Kopi fra Christensen (2006:367)

Den formative evaluering omfatter således flere delprocesser fra dataindsamling, refleksion, forandring til feedback i en iterativ proces. Den formative evaluering kan sigte helt direkte mod det faglige stof, læringsproduktet, som skal læres ifølge en given læreplan. Den kan også sigte mod de bredere almene kompetencer som læringen har til formål at fremme. Test må i sig selv siges at være summative, men som ovenstående figur viser, kan den ses i en formativ sammenhæng, hvilket er det

mål vi forfølger i specialeprojektet. Der er som tidligere nævnt en lidt flydende grænse mellem formativ og summativ evaluering. Er Folkeskolens nationale test summative, når deres formål er at være et pædagogisk redskab for læreren? Det afgørende er, hvordan den enkelte evalueringsform benyttes i den konkrete undervisningssammenhæng. Her ligger det store ansvar og udfordringen for læreren eller lærerteamet i evalueringsøjemed. Den summative evaluering har faldgruber som kan lukke for den åbne dialog, idet den indbefatter et stort kontrolelement. To spørgsmål som, foreløbigt, står uafklarede er om den summative evaluering, herunder test, kendetegnet med den faglige målstyring, står i modsætning til socialkonstruktivistiske grundprincipper, hvor viden forhandles i den sociale kontekst? og i hvilken grad det er muligt at integrere formative og summative elementer i den pædagogiske praksis?

### ***2.4 Opsamling***

Vi har med dette kapitel forsøgt at beskrive og problematisere begrebet evalueringsskultur. Med henvisning til Peter Dahler-Larsen, har vi redegjort for, at begrebet kan opfattes som et paradoks idet evaluering og kultur, hver for sig kan betragtes som modsætninger. Vi har nævnt de nationale test i folkeskolen som værende et konkret eksempel på en bestemt evalueringsskultur. Med henvisning til de formelle krav i gymnasireformen, har vi kontekstualiseret begrebet ind i den verden vi beskæftiger os med, dvs. undervisning i gymnasieskolen. Vi har endvidere operationaliseret begrebet ved at beskrive forskellige evalueringsformer herunder forskellen på summativ og formativ evaluering og i den forbindelse peget på test, som et muligt værktøj i den formative evaluering.

### 3. IKT-baserede test i evalueringsfeltet

#### 3.1 Testning i historisk perspektiv, placering og typologisering

##### IQ test og racisme

Det er forståeligt at flere i dagens samfund med et historisk indblik rykker sig lidt ved tanken om testning og begrebet test. Rust & Golombok (1999) beskriver i en kort gennemgang af psykometriens historie intelligens-testbevægelsen i forrige halvdel af 1900-tallet som værende ”not simply like Nazism in its racist aspects – it was its ideological progenitor” (Rust & Golombok 1999:8). IQ-test blev både i USA og Europa brugt i racehygiejnisk sammenhæng, hvor personer med lav IQ blev steriliseret og fik indsnævret deres bevægelsesfrihed i samfundet og mellem stater og lande (ibid:9). At der var tale om ekstreme synspunkter og antagelser i denne periode illustreres af nedenstående citat, hvor irere og skotter sammenlignes:

The careless, squalid, unaspring Irishman, fed on potatoes, living in a pig-sty, doting on a superstition, multiplies like rabbits or ephemere: - the frugal, foreseeing, self-respecting, ambitious Scot, stern in his morality, spiritual in his faith, sagacious and disciplined in his intelligence, passes his best years in struggle and in celibacy, marries late, and leaves few behind him (ibid:14).

Rust & Golombok understreger, at den viden som findes i dag indenfor områder som informationsvidenskab kraftigt anfægter den snævre opfattelse af, at ”kvaliteten” af vores bevidsthed eller intellektuelle formåen kan måles som vores evner til at løse simple logiske, aritmetiske opgaver gennem typiske IQ-test (ibid:18). De nævner dog også (ibid:19), at nyere opfattelser af intelligens som Howard Gardner og Daniel Goleman (Emotional Intelligence), har haft en begrænset indflydelse på psykometrien. Selv om disse intelligensformer også kan måles psykometrisk har testresultaterne vist sig ikke at være særlig fjernt fra eksisterende personlighedstest og test af evner (ability). Men konkluderende kan det siges, at psykometrien har bevæget fokus hen imod en mere differentieret forståelse af menneskelige evner. Evner ses nu mere som en forudsætning for særlige færdigheder indenfor områder som verbal og abstrakt ræsonnering, stavning og sprogbrug o.a. (ibid:21). Disse delelementer definerer så en persons evner, men der er efter vores mening stadigvæk tale om testformer som kan kaldes intelligens eller personlighedstest, selv om det bagvedliggende menneskesyn er forandret med tiden. Ud fra bogens indledende fremstilling må vi konkludere, at nutidens psykometrikere har gjort op med dette forskningsfelts fortid. Det er nok heller ikke en velbegrundet og

retfærdig opfattelse af forskningen indenfor psykometri i dag, at tro at området skulle have en lavere moralsk kodeks end andre felter i det moderne samfund per se.

Haywood et al nævner, at der har været kritik af psykometrisk testning heriblandt bias overfor minoritets- og specialundervisningsgrupper, selektiv fortolkning af resultater indenfor disse grupper og manglende overvejelse over motivationsmæssige, personlighedsmæssige, og sociale faktorer som afgørende for menneskelig gøren (Haywood et al 1992:6).

### Nyere tilgange

Haywood og hans forskningsteam plæderer for et "transactional perspective" i forbindelse med "Psychoeducational Assessment". Dette perspektiv kan kort beskrives som en opfattelse af intelligens som flerfacetteret, i forhold til generel intelligens (g-faktor), og en skarp opdeling mellem intelligens og kognitive processer.

Tabel 3.1. Sammenligning af intelligens og kognitive processer. (Efter Haywood et.al 1992:47)

Dimension	Intelligence	Cognitive processes
Source	Largely genetic	Must be taught/learned
Modifiability	Modest, with great effort	High, with teaching
Character	Both global and specific; equals ability to learn	Generalized across content domains
Assessment	Achievement; product of past learning	Process assessment; learning in teaching situations; dynamic
Composition	Intellectual aptitudes (verbal, spatial, memory, quantitative, etc.)	Mix of "native" ability, work habits, attitudes, motives, strategies
Parents' role	Genes, nutrition, health, safety	Mediated learning; active, directed teaching

Intelligens er ifølge Haywood et al fortrinsvis genetisk bestemt, mens kognitive processer erhverves. Det er kun i moderat form muligt at modificere intelligens, mens de kognitive processer i stor grad kan modificeres, idet de kognitive strukturer er etableret ved læring. Kort kan siges, at denne indfaldsvinkel peger på intelligens begrebet som udækkende og bør erstattes med en fokusering på kognitive processer. Vi tilslutter os således opfattelsen af intelligens betragtet som en genetisk evne, som værende et utilstrækkeligt redskab til at forklare individuelle forskelle i perception, læring, problemløsning og social interaktion.

Dette har relevans i forhold til vores empiri, idet vi foretager test med udgangspunkt i Gardners Mange Intelligenser og Dunn & Dunns læringsstile, hvor Gardner taler om intelligenser og Dunn & Dunn om læringsstile som delvis bygger på kognitive processer.

### **Testområder og typer**

De overordnede testtyper som vi bruger i vores empiriske undersøgelser er selvvurderingstest og faglige test. Vores test af Multiple Intelligens og Læringsstil-test er at betragte som selvvurderings- eller personbaserede test, hvor spørgsmålene går på personen, og skal tage stilling til, hvordan forskellige spørgsmål, udtalelser eller påstande harmonerer med ens opfattelse af sig selv og sine omgivelser. De faglige test kan kaldes vidensbaserede og går ud på at undersøge testpersonens kunnen eller færdigheder indenfor et udvalgt fagområde.

En stor del af forskningslitteraturen indenfor testning er engelsk, og det skyldes, at forskning i test er noget mere udbredt i Angelsaksiske lande end i Skandinavien. I bogen *Psykologisk Testning* (Jackson 2002) giver Jan Ivanouw et samlet overblik over de områder, hvor der er en tradition for at bruge test i Danmark (ibid:183). De områder han nævner er: det klinisk psykologiske område, militæret som bruger psykologiske prøver, erhvervslivet, hvor der ofte bruges oversatte amerikanske test og det pædagogiske område, hvor DPU har udviklet eksempelvis stave- og læseprøver til brug i folkeskolen. Han nævner også, at der inden for voksenpædagogik ikke har været tradition for at anvende test (ibid:187).

Der skelnes ofte i litteraturen mellem klassisk testteori og Item Response Theory (IRT) (Williams et al 2006:181), hvor IRT har haft stor indflydelse i forbindelse med uddannelsesmæssig testudvikling. En af de mange forskelle mellem teorierne er, at IRT giver teststatistikere mulighed for meningsfuldt at sammenligne et testspørgsmåls sværhedsgrad og testtagerens evner<sup>9</sup>. Det er dog udenfor nærværende projekts formål at forsøge at beskrive og sammenligne disse teorier, men de nævnes her dels fordi IRT er grundlag under ”assessment”-forskningen og at teorien betragtes som en grundlæggende forudsætning for computerbaserede adaptive test<sup>10</sup>. Det er også på dette område at psykometrikere sætter sin lid til udviklingen indenfor test udviklingen:

Adaptive testning seems such an obvious way forward that it is surprising that it has not become more widely used. There are several studies that have shown it to be more successful than classical techniques.....Reasons for the technique not becoming universally popular may include the lack of software, the sophistication needed to interpret the item response theory

---

<sup>9</sup> kilde: [http://en.wikipedia.org/wiki/item\\_response\\_theory](http://en.wikipedia.org/wiki/item_response_theory)

<sup>10</sup> som fodnote 9

models involved, and the lack of faith of educational and psychological professionals in models they cannot understand (Rust & Golombok 1999:187).

Vi vender i kapitel 6 tilbage til adaptive test.

Haywood et al. nævner Vygotsky og Feurstein som personer, der har haft afgørende indflydelse på at forandre den traditionelle tilgang til testning og vurdering. "Both Vygotsky and Feurstein have responded to social need and cultural-historical changes for assessing human cognitive potential rather than assessing only contemporary performance" (Haywood 1992:7). Det nævnes også hos Haywood et al., at der har været forsøg på at kombinere psykometriske principper med ideen om at fastsætte (assess) zonen for den nærmeste udvikling (ibid:70). Denne tilgang forfølges i kapitel 6.

### ***3.2 Folkeskolens Nationale Test***

Som beskrevet i afsnit 2 er Folkeskolens obligatoriske nationale test en del af et større initiativ fra undervisningsministeriet på evalueringsområdet som bygger på et politisk forlig og en ændring af folkeskoleloven (marts 2006). Der har været udtrykt store forventninger til evalueringsinitiativet hvor fx. Bertel Haarder udtaler at: "Vi træder helt nye stier med disse test. Det er første gang, at man laver elektroniske test, der ligesom en mundtlig prøve kan kortlægge en elevs niveau. De tester netop ikke paratviden, men færdigheder som fx at kunne bruge forholdsord rigtigt"<sup>11</sup>. Nogle af lovordene fra undervisningsministeriet er at testene er brugervenlige, meget driftssikre, valide og reliable, og kan give tilbagemeldinger på flere niveauer. Ifølge en pressemeddelelse fra 9. maj 2007, gik det dog galt med driftssikkerheden i dette tilfælde:

#### **DIGITALE AFGANGSPRØVER I GEOGRAFI I MORGEN AFLYST – ELEVERNE SKAL I STEDET TIL PRØVE PÅ PAPIR**

Det er andet år, at afgangsprøverne i biologi og geografi kan tages elektronisk. Sidste år var der ingen problemer. Og leverandøren bag, Cowi, har garanteret, at der ikke ville være problemer i år. Skolestyrelsen har derfor bedt leverandøren, Cowi, om en redegørelse for problemerne.<sup>12</sup>

Svend Kreiner, lektor i statistik ved Københavns Universitet, vurderer, at det omfattende projekt med udvikling af såkaldte adaptive test er særdeles ambitiøst, både fra en international og især fra

---

<sup>11</sup> [www.evaluering.uvm.dk](http://www.evaluering.uvm.dk)

<sup>12</sup> Pressemeddelelse fra Skolestyrelsen 9. maj 2007, <http://www.uvm.dk/07/problemer.htm?menuid=6410>

en dansk synsvinkel. Han forudser, at metoden vil vække international opmærksomhed blandt førende testteoretikere<sup>13</sup>.

Der har været, som det fremgik af afsnit 2.4, en heftig debat blandt danske skolefolk om indførelsen af test. Det følgende vil dog fokusere på den indholdsmæssige side af de nationale test, både hvad angår formål, opbygning, teoretisk baggrund. Det vil hovedsageligt være en beskrivende fremstilling, men nogle kritiske tvivlsspørgsmål vil også blive stillet.

### **Hvad er formålet med Folkeskolens nationale test?**

Ejerne af projektet, undervisningsministeriet, udtaler om formålet med de nationale test at de er "tænkt som et redskab til den formative eller mere processuelle evaluering – som et værktøj til at samle informationer og på den baggrund tilrettelægge en differentieret undervisning" (Margit Holm Nielsen<sup>14</sup>) og siger endvidere at "det er vigtigt, at læreren er bevidst om, hvilke områder testen ikke belyser, og som derfor må evalueres på en anden måde. Det gælder de aspekter af elevens udvikling, der vedrører den sociale og personlige udvikling, for eksempel selvstændighed, samarbejde og initiativ". Her foregribes en del af den skepsis som har været omkring testene, uden at det af den grund har nedtonet debatten. Testene ses her som et pædagogisk værktøj for læreren, som skal give et nuanceret billede af den enkelte elevs faglige styrker, svagheder og potentiale<sup>15</sup>.

Der påpeges flere steder af de ansvarshavende, at testene ikke må bruges til offentliggørelse eller rangordning af skoler og elever og at alle resultater er strengt fortrolige<sup>16</sup>. Samtidig er det alligevel meningen at landsresultater, landsgennemsnit og national præstationsprofil, skal offentliggøres. Desuden ligger det i planerne, at testopgaverne skal genanvendes med henblik på at give sammenlignelige resultater år efter år.

Hvad skal så testes? I folkeskolens bekendtgørelse er der beskrevet slutmål for fagene, som er langsigtede mål for kundskaber og færdigheder, og trinmål som er de kortsigtede på det enkelte klassetrin. Disse "angiver de nationale mål for, hvilke kundskaber og færdigheder eleverne forventes at have tilegnet sig ved henholdsvis afslutningen af undervisningen i et fag i folkeskolen og ved afslutningen af bestemte klassetrin" (Undervisningsministeriet 2006). Der er vedrørende det faglige indhold taget udgangspunkt i udvalgte trinmål, angående det berørte fag og klassetrin, som er testegnede. Eksempler på åbenlyse ikke egnede trinmål er mundtlige og sproglige mål. Om testegnet-

---

<sup>13</sup> [www.evaluering.uvm.dk](http://www.evaluering.uvm.dk)

<sup>14</sup> kilde: <http://evaluering.uvm.dk> "Test er et værktøj for lærerne"

<sup>15</sup> [www.skolestyrelsen.dk](http://www.skolestyrelsen.dk)

<sup>16</sup> Steen Harbild undervisningsministeriet

heden står der i kravspecifikationen, at ”De it-baserede test kan kun omfatte en del af fagets trinmål. For en dækkende vurdering af elevens kundskaber og færdigheder inden for alle fagets trinmål må yderligere evalueringsmetoder også tages i brug. Testen kan således ikke stå alene – heller ikke for de trinmål, som testen omfatter” (Undervisningsministeriet 2006:48). Der kan så drages tvivl om dette matcher med udtalelserne i starten om valid og reliabel vurdering af elevens standpunkt? Det er vel kun gældende indenfor det afgrænsede område som bliver testet og heller ikke mere.

Som led i opbygningen af testdesignet er der formuleret faglige profilområder. Der er blandt de involverede fagpersoner sket en udvælgelse af egnede trinmål til de syv it-baserede test og en fastlæggelse af profilområder, som har dannet det faglige grundlag for udviklingen af testopgaver. To forskellige profilområder i matematik kunne være Algebra, og Geometri.

### **Testopbygning**

Vi har ikke mulighed for at foretage en analyse af funktionaliteten, da vi ikke har adgang til testene. Derfor vil vi kun omtale funktioner og oplysninger, vi har fået andetsteds fra, hvilket gør at det er større mulig for at overse centrale ting ved systemet og intentionerne bag.

Testene er selvscorende, således at lærerne ikke selv skal rette dem, men får leveret og fortolket resultaterne. Eleven får ikke testscoren efter gennemførelse. Læreren får en rapport som indeholder en profilbeskrivelse og testrapport for hver elev, og en testrapport og profiloversigt for klassen.

Der er 3 hoved testtyper: Multiple Choice, indsæt kort tekst, drag & drop.

Testene er adaptive hvilket vil sige at de undervejs tilpasses den enkelte elevs faglige niveau. Intentionerne er at give et mere udfordrende og effektivt testforløb for den enkelte elev. Der arbejdes med 5 forskellige sværhedsniveauer, hvor der lægges ud med et middelsvært spørgsmål. Hvis der så svares rigtig, kommer man henholdsvis videre til en lidt sværere eller lettere opgave afhængigt om der svares rigtigt eller forkert. Spørgsmålet er hvor grovkornet niveaudelingen er, da kun 5 forskellige niveauer ikke er særlig nøjagtig, men et spørgsmåls sværhedstal vil blive normeret efter en numerisk skala fra 0 til 100. Et spørgsmål med sværhedstallet 0 kan alle besvare og et med 100 kan ingen elev på klassetrinet besvare korrekt (Undervisningsministeriet 2006:65). Der må således blive tale om ret mange testopgaver til hvert område.

Christen Sørensen (Sørensen 2006) påpeger at pga. at testene er baseret på multiple choice, er gættetsandsynligheden ved 5 svarmuligheder 20 %. Eleven kan således svare korrekt uden at kunne redegøre for svaret! Dette har indvirkning på testens reliabilitet, men reliabiliteten kan øges igen med et stort antal spørgsmål, som gør det usandsynligt at få en høj score som samlet resultat ved gætning. Sværhedstallet bruges til beregning af scoren og til opdeling af spørgsmål i sværhedsgrader. I



kravspecifikationen til testene nævnes også at ”Opgaverne, der skal indgå i testene, skal fordele sig omtrentlig symmetrisk med middelsværdighedstal på ca. 50. Klassificering sker på en 5-trinsskala, hvor opgaverne efter sværdhedstal fordeles i 5 grupper, således at opgaverne med sværdhedsniveauet 1 udgør de ca. 20% letteste (mål ved sværdhedstallet) og opgaverne med sværdhedsniveauet 5 udgør de ca. 20% sværeste” (Undervisningsministeriet 2006:65). Dette indebærer at sværdhedsgraden (1-5) kan ændres, uden at sværdhedsniveauet ændres. Samtidigt med, at der i testene benyttes en opdeling i profilområder i fagene, kan dette indebære mere nuancerede resultater end bare en sværdhedsgrad fra 1-5.

### **Teoretisk baggrund for testene**

Folkeskolens nationale test er baseret på forskning samt nationale og internationale erfaringer<sup>17</sup>, hvor der både har været involveret it-folk, designere, statistikere og pædagogisk/fagligt forankrede institutioner (DPU, Bergen Universitet). Michael Lund-Larsen fra @ventures i Århus som har været med i konsortiet bag udviklingen af testen<sup>18</sup> nævner, nogle læringsteoretikere som har været inde i billedet ved udvikling af test i deres regi og sandsynligvis også ved de nationale test. Disse er Bloom, Dreyfus & Dreyfus, Klafki og Vygotsky. I kapitel 4 vil vi i forbindelse med vores faglige test komme ind på Bloom. Vi vil desuden i det perspektiverende kapitel 6 komme med et bud på én måde at fortolke de faglige IKT-baserede test ud fra Vygotskys udviklingszone (Nuzo) til benyttelse i pædagogisk/faglig sammenhæng.

### **Kritik af testformen**

Vi har tidligere været inde på, at et grundlæggende spørgsmål i forbindelse med udvikling af faglige test er: Hvilke dele af faget kan evalueres med test og IKT? Lisa Steinmuller fra @ventures har en opstilling af hvad hun mener der kan afvikles ved en test og hvad der ikke kan. Hun definerer test som en række af spørgsmål stillet ud fra objektive kriterier, hvor de tilhørende svar kan enten være rigtige eller forkerte:

---

<sup>17</sup> Jakob Wandall skolestyrelsen

<sup>18</sup> Foredrag ved UFO 2006 i Odense

<ul style="list-style-type: none"> <li>• Hvad kan ikke testes             <ul style="list-style-type: none"> <li>– argumentation</li> <li>– fritekst</li> <li>– analyser</li> <li>– forsøg</li> <li>– mundtlighed</li> <li>– ikke målbar viden</li> <li>– sociale kompetencer</li> <li>– arbejdsformer</li> <li>– kreative løsninger</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Hvad kan testes             <ul style="list-style-type: none"> <li>– paratviden</li> <li>– faktuelle oplysninger</li> <li>– nødvendig basisviden</li> <li>– grammatik</li> <li>– sætningskonstruktioner</li> <li>– visse former for bevisførelse</li> <li>– aflæsning af oversigter</li> <li>– stavning</li> <li>– formler</li> <li>– regne og matematikstykker</li> </ul> </li> </ul>
---	---

Fig. 3.1 fra Lisa Steinmuller<sup>19</sup>

Vi har ovenfor henvist til Bertel Haarders bud. Men der er tale om et ”paratvidens testsystem” siger Christen Sørensen (Sørensen 2006:2), hvor han betvivler den grundlæggende betragtning, eller måske rettere grundlaget, bag beslutningsprocessen om, at testene skulle medføre et højere fagligt niveau, idet klarheden omkring formålet med undervisningen ikke er tydeligt nok. Han har analyseret på, hvordan testudformningen vil påvirke læringsituationen i folkeskolen. Det er dog vigtigt at påpege, at Sørensens vinkel er på det IKT-baserede testdesign og ikke på læring. Desuden skal det pointeres, at hans analyse bygger på kravspecifikationen<sup>20</sup> for designet og ikke på den endelige testudformning. Han siger at det drejer sig om paratviden fordi 50-60 spørgsmål skal besvares på 45 minutter, og fordi der er tale om multiple choice spørgsmål, foruden at it-mediet lægger op til hurtig respons<sup>21</sup> (ibid:4). Muligheden for fordybelse er yderst begrænset, hvilket peger i modsat retning af fundering og refleksion, som efter vores mening er læringsmæssige tilstræbelsesværdige forudsætninger i den pædagogiske praksis<sup>22</sup>. ”Tilhængere” peger, i modsætning til Christen Sørensen, på at testene skal ses som et led i den formative evaluering. Han rejser her et grundlæggende spørgsmål som han ikke forfølger i resten af fremstillingen, til gengæld fokuserer han på en analyse af grundstrukturen i it-testsystemet. Vi vil afslutningsvis i specialet konkludere på denne problemstilling.

<sup>19</sup> [http://www.fluid.dk/arrangementer/LAER271005/Test%20i%20ungdomsudd\\_Fluid\\_Del\\_271005.pps](http://www.fluid.dk/arrangementer/LAER271005/Test%20i%20ungdomsudd_Fluid_Del_271005.pps)

<sup>20</sup> undervisningsministeriet 2006. Bilag 2. Kravspecifikation og testafvikling

<sup>21</sup> Det er uklart for os om det er muligt at undlade at svare på spørgsmål i testen

<sup>22</sup> Hvis man i testdesignet har spørgsmål med delspørsmål ’items’ er der lidt større mulighed for fundering.

### 3.3 Testværktøjet Quia

Vi har valgt at benytte det webbaserede testværktøj Quia til vores undersøgelser og faglige test. Der findes andre programmer som er tilgængelige på internettet, som dele af LMS-systemer (Frontier, Lectio etc). Foruden professionelle programmer som vi ikke umiddelbart kan få adgang til. Det er bl.a. blevet benyttet af mange sproglærere i gymnasiet.

#### Beskrivelse – generelt

Quia er et amerikansk online testværktøj, som er serverbaseret, hvor interaktionen foregår online. Programmet kræver licens som kan købes på hjemmesiden [www.Quia.com](http://www.Quia.com), men der er en fuldt funktionel 30-dages prøvelicens som er gratis.

Quia koster 49 \$ om året, men man kan starte med en prøveversion, og de ting, man laver i denne periode forsvinder ikke. Det er muligt for nogle kolleger at gå sammen om en konto, man kan oprette mapper i systemet, så kan prisen siges at være overkommelig. Desuden findes der skoler som har licens på systemet.

De vigtigste områder i Quia er:

Quiz programmer, Aktiviteter og øvelser, Spørgeskema, Klasseadministration

Her introducerer vi kort de forskellige områder. Vi vil særlig beskrive Quizzes og Surveys, idet det er specialets fokus. Aktiviteter og øvelser har dog nogle ekstra funktioner som java-appletter, som mangler i Quizzes.

#### Quiz-delen

Der opereres med begreberne Quizzes og Surveys i Quia-jargon. Der skelnes mellem Quizzes og Surveys på den måde, at den først type kræver et bestemt svar, således at når du opretter et spørgsmål skal du også opstille svarmuligheder og svarstype. Spørgsmålstyperne bliver gennemgået i det følgende:

Graduerede spørgsmål:

- multiple choice,
- true-false
- pop-up (i princippet som multiple choice, men som dropdownliste)
- multiple correct (forskelligt fra multiple choice, fordi respondenten kan vælge flere svar)

- fill-in (er velegnet til spørgsmål som kræver, korte, et- til to-ords svar, de kan være auto-graduerede, hvor der må præfabrikeres de rigtige svarmuligheder, eller manuelt-graduerede, hvor man efterfølgende selv skal tage stilling til om svaret er rigtigt eller forkert)
- initial answer (ofte brugt som et tekststreng med indbyggede fejlstavninger, som eleven så kan rette, kan auto- eller manuelt gradueres), short answer (åbent spørgsmål som ikke kan autogradueres), essay (åbent spørgsmål som ikke kan autogradueres),
- matching (to eller flere par af begreber som står i vilkårlig rækkefølge overfor hinanden, som så identificeres og matches korrekt, det er muligt at give automatisk partiel kredit, hvis man har nogle men ikke alle matches korrekte)
- ordering (sætte to eller flere begreber, som står vilkårligt, i rigtig rækkefølge)

### Ikke-graduerede spørgsmål:

- multiple select (samle meninger og demografisk information fra respondent)
- yes-no (etc)
- pop-up (etc)
- free response (frie svar eller meninger)

En tredje mulighed er åbne spørgsmål (med graduering af udsagn) eller rent åbne svarmuligheder. De forskellige spørgsmålstyper kan sammensættes uafhængigt af hinanden. Man kan tilføje billeder og lydseksempler i Quia. Desuden er det overordnet muligt at lave tilfældig rækkefølge på spørgsmålene, således at rækkefølgen bliver forskellig fra elev til elev, der kan vælges mellem anonymitet og navngivning, der kan vælges om respondenten løbende skal se sin score, der kan sættes kode på for at undgå uvedkommende respondenter, der er muligt at tilføje links i systemet. Endelig kan testen tidsbegrænses, så eleven kun har et bestemt antal minutter til at besvare testen, f.eks. 30 minutter.

### Survey-delen

Surveys er spørgeskemaer, som er velegnede til tilfredshedsundersøgelser og holdningsundersøgelser, hvor svaret ikke er defineret på forhånd. Denne type er velegnet til evaluering af undervisningen, og elevernes evaluering af egen indsats og standpunkt. Spørgeskemaerne kan være anonyme eller man kan bede eleverne skrive deres navn i det første felt. Spørgeskemaprogrammet tæller sva-

rene sammen, men der kan ikke laves avancerede beregninger på svarene. Det tager ikke ret lang tid at lave et spørgeskema. Nedenstående link viser et evalueringsskema anvendt efter et forløb i kultur og samfundsfag.

<http://www.Quia.com/sv/72701.html>, <http://www.Quia.com/sv/57388.html>

Når man er færdig med at konstruere et spørgeskema oprettes et unikt link som kan formidles til eleverne via konference eller mail. Det er muligt at angive et tidsinterval for, hvornår og i hvor lang tid en test kan besvares, eller det kan gøres frit, hvornår eleverne vil svare på det. Programmet kan udskrive forskellige data fra besvarede Surveys: F.eks. hvor mange respondenter har besvaret hvert spørgsmål (Question Summary), informationer og statistik om hvert enkelt spørgsmål (Question Details), hver enkelt respondents svar på hvert spørgsmål (View by respondent)

Der er 7 forskellige typer af spørgsmål til Surveys:

- **Multiple Select og Multiple Choice:** Der stilles et spørgsmål med to eller flere svarmuligheder. Ved denne type er det muligt at have flere korrekte svarmuligheder, som respondenteren kan krydse af.
- **Yes/No:** også kaldet true/false Her kan man stille et ja-nej spørgsmål, enten/eller enig/uenig etc.
- **Pop-up:** Som ved multiple choice, men figurerer som en drop-down list. Det særlig ved denne type er at den kan placeres midt i teksten således, at den kan bruges som fill-in ud fra prædefinerede svarmuligheder. Så undgår man at få svar som der ikke er taget højde for i forvejen.
- **Ranking:** Her sættes nogle attributter op og en karakteristisk på, hvordan de skal opstilles, hvorefter respondenteren skal ordne dem i rigtig rækkefølge.
- **Rating scale:** Respondenteren afgiver en talværdi som svar, hvor værdien svarer til en bestemt svarkategori. Vi har brugt denne type i holdningsundersøgelsen (se afsnit 4.3).
- **Rating grid:** Kan beskrives som flere rækker af rating-scale typen i én
- **Free response:** Respondenteren læser spørgsmålet og indtaster et svar i tekstfeltet. Det kan så prædefineres, om der skal være plads til ét ord, en sætning, eller et helt afsnit, afhængigt af spørgsmålstype

### **Aktivitets-delen**

Et meget væsentligt træk ved Quia er vidensdeling. Der findes et utal af delte aktiviteter (shared activities), som er udviklede forløb/eksempler som man kan importere til ens egen Quia konto og derefter redigere dem, så de passer til det, man har brug for. Der findes 16 typer af prædefinerede spil og læringsaktiviteter (se bilag 1)

### **Klasseadministration**

Når man beder eleverne svare på en quiz kan man nøjes med at bede dem skrive deres navn inden de starter, herefter kan man som lærer identificere elevernes svar. Man kan desuden samle test, øvelser mv. på en klassehjemmeside. Se eksempel på nedenstående link:

[www.quia.com/pages/evadokit.html](http://www.quia.com/pages/evadokit.html)

Man kan også oprette klasser med navne og e-mailadresser og automatiske koder. Dette giver mulighed for at se, hvordan den enkelte elev har klaret sig i forskellige quiz'er, læreren kan give en personlig tilbagemelding. Eleven skal så logge sig på med koderne. Du kan se, hvor lang tid han/hun har brugt på quiz'en.

### **Databehandling**

En af fordelene ved elektroniske programmer til test og spørgeskemaundersøgelser er, at databehandling og rapportering er automatiseret så læreren undgår det manuelle arbejde med efterbehandlingen, sammentælling, strukturering og arkivering af data. Quia har en omfattende rapporteringsdel, hvor der kan trækkes data på enkeltpersoner, og statistik, grade respondenterne etc.

I programmet er der automatisk optælling af rigtige/forkerte svar, statistik for det samlede og det enkelte spørgsmål, hvor du får et godt overblik på skærmen.

Der er også muligt at eksportere data fra systemet til regnearks programmer. Desuden kan der i systemet oprettes en elektronisk klasse, med tilgængelige testresultater for den enkelte elev og klassen.

## ***3.4 Psykometriske principper***

Psykometriens indflydelse i samfundet er større end vi umiddelbart tror, på godt og ondt. Mange undersøgelser i dagligdagen bygger ikke på psykometriske principper. Et eksempel på en test fra dagligdagen som man kan problematisere er "integrationstesten" i Danmark hvor Simon Emil Ammitzbøll udtaler, at "Det siger ikke noget om, hvor dansk man er, hvis man kan svare på de her spørgsmål. Hele prøven er gennemsyret af en nationalromantisk forestilling om det at være dansk.

Man er hverken mere eller mindre dansk, hvis man kan svare på, hvornår kvindelandsholdet vandt VM i håndbold”<sup>23</sup>. Testen bygger på paratviden, idet den består af 35 spørgsmål fra en bank på 200 som er frit tilgængelige på nettet. Man kan med rette spørge hvorfor denne test? og hvad kan den bruges til? Medierne refererer ofte, efter vores påstand, meget ukritisk om undersøgelser, hvor modtageren ikke har mulighed for at vurdere om disse undersøgelser bygger på velbegrundede kriterier og pålidelige oplysninger. Mediehistorier af denne karakter har stor indflydelse i det offentlige rum, men også i undervisningssituationer har de test og vurderinger underviseren foretager ret stor indflydelse på den enkelte elevs tilværelse. Det er derfor vores opfattelse, at hvis man vælger at benytte sig af test, så skal det ske på et så retfærdigt og præcist grundlag som muligt. Det vil sige man bør følge nogle procedurer og principper, som gør testudformningen og konsekvenserne ved benyttelse af denne, til en reflekteret proces, og ikke blot forlade sig på et intuitivt testdesign. Det ansvar, der ligger i at benytte test kan ikke uddelegeres til et computerprogram.

Vi vil i det efterfølgende beskrive og forklare psykometriske principper, relevans for undervisningsfeltet og den empiri vi indsamler. Fremstillingen skal derfor ikke betragtes som en fyldestgørende beskrivelse af området, men som udvalgte nedslag.

Vores indfaldsvinkel til psykometriske test er funktionel, i den forstand at den er bredere end traditionel psykometri, som var under stor indflydelse at intelligens test “bevægelsen” (omtalt i afsnit 3.1). Den funktionelle betragtning på psykometrien gør, at vi kan bruge teorier fra andre områder såsom interview-teori med lukkede og halvåbne spørgeskemaer. (Rust & Golombok 1999:34). Vi betragter således distinktionen mellem spørgeskema og test som flydende.

Når vi i vores tilfælde vil lave tests som en del af den formative evaluering, er der ikke de samme krav til de bagvedliggende principper som ved mere formelle test. Torben Spanget Christensen formulerer det på denne måde: “...vi skal stille lavere metodekrav til spørgeskemaundersøgelser, lærere laver i deres egen undervisning. Og de skal alligevel ikke udsættes for statistiske analyser, idet der som regel er tale om optælling (alle eleverne er med i undersøgelsen) og ikke om stikprøver” (Christensen 2006: 379).

I psykometrien skelner man mellem objektive test som er dem med faste utvetydige svar, hvor der enten er rigtige eller forkerte forhåndsdefinerede svar som true/false eller multiple choice. I åbne essaysvar kan man ikke udelade det subjektive element i vurderingen. (Rust & Golombok 1999:47) Der skelnes i psykometrien mellem norm-referentielle og kriterie-referentielle test. Vores faglige tests er kriterie-referentielle, idet vi sammenholder elevernes formåen med objektiv definerede kriterier (fagmål), og ikke ud fra sammenligning med større befolkningsgrupper, hvilket kendetegner

---

<sup>23</sup> <http://avisen.dk>: Dansktest hindrer god integration (27. februar 2007)

norm-referentielle test (Rust & Golombok 1999:40ff). Folkeskolens nationale test kan siges at indeholde begge elementer da sværhedsgraden (og deraf testscoren) er defineret ud fra de tidligere nævnte fagmål og hvad man kan forvente (ved at gennemføre omfattende pilottestning af en større gruppe elever,) at de formår på deres klassetrin.

Man kan opdele spørgeskemaer/test i vidensbaserede, som dækker f.eks. færdigheder, og kundskaber og personbaserede som undersøger bl.a. personlighedsrelaterede forhold og holdninger (Rust & Golombok 1999:196). I afsnit 4 beskriver vi de videns- og personbaserede undersøgelser (faglige test og læringsstiltest) vi foretager.

Nedenfor har vi udarbejdet et eksempel på en fremgangsmåde i forbindelse med oprettelse/design af test/spørgeskema, som bygger på en opstilling fra (Rust & Golombok 1999:196ff).

Man skal klargøre for sig selv, hvad formålet med spørgeskemaet klart og præcist er. Det foreslås også at man bruger en test specifikation, kaldet blueprint. Et konstrueret eksempel kunne være en faglig test i matematik:

	Elevens kendskab til fag- termer	Elevens forståelse af emnet	Eleven kan overføre sin viden til rela- terede områder
Algebra	5	2	1
Geometri	2	3	x

Blueprintet kan i simpel form bestå af fx. fagområder og taksonomiske niveauer, som er med til at klargøre hvilke områder af faget man vil dække, hvordan det enkelte område skal vægtes med antal spørgsmål (items) for hvert område og niveau. Ved at systematisere på denne måde får man mulighed for at se, hvor eleven har stærke/svage sider og kan sætte fokus på indlæring af bestemte mangler i elevens faglige viden.

Der ligger et stort arbejde i at skrive items (der er forskellige typer items som er nærmere beskrevet i afsnit 4.2). Vi mener ikke at ja/nej svarmuligheden er egnet til andet end paratviden, og her er der desuden 50% mulighed for at gætte det rigtige svar. I multiple choice typen bruger man svarmulighederne korrekt og så et antal såkaldte distractors, typisk 4-5 svarmuligheder. Det kræver et vist arbejde at lave sandsynlige distractors, således at de ikke udelukkes på forhånd af eleven.

Man skal også overveje baggrundsinformationer (overskrift, navn, dato) og instruktioner eller vejledning til eleven, et eksempel på vejledning:

- For hvert item skal du vælge det rigtige svar blandt de 5 mulige svar, ABCD eller E. Læs spørgsmålene grundigt og afgør hvilken svarmulighed du vælger. Du vil, når du har svaret på al-



le spørgsmål få din samlede score i testen at vide - i forhold til det maksimale som er 100 point. Husk at skrive dit navn på testen. Det er kun læreren som får adgang til resultatet.

Layout og design vil have betydning og kan sandsynligvis påvirke elevens motivation og opfattelse af hvor seriøs testen skal betragtes.

Testscore: Det skal afgøres hvor meget hvert rigtigt svar er værd. Skal rigtige svar give 1 point, hver gang eller skal det afhænge af sværhedsgrad?

Oftest anbefales det at pilotteste undersøgelser ved en repræsentativ gruppe (Rust & Golombok 1999:199). Pilottestningen er grundlag for item analyser, hvor man analyserer de data man har fået gennem pilottesten. Analysen foretages blandt andet med henblik på at identificere 1) distractors som overhovedet ikke vælges i pilottesten og 2) items, hvor alle svarer korrekt. Disse tilfælde skal så vidt muligt ikke forekomme i den endelige test. (Rust & Golombok 1999:209ff).

### **Realibilitet og validitet**

Reliabilitet og validitet er begreber som ikke kun vedrører testning men alle former for dataindsamling eller dataproduktion.

“Reliability has been defined as the extent to which the test is effectively measuring anything at all, and validity as the extent to which the test is measuring what it is supposed to measure” (Rust & Golombok 1999:64)

Realibilitetet er afhængig af antallet af items, og herunder nævnes i punktform nogle forhold som har indflydelse på reliabiliteten:

- Tidspunkt for testning (dag og tid på dagen)
- Hvilket slags evaluering/vurderingstype der bruges (fx. antal svarmuligheder)
- Har vedkommende svaret rigtigt på et svært spørgsmål og forkert på et nemt spørgsmål som er relaterede?
- Flere items desto større reliabilitet. (afhænger også af kvaliteten)
- Variationer i sværhedsgraden styrker reliabiliteten

Validitet dækker spørgsmålet om vi måler det som vi intenderer (Rust & Golombok 1999:215). To underkategorier er “Face validity” som handler om udformning og designet og indholdsvaliditet, hvilket i vores faglige test indebærer en sammenligning mellem læseplanen mål og indholdsbeskrivelse og testens indhold/opbygning.

Det er selvfølgelig langt flere overvejelser man skal gøre sig. Herunder følger et eksempel på en tjekliste med nogle 'common-sense' overvejelser.

- Er spørgsmålene forståelige og gennemskuelige?
- Er spørgsmålene for tekstnære?
- Er der fare for fortolkningsproblemer?
- Kan spørgsmålene svares entydigt?
- Er der ironi, sarkasme eller humor i spørgsmålene?

### **3.5 Opsamling**

Vi har i dette kapitel behandlet test i evalueringsfeltet ud fra en historisk såvel som en aktuel synsvinkel. Vi har redegjort for det test-teorigrundlag som vi anvender i vores empiri. Vi har introduceret det IKT-værktøj vi anvender til indsamling af den del af vores empiri, der omhandler faglige test, Multiple Intelligenser, Læringsstile og holdningsundersøgelse. Vi har desuden problematiseret en række forhold vedrørende test generelt som vi vender tilbage til i de følgende afsnit.

## 4. Metode - teori

### 4.1. Multiple intelligenser og læringsstile

En af hypoteserne i afhandlingen bygger på den antagelse, at der er en sammenhæng imellem elevers præstationer i test og samme elevers intelligenser og læringsstile. Antagelsen er dog kompleks i den forstand, at elevernes præstationer i en konkret test må forventes, at bero på mindst to forhold, nemlig 1) deres intelligenser og læringsstile i forhold til den undervisning og læringsvirksomhed, der er gået forud for testseancen og 2) deres intelligenser og læringsstile som forudsætninger for at besvare selve testen. Vi har valgt at afgrænse os til alene at undersøge det sidste forhold. Ud fra den betragtning, at det er selve testmetoden med anvendelse af IKT-værktøjer og det deri liggende potentiale, der har vores interesse. I det følgende vil vi redegøre for det teoretiske grundlag for de intelligenser og læringsstile vi mener har betydning for elevers præstationer i faglige test. Vi vil i fremstillingen og i diskussionen i kapitel 6 også komme ind på de pædagogiske konsekvenser af inddragelse og anvendelse af viden om elevernes MI-præference og læringsstile.

#### 4.1.1 Teoretisk udgangspunkt – Multiple Intelligenser – Howard Gardner

Vi skriver intelligenser i flertal og har dermed allerede antydnet, at vi operer med et flertydigt intelligensbegreb eller mere præcist: de mange intelligenser (MI) med udgangspunkt i Howard Gardners teori. Der er tale om en kognitiv teori jf. nedenstående: "As the name indicates, we believe that human cognitive competence is better described in terms of a set of abilities, talents or mental skills, which we call intelligences" (Gardner 1993:15). Han beskrev oprindeligt 7 intelligenser, men der er siden tilføjet intelligenserne Naturalistisk og Eksistentiel intelligens. I det følgende opstilles de oprindelige 7 intelligenser:

Lingvistisk/sproglig:	Man er god til at behandle tegn/ord/begreber i både skrift og tale. Man har sans for lyde, ord, sproglige strukturer og sans for sprogs betydninger og funktioner.
Logisk/matematisk:	Man er god til at behandle tal og logiske mønstre. Man kan abstrahere og generalisere, deducere og inducere. Kan behandle lange årsagssammenhænge.
Spatial:	Man kan opfatte den visuelt rumlige verden: flader, farver, mønstre, afstande og rumlighed – og kan transformere sine perceptioner, handle i det rumlige
Kinæstetisk-kropslig:	Man har en fornemmelse af og kan kontrollere sin krop og dens bevægelser og dens håndtering af genstande.
Interpersonel:	Man kan skelne mellem og reagere hensigtsmæssigt på andre menneskers sindsstemninger, følelser, ønsker og motivationer. Man kan tage hensyn til andre.
Intrapersonel:	man kender sine egne følelser og motiver til at gøre det, man gør – og man kan arbejde med sig selv, sine følelser og sit intellekt.
Musikalsk:	Man er i stand til at adskille toner, høre rytmer, klange og lydmønstre, og man har sans for musikalsk udtryk – kan orientere sig lydligt i tid og rum. Man er i stand til selv at frembringe musik.

Fig. 4.1.1 (Opstilling efter Damberg et al 2006:344-345)

Ved den lejlighed beskrev han dem også som en teori. Men han tager dog også et forbehold, idet han konstaterer at "Even so, by featuring the word idea I want to underscore that the notion of multiple intelligences is hardly a proven scientific fact: it is, at most, an idea that has recently regained the right to be discussed seriously" (Gardner 1983:11). Mogens Hansen følger op ved at konstatere at Gardners teori "ikke kan leve op til de krav, der stilles til en traditionel teori om indre konsistens, hypoteser der kan afprøves og definitioner, der præcist afgrænser begreber..Der er derfor nok mere tale om et teori-udkast end en teori i sædvanlig positivistisk videnskabelig forstand" (Hansen 2005:35). Positivistisk videnskab indebærer, at der indsamles data til støtte for teorien ud fra strukturerede og kontrollerbare eksperimenter og/eller spørgeskemaer. De indsamlede data skal kunne registreres med sanserne dvs. måles, vejes, tælles el. lign. Holdninger, tro mv. er ikke gyldige data indenfor det positivistiske videnskabsparadigme (Pedersen & Land 2001:30). En teoris indre konsistens fremkommer ved, at der anvendes en metode, som giver det samme resultat når metoden anvendes på et andet objekt, eller samme objekt til et andet tidspunkt. Metoden må således ikke give anledning til indre modsigelser, men give et entydigt resultat for de, der anvender den (Dolin 2001)<sup>24</sup>. Desuden er det ofte indenfor psykologien vanskeligt, at afgrænse begreber, fordi det ikke lader sig gøre at fastlægge begrebsindholdet. Det gælder således også intelligens, bevidsthed og selvet som ikke har genstandskarakter og ikke entydigt kan afgrænses (Hansen 2005:36).

Ikke desto mindre fastholder Gardner betegnelsen teori. Det må siges, at være tankevækkende, at brede kredse i det videnskabelige samfund accepterer betegnelsen teori om noget som ikke lever op til rodfæstede videnskabelige normer. Det kan selvfølgelig have noget at gøre med, at der er stort ønske om, at slippe af med den klassiske endimensionelle intelligensopfattelse, som af mange opfattes som værende for snæver i sin vurdering af menneskers intelligens. En forklaring kan være, at teorien om MI har fået stor udbredelse i pædagogiske kredse og i undervisningssektoren, særligt i grundskolen (Hansen 2005:17). Ifølge Mogens Hansen er der i øjeblikket tale om et opbrud både pædagogisk og forskningsmæssigt, hvor to teorier kæmper om det gældende paradigme og måske bliver der på længere sigt tale om et paradigmeskift (Hansen 2005:17).

Grunden til at MI teorien har haft så stor gennemslagskraft i undervisningssektoren er, at den giver grundlag for en udviklingsoptimisme, forstået på den måde at vi kan tro på, at undervisning og læ-

---

<sup>24</sup> <http://pub.uvm.dk/2001/fysik/index.html>

ring nytter noget, i modsætning til den endimensionelle intelligens opfattelse, der siger at intelligensen er fastlagt genetisk og dermed er relativt uforanderlig gennem tiden (Hansen 2005:27ff.). Når vi skriver, at teorien har haft stor gennemslagskraft i undervisningssektoren, kræver det den tilføjelse, at teorien næppe har haft så stor betydning for udviklingen i gymnasieskolen som i grundskolen (Damberg 2006:351). En forklaring kan være at gymnasiet traditionelt har rekrutteret en mere homogen elevmasse end grundskolen, og de fremherskende intelligensformer har været de verbalsproglige og matematisk logiske. Behovet for at se bredere på intelligensproblematikken og de deraf afledte konsekvenser for pædagogikken, har derfor ikke været så påtrængende i gymnasiet. En anden forklaring kan være, at gymnasiet traditionelt og historisk har opfattet sig som først og fremmest en fag(videnskabelig) institution og i anden række en pædagogisk institution. Opmærksomheden på pædagogisk teori og metode har derfor også været derefter. Erik Damberg formulerer det på følgende måde: ”Måske er elevsammensætningen i de gymnasiale uddannelser væsentligt anderledes, end den er i grundskolen - men meget tyder på, at denne påstand ikke holder helt stik længere: Nogle gymnasieskoler arbejder med elitehold og andre skoler og kurser føler sig magtesløse i forhold til elevmaterialet og synes, at der er vanskeligt at få læreprocesserne til at foregå, som de burde, eller som de plejede. I alle tilfælde er de ’forskellige elever’ en opgave, de gymnasiale uddannelser ikke slipper for at tage stilling til” (Damberg 2006:351).

Den endimensionelle intelligens (g-faktoren) har stort set udelukkende hentet empirisk støtte fra standardiserede ”papir og blyant” test, som næsten udelukkende baserer sig på matematisk-logisk og sproglig intelligens. Et forhold som Gardner er meget kritisk overfor, idet han mener at intelligenserne er kontekstbundne og derfor vil de intellektuelle funktioner være af ringere kvalitet under dekontekstualiserede sammenhænge f.eks. i test situationer (Gardner 1993:31, Hansen 2005:28). Jean Lave definerer dekontekstualisering på denne måde: ”To dekontextualize knowledge is to formalize it (to contain it, pour it into forms) at a more inclusive level. To formalize is to contain more forms. It follows that abstraction from and generalization across ’context’ are mechanisms that are supposed to produce decontextualized (valuable, general) knowledge” (Lave 1993:23). Lave er kritisk overfor den position – det er næppe for meget sagt, at han afviser læring i den dekontekstualiserede form. Han siger videre, at “further, it is assumed that what is learned is of a general nature and powerful because it is not embedded in the particularities of specific practices. These assumptions are brought into question simultaneously once learning is conceived as situated practice..Also according to the authors of this in this book, The dualistic division and categorization of experience

that privileges ‘decontextualization’ does not offer an adequate explanation of thought or action” (ibid:24).

Vi er enige i, at læring i dekontekstualiseret form ofte medfører vanskeligheder for eleven/den lærende, fordi det kan være svært at se relationen til egen hverdag eller livsverden. På den anden side vil vi ud fra vores faglige position i biologi - henholdsvis musikfaget, forfægte det synspunkt, at formaliseret og generaliseret viden på metaniveau er et både frugtbart og meningsfuldt læringsmål. Et eksempel fra biologi: Viden om arveanlæg på et abstrakt formaliseret grundlag er ganske nyttig, hvis man skal kunne tage stilling til komplicerede spørgsmål vedr. moderne bioteknologi. Som følge af Gardners kritiske position i forhold til test under dekontekstualiserede sammenhænge foreslår han derfor, at test skal baseres på problemløsningsevner og handleevner på tværs af forskellige intelligensområder (Gardner 1993:31).

Spørgsmålet om den logisk-matematiske intelligens status i forhold til de øvrige intelligenser diskuteres indgående af Gardner uden, at han når frem til en endegyldig afklaring. I ‘Frames of mind’ fra 1983 siger han: ”one could conclude that logical-mathematical ability is not as ‘pure’ or ‘autonomous’ a system as others reviewed here, and perhaps should count not as a single intelligence but as some kind of supra- or more general intelligence” (Gardner 1983:159). Senere stiller han spørgsmålet: “Does it still make sense to group all of logic and mathematics together, as one form of intelligence, and to set it off sharply from other forms? Only time can tell whether the grouping that I have proposed here has long term validity” (ibid. s. 168).

Hansen behandler også problemstillingen og henviser til den amerikanske psykolog Rex Li, der peger på at ”der i Gardners beskrivelse af den logisk-matematiske intelligens, er tale om en generel funktion, hvis kerne består af begrebslig og abstrakt tænkning” (Hansen 2005:34ff.). Gardner erkender endvidere, at han ikke principielt vil afvise g-faktoren, men blot pege på, at dens basis er testopgaver af især verbalt-sproglig og logisk art og altid bygger på korte svar og ikke på mere omfattende udredninger eller problem- og handleforløb (Hansen 2005:35).

Vi opfatter problemstillingen som meget central i diskussionen om intelligens og intelligensformer og undersøgelser heraf. Hvis den logisk-matematiske intelligens har status af en generel funktion, er det vel en ganske rimelig begrundelse for, at de klassiske IQ tests tager udgangspunkt i denne. Det er selvfølgelig så stadig et udestående med den sproglige intelligens som øjensynlig ikke har nogen særstatus i forhold til andre intelligensformer. I forhold til nærværende projekt, er det selvsagt af

betydning, hvordan den logisk-matematisk intelligens betydning skal tolkes i forhold til de andre intelligensformer, selvom det ikke er projektets formål at foretage intelligens test.

Fibæk Laursen påpeger at ”Gardner interesserer sig for, hvordan hjernen forarbejder information, mens de konkrete omstændigheder, hvorunder man lærer ikke interesserer ham. Han siger f.eks. at vores hjerne forarbejder sprog som sprog, uanset det kommer ind af øret (det talte sprog), af øjet (skriftsproget) eller via fingerspidserne (hvis man læser blindskrift). Det forarbejdes på samme måde, af samme intelligens og i samme center i hjernen” (Fibæk Laursen 2005:64).

### **Gardner om tests**

Gardner påpeger betydningen af den kulturelle kontekst i forbindelse med testning og testudformningen, hvor vurderingsmaterialet er afhængigt af den kulturelle sammenhæng. I ’Den intelligente skole’ (Gardner 1999) nævner han at ”Næsten alle formelle test forudsætter, at deres brugere er fortrolige med kulturens avancerede, boglige symbolsystemer. Disse test frembyder derfor særlige vanskeligheder for enkelte, som af en eller anden grund har haft svært ved at tilegne sig det avancerede symbolkendskab eller ikke kan få dette kendskab til at passe sammen med tidligere former for mentale forestillinger” (Gardner 1999:199). Han lægger også op til et skift i fokus væk fra standardiserede test hen imod vurdering (assessment) og problemløsning, hvilket også er mere i tråd med en skandinavisk tradition. Han plæderer også for mere sammenhæng mellem vurderingen og persons daglige aktiviteter og omgivelser (Gardner 1993:31-32).

#### **4.1.2 Teoretisk udgangspunkt læringsstile – Dunn & Dunn**

##### **Læringsstile generelt**

Rita Dunn nævner i en artikel<sup>25</sup>, at der er mindst fem forskellige måder at undervise på, foruden forelæsning og diskussioner, som er velegnede til ny og vanskelig information. Hun har sammen med andre forsket i ’Learning Style Inventories’, hvor de udvikler undervisningspakker der tiltaler forskellige typer af elever. Det minder ret meget om Gardners efterfølgende udtalelse, hvor han siger at hans forskning har antydnet, at ”man kan gå til ethvert godt, givende emne – ethvert begreb, der egner sig til undervisning – på mindst fem forskellige måder, som groft sagt tager højde for de mange intelligenser” (Gardner 1999:237). Han nævner disse indgange: fortællingsindgang, logisk-kvantitativ indgang, grundlagsorienteret indgang, æstetisk indgang, empirisk indgang (ibid:238-239). Han belyser den ’effektive’ lærer som den engagerede formidler, der varierer undervisningen

---

<sup>25</sup> artiklen ”The Dunn and Dunn Learning-Style Model and Its Theoretical Cornerstone”

ved at inddrage diverse pædagogiske hjælpemidler – tekster, film, software – under hensyntagen til elevernes indlæringsstil, hvilket bl.a. leder tankerne over på undervisningsdifferentiering (ibid:239). Der er således fælles træk mellem teorierne men også forskelligheder, som vi vil tage op i et efterfølgende afsnit. Kort kan der siges her, at Gardner ikke har beskæftiget sig direkte med hvordan og under hvilke omstændigheder folk lærer, hvilket har været fokus hos Dunn & Dunn.

En læringsstil kan beskrives som en sammensætning af kognitive, affektive og fysiske faktorer som er karakteristiske for en person og indikerer hvordan vedkommende opfatter, interagerer og reagerer på og i læringsmiljøet (**Bates & Leary:45**). I dansk sprogbrug bruges også begrebet indlæringsstil om en persons foretrukne måde at lære på eller at processere information på. Læringsstilen kan siges at have sammenfald med personlighedstræk hos personen og påvirker dennes holdninger, værdier og den sociale interaktion.

Der er megen forskning indenfor læringsstile og mange forskellige retninger også. Vi har i vores uddannelse arbejdet med Kolb, Honey & Mumford og Dunn & Dunn, men der er mange flere. I en rapport af Coffields et al 2004 (her efter Mayes), er der en klassificering af læringsstilmodeller, hvor Dunn & Dunn modellen kaldes “Konstitutionelt-baserede læringsstile og præferencer” (Mayes:9). Den kunne også kaldes undervisningspræference, idet den fokuserer meget på undervisningsmiljøet. Der siges endvidere om Dunn & Dunn, at selv om de har et stort og omfattende forskningsprogram, påpeger de kraftigt en mangel på uafhængig forskning og begrænsninger i de empiriske undersøgelser.

Kolb (1984) var en af de første der udviklede en decideret teori om læringsstile. Den er baseret på den kognitive del af læringsprocessen. Coffield et al påpeger problemer vedrørende reliabilitet, validitet og selve læringscirklen hos Kolb (Mayes:10). Dette nævnes også hos Bates & Leary, at Kolbs model “lacks psychometric rigour and empirical studies indicate a lack of verifiability using the measures in a pilot study” (Bates & Leary:46). Der sættes også spørgsmålstegn ved om modellen er en beskrivelse af en læringscyklus eller individuelle forskelle hos den lærende.

Honey & Mumford som bygger på Kolb er blevet benyttet i virksomhedsregi, men også i uddannelsessammenhæng (Malberg 2003).

### **Dunn & Dunn**

Selv om Gardner ikke fokuserer på, hvordan man lærer eller under hvilke omstændigheder man lærer, så interesserer det i høj grad Dunn & Dunn. Deres teorier handler om måder at lære på (Dunn 2006:2, Fibæk Laursen 2005:62, Damberg et al 2006:349). ”Grundantagelsen er, at vi har forskellige præferencer for, hvordan vi gerne vil lære og hun hævder med baggrund i en række empiriske



undersøgelser, at vi lærer mere, når vi får mulighed for at arbejde i overensstemmelse med vore præferencer” (Fibæk Laursen 2005:62).

Gardner og Dunn er enige om, at vi intellektuelt og læringsmæssigt er kvalitativt forskellige, og at der ikke er nogen bestemt intelligensprofil eller læringsstil, der er bedre end andre (Fibæk Laursen 2005:62). Man kan måske sige, at Gardners intelligens teori er blevet operationaliseret gennem Dunn og Dunns læringsstile og det kan måske være en forklaring på den store udbredelse teorierne har fået bl.a. i grundskolen?

I en artikel<sup>26</sup> opstiller Rita Dunn 5 teoretiske ‘hjørnestein’ i deres læringsstilmodel, som kunne kaldes hypoteser, hvoraf 3 af dem er: 1) forskellige undervisnings omgivelser, uv-ressourcer og tilgange svarer til forskellige læringsstil styrker 2) individuelle undervisnings præferencer eksisterer og kan måles pålideligt 3) med tilsvarende omgivelser, ressourcer, og tilgange, kan elever statistisk opnå en højere færdigheds- og holdnings-test score når disse er i overensstemmende, heller end i uoverensstemmende med deres respektive præferencer (efter Dunn 2006:2).

Ligesom for Gardners teori gælder det om Dunns, at man ikke kan betragte den som empirisk bevist, selvom der dog er en del empirisk støtte til teorien (Fibæk Laursen 2005:63). Man kan blandt andet kritisere Dunn & Dunn for, at den empiri de underbygger deres teori med stammer fra test, der i overvejende grad bygger på viden, som befinder sig på de nederste trin i Blooms taksonomi. Det vil vi jo i hvert fald ikke Danmark betragte som et fuldstændigt udtryk for et individs intellekt (Hildebrand Jensen et al 2005).

Dunn & Dunns læringsstil indeholder mange elementer, som kan kategoriseres under fem forskellige stimuli:

Miljømæssige	lyd	lys	temperatur	indretning	komfort	
Emotionelle	motivation	udholdenhed	ansvarsfølelse	struktur		
Sociologiske	solo	parvis	jævnaldrende	teams	voksne	varieret
Fysiologiske	perceptuelle	indtagelse	tid	mobilitet		
Psykologiske	globalt	analytisk	hemisfærisk	impulsiv	reflektiv	

Der er talen om fem hovedstimuleringer som opererer med forskellige elementer:

de fysiske omgivelser (lyd, lys, varme, indretning, komfort), det følelsesmæssige (motivation, udholdenhed, ansvarsfølelse, behovet for struktur), det sociologiske (solo/team), det fysiologiske/-perceptuelle (visuelt, auditivt, taktilt, kinæstetisk) og psykologiske (global/analytisk, impulsiv/-reflektiv). Disse har betydelig indvirkning på hvordan individer lærer. Derudover er der endnu flere elementer, idet flere af elementerne er opdelt i underkategorier. F.eks. er det perceptuelle element,

---

<sup>26</sup> ”The Dunn and Dunn Learning-Style Model and Its Theoretical Cornerstone”

som hører under den fysiologiske præference, opdelt i 4 underkategorier: visuel, auditiv, taktil, og kinæstetisk. Det kan virke noget uoverskueligt at der er så mange elementer der indgår i læringsstil, men den enkelte person er kun påvirket af en del af disse.

De psykologiske elementer analytisk og holistisk, i modellen, læner sig op ad kognitiv læringsteori. Piaget taler om figurativ (det umiddelbart oplevede/perceptuerede) og operativ (sammenhæng, struktur, væsen) erkendelse, hvor det første griber om helheden, det sidste om elementer eller analysen (Hermansen 2001:44). Bruner bruger begrebet intuitiv tænkning som er holistisk, den forholder sig til helheden, i modsætning til umiddelbar tænkning og analytisk tænkning som er tidslig og logisk, deduktiv, problemløsende og middelbar (ibid:50). Desuden er der tilnærmelser mellem områderne når Hermansen behandler Batesons læringsniveauer og siger at "I spændingsfeltet på læringsniveau 2 har vi nu fået placeret det kognitive arbejde med at gå fra den konkrete repræsentation (det billedlige) til den abstrakte repræsentation (det symbolske)" og endvidere at "læringsniveau 2 har således alt at gøre med det, der foregår eller burde foregå i de forskellige læringsrum vi har indrettet os med samfundsmæssigt" (ibid:148-149).

Hvis vi går tilbage til Dunn, har vi den analytiske type som arbejder med detaljer, procedurer, linært og med en logisk struktur, og så påstår Rita Dunn<sup>27</sup>, at selv om undervisningen for det meste er tilpasset den analytiske type, er majoriteten af elever på alle akademiske niveauer globale snarere end analytiske. De globale kan klare nyt og svært stof når det introduceres med en kort historie eller anekdote, som viser hvordan indholdet i emnet relaterer til deres liv/virkeligheden. Der skal tilføjes tegninger, suppleret med illustrationer, oversigter, figurer, tabeller og farver etc. Undervisningsforløb eller gennemgang af emner skal opdeles i korte bidder, fordi den globale elev har brug for (tænke)pauser til at internalisere stoffet i. Emnet skal gerne kunne introduceres taktil/visuelt eller kinæstetisk/visuelt hellere end auditivt/visuelt og de skal kunne vælge hvor, hvornår, med hvem, med hvilket/hvad, og hvordan de må arbejde i klasserummet (Dunn note 27). Dette er ikke fremmede tanker, idet det er gængs pædagogisk praksis at påpege behovet for variation af undervisning og uv-materiale og tilgange til emner. Nøglebegreber som, at undervisningen skal være varieret, engagerende og vedkommende er vanlige.

### **Kritik af teorierne**

Fibæk Laursen er skeptisk overfor Dunn og Dunns læringsstile, i hvert fald har han mere tiltro til Gardners teori. Han er bl.a. kritisk overfor om alle de læringsstilelementer de operer med er lige

---

<sup>27</sup> [www.learningstyles.net](http://www.learningstyles.net) "The Dunn and Dunn Learning-Style Model and Its Theoretical Cornerstone" af Rita Dunn

gyldige, f.eks. forekommer det tvivlsomt om de miljømæssige forhold skulle have nogen væsentlig indflydelse på vores læringsvirksomhed. Desuden frygter han, at disse mange læringsstile kan give anledning til en vidtgående individualisering af undervisningen som vil gå udover de kvaliteter der ligger i fællesskabet. Her er der også en forskel mellem Gardner og Dunn. Dunn interesserer sig ”kun” for, hvordan der læres i modsætning til Gardner, der også interesserer sig for hvad der læres (Fibæk Laursen 2005:65).

Bricheno & Younger<sup>28</sup> referer til en undersøgelse, hvor der er foretaget en VAK- (Visual Auditiv Kinæstetisk) undersøgelse, hvorefter der er foretaget interviews af de samme elever, hvor der ikke var en klar sammenhæng mellem data fra spørgeskema og interviews af samme elever om deres foretrukne læringsstil (Bricheno & Younger:9). Dette kunne tyde på at test og en efterfølgende kategorisering kunne være i uoverensstemmelse med elevernes læringspræference. Dette er forhold som man ikke kan ignorere. De kritiserer også MI forskningen og nævner at “there is little independent empirical research which supports the wide-ranging claims made by protagonists of accelerated learning and multiple intelligences” (ibid:3).

Coffield et al<sup>29</sup> går så langt som at sige om Dunn & Dunns model at “Despite a large and evolving research programme, forceful claims made for impact are questionable because of limitations in many of the supporting studies and the lack of independent research on the model. Concerns raised in our review need to be addressed before use is made of the model in the UK” og at “In many ways, the use of different inventories of learning styles has acquired an unexamined life of its own, where the notion of learning styles itself and the various means to measure it, are accepted without question” (Bricheno & Younger:2-3).

Der er en vis mismatch mellem den popularitet og den udbredelse som læringsstil-teorierne har fået hos praktikere og så kritikken fra pædagogiske forskningskredse, hvor der bl.a. påpeges mangel på uafhængige empiriske studier. Forskningen og litteraturen indenfor læringsstilområdet er omfattende og disse udtalelser tyder på, at der i praksis sker en forenkling af et komplekst område, og at en del af antagelserne fra læringsstilteoretikere og praktikere kan kaldes noget tvivlsomme.

En anden grundlæggende kritik som bliver rejst hos Terry Mayes<sup>30</sup> som går på selve undersøgelsesformen er, at “All of these are based on self-reports of preferences in learning. The validity of these rests largely on the assumption that learners have insight into, and awareness of, their own learning processes.” (Mayes:9). Dette rammer faktisk hele læringsstilsbranchen, som i høj grad bygger på afkrydsning i spørgeskemaer, hvor personen skal udtale sig om sin egne læringspræferencer. Disse

---

<sup>28</sup> online artikel se litteraturliste

<sup>29</sup> her efter Bricheno & Younger

<sup>30</sup> Terry Mayes Stage2 LearningStyles.pdf

overvejelser er vigtige at have med, når man vil i gang med at lave sine egne undersøgelser på lokalt lærerplan. Kritikken retter sig dog mod de professionelle undersøgelser indenfor området.

#### 4.1.3 Sammenligning af teorierne

Teorierne om Multiple Intelligenser og Dunn & Dunns Læringsstile er forskellige vedrørende spørgsmålet om, hvordan vi lærer. Indenfor læringsstil-teorierne påpeges det, at ikke alle lærer intuitivt og at mange elever har brug for struktur og supervision, men i MI-teorien antydes at læring er intuitivt (Dunn et al 2001).

Men der er forfattere som mener at bestemte intelligenser som lingvistisk/sproglig, visuel-spatial og kropslig-kinæstetisk har fælles karakteristika med forskellige perceptuelle styrker, nemlig de visuelle, auditive, kinæstetiske og taktile, og at disse intelligenser skal henvendes eller adresseres til den tilsvarende perceptuelle styrke. Ved denne fremgangsmåde får vi en kobling mellem teorierne. Den efterfølgende sammenligning bygger på (Chau 2006<sup>31</sup>) og er et bud på en sammenligning af perceptuelle styrker i læringsstile og tilsvarende MI-intelligenser.

Hvis man sammenligner den visuelle/spatiale læringsstil og spatial intelligens, så lærer den **visuelt-spatiale lærende** holistisk hellere end skridt for skridt. Den visuelle forestillingsevne er vigtig i læreprocessen. Lineær sekventiel tænkning er særlig svær for denne type. Grafiske værktøjer spiller en stor rolle for at facilitere denne type. Den **spatiale intelligens** er som nævnt evnen til at tænke rumligt, i billeder, og evnen at opfatte den synlige verden på en præcis måde, og kunne transformere sine perceptioner billedligt i hovedet eller på papir (er typisk arkitekter, kunstnere, designere eller skulptører). Til sammenligning er det gældende for begge typer at de er visuelt orienterede og kan se informationer for det indre øjne og kan tænke abstrakt.

Vedrørende den auditive læringsstil og lingvistisk intelligens, så lærer den **auditive lærende** bedst ved at høre, og forstår og husker mundlig formidlet information godt. De har gode sproglige færdigheder og er gode til at formulere ideer tydeligt og får stort udbytte ud af verbalt/sprogligt-orienterede/formidlede undervisningssituationer. Den **lingvistiske intelligens** indebærer god sans for lyde, det talte og skrevne ord, evnen at lære sprog, sproglige strukturer, betydninger og funktioner, og at bruge sproget som udtryksmiddel (er typisk skribenter, poeter, advokater, foredragsholdere). Det er gældende for begge typer, at de foretrækker at bruge sproget i læreprocesser og husker godt hvad de hører.

I en sammenligning af den taktile-kinæstetiske læringsstil og kropslig intelligens, skal det bemærkes, at Dunn & Dunn skelner mellem taktile og kinæstetiske lærende, hvilket Gardner ikke gør. Den

---

<sup>31</sup> <http://libres.curtin.edu.au/libres16n1/Chau.htm>

**taktile lærende** er den som bruger fingre og hænder for at koncentrere sig og husker ved at skrive, tegne kruseduller og bevæge fingrene. Den **kinæstetiske lærende** lærer bedst ved en kombination af taktile og kinæstetiske erfaringer, eksperimenteren, lave ting og være involveret. **Den kropslig/kinæstetiske intelligens** kendetegnes ved brugen af de mentale evner til at koordinere og kontrollere krop og bevægelser og dens håndtering af genstande. Det kendetegner begge typer, at de relaterer kropslig bevægelse til læring, hvor enten hele eller dele af kroppen bliver brugt.

Vi har i det foregående set, at der er visse fællestræk mellem teorierne, men også forskelle, hvor det grundlæggende nok er forskellen i fokus på de 'indre' intelligenser og så de 'synlige' læringsstile og disses samspil med omgivelserne.

Vi finder det spændende at forfølge tanken om hvordan disse teorier rent praktisk kan benyttes, hvordan finder vi ud af hvilke præferencer den enkelte elev har? hvordan kan vi bruge denne viden? og hvad er i den praktiske undervisningssituation forskellen på disse to teorier? Teorigrundlaget skitseret ovenfor danner således udgangspunkt for de empiriske undersøgelser af MI og læringsstile som vi foretager i dette projekt. I det følgende afsnit redegør vi for, hvordan vi har indsamlet denne empiri.

#### **4.1.4 Udformning og gennemførelse af undersøgelse – MI og læringsstil**

Ved undersøgelserne af elevernes læringspræferencer har vi benyttet spørgeskemaer. I forberedelsesfasen har vi brugt internet-tilgængelige ressourcer som grundlag for den indholdsmæssige udformning af undersøgelsen. Disse er opstillet og annoteret i bilag 2. En analyse af disse har dannet udgangspunkt for arbejdet med det endelige indhold i spørgeskemaerne.

##### **MI-undersøgelsen (MI)**

Med baggrund i de 7 intelligensformer beskrevet i afsnit 4.1.1 undersøger vi, i hvilket omfang de forskellige intelligensformer forekommer hos eleverne (MI-undersøgelse).

##### **Udformning**

I den endelige MI-undersøgelse benyttede vi 42 spørgsmål. (se spørgsmålene i bilag 4).

Spørgeskemaet indledes med en kort tekst:

Der er 42 spørgsmål i denne undersøgelse. Venligst svar på dem alle sammen efter bedste overbevisning. Husk at der ingen rigtige eller forkerte svar er.
--

Spørgsmålene er opstillet som udsagn som eleven skal erklære sig enig eller uenig i:

1. Bøger er meget betydningsfulde for mig.

- Enig
- Uenig

Dette spørgsmål gik på den lingvistiske/sproglige intelligens, mens det følgende går på den spatiale/visuelle:

36. Jeg kan godt se klare visuelle billeder for mig selv når jeg lukker øjnene.

- Enig
- Uenig

Der er så konstrueret 6 spørgsmål i hver intelligenskategori. I nogle af de undersøgelser som vi har undersøgt var der op til 18 spørgsmål pr. kategori, hvilket ikke var muligt for os, idet vi havde begrænset tid hos eleverne. Vi havde tilsammen 90 minutter til den faglige test og MI- og LS- undersøgelserne. Det har desværre stor betydning for reliabiliteten af undersøgelsen, hvilket vi kommer tilbage til i kapitel 5. Vi overvejede at bruge 5 svarmuligheder fra 'meget enig' til 'meget uenig', men gik fra det idet 'meget enig' og 'enig' alligevel dækker nogenlunde det samme og kan analyseres som betydningsmæssigt ens. Nogle persontyper svarer mere på de ekstreme svarmuligheder og andre på de mere neutrale. Vi har desuden forsøgt at gøre spørgsmålene så korte og præcise som muligt.

### **Gennemførelse**

Vi lavede en pilotundersøgelse i to klasser, hvor vi lavede en papirbaseret undersøgelse med 10 elever i alt med 35 spørgsmål og svarmulighederne sand, falsk, blank. 72 elever har deltaget i den endelige undersøgelse, der er foretaget ikke-anonymt. Vi har brug for at kende respondentens identitet, med henblik på at kunne sammenholde data fra forskellige dele af vores empiri. Spørgsmålene er administreret elektronisk ved hjælp af Quia.

### **Læringsstil-undersøgelse (LS)**

Vi fokuserer i undersøgelsen på hvilke psykologiske og perceptuelle præferencer eleverne foretrækker når de vil lære eller bearbejde informationer. Vi undersøger om eleverne er globale eller analytiske og auditive, visuelle, taktile eller kinæstetiske. De 2 sidste kategorier har vi valgt at lægge sammen i en taktil/kinæstetisk kategori (herefter benævnt kinæstetisk). Vi har valgt at fokusere på de elementer i Dunn & Dunn's model som har med kognition og perception at gøre, idet vi mener, at det er disse elementer der er de mest centrale i dette projekt og danner det bedste sammen-

ligningsgrundlag med MI-undersøgelsen. Det er også de elementer som vi IKT-mæssigt har mere direkte indflydelse på, desuden er de andre elementer i modellen også relateret til disse udvalgte kategorier.

Undersøgelsen er således opdelt i to undersøgelser i én. De første 26 spørgsmål (13 i hver kategori) undersøger om eleverne er holistiske/analytiske (HA-undersøgelse) og de sidste 36 (12 i hver kategori) om de er visuelle/auditive/kinæstetiske (VAK-undersøgelse). Vi har således 62 spørgsmål i alt (se spørgsmålene i bilag 3). Disse undersøgelser kan så fortolkes i samme ramme, hvor vi har 6 kombinationsmulighederne:

Holistisk/Visuel	Holistisk/Auditiv	Holistisk/Kinæstetisk
Analytisk/Visuel	Analytisk/Auditiv	Analytisk/Kinæstetisk

For eksempel vil den holistisk/kinæstetiske elev have sværere ved at sidde stille samme sted i længere tid og arbejde selvstændigt med opgaveløsning end en analytisk elev med en anden perceptuel præference fx. Holistisk/Visuel.

I HA-undersøgelsen er svarmulighederne som i MI-undersøgelsen. Et spørgsmål som går på det holistiske fra undersøgelsen er:

15. Jeg foretrækker at lave lektier med baggrundsmusik.

- Enig
- Uenig

Et eksempel på et spørgsmål fra VAK-undersøgelsen som går på det auditive:

53. Jeg husker bedst noget når jeg får det fortalt, heller end selv at læse om det.

- Enig
- Uenig

Vi er begge enige om at det har været svært at skelne mellem spørgsmålene til MI-undersøgelsen og LS-undersøgelsen og vi har taget os selv i at bruge spørgsmål som er nærmest identiske i begge undersøgelser. Problematikken illustreres med følgende spørgsmål, der stammer fra HA-undersøgelsen, men som også vedrører den logisk/matematiske intelligens og dermed MI-undersøgelsen:

25. Jeg bruger mest logik og ser på fakta, tal og data for at hjælpe mig til at tage beslutninger hellere end følelse og fornemmelse for tingene.
- Enig
  - Uenig

Det er fra HA-undersøgelsen, men går i særdeleshed også på den logisk/matematiske intelligens.

### **Gennemførelse**

LS-undersøgelsen er også pilot-testet. Spørgsmålene er administreret ikke-anonymt, idet vi har brug for at kende respondentens identitet, med henblik på at kunne sammenholde data fra forskellige dele af vores empiri. Spørgsmålene er administreret elektronisk ved hjælp af Quia. 72 elever har deltaget i undersøgelsen.



## 4.2 Faglige Test

### 4.2.1 Målbeskrivelse

Som redegjort for tidligere, er empirien indsamlet i gymnasieklasser der undervises i henholdsvis biologi og musik. Nedenfor redegøres for den gymnasiekontekst og det fagdidaktiske grundlag, der danner grundlag for de faglige test som vi har afprøvet i de to fag. Uanset hvordan man evaluerer, må evalueringen selvsagt betragtes i lyset af de mål man har med undervisningen og dermed også de mål, der overordnet er formuleret for det almene gymnasium (stx). Vi vil derfor indlede afsnittet med, at beskrive det overordnede mål for gymnasiet efter gymnasireformen. Dernæst redegør vi for de fagdidaktiske mål som gælder for henholdsvis musik og biologi.

Det almendannende mål har traditionelt og historisk været konstituerende for det almene gymnasium. Men dannelsesbegrebet er et dynamisk begreb som oprindeligt var inspireret af tysk dannelses-tænkning og som har dannet grundlag for den såkaldte didaktiktradition<sup>32</sup>, der har været fremherskende i de nordiske lande i modsætning til de angelsaksiske landes curriculumtradition<sup>33</sup> (Krogh 2006:187). Harry Haue opfatter dog dannelse og almendannelse som forskellige begreber og knytter almendannelse til undervisning: ”Almendannelsen er knyttet snævert til den organiserede undervisning og er således noget andet end dannelse. Almendannelsen er et begrænset projekt, der knytter sig til et vist antal skoleår, mens dannelse er noget, der kan arbejdes med hele livet” (Haue 2004:9). Efterfølgende redegør han for, hvor og hvornår det udfoldes: ”Almendannelsen kan udvikles i en undervisning, der omfatter de almene dele af de videnskaber og fag som samfundet har brug for med henblik på at udvikle elevernes personlige myndighed til at reflektere over deres eget forhold til andre medmennesker, natur og samfund” (2004:9ff). I modsætning her til ser Klafki ikke nogen skel mellem dannelse og almen dannelse. Han ser ”almendannelse som dannelse for alle” (Klafki 2002:69). Klafkis forståelse af almendannelse bygger på 3 evner som hvert enkelt menneske stræber frem mod og tager ansvaret for: a) selvbestemmelsesevne, b) medbestemmelsesevne og c) solidaritetsevne (ibid:69) og han skærper sin forståelse yderligere ved at fastslå, at ”Almendannelse må netop i dag som modvægt mod de nye afpolitiseringsbestrebelse, der dukker op, opfattes også som

---

<sup>32</sup> En central skikkelse i denne tradition er Wolfgang Klafki, der udviklede begrebet kategorial dannelse (Hiim og Hippe 1997:50ff) og senere videreudviklede han begrebet til den kritisk konstruktive didaktik (Klafki 2002:101ff)

<sup>33</sup> Traditionen er teoretisk inspireret af John Dewey i praksis udmøntes den i, at indholdet som det fremtræder i lærebøger og undervisning præsenteres som objektive strukturer og ”undervisningen opfattes som en samfundsmæssig serviceydelse, der måles gennem test af elevpræstationer” (Krogh 2006:188).

politisk dannelse til en aktiv deltagelse i udformningen af den demokratiseringsproces, som skal fremmes yderligere” (ibid:55). Så vidt vi forstår Klafki, ser han derfor ikke almindelse som noget der kun finder sted et vist antal år i en bestemt uddannelseskontekst, men derimod noget, der finder sted som en livslang bestræbelse. Men det er samtidigt klart, at skolen/uddannelsessektoren har meget stor betydning for, hvor godt det lykkes det enkelte menneske at blive (almen)dannet.

Uanset disse forskellige opfattelser af dannelsesbegrebet ser det ud som Haues forståelse af dannelsesbegrebet har været inspirationskilde i forbindelse med gymnasireformens tilblivelse. Af den politiske aftaletekst om gymnasireformen fremgår det: ”Almindelse og viden kan ikke skilles ad i de gymnasiale uddannelser. Viden skal give almindelse et indhold, og almindelse skal placere viden i en sammenhæng, som viden om de enkelte fagområder i sig selv ikke giver” (Undervisningsministeriet 2003:1). Almindelse skal altså opnås ved fagsamarbejde - samspillet mellem fagene skal styrkes – hvor fagene bidrager med deres vidensindhold. Desuden fremgår det, at naturvidenskaben nu skal have en mere fremtrædende placering i almindelsen: ”Traditionelt er almindelsen overvejende humanistisk og samfundsvidenskabelig i sin orientering, og begge tilgange skal fastholdes og uddybes. Det er imidlertid afgørende for fremtidens demokratiske beslutningsprocesser, at borgerne får en øget forståelse for den naturvidenskabelige og teknologiske udvikling. Derfor skal dette aspekt have en styrket position i dannelsesdimensionen i de gymnasiale uddannelser” (Undervisningsministeriet 2003:2).

Konkret sikres dette ved:

- indførelse af det nye forløb, almen studieforbereelse, hvor der lægges vægt på, at også naturvidenskabelige dannelseselementer indgår som en væsentlig baggrund for en sammenhængende videnskabelig og historisk forståelse.
- indførelse af et naturvidenskabeligt forløb i grundforløbet med fagelementer fra biologi, fysik, naturgeografi og kemi.
- øgede muligheder for samspil mellem de naturvidenskabelige fag. Forsøg med særlige naturvidenskabelige klasser og med fagpakker inden for matematik, fysik og kemi har vist, at et øget samspil på fagenes høje niveauer styrker mulighederne for faglig fordybelse.
- bedre muligheder for, at særligt interesserede elever kan tone deres uddannelse i naturvidenskabelig retning.

(Undervisningsministeriet 2003:9ff).

Almendannelsen er altså stadig konstituerende for stx, men med reformen er der kommet mere til: Af gymnasieloven<sup>34</sup> fremgår det af §2 stk.2 at: ”Formålet med uddannelsen er at forberede eleverne til videregående uddannelse, herunder at de tilegner sig almindannelse, viden og kompetencer gennem uddannelsens kombination af faglig bredde og dybde og gennem samspillet mellem fagene”. Hvilket må tolkes på den måde at viden og kompetencer er sidestillet med almindannelsen. Yderligere nævnes det i § 3: ”Uddannelsen til studentereksamen gennemføres med fokus på det almindannende og studieforberevende. Fagligheden er nært forbundet med sider af videnskabsfagene, og eleverne skal opnå almindannelse og studiekompetence inden for humaniora, naturvidenskab og samfundsvidenskab med henblik på at kunne gennemføre videregående uddannelse”. Det kan tolkes på den måde, at fagligheden er det begreb, der skal knytte almindannelse, viden og kompetencer sammen. Dolin (2006) har konstrueret en model, der beskriver sammenhængen mellem begreberne på følgende måde:

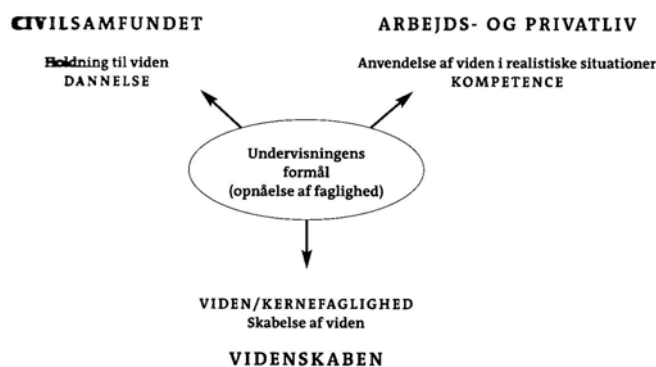


Fig. fra Dolin 2006:76ff

Imidlertid er der et problem med at skelne begreberne fra hinanden, et forhold som Dolin også diskuterer. Problemet opstår fordi lovteksten er uklar med hensyn til at definere begreberne. Hvordan er for eksempel sammenhængen mellem viden og faglighed? Dolins skelnen mellem kernefaglighed og faglighed er heller ikke helt tydelig.

I praksis er de overordnede mål blevet operationaliseret i de enkelte fag på den måde, at undervisningen skal beskrives og dokumenteres med hensyn til faglige mål, faglige kompetencer og personlige kompetencer. Evalueringen, hvordan den end foregår, skal derfor forholde sig til disse mål.

<sup>34</sup> Fra webadressen: <http://147.29.40.90/DELFIN/HTML/A2004/0009530.htm#K3> Undervisningsministeriet 2004

#### 4.2.2 Faglige mål, vidensniveauer og taksonomi.

I bekendtgørelserne for fagene, står der hvilke mål der er med undervisningen i det enkelte fag. Disse mål er oftest beskrevet i store og brede vendinger, således at det er op til den enkelte lærer at vælge hvordan han vil opfylde disse overordnede mål. Faglige mål<sup>35</sup> kan operationaliseres af den enkelte lærer som udsagn/læringsmål om hvad vi forventer at eleven har lært efter at have deltaget i et undervisningsforløb.

#### Beat og offbeat

Vi vil her tage et konstrueret eksempel på et udspecificeret fagligt mål fra musikteorien. Vores læringsmål er, at efter det gennemførte forløb antager vi at: Eleven kan skelne mellem beat og offbeat. Dette læringsmål kan så foregå/indlæres på forskellig vis. Det kan indlæres kinæstetisk/motorisk gennem rytmeøvelser hvor eleven kropsligt "føler" forskellen og kan udføre/klappe beat og offbeat, det kan foregå gennem notationssystemer (noder eller andre grafiske fremstillinger), hvor man kan inddrage logisk/matematiske forklaringsmodeller, eller auditivt præsenteret ved lydseksempler eller imitation, eller sprogligt/visuelt forklaret i tekstform med illustrationer. Desuden ligger der forud for denne kundskab/færdighed visse faglige forudsætninger som kan være nødvendige for at kunne mestre dette snævre mål, alt efter hvilken forståelsesform/kognitive processer der ønskes. Disse kunne være: puls, tempo, betoning, rytmeværdier, nodekendskab.

I musik stræbes der ofte efter at eleven opnår en musikalsk erkendelse (svært definerbar), som både er en begrebslig og kropslig erkendelse således at talen er om både symbolsk viden<sup>36</sup> og kropslig viden<sup>37</sup>. Dette bringer os hen til læringsmålet igen. Når vi i en test ønsker at undersøge om eleven kan skelne mellem beat og offbeat, bliver det afgørende hvordan vi ønsker at eleven skal ræsonnere sig frem til svaret. I det følgende kommer der flere metoder, der undersøger de samme begreber fra forskellige vinkler.

Ønsker vi et sprogligt udsagn spørger vi "Hvad er forskellen mellem beat og offbeat?"

Elevens svar kunne så være: "Beat er de rytmiske betoninger som ligger på 1/4 slaget i takten og offbeat er de ubetonede slag på 1/8 niveau som ligger på og-slagene i takten"

Ønsker vi at eleven bruger et notationssystem (forudsætning at eleven kan læse noder) kan vi spørge: "Hvilke noder i eksemplet ligger på beatet og hvilke på offbeat?"

Eller hvis vi ønsker at eleven selv noterer et nodeeksempel: "Skriv 2 takter med noder hvor du bruger henholdsvis 4 noder på beatet og 4 på offbeat"

---

<sup>35</sup> I angelsaksisk litteratur bruges begreber som objectives og claims om faglige mål

<sup>36</sup> Gardner om symbolsystemer (Den intelligente skole – Gardner i praksis s.198)

<sup>37</sup> I betydningen tavs viden

Her skifter vi fra spørgeformen til en instruks, men det som vi undersøger er det samme, bortset fra at eleven her i større grad skal tage stilling til hvordan han skal udforme sit svar. Kravet til forudsætningerne er større hvor han nu skal kunne skrive noder, ikke bare som afskrift, men skal selv finde på hvordan svaret skal udformes.

Ønsker vi at undersøge om eleven auditivt kan beherske det opstillede faglige mål kan vi bruge et lyd eksempel: “Lyt til eksemplet og afgør hvor mange klap ligger på beatet og hvor mange der ligger offbeat”

Hertil kommer også at selve lyd eksemplet kan have ret mange forskellige sværhedsgrader.

Et sidste eksempel er at undersøge om eleven kan udføre det faglige mål kropsligt (klappe med puls, spille et instrument solo eller i et ensemble): “Du skal nu (når vi har sat tempoet og talt til fire) klappe alle offbeat på 1/8 dels niveau i 2 takter”

Alle eksemplerne er faglige elevvurderinger på forskellige niveauer og illustrerer en problematik med det at isolere ‘items’ som man kan teste på. I testsammenhæng er det nødvendigt med entydige læringsmål for at kunne afgrænse/udskille det som man vil teste så klart som muligt af hensyn til validitet. Hvis du ikke har helt klart definerede læringsmål vil testspørgsmålet kunne fortolkes på forskellige måder som man ikke har taget højde for i designet af testen. Det kan synes omstændeligt at skulle formulere for hver enkelt aktivitet, men det har som sagt betydning for validiteten af test, og der også faglige grunde til at der er en tydelig sammenhæng mellem faglige mål og undervisningsaktivitet.

De 4 efterfølgende eksempler handler om samme læringsmål som vi har udepenlet, men der er nuanceforskelle mellem udtalelserne, som har betydning for både sværhedsgraden, hvordan det skal læres og hvordan det skal evalueres:

Eleven forstår forskellen på beat og offbeat.

Eleven kan skelne mellem beat og offbeat.

Eleven kan beskrive forskellen mellem beat og offbeat.

Eleven kan høre og afgøre hvilke rytmer er beat og offbeat.

### **Blooms Taksonomi**

En måde at kvalificere operationaliseringen af de faglige mål er at tage udgangspunkt i en taksonomi, som i uddannelsessammenhæng er en opstilling af målkategorier for undervisningen (Damgaard et al 2006:332). Blooms kognitive taksonomi er blevet brugt siden sidst i 1960’erne og “Pædago-

gisk forskning har gennem de sidste 40 år bekræftet taksonomiens gyldighed hvad angår hierarkis beskrivelse af pædagogiske mål, undtagen hvad angår de to sidste trin” (ibid:335). Der har været diskussion om hvilke af de kognitive processer vurdering eller syntese opererer på et højere plan og desuden har det i praksis været svært at skelne mellem niveauerne forståelse, anvendelse og analyse. Ofte ses en praksis hvor Blooms model forsimples til tre niveauer nemlig: beskrivelse, forklaring og vurdering (ibid:335). Det kan måske hævdes at Blooms taksonomi har nogle behavioristiske kendetegn, idet den er udviklet med henblik på at opstille så klare og snævre faglige læringsmål og ofte bliver brugt sådan, men taksonomien rummer også muligheder for eleverne at lære og tilegne sig viden og færdigheder på forskellige vidensniveauer, som ikke er enten rigtige eller forkerte. Denne diskussion forfølges i kapitel 6.

Vi vil nu komme med et bud på hvordan man kan kategorisere/formulere faglige mål i forbindelse med opstilling af disse i sværhedsgrader, således at disse kan placeres taksonomisk.

I forlængelse af eksemplet i det forrige afsnit kan der udledes at “samme” faglige mål kan stilles på forskellig vis og på forskellige niveauer. En mindre omskrivning af spørgsmålene kunne se således ud:

Ex.1. Beskriv forskellen mellem beat og offbeat (taksonomisk niveau 1)

Ex.2 Bestem hvilke noder i eksemplet ligger på beatet og hvilke på offbeat (taksonomisk niveau 2)

Ex.3 Demonstrer at du kan det faglige mål ved at skrive 2 takter med noder hvor du bruger henholdsvis 4 noder på beatet og 4 på offbeat (taksonomisk niveau 5—6)

Ex.4 Afgør/undersøg hvor mange klap der ligger på beatet og hvor mange der ligger offbeat ved at lytte til eksemplet (taksonomisk niveau 4)

Ex.5 Overfør din teoretiske viden om det faglige mål/begreberne til praksis ved at du demonstrerer (når vi har sat tempoet og talt til fire) at du kan klappe alle offbeat på 1/8 dels niveau i 1 takt og derefter skifte til beatslag i næste takt etc. (taksonomisk niveau 3,4,6)

Disse verber hører til forskellige taksonomiske niveauer i Blooms taksonomi, og beskriver forskellige typer af kognitive processer særlig at huske, forstå, anvende.

I 2001 kom der en revideret udgave af Blooms taksonomi (Anderson et al 2001<sup>38</sup>), hvor forskergruppen tilføjede en yderligere dimension til taksonomien, foruden ændringer i rækkefølgen og begrebsapparatet. Figuren nedenfor viser et enkelt begreb som eksempel for hver kategori:

### Bloom's Taksonomi (revurderet 2001)

	Kognitive processer					
Vidensniveauer	Huske	Forstå	Anvende	Analysere	Evaluerer	Skabe
Faktuel viden	opremse	sammenfatte	Klassificere	strukturere	rangere	kombinere
Konceptuel viden	beskrive	udlægge/forklare	eksperimentere	forklare	vurdere	planlægge
Procedural viden	tabulere	forudsige	Beregne	adskille	konkludere	komponere
Metakognitiv	passende brug	udføre	Konstruere	udføre	handling	virkeliggøre

Fig. 4.2 (efter Anderson et al 2001)

De tilføjede 4 vidensniveauer, faktuel, konceptuel, procedural og metakognitiv viden. Denne dimension differentierer mellem disse niveauer således at man på de forskellige kognitive taksonomiske niveauer kan bevæge sig på forskellige vidensplaner som operationaliseres ved at man bruger forskellige 'kognitive' verber for at beskrive et taksonomisk niveau afhængigt af hvilket videnstype talen er om. I det følgende er der for en ordens skyld en gennemgang af den reviderede taksonomi, med vægt på den 'nye' dimension.

Anderson et al's taksonomi består af to grundlæggende dimensioner. Den første kaldes vidensdomænet og indbefatter 4 typer af viden: faktuel, konceptuel, procedural, og metakognitiv. Faktuel viden omfatter 'basic elements students must know to be acquainted with a disciplin or solve a problem in it' (Anderson et al 2001:29). Konceptuel viden omfatter 'the interrelationships among the basic elements within a larger structure that enable them to function together' (ibid:29). Procedural viden omhandler 'how to do something, methods of inquiry, and criteria for using skills, algorithms, techniques, and methods' (ibid:29). Metakognitiv viden omhandler 'knowledge of cognition in general as well as awareness and knowledge of one's own cognition' (ibid:29). Den anden dimension kaldes domænet for kognitive processer og den omfatter seks måder at tænke/ræsonere på. Den er tæt på den klassiske Bloom version. At erindre/huske (remembering) omfatter at drage 'relevant knowledge from long-term memory' frem (ibid:31). Niveau 2 at forstå omfatter det at konstruere 'meaning from instructional messages, including oral, written, and graphic communication' (ibid:31). At anvende (applying) omfatter at udføre eller bruge 'a procedure in a given situation' (ibid:31). At analysere omfatter at dele et emne eller lignende i 'constituent parts' og afgøre 'how

<sup>38</sup> Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of educational objectives: Complete edition*, New York : Longman.

parts relate to one another and to an overall structure or purpose' (ibid:31). At evaluere omfatter at tage 'judgements based on criteria and standards' (ibid:31). At skabe (create) omfatter at sætte 'elements together to form a coherent or functional whole' og at reorganisere 'elements into a new pattern or structure' (ibid:31).

I bilag 5 er der opstillet en mere fyldig taksonomisk oversigt med flere begreber under den enkelte kategori og uddybning af kategoriernes indholdsmæssige betydning.

Også i den danske litteratur antydes der en lignende opdeling, dog ikke i forbindelse med Bloom. I Gymnasie-pædagogik (Dolin 2006:73-75) omhandles begrebet faglighed, hvor der tales om forskellige dimensioner henholdsvis en indholds-, kognitiv og metakognitiv dimension. Her står også (ibid:74-75) at "faglighed i skolemæssige sammenhæng vil kunne defineres som bestående af følgende dimensioner":

- indholdsdimension (at lære fagets indhold og kultur)  
faktaforståelse, procedureforståelse, kontekstforståelse
- Kognitiv dimension (at lære at ræsonnere inden for faget)  
Huske, viderebearbejde, omstrukturere
- Metakognitiv dimension (at lære at lære faget)  
fokusere opmærksomhed, bevidsthed/viden om læreprocesser i faget, kontrollere sine læreprocesser i faget

(Dolin 2006:74-75)

Vi har valgt at introducere den reviderede version af Bloom, selv om den ikke har været benyttet i vores faglige kredse, fordi den er interessant på den led at læringsmål kan placeres noget mere præcis end tidligere, hvilket vel at mærke også går det noget sværere at placere læringsmålene. Vi har benyttet den som eksempel i afsnittet om den faglige test i musik (kap.2.3.4). I kapitel 6 vil vi tage taksonomien op i et diskuterende afsnit.

#### 4.2.3 Gennemførelse – faglige test biologi

Den faglige test i biologi tager udgangspunkt i et undervisningsforløb, hvor temaet er "Ernæring og fordøjelse". Forløbet er beskrevet i skemaet nedenfor:

Fag: Biologi	Titel: Ernæring-fordøjelse	Kommentar til udførelsen
Mål og indhold	Forståelse for opbygningen og funktionen af: Proteiner, kulhydrater, fedtstoffer, vitaminer og mineraler.	



	Eleverne skal opnå kendskab til fordøjelsessystemets opbygning og funktion Dataindsamling og bearbejdning af data, anvende naturvidenskabeligt sprog, indhente og vurdere naturvidenskabelig viden fra forskellige kilder. Overføre teoretisk viden til praksis	
Faglige kompetencer og metoder	Fagligt arbejde på de nederste taksonomiske niveauer: Beskrivelse og redegørelse, faglig metode	
Studiekompetencer	Notatteknik, elementære studieteknikker Skriftlig fremstilling/formidling Laboratoriemetodik,	
Arbejdsformer	Klasseundervisning og gruppearbejde, herunder matrixgrupper, laboratorieøvelse	
Materiale	Lærebog: "Ram på biologien", Nucleus 1999, Side 11-22 Tabeller over energiforbrug for voksne og børn. Den nye kostpyramide (Samvirke, ) Øvelser: Bestemmelse af fedtindhold i en fødevarer (NV) Bestemmelse af en Mc. Donald menu (NV)	
Evaluering	Mundtlig formidling, hurtig skrivning	Afleveres i uge 4

Forløbsbeskrivelsen er udarbejdet af læreren som underviser de to klasser vi undersøger. Det skal for en god ordens skyld bemærkes, at undervisningsbeskrivelsen er udarbejdet inden det blev besluttet at afprøve testen i dette forløb og i denne klasse, derfor er der ikke nævnt noget om test under evaluering.

Den faglige test vi afprøver, er således designet med henblik at evaluere elevernes udbytte af undervisningen i dette tema, under skyldig hensyntagen til mål og indhold. Det fremgår af afsnittet om faglige kompetencer og metoder, at der hovedsagligt arbejdes på de nederste taksonomiske niveauer. Dog kan det også ses af afsnittet mål og indhold, at eleverne skal kunne vurdere naturvidenskabelig viden fra forskellige kilder og overføre teoretisk viden til praksis. Testdesignet er derfor indrettet på, at tage hensyn til disse betingelser. Testen består af 15 spørgsmål og afprøves i en prototype (bilag 6) i en klasse som har gennemført det samme undervisningsforløb, men som ikke indgår i den egentlige test.

Inden den egentlige test afvikles en prototype i to klasser. Prototypen afvikles i fronters prøvefacilitet. Da fronter bruges som LMS system på Gymnasiet har det nogle praktiske fordele at anvende dette prøveværktøj. I 6 af spørgsmålene skal svaret gives som afkrydsningsboks (multiple respos), 3 spørgsmål som radioknapper (multiple choice), 2 spørgsmål som matricer, 3 spørgsmål som nedtrækningsfelt (rullegardin), 1 som tekstfelt (fill-in). Spørgsmålene kommer i tilfældig orden. Som nævnt ovenfor er testen designet med baggrund i mål og indhold for temaet.

Nedenstående skema viser, hvordan de enkelte spørgsmål passer til mål og taksonomiske niveauer.

Spørgsmål	Mål	Taksonomisk niveau
1	Overføre teoretisk viden til praksis.	Paratviden, forståelse, analyse, vurdering
2	Forståelse for opbygningen og funktionen af: kulhydrater.	Beskrivelse, paratviden
3	Forståelse for opbygningen og funktionen af: fedtstoffer.	Beskrivelse, paratviden
4	Forståelse for opbygningen og funktionen af: fedtstoffer.	Beskrivelse, paratviden
5	Forståelse for opbygningen og funktionen af: fedtstoffer.	Beskrivelse, paratviden
6	Forståelse for opbygningen og funktionen af: kulhydrater. Eleverne skal opnå kendskab til fordøjelsessystemets opbygning og funktion.	Beskrivelse, paratviden, forståelse, analyse
7	Eleverne skal opnå kendskab til fordøjelsessystemets opbygning og funktion.	Beskrivelse, paratviden, forståelse
8	Eleverne skal opnå kendskab til fordøjelsessystemets opbygning og funktion. Overføre teoretisk viden til praksis	Paratviden, forståelse, analyse, vurdering
9	Forståelse for opbygningen og funktionen af: proteiner	Beskrivelse, paratviden
10	Overføre teoretisk viden til praksis	Beskrivelse, paratviden, forståelse
11	Forståelse for opbygningen og funktionen af: Proteiner, kulhydrater, fedtstoffer. Overføre teoretisk viden til praksis	Forståelse, analyse, vurdering
12	Eleverne skal opnå kendskab til fordøjelsessystemets opbygning og funktion.	Beskrivelse, paratviden,
13	Eleverne skal opnå kendskab til fordøjelsessystemets opbygning og funktion.	Beskrivelse, paratviden
14	Eleverne skal opnå kendskab til fordøjelsessystemets opbygning og funktion. Overføre teoretisk viden til praksis. Indhente og vurdere naturvidenskabelig viden fra forskellige kilder.	Paratviden, forståelse, analyse, vurdering
15	Eleverne skal opnå kendskab til fordøjelsessystemets opbygning og funktion.	Beskrivelse, paratviden

25 elever besvarede testen og efterfølgende bliver 3 spørgsmål taget ud og erstattet af 3 nye i den endelige test. De tre spørgsmål, der tages ud er: 3. Hvor meget energi er der i 1 gram fedt? 4. Triglycerider er opbygget af? og 10. Er der DNA (gener) i den mad vi spiser? Alle elever svarer rigtigt på sp. 3 og 24 elever svarer rigtigt på sp. 4 og 10. På den baggrund vurderes de at være for nemme at besvare. I Spørgsmål 6 indgår en animation i spørgsmålet. I spørgsmål 7, 13 og 14 indgår illustrationer i spørgsmålet.

Eleverne får en halv time til at besvare testen (brugergrænsefladen lukker efter 30 min.) og de har hjælpemidler til rådighed (lærebøger, noter mv.). Under afvikling af prototypetesten viser det sig uventet, at fronters prøve-brugergrænseflade ikke kan vise de illustrationer, der er lagt ind i de ovennævnte spørgsmål. Problemet afhjælpes efterfølgende ved at vise illustrationerne på en stor-skærm. På grund af dette problem besluttet det, at gennemføre den egentlige test i Quia.

### Den egentlige test

De 15 spørgsmål der indgår i testen er vist i bilag 7. Som nævnt ovenfor er der kommet tre nye til i forhold til prototypen:

Spørgsmål	Type	Antal point	Grafik
1	Multiple choice	1	
2	Matching	3	
3	Ordering	3	illustration
4	Pop-up	1	
5	Matching	3	
6	Multiple choice	1	Animation <sup>39</sup>
7	Multiple correct	3	illustration
8	Multiple correct	4	
9	Multiple choice	1	
10	Multiple correct	2	
11	Multiple correct	3	
12	Multiple correct	2	
13	Pop-up	1	illustration
14	Multiple correct	4	illustration
15	Multiple choice	1	

De tre nye spørgsmål kan beskrives med følgende mål, indhold og taksonomiske niveau:

Spørgsmål	Mål	Taksonomisk niveau
3	Forståelse for opbygningen og funktionen af: Proteiner, kulhydrater, fedtstoffer. Overføre teoretisk viden til praksis. Indhente og vurdere naturvidenskabelig viden fra forskellige kilder.	Paratviden, forståelse, analyse, vurdering
4	Forståelse for opbygningen og funktionen af: fedtstoffer.	Beskrivelse, paratviden
10	Eleverne skal opnå kendskab til fordøjelsessystemets opbygning og funktion. Overføre teoretisk viden til praksis.	forståelse, analyse, vurdering

Efter hver besvarelse af de enkelte spørgsmål får eleven feedback på, om svaret er korrekt eller forkert og hvor mange point besvarelsen giver. Der er som i prototypen 30 min. til at besvare testen i. I brugergrænsefladen tæller et ur ned fra 30 til 0 minutter. Testen afvikles som tidligere nævnt i to klasser. Problematikken omkring feedback og hvordan det indvirker på elevernes holdning til testen behandles i afsnit 4.3 og vil endvidere blive diskuteret i afsnit 5.

#### 4.2.4 Gennemførelse – faglige test musik

Selv om musikfaget fakultetsmæssigt på universiteterne traditionelt har hørt hjemme på de humanistiske uddannelser, integreres elementer fra det humanistiske, det naturvidenskabelige og det sam-

<sup>39</sup> Quia kan ikke håndtere links i brugergrænsefladen, derfor skal eleverne copy-paste web-adressen over i et nyt browser-vindue for at åbne animationen

fundsvidenskabelige område med henblik på almindannelse, identitetsudvikling og studiekompetence. Det vil sige, at der er elementer fra de andre didaktiske positioner indeholdt i faget også. Musikfaget har den særlige form, at det forener en teoretisk-videnskabelig, med en kunstnerisk og en performativ tilgang til musik.

I bekendtgørelsen benyttes der under de faglige mål indenfor et af de to hovedområder musikkundskab (-teori og -historie) begreberne: anvende, anlægge synsvinkler, demonstrere kendskab, udfærdige, opsøge, bearbejde. Disse er sammenlignelige med Bloom, og indenfor det andet område musikudøvelse (solospil og -sang, sammenspil og korsang) står begreberne: udfylde, reflektere, samarbejde, fremføre, tilrettelægge, arrangere og realisere, som umiddelbart også er forenelige med Bloom. Det står dog også i bekendtgørelsen, at musikundervisningen beskæftiger sig med følgende niveauer i læreprocessen: **perception** (lytning, oplevelse), **reproduktion** (udførelse af eksisterende kompositioner), **imitation** (eftergørelse), **produktion** (komposition, improvisation), **interpretation** (fortolkning) og **refleksion** (perspektivering), som udgør en taksonomisk læringsproces, en opstilling som har nogle elementer der er forskellige i forhold til Bloom.

Der står ikke meget om den løbende evaluering i musikbekendtgørelsen. Det står kun kort at “for at sikre, at eleverne opnår de beskrevne faglige mål for undervisningen i musikkundskab, gør læreren brug af mindre test undervejs”. Desuden evalueres musikudøvelse og den skriftlige dimension i faget selvfølgelig også.

Følgende figur viser en tematisk opstilling af emnet ‘Populærmusik gennem tiderne: Jazz frem til 1970’ indenfor hovedområdet musikkundskab, som er den undervisningsmæssige kontekst testen er designet ud fra.

Stikord/faglige mål	Underemner	Indhold
Musikhistorisk placering og sammenhæng	Historisk baggrund	Gennemgang af litteratur og lydseksempler, og filmklip
Musikanalytisk perspektivering	Præsentation af årtierne	Opdeling af jazz-historien i stilarter
Forudsætninger	Jazzens rødder	Blues, folk, gospel og songwriter traditionerne
Instrumentlære, hørelære, (kernestof)	Instrumentering	De forskellige jazzorkestre og deres instrumenter
Partiturlæsning, form, notation, forståelse/analyse	Jazz-analyse	3 forskellige stilmæssigt spredte modeleksempler
Musikalske parametre, (kernestof)	Jazz-teori	akkorder, alterationer, skalaer, harmonik og melodik, rytmik
Informationsanvendelse	Fremlæggelser og gruppearbejde	Fordeles på de forskellige jazz-perioder

Eleverne får ikke feedback i denne test under eller umiddelbart efter besvarelsen. Testen skal afvikles på 30 minutter og der er et spørgsmål på skærmen ad gangen. Desuden er testen indstillet således at spørgsmålsrækkefølgen blandes tilfældigt fra elev til elev for at mindske risikoen for at de skal kigge hinanden over skulderen (skærmene står meget tæt i computerlokalet). Testen afvikles i to klasser.


Efterfølgende skema er en opstilling i faglige mål for hvert spørgsmål (se testspørgsmål i bilag 8) og dets taksonomiske niveau. Vi har forsøgt at bruge Blooms reviderede taksonomi, men fandt det svært at placere særligt de auditivt baserede spørgsmål i vidensniveauerne i taksonomien.

Spm.	Mål	Taksonomisk niveau	Sværhed
1	Kendskab til jazz-genrer	huske, faktuel viden	1
2	Kendskab til musikalsk form	huske, faktuel viden	1
3	Kendskab til musikalsk form	huske, faktuel viden	2
4	Forstå rytmisk notation	Huske, forstå og anvende, procedural viden	1
5	Sammensætte genrer med tilsvarende kunstnere	Klassificering, faktuel viden	5
6	Sammenkoble begreber med genrer	Klassificering, konceptuel viden	4
7	Afgøre et musikstykkets tonalitet ved lytning	Vurdere, konceptuel viden	3 (auditivt)
8	Kendskab til musikalsk form ved lytning	Vurdere, konceptuel viden	4 (auditivt)
9	Kendskab til musikkens historiske rødder	huske, faktuel viden	3

10	Knytte en musikhistorisk person med et instrument	huske, faktuel viden	2
11	Kendskab til en genres musikalske udtryksformer	vurdere, konceptuel viden	3
12	Kendskab til musikalske begreber	forstå, konceptuel viden	2
13	Kendskab til musikalsk form	huske, faktuel viden	2
14	Grundlæggende kendskab til notationssystemet og skalaer	udlægge og analysere, konceptuel viden	2 (visuelt)
15	Grundlæggende kendskab til noder og intervaller	udlægge og analysere, konceptuel viden	1 (visuelt)
16	Kategorisere jazz-genrer historisk	Anvende, faktuel viden	4
17	Kendskab til musikalsk genre ved lytning	Vurdere, konceptuel viden	2 (auditivt)
18	Overføre viden om en notationsform til en anden	Forstå og anvende, procedural viden	3
19	Genkende et instruments klang ved lytning	klassificere, faktuel viden eller procedural (udføre 'aktiv' lytning)	3 (auditivt)
20	Genkende et instrument ikonisk	klassificere, faktuel viden	2 (visuelt)

Det er fortrinsvis kernestof fra musikkundskabsdelen af faget som bliver benyttet i testen, mest fra musikteorien: musikalske parametre (fagterminologi, hørelære, musikhistorie, stilkendskab, musiklære, nodelære, rytmelære etc.). I den efterfølgende opstilling er spørgsmålene kategoriseret i den taksonomiske matrix, som blev gennemgået i afsnit 4.2.2. Læg til mærke at nogle spørgsmål er placeret flere steder idet nogle testspørgsmål opfylder flere niveauer i taksonomien. Oversigten skal tages med noget forbehold, idet nogle af spørgsmålene har været svære at placere, men tendensen er at der er flest spørgsmål på de laveste taksonomiske niveauer og der er både fakta og forståelse, som bliver testet.

### Taksonomi over læringsmålene i musiktesten

	Kognitiv Dimension 					
	Huske	Forstå	Anvende	Analysere	Evaluerer	Skabe
<b>Faktuel viden:</b>	1,2,3,9,10,13,	5,19,20	16			
<b>Konceptuel/begrebslig viden:</b>		6,12		14,15	7,8,11,17	
<b>Proceduremæssig viden (know-how):</b>	4	4, 18,19	4,18			
<b>Metakognitiv viden:</b>						

Det ovenstående siger noget om hvilke kognitive processer der skal tages i brug af eleven og ikke noget om det enkelte spørgsmåls sværhedsgrad fagligt set, som afhænger af diverse faglige faktorer. Dette vil blive diskuteret i kapitel 6.

Efterfølgende figur viser hvilken spørgsmålstype de 20 spørgsmål har og en vurdering af sværhedsniveauet fra 1-5 hvor 5 er den sværeste. Det fremgår også i hvilke af spørgsmålene der er brugt illustrationer og lydseksempler:

Spørgsmål	Type	sværhedsniveau	Medie
1	Multiple choice	1	
2	Fill-in	1	
3	Fill-in	2	
4	Fill-in	1	
5	Matching	5	
6	Matching	4	
7	Multiple choice	3	lydeksempel/illustration
8	Multiple choice	4	lydeksempel
9	Multiple choice	3	
10	Multiple choice	2	
11	Multiple choice	3	
12	Multiple choice	2	
13	Multiple choice	2	
14	Multiple choice	2	nodeksempel
15	Multiple choice	1	nodeksempel
16	Ordering	4	
17	Multiple choice	2	lydeksempel
18	Multiple choice	3	
19	Multiple choice	3	lydeksempel
20	Multiple choice	2	billede

Analyserne af de faglige test foretages i kapitel 5.

### 4.3 Holdningsundersøgelse – elevsurvey

#### 4.3.1 Teoretisk grundlag og tilrettelæggelse

Efter at eleverne har besvaret de faglige test, beskrevet i foregående afsnit, skal de besvare et holdningsurvey, bestående af 21 spørgsmål, der undersøger elevernes holdning til den test, de netop har gennemført og deres holdning til test generelt.

Formålet er at finde indikationer på om eleverne ser et læringspotentiale i at besvare en faglig test. Desuden ønsker vi at undersøge om der er forskellige præferencer for hvordan en test skal udformes, for at den kan indeholde et læringspotentiale – altså være formativ.

Spørgsmålene i spørgeskemaet formuleres så de kan bruges i begge fag. Et særligt problem med hensyn til at spørge til elevernes holdning til test er, at den måske vil være præget af det testresultat, de lige har opnået. Vi vil derfor lave en prototype-undersøgelse for at afklare dette. En anden grund til at lade eleverne kende testresultatet inden de besvarer holdningsspørgsmålene er, at vi vurderer, at respondenterne kan have vanskeligt ved at vurdere værdien af testen, hvis ikke de ved, hvor godt eller dårligt de har klaret testen. De vil f.eks. have vanskeligt ved at svare meningsfuldt på følgende spørgsmål, hvis ikke de kender deres testresultat: ”En test som denne gør, at jeg bedre kan vurdere

mit faglige standpunkt". Omvendt er det også et problem for undersøgelsen, hvis de svar de afgiver er præget af, om de har klaret sig godt eller dårligt i testen. Vi undersøger om dette er tilfældet ved at sammenholde den enkelte elevs testresultat med vedkommendes svar i spørgeskemaet. Viser det sig, at der er en tendens til, at lav testscore falder sammen med negativ holdning til test og omvendt så vil vi ændre proceduren. En mulighed kunne være at tilbageholde testresultatet og til gengæld udelade spørgsmål i spørgeskemaet som kræver kendskab til testresultatet.

Spørgsmålene besvares ud fra 5 svar-kategorier (helt enig, noget enig, hverken enig eller uenig, noget uenig helt uenig). Ifølge Henning Olsen findes er kun i begrænset omfang konsensus om, hvilke tommelfingerregler der bør gælde når det gælder spørgeskemakonstruktion (Olsen 2005:179ff). En af de forholdsregler der dog synes enighed om er, at man skal undgå dobbelt negation (negation i både spørgsmål og svar) og at svarkategorierne skal være balancerede (ibid:44ff). Derimod er der forskellig opfattelse af om "på midten" og "ved ikke" bør indgå i spørgeskemaer (ibid:81ff). På trods af, at metodelitteraturen indeholder eksempler på, at "på midten" kategorier ændrer svarfordelingen (ibid:82), har vi alligevel valgt at have en "på midten" kategori med i nærværende spørgeskema idet, vi mener der skal være mulighed for at respondenter som ikke har gjort sig deres holdning klar skal have mulighed for at kunne afgive et svar. Til gengæld undlader vi så en "ved ikke" kategori, idet vi mener den overflødiggøres. Reliabiliteten af survey-undersøgelsen kan undersøges ved at gentage undersøgelsen flere gange, såkaldt test-retest (ibid:24ff). Imidlertid er det ikke en uproblematisk metode idet tidsfaktoren kan undervurdere, hvis der går for lang tid, henholdsvis overvurdere reliabiliteten, hvis der går for kort tid (ibid:25). Vi ser derfor bort fra denne mulighed. Til gengæld mener vi, at det forhold, at vi gennemfører undersøgelsen på to forskellige gymnasier i to forskellige fag og i flere klasser med de respektive fag, gør det muligt at give en tilnærmet vurdering af reliabiliteten. Det skal i parentes bemærkes, at Henning Olsen fremfører at der i litteraturen ikke er konsensus om, hvad reliabilitet er, og hvordan det måles (ibid. s. 25).

Hvad angår validiteten, synes der derimod at være nogenlunde enighed om, at validitet er udtryk for, hvor godt en undersøgelse måler det, den foregiver at måle (Jacobsen et al 1999:127, Rust and Golombok 1999:215). Imidlertid citerer Henning Olsen Fowler (Fowler 1993:80) for følgende "Validitet er det omfang i hvilket et svar er en sand måling og betyder, hvad forskeren ønsker eller forventer, det skal betyde", hvorefter Olsen tilføjer: "Men han tilføjer at besvarelse af subjektive spørgsmål – f.eks. holdningsspørgsmål ikke lader sig validere fordi der ikke findes sande størrelser" (Olsen 2005:26). Det kunne jo tyde på, at vi efterlades med et problem som vi vanskeligt kan løse. Vi har imidlertid konstrueret spørgsmålene på en måde så nogle spørgsmål spørger til det samme forhold. Eventuel modstrid eller sammenfald i svarene kan derved give et fingerpeg om validiteten.



Spørgsmålene er efterfølgende tematiseret med henblik på at kunne analysere elevernes besvarelse af spørgeskemaet. Tematiseringen er sket ved at gruppere spørgsmålene efter, hvilken problemstilling de spørger til:

**Tema 1: Fagligt standpunkt og egen vurdering af faglige kompetencer**

1. En test, som den jeg lige har gennemført, kan fortælle mig om mine faglige stærke og svage sider
2. En test som den jeg lige har gennemført gør, at jeg bedre kan vurdere mit faglige standpunkt/niveau
3. En test, som den jeg lige har gennemført, kan fortælle mig om jeg skal øge/justere min indsats i faget
4. Jeg vidste godt på forhånd, hvor mine stærke og svage sider var i forhold til det faglige emne
5. En test, som den jeg lige har gennemført, giver mig mulighed for at vise, alt hvad jeg virkelig kan i faget
10. Testen giver mig indblik i mine faglige kompetencer indenfor dette emne i faget

**Tema 2: IKT kontra papir test**

6. Jeg er mere motiveret for at deltage i en test når den foregår på Pc end hvis den foregår på papir
8. Jeg foretrækker, at besvare en test, som den jeg lige har gennemført, på almindeligt papir
19. Jeg foretrækker testspørgsmål, hvori der indgår billeder, lyd og/eller animationer
21. Jeg synes det skærbillede som den faglige test foregår på, er for uoverskueligt

**Tema 3: Motivation**

7. At besvare en test, som den jeg lige har gennemført, er som at deltage i en konkurrence
9. Jeg bliver motiveret af at deltage i test som den jeg lige har gennemført

**Tema 4: Holdning til test, test positivt eller negativt**

11. Jeg synes det er positivt at få kontrolleret mine faglige kompetencer ved brug af test
16. Jeg vil gerne have flere af denne slags faglige test
17. Jeg foretrækker test opgaver, hvor jeg selv skal/kan formulere svarene

**Tema 5: Spørgsmålenes karakter**

12. Spørgsmålene i den test jeg lige har gennemført, var svære at forstå
13. De fleste spørgsmål i den test jeg lige har gennemført, var lette at besvare

**Tema 6: test som stressfaktor**

14. Jeg følte, jeg var under tidspres, da jeg skulle besvare spørgsmålene i den test jeg lige har gennemført
15. Jeg bliver usikker, når jeg skal besvare test, som den jeg lige har gennemført

**Tema 7: Feedback**

18. Den tilbagemelding jeg fik efter, at jeg havde besvaret den faglige test, var tilfredsstillende
20. Når jeg besvarer en test, vil jeg gerne have lærer-kommentarer til mit (test)resultat

### **4.3.2 Gennemførelse**

Som omtalt ovenfor foretages der en prototype-undersøgelse. Denne administreres som et word-dokument via fronter til 10 elever på det jyske gymnasium, der forinden har besvaret den faglige test i biologi. Efterfølgende evaluerer vi de enkelte spørgsmål og svar og når frem til at de kan anvendes uforandret. Der synes ikke, som omtalt ovenfor, at være en sammenhæng mellem elevernes testresultat og deres holdning til test. F. eks. svarer en elev, der har lidt under 50 % rigtige i den faglige "næsten enig" på sp. 11: Jeg synes det er positivt at få kontrolleret mine faglige kompetencer ved brug af test. Og samme elev svarer også "næsten enig" på sp. 9: Jeg bliver motiveret af at deltage i test som den jeg lige har gennemført. Omvendt svarer en elev, der har mere end 90% rigtige i den faglige test, "helt uenig" på sp. 11.

I den endelige holdningsundersøgelse administreres spørgsmålene ligesom de andre undersøgelser af Quia. I den endelige version erstattes svarkategorierne af numeriske kategorier fra 1 til 5, hvor 1 er lig med "meget enig" og 5 er lig med "meget uenig".

## **4.4 Fokusgruppeinterview - lærere**

Formålet med interviewet er at afdække gymnasielæreres brug af og syn på test som et pædagogisk værktøj og om test kan indgå som en del af en evalueringskultur. De to fokusgruppeinterview vi foretager, udgør således den kvalitative del af vores empiri.

### **4.4.1 Teoretisk grundlag**

Fokusgruppeinterviewet, betragtet som et kvalitativt forskningsinterview, har ifølge Steinar Kvale "til formål at indhente beskrivelser af den interviewedes livsverden med henblik på at fortolke betydningen af de beskrevne fænomener" (Kvale 1997:41). Eftersom Kvales definition er formuleret med udgangspunkt i et individuelt interview, skal det bemærkes, at vi går ud fra, at definitionen kan overføres til en situation, hvor en gruppe af personer interviewes. Med den modifikation, at det så i højere grad bliver gruppens livsverden, mere end den enkelte deltagers livsverden, der træder frem selvfølgelig alt afhængig af, hvordan det konkrete interview kommer til at forløbe.

Vi har valgt interviewmetoden i en semistruktureret udgave. Den semistrukturerede interviewform har sin styrke i sin åbne udforskende form, hvor samtalen mellem interviewer og de interviewede kan udvikle sig og inddrage f.eks. nye uventede aspekter eller fordybe sig i mere komplekse forhold, der kommer frem under interviewet. Fokusgruppeinterviewet kan med andre ord bruges til at lokalisere evt. holdninger og problemstillinger, men ikke til hvor udbredte de er. Fokusgruppeinter-

viewets styrke ligger i, at det er en forholdsvis let tilgængelig måde at indsamle en stor mængde data på, med det formål at få deltageres egne opfattelser og vurdering frem i lyset. Fokusgruppen supplerer hinandens udtalelser, og på den måde sker der en form for verificering og nuancering af deltageres udsagn, da deltagerne på sin vis holdes i skak af de øvrige deltagers udtalelser (Kvale 1997:108). Informanterne har ligeledes mulighed for at nuancere deres udtalelser, da de gennem interviewet kan blive inspireret af de andre deltagere, og ligeledes har de også mulighed for at få “frigjort” skjult eller glemt viden gennem dialog med de andre deltagere og interviewer.

Kvale hævder endvidere, at “samtalen er efter vor opfattelse ikke kun en specifik, empirisk metode. Den er tillige forbundet med en grundlæggende form for konstituering af viden, og menneskets verden er en samtalevirkelighed” (Kvale 1997: 47).

I forlængelse heraf producerer fokusgrupper, efter Bente Halkiers opfattelse, “accounts in action”. “Det vil sige deltagerne udveksler beretninger om handlinger og forståelser, som en del af interaktionen i en social, hverdagslig genkendelig kontekst rundt omkring forskeren” (Halkier 2003:12). Halkier bruger bevidst udtrykket “producere” om den viden eller mening som fokusgruppen etablerer. Halkier siger, at “når man arbejder inspireret af socialkonstruktivismen, så indsamler man ikke viden og data” (ibid:28). Efterfølgende tilføjer hun, med henvisning til Guba og Lincoln, at “en sådan antagelse passer dårligt med konstruktivismens opgør med en positivistisk opfattelse af data som noget der allerede ‘er derude’ uafhængigt af forskeren”. Set i et socialkonstruktivistisk optik sker der altså en “produktion” af viden i forbindelse med samtalen og meningsforhandlingen deltagerne i mellem – og interviewer er en del af denne. En anden vigtig pointe hos Halkier er, at interaktionen foregår i en hverdagslig kontekst. Vi får med andre ord indblik i gruppens livsverden, for at bruge Kvales ord, når vi interagerer med deltagerne.

Denne videnskabsteoretiske forståelse giver god mening i forhold til den del af vores projekt, der handler om evalueringskultur. Vi har nemlig en forforståelse af, eller antagelse om, at evalueringskulturen i gymnasiet til dels i hvert fald, manifesterer sig i kraft af den diskurs eller de diskurser der er en del af hverdagen i gymnasieskolen. Evalueringskulturen, eller diskursen om den, antages dermed at etablere sig som et produkt af meningsforhandling mellem interessenterne.

#### 4.4.2 Tilrettelæggelse af interview

Vi har udarbejdet en interviewguide (bilag 9) med 4 forskningsspørgsmål og dertil hørende interviewspørgsmål i henhold til Kvale (1997:133ff). Fokusgruppen på det jyske gymnasium består af 3 personer: en mandlig lærer (B), der underviser i biologi, en kvindelig lærer (Bk), der underviser i biologi og kemi og en mandlig lærer (Fh), der underviser i fysik og historie. De tre interview personer dækker således, de i gymnasiesammenhæng traditionelle naturvidenskabelige fag. Dog er (natur)geografi ikke repræsenteret. Bk er samtidig lærer for de to klasser, der indgår i det empiriske materiale jvf. afsnit 4.1, 4.2 og 4.3. B er lærer for den klasse, der indgik i en pilotundersøgelse forud for de egentlige undersøgelser. Alle tre er tilmed kolleger til interviewer. Interviewet foregår i en "fritime", hvor lærerne alligevel er på gymnasiet. Der er således afsat 1,5 time til interviewet. Fokusgruppen på det færøske gymnasium består af 7 undervisere dækker 2 fys/mat 2 historie/idræt og 3 sproglærere, som ikke har nogen tilknytning til klasserne i undersøgelsen.

#### 4.4.3 Gennemførelse

På grund af den geografiske afstand mellem forfatterne foretages interviewet af kun én interviewer. Denne er som nævnt ovenfor også kollega til de interviewede. Det er således åbenbart, at der er et afhængighedsforhold mellem interviewer og interviewede, som rækker ud over det konkrete projekt. Dette afhængighedsforhold kan få indflydelse på validiteten og realibiliteten i de udsagn, der fremkommer. Man kan for eksempel forestille sig, at interviewpersonerne giver udtryk for en positiv vurdering af problemfeltet for at tilgodese interviewer. Omvendt kan interviewerens forhåndsopfattelse af interviewpersonerne komme til at påvirke interviewforløbet og analysen af denne. Forhåndskendskabet mellem interviewer og de interviewede kan imidlertid også være en fordel i forbindelse med kontrol af realibilitet og validitet, idet de udsagn der fremkommer, kan analyseres og vurderes i forhold til det kendskab interviewer har til de interviewede.

Det tilstræbes (jvf. ovenstående), at interviewer intervenserer så lidt som muligt i interviewet. Hensigten er at interviewet med fokusgruppen forløber således, at spørgsmålene åbner op for en samtale, primært mellem deltagerne, og at interviewer kun indgår i dialogen når/hvis denne stopper og har brug for moderering. Intervieweren stiller desuden uddybende eller supplerende spørgsmål efter behov.

Interviewet optages på digital diktafon med henblik på senere transskription. Fokusgruppen orienteres ved interviewets begyndelse om, at de vil blive anonymiseret i forbindelse med transkriberingen og afrapporteringen.

#### **4.5 Opsamling**

Det empiriske grundlag i projekt bygger på både kvalitative og kvantitative data. Man kan sige at vi har en pragmatisk tilgang til metodevalg og deraf følgende dataproduktion. Samfundsvidenskabelig forskning kan i følge Olsen ”gennemføres ved anvendelse af kvalitative og/eller kvantitative metoder, der undertiden betegnes som to metodologiske spor: et meningstydende og et meningsskabende” (Olsen 2005:10). Petersen og Land siger, at mens kvalitative metoder sætter fokus på sociale eller andre fænomeners beskaffenhed, sigter anvendelse af kvantitative metoder mod fænomeners udbredelse og statistiske sammenhænge. Videnskabsteoretisk har den kvantitative metodes grundlag været det positivistiske videnskabssyn, hvorimod den kvalitative forskningsmetode er blevet operationaliseret ud fra flere forskellige paradigmer herunder et konstruktivistisk (Pedersen & Land 2001:17,29). Vores pragmatiske tilgang ekspliciteres dermed i at vi anerkender legitimiteten af begge metodiske tilgange. Vi betragter ikke den ene som hævet over eller bedre end den anden. Vi er meget på bølgelængde med Henning Olsen, når han lidt populært forskellen fremstiller forskellen på de to metoder således: ”Kvantitative forskere er optaget af at tælle æbler og pærer, mens kvalitative forskere sætter fokus på forskellen mellem frugterne” (Olsen 2005:10). Pointen er her, at det er to forskellige ting der ”måles”.

Opsummerende består de konkrete empiriske undersøgelser i kvantitative undersøgelser i 4 gymnasieklasser: to klasser på et jysk gymnasium og 2 klasser på et færøsk gymnasium. Begge klasser besvarer en faglig IKT-baseret test i de respektive fag. Derudover besvarer klasserne 21 holdnings-spørgsmål med relation til den faglige test. For det tredje stilles 36 henholdsvis 42 spørgsmål til eleverne med det formål, at undersøge multiple intelligenser (Gardner) og læringsstile (Dunn & Dunn) hos eleverne. De kvalitative undersøgelser består i to fokusgruppeinterview. Fokusgrupperne består en gruppe lærere med både humanistisk og naturvidenskabelige fag på det færøske gymnasium, og en gruppe lærere på det jyske gymnasium med naturvidenskabelige fag.

## 5. Empiri – analyse

I dette kapitel fremlægger vi analyser af projektets empiri. Analyserne fremlægges forholdsvis struktureret med udgangspunkt i de deelelementer empirien består af. Først analyseres de to fokusgruppeinterview vi har foretaget, dernæst analyseres de kvantitative data. I det sidste afsnit foretager vi nogle korrelationsanalyser af det kvantitative materiale med henblik på, at undersøge om der findes sammenhænge mellem elevernes besvarelse af de faglige test og deres multiple intelligenser henholdsvis læringsstile og om der er sammenhænge mellem deres holdning til test og deres besvarelse af testene.

### 5.1 Fokusgruppeinterview analyse

#### 5.1.1 Analysemetode

Analyse arbejdet kan angribes på forskellig måde. Kvale nævner 5 forskellige metoder: Meningskondensering, Meningskategorisering, Narrativ strukturering, Meningsfortolkning og skabelse af mening gennem ad-hoc metoder (Kvale 1997:188ff.).

I vores analyse af fokusgruppeinterviewene tager vi udgangspunkt i Ad-hoc metoden. Metoden er eklektisk (ibid:191), dvs. udvælgende eller udsøgende i analysen af interviewteksten. I følge Kvale er det formentlig den mest almindelige metode til skabelse af mening. Der bruges ikke nogen standardmetode til analyse af interviewmaterialet, derimod er der tale om et frit samspil af teknikker under analysen (ibid:188ff). At vi vælger denne form for analysemetode beror til dels på en efterrationalisering, i og med en gennemlæsning af interviewmaterialet tilsiger os, det vil være netop det mest rationelle at gøre, samtidigt med, at denne metode formentlig vil være mindre tidskrævende end de øvrige. Målet med analysen er således at udfolde interviewenes meningsindhold, via meningskondensering og meningsfortolkning.

I det følgende analyser vi den foreliggende (transskriberede) interviewtekst med udgangspunkt i de fire forskningsspørgsmål vi stiller i interviewguiden:

1. Hvordan ser evalueringskulturen i gymnasiet ud i dag set fra henholdsvis naturvidenskab/humaniora?
2. Hvilke muligheder ser undervisere i gymnasiet ved IKT-baseret evaluering?
3. Hvad er holdningen til faglige test hos naturvidenskab/humaniora?
4. Hvad er holdningen til læringsstiltest hos naturvidenskab/humaniora?

### 5.1.2 Databehandling – det jyske gymnasium

Det transskriberede interview foreligger i (bilag 10). Der er ikke tale om en fuldstændig transskription. De første 10 minutter er udeladt, fordi det der siges her, er interviewerens beskrivelse af formålet med interviewet og dels udvikler der sig en diskussion om nogle tekniske forhold vedr. de afviklede faglige test i biologi og de supplerende undersøgelser vedr. læringsstile og multiple intelligenser som ikke direkte indgår i formålet med interviewet. Udover de første 10 minutter er der udeladt mindre passager, hvor det enten er utydeligt, hvad der siges eller, hvor det der siges, vurderes at være irrelevant i forhold til formålet med interviewet.

Det skal derudover tilføjes, at transskriptionen har været besværlig, fordi informanternes udsagn bærer præg af mange øh'er, gentagelser, ufuldendte sætninger og talen i munden på hinanden. Så for nogle udsagns vedkommende kan det være vanskeligt at forstå meningen med det sagte.

Interviewspørgsmålene i interviewguiden følges ikke slavisk under interviewet, idet interviewet udvikler sig på en måde, hvor nogle svar/udsagn indeholder elementer som samtidigt dækker flere af spørgsmålene.

<p><b>1. Forskningsspørgsmål: Hvordan ser evalueringkulturen i gymnasiet ud i dag set fra henholdsvis naturvidenskab/humaniora?</b></p>
---

#### Kondensering

De tre informanter har, hver deres måde at evaluere på: B bruger test, på papir, systematisk i forbindelse med første karaktergivning. Fh bruger typisk skriftlige afleveringer, og har ikke været særlig testorienteret i sit virke, men kan godt se nogle fordele ved IKT-baserede test:

*....fordi jeg har heller ikke været særlig test orienteret, men jeg vil da nok synes, at med IKT er det blevet lidt nemmere og få det der overblik, det synes jeg er rigtigt – det tror jeg også du nævnte B – altså på den måde er jeg sådan set blevet mere test orienteret end jeg har været, det tror jeg nok jeg kan sige.*

Bk bruger ikke test, men kan godt se en idé i at bruge test. Men er dog bange for det store arbejde med at lave testspørgsmål (som Fh er enig i). Betegner desuden brugen af test som amerikansk! Og siger *.... fordi det er ikke vores skolekultur at gøre det her* og senere siger hun, *skolen har ikke nogen god evalueringkultur* (B og Fh tilslutter sig dette synspunkt).

I det efterfølgende interviewspørgsmål: **Hvordan vil I karakterisere evalueringkulturen her på skolen?** Bliver dette yderligere udfoldet, hvor Bk siger at evalueringkulturen er mangelfuld – uden konsensus, sammenhæng og overblik. Hertil bemærker B at *man får ikke en evalueringkultur bare*

*ved at sige man skal evaluere – det skal specificeres om det er undervisningen eller eleverne, der skal evalueres? Og efterlyser en norm for eller specifikation af hvordan, hvor meget og hvornår.*

Informanterne giver hinanden ret i at, det står i bekendtgørelsen, at skolen skal have en evalueringskultur.

I Spørgsmålet om evalueringen af undervisningen anes nogen uenighed. B er skeptisk overfor værdien af undervisningsevaluering, hvorimod Bk godt kan se pointen i dette: *jamen vi bliver jo til hele tiden nødt til at arbejde med os selv som lærer ikke. Det er jo svært, det er jo nogle store tankskibe, vi skal have vendt, fordi vi har en bestemt undervisningsstil vi har lært og den virker godt afhængig af temperament ikke. Og så har vi svært ved at trække de andre ind, men det er sgu da vigtigt, at vi er bevidste om det.*

Herefter konkludere Bk på interviewspørgsmålet og siger, at evalueringskulturen er mangelfuld og individuel (og modsiges ikke de to øvrige informanter).

Senere hen i interviewet dukker evalueringsspørgsmålet op igen og

Bk siger da: *...men altså jeg er også sådan en jeg lader eleverne selvevaluere og stiller begavede lidt kringledede spørgsmål og linker indtil nogle hjemmesider og så for de et modul, hvor de simpelt hen bare sidder og arbejder.* B supplerer og siger han gør nogen lignende, men med for lange mellemrum. Fh indskyder så: *Det du sagde lød for mig mere som indlæring, jeg mener de har arbejdet med noget i bogen og så supplerer du med et eller andet på nettet eller en webquest. Det er vel ikke en test eller evaluering eller hvad? ....*

### **Fortolkning:**

Besvarelsen af det første forskningsspørgsmål går altså gennem to interviewspørgsmål dels et spørgsmål, der undersøger, hvordan de tre informanter bruger evaluering i deres undervisning dels, hvordan de ser på evalueringskulturen på skolen. Informanternes forskellige udsagn om deres evalueringspraksis viser, at konstateringen af den manglende evalueringskultur har en vis pålidelighed. Den diskussion, der senere dukker op, indikerer også, at der er tvivl om, hvad man egentlig forstår ved evaluering. Det skal dog bemærkes, at der kan være noget bias i udsagnene om, hvordan informanterne bruger evaluering, idet de på forhånd er sporet ind på brugen af test i denne sammenhæng. Konteksten for interviewet er jo netop brugen af test i evalueringen. Der kan jo evalueres på andre måder end ved brug af test, men det siges der ikke ret meget om.

<b>2. Forskningsspørgsmål: Hvilke muligheder ser undervisere i gymnasiet ved IKT-baseret evaluering?</b>
--



### **Kondensering:**

I forbindelse med interviewspørgsmålet **Har I eller skolen de nødvendige ressourcer og kompetencer til at gennemfører IKT-baseret evaluering?**

Svarer B at ...*Altså som sagt tror jeg det er lige så godt på papir.* Han siger dog også at fronters prøveværktøj er nemt at bruge, men ikke overbevisende fantastisk. Fh har til gengæld tidligere givet udtryk, at han ser nogle fordele ved IKT-baserede test (se ovenfor).

Bk og Fh nævner muligheden for at bruge multimedie virkemidler (billeder, lyd, animationer) i test. B spørger Fh om han rent faktisk gør det? Hvortil Fh svarer: *Nej vi bruger det bare i undervisningen. Jeg siger bare den mulighed er der, så derfor synes jeg, det er oplagt os og gøre det, så det vil jeg bestemt overveje.*

Bk ser et behov for efteruddannelse i brugen af multimedie virkemidler: *Det vil så også kræve, hvis vi skal bruge det der IKT mere til evaluering og vil gerne have den der animationsdel og billeder med så kræver det os noget, at vi kursus altså, at vi skal lære det.*

Der fremkommer efterfølgende nogle udsagn som indikerer, at IKT infrastrukturen på skolen ikke er optimal. B bemærker hertil: ...*Fordi nu har vi efterhånden fået dem ud i alle lokalerne, så oplever vi nu, at til sådan noget der, har vi måske brug for flere computerrum, man kan næsten ikke forestille sig det afviklet uden, at man sidder i et rum.*

### **Fortolkning:**

Der er ikke entydig tilslutning blandt informanterne til, at IKT-baserede test er et bedre værktøj end papir og blyant. Men Bk og Fh argumenterer dog for, at multimedie virkemidler bidrager til brugbarheden af ikt-baserede test. Det indikeres dog, at der kan være problemer med lærernes kendskab og færdigheder i at bruge IKT-værktøjerne ligesom utilstrækkelig IKT-infrastruktur kan være en begrænsning. Som nævnt ovenfor gælder det også her, at der kan være bias i informanternes udsagn idet det, der udelukkende tales om IKT-baserede test og ikke IKT-baseret evaluering i bredere forstand.

<b>3. Forskningsspørgsmål: Hvad er holdningen til faglige test hos naturvidenskab/humaniora?</b>
--

### **Kondensering:**

Til dette 3. Forskningsspørgsmål konstaterer B: *Vi er vel alle sammen indforstået med faglige test ikke?* (og modsiges ikke af de to øvrige informanter) og B replicerer: *altså jeg synes i mange af de eksempler vi nævner, det handler jo netop om, hvordan vi på sin vis af alt magt forsøger at bruge det i undervisningen. Og ikke bare som bedømmelse af eleverne.*

Herefter modificeres interviewspørgsmålet til **kan man stille anvendelsesorienterede spørgsmål i en test?** Intervieweren motiverer spørgsmålet med at kravet om anvendelsesorienteret undervisning er kommet ind i fagenes læreplaner. Det giver anledning til en diskussion om, hvad der forstås ved anvendelsesorienteret og hvad undervisningsministeriet mener med dette. Fh og B når frem til at det må være noget med et ”samfundsperspektiv”. Informanterne når så frem til, at det bliver svært at lave multiple-choice spørgsmål inkluderende dette perspektiv. Supplerende spørges fokusgruppen om det er nok at teste paratviden? Hertil svarer B: *Nej det er ikke nok, det også derfor de test og den type de udgør også kun en lille del af min undervisning og min bedømmelse af eleverne, det skulle meget nødtigt blive for dominerende, at jeg kun er optaget af, at de har en paratviden af en vis størrelse, men samtidig så vil jeg nok vove at påstå, at det er svært at have forståelse uden paratviden i et eller andet omfang, der er jeg nok lidt gammeldags på den måde – paratviden må der til.* Det sidste synspunkt – at paratviden er en nødvendighed bakkes op af Bk og Fh.

Fokusgruppen stilles derpå overfor spørgsmålet om **faglige test er et brugbart redskab for eleverne? Altså kan eleverne bruge det til noget eller er det primært noget lærerne har gavn af?**

Som svar her på udtaler B og Fh sig om værdien af den feedback de giver til eleverne og betvivler i den forbindelse om eleverne bruger denne feedback konstruktivt. Fh nævner efterfølgende IKT som en mulighed og henviser her til de nationale test i folkeskolen som et brugbart redskab. Uden han nævner det, tyder det på, at det især er det adaptive element han tænker på.

Bk siger hertil: *Ja og så der går lidt leg i det ik... altså kan det laves som et computerspil vil det være optimalt.* Og Fh replicerer:

*Så kan man jo også tilgodese læringsstile i det og så nogen ting.*

Da informanterne efterfølgende konfronteres med spørgsmålet: **Vil det være en god ide at indfører IKT-baserede nationale test i gymnasiet?** Bakker de med følgende udtalelser tilbage:

B: *Vi vil kunne få os nogle flade grin over det*

Bk: *Vi skal leve af at være innovative, det bliver vi kraft æde mig ikke innovative af.*

B: *Der vil lynhurtigt ske det, at vi vil rette ind efter den test (Fh: ja, ja det er jo det)*

Fh forsøger så alligevel, at argumentere for ideen:

Fh: *Sådan er det jo i for vejen ved eksamen i et eller andet omfang.....*

Bk: *Jo men vores eksamen er stadigvæk tænke opgaver det er sådan koble kan du koble*

Fh: *Det kunne man jo altså også lave i sådan en ikt-test (B og Bk: kan man det?) hvorfor skulle man ikke kunne det? (B: det vil jeg fandme gerne se et eksempel på)*

### **Fortolkning:**

Det billede, der tegner sig i henhold til faglige test er ikke entydigt. Indledningsvis ser det ud som om der er tilslutning til faglige test. Og i den forbindelse er der også enighed om nytten af paratviden. Men efterfølgende bliver der så problemer med at få det anvendelsesorienterede ind i testene. De nationale test i folkeskolen bliver så trukket frem som et eksempel, der rummer et potentiale og kunne det suppleres med computerspil vil det være optimalt. Den enighed, der er om brugbarheden af faglige test hører så op, når det kommer til indførelsen af nationale test i gymnasiet. Her falder to af informanterne fra og selvom argumenterne mod de nationale test kan være relevante nok, kan det også tænkes, at informanternes forforståelse omkring de nationale test, der har været udsat for megen kritik fra lærerkredse spilder ind her. Altså på grund af den megen kritik, der har været, har informanterne en kritisk forståelse af nationale test.

#### 4. Forskningsspørgsmål: Hvad er holdningen til læringsstilttest hos naturvidenskab/humaniora?

##### **Kondensering:**

Informanterne bliver spurgt om de bruger test af læringsstile i undervisningen. Bk og Fh svarer direkte nej til det, men Bk siger så: *Men indirekte bruger vi det i vores to fag. Fordi vi har jo som sagt, der er nogen, der er nogen taktiler, derfor laver i laboratoriearbejde* og B siger næsten noget tilsvarende. Intervieweren replicerer hertil **Men man laver vel ikke eksperimentelt arbejde af hensyn til "taktilerne". Man laver det vel fordi det er en del af faget?** Men B forsvarer synspunktet med at man dybest set kunne læse det hele i bogen og så lave "tankeeksperimenter", men det vil ikke være det samme som at have det i hånden. Og Bk supplerer med, at *Det er noget med at have det i hænderne, vi tror på det fremmer forståelsen ik.*

B henviser til første gang han læste om læringsstile og blev meget interesseret i det, men blev så senere i tvivl om, hvad han kan bruge det til. Efter hans opfattelse baserer ideen om læringsstile sig på undervisning i mindre grupper evt. projektgrupper. Og gruppesammensætning efter læringsstile er efter hans opfattelse problematisk, hvis/når eleverne skal være medbestemmende. Som en bemærkning hertil kæder Bk spørgsmålet om læringsstile sammen med elevfrafaldet på skolen: *Men et eller andet sted, tror jeg sgu det er vigtigt, fordi det er da ikke tilfældigt, at vi år efter år mister elever fordi vi står og kæfter op og skriver ik.*

B tilslutter sig dette synspunkt, og ser desuden en mulighed for, at eleverne kan have gavn af, at kende deres foretrukne læringsstil: *Men hvis to elever finder ud af, de har faktisk nogen læringsstile, der ligner hinanden meget og som kunne betyde, at det er tænkeligt de kunne få et konstruktivt samarbejde ud af det. Det kunne bruges til noget.*

Intervieweren stiller derefter spørgsmålet: **Vil der ikke også være nogle af de her læringsstile, der ville lide en krank skæbne ift. At når alt kommer til alt så er det vi gør i gymnasiet at dosere boglig, intellektuel, logisk-matematisk viden.**

Informanterne giver lidt forskellige svar på spørgsmålet. Bk mener ikke vi til gode ser dem med musikalsk intelligens nok. B mener, at vi i undervisningen, med forskellige virkemidler, gør hvad vi kan for, at til gode se, i hvert fald nogle intelligenser. Fh er inde på, at man i visse tilfælde godt kan bruge drama som virkemiddel.

På spørgsmålet om: **eleverne kan have gavn af at kende deres læringsstil?** Vender Fh tilbage til evalueringskulturbegrebet og siger: *Nej der mener jeg, det vi sagde om evalueringskultur før altså eleverne det kan godt være, vi skal have en evalueringskultur, men eleverne skal jo også have en evalueringskultur, de skal jo have en forståelse for det som redskab, de kan bruge for altså det er ikke bare fordi vi skal give en karakter, men fordi de skal finde ud af, hvordan de kan lære bedre i nogen sammenhænge ik. Det er vi jo nødt til at hjælpe dem med på en eller anden måde så på den måde, synes jeg i virkeligheden det er vigtigere end alt muligt andet dybest set, fordi har de først fået fod på det, så er de jo selvhjulpne langt hen ad vejen. Så kan vi gøre meget skade, men alligevel kan de selv redde den. Så der er da helt centralt den del der, synes jeg faktisk. B gentager sit tidligere synspunkt om, det kan være meget nyttigt for eleven.*

Foranlediget af spørgsmål fra intervieweren diskuteres det, om undervisning efter læringsstile i sin yderste konsekvens ikke kan medføre en stærk individualisering af undervisningen og om det er ønskeligt? informanterne når ikke rigtig frem til en klar stillingtagen til spørgsmålet. B giver dog udtryk for følgende: *Men det som Dunn og Dunn jo kritisere, det er jo, at vi i den vestlige verdens undervisningskultur, der har vi så at sige tvunget folk ind, presset folk ind i den ramme og så kan man sige jamen det vil vi gerne gøre, men så er prisen, at der er nogen der bliver hægtet af.... Og så kan man sige, hvis ideen så er vi skal have nogen færre hægtet af, så er det eneste korrekte svar på det, at vi må gøre både og.*

### **Fortolkning:**

Der synes at være enighed om, at eleverne kan have glæde af kende til deres foretrukne læringsstil. Men det fremgår ikke helt klart om de tre informanter er enige, om hvor brugbart et redskab det er for læreren. I øvrigt er der et tolknings- og analysemæssigt problem i det forhold, at der tales i flæng om læringsstile og multiple intelligenser. Det er ikke altid klart om informanterne taler om det ene eller det andet, og det kan jo så igen have noget med forforståelsen at gøre.

### 5.1.3 Databehandling – det færøske gymnasium

Det transkriberede interview kan ses i bilag 11.

**1. Forskningsspørgsmål: Hvordan ser evalueringskulturen i gymnasiet ud i dag set fra henholdsvis naturvidenskab/humaniora?**

#### **Kondensering:**

Der fremkommer en række forskellige udsagn fra informanterne, der både direkte og indirekte fortæller noget om evalueringskulturen for eksempel følgende:

*H: Arbejder mere resultatorienteret i stedet for evalueringsorienteret, jeg kører efter resultater i timen. Søger i den daglige undervisning efter andre former for værdier, som ikke er evalueringsværdier, en didaktiker siger når vi ser på Gardner der er ingen evalueringsformer som duer, skab læringsmiljø der er vores mål.*

*Vi har en klassisk evalueringskultur til den rigtige eksamen.*

*....ingen fælles strategi for hvordan evalueringskulturen skal være på vores skole, vi mangler den pædagogiske inspektør.*

*M: snakker med eleverne hver især efter de har fået karaktererne, det får man meget ud af.*

*Vi er bundet på hænder og fødder ovenfra, karaktersystem, datoer, terminer, prøver. Der er en årsag til at vores kultur er udviklet som den er og er slet ikke trit med udviklingen udenfor skolen.*

Og til sidst i interviewet fremkommer så følgende udsagn i forbindelse spørgsmålet om læringsstile:

*Mht. dokumentation og evaluering, det er et princip om gennemsigtighed og åbenhed at nogen personer arbejder på en meget åben måde og lægger alting ud til eleverne også med hensyn til evalueringen at man lægger uv-forløb på nettet og skriver om det gik godt eller dårligt, så kan man så tage og genbruge og det ligger frit for alle at genbruge. Det ville være en evaluerings-kultur, og en modsætning af den privatpraktiserende. Når man starter op som gymnasielærer her på stedet har man ikke adgang til ens kollegers skrivebord. Man har ikke mulighed for at evaluere på deres forløb, men har heller ikke mulighed for at se hvad de synes.. man kender dem ikke..så det princip om gennemsigtighed og åbenhed ville være interessant.*

#### **Fortolkning:**

Udsagnene indikerer, at lærerne efterlyser en evalueringskultur måske lige bortset fra H, der øjensynlig har en anden opfattelse af, hvad evaluering indebærer. I virkeligheden er der nok tale om, at den evalueringskultur som er ekspliceret i form af den klassiske, der udmøntes i den rigtige eksamen, som det siges i citatet ovenfor forekommer utilstrækkelig eller ude af trit med virkeligheden.

Det sidste citat, der taler om åbenhed og gennemsigtighed og bruger det som modstilling til den "privatpraktiserende" peger på en af hovedhjørnestene i gymnasireformen som også er et opgør med den "privatpraktiserende lærer". Konkret sker det ved at stille krav om tværfagligt samarbejde, såvel i grundforløb som i studieretningsforløb (se afsnit 4.2). I virkeligheden er det måske ligeså meget den privatpraktiserende lærer, der har været en hindring for etablering af en evalueringskultur som det er den overordnede eksamensrelaterede evaluering.

<b>2. Forskningsspørgsmål: Hvilke muligheder ser undervisere i gymnasiet ved IKT-baseret evaluering?</b>
--

**Kondensering:**

Der tegner sig forskellige temaer i de udsagn der forekommer: På den ene side IKT-basering indebærer nogle fordele, på den anden side IKT kan ikke erstatte samtalen, den menneskelige kontakt.

Den IKT-baserede evaluering er en tidskrævende proces. Der mangler ressourcer/software.

*Der er bestemt noget det kan bruges til og bestemt noget som er tidsspilde at begynde at skruer evaluering sammen med IKT. Man kan spørge om nogen ting og gøre nogen ting men ikke alting.*

*Computeren er stærk til multiple choice, fx, alle elever kan forholde sig til nogle spørgsmål som man kan lave statistik på, nogen ting ville være hurtige at undersøge. fx spørgsmål fra ledelsen om noget som vedrører alle, så kan man få hurtige svar, og quick poll på en hjemmeside.*

*Det smart at kunne graduere o.a. man får også meget ud af en løs snak i timen, det andet kan man måle, det er bedst, at de også kan lave kommentarer (i spørgeskema) og ikke bare afkrydse (fritekst, essay).*

*IKT kan ikke slå den månedlige samtale med eleven, det giver en ansvarlighed som ikke kan måles, det giver større udbytte.*

*Jeg mener den menneskelige kontakt er meget vigtig, it er egnet til bestemte ting....*

*Vi mangler integration imellem systemerne, sådan at tingene arbejder sammen. (det kan man andre steder med fx, fronter, spørgeskema, gruppedokumenter, det har vi ikke her). Har kompetencerne, men ikke ressourcerne til at betale for den enkelte gør det.*

*Hvis vi havde en fast politik ville hjælpe på evalueringsområdet, hvis vi fik lidt mere introduktion til Quia, hvordan skal man gøre, metodemæssigt og det faglige. vi har ingen reduktion, årsværk*

### **Fortolkning:**

Det indtryk, der står tilbage er, at informanterne ikke er vildt imponerede over de muligheder IKT-baseringen kunne tilføjer evalueringen. Informanterne anerkender, at IKT kan gøre nogle ting lettere, men på den anden side er det også tidskrævende. Informanterne ser det fysiske møde, samtalen, som et stærkere medie i evalueringen. Tankevækkende er det, at det ikke fremgår, om der kunne være nogen problemer ved den mundtlige evaluering, samtalen. Endelig er der et ressourceproblem med mangelfulde IKT-systemer og manglende integration. I det sidste citat antydes evalueringskultur diskussionen igen. Med en mere veldefineret evalueringskultur kunne problemerne med software og ressourcer måske undgås.

<b>3. Forskningsspørgsmål: Hvad er holdningen til faglige test hos naturvidenskab/humaniora?</b>
--

### **Kondensering:**

Der fremkommer nogle udsagn, hvor der direkte siges noget om, hvad IKT-baserede test kan bruges til, men der fremkommer også nogle udsagn, der ikke handler om IKT-baserede test, men derimod om nogle udviklingstendenser som peger i retning af, at der skal arbejdes mere med konkrete faglige kompetencer, der kan måles ved hjælp af test.

*Faglige test er sikkert brugbare i sprogfag og naturvidenskabelige fag, men jeg ikke kan forestille mig det i historie, dansk, færøsk, samfundsfag i særlig stor stil. Jeg kan forestille mig det i sproget dansk stil, men ikke indholdsmæssigt.*

*Kunne bruges i fysiologien i idræt, hvor der er et bestemt fastlagt indhold, der ikke er til diskussion. Efter reformen er 20 % at karakteren direkte afledt af, hvor god eleven er til at fremstille sine tanker. Det er en betydelig del. Der er it overhovedet ikke egnet, heller ikke i de matematiske spørgsmål, er nemmere at håndtere i hånden, umuligt for 99% af eleverne at gøre på computer.*

*Der sker en tilnærmelse mellem naturvidenskab og humaniora. På det sidste møde dansklærerne havde med fagkonsulenten, blev der præsenteret en ny type opgave, hvor der er mere stringent faglige for, at de skal kunne måles, det skal ikke være sådan, at det kun er almindelse, der skal være en faglig kompetence som man også skal kunne dokumentere ved hjælp af nogle konkrete spørgsmål.*

*Det er evalueringen der kommer til at styre undervisningen.*

*Dem der præger vores kultur engelsk, amerikansk, fransk har en benhård evalueringskultur (test) ikke baseret på læring, men at klare de prøver, edb er vejen til at skabe standard (på godt og ondt).*

*Den skandinaviske tradition er problemorienteret i samfundsfag, hvis det andet kommer ind er det ikke den samme refleksion over det du arbejder med.*

*Systemet er under angreb.*

*Bliver testet i amerikansk kultur og ikke skandinavisk, internationalisering og globalisering.*

### **Fortolkning:**

Der tegner sig et billede af, at sprogfagene og de naturvidenskabelige fag kan have gavn af faglige test, og det antydes, at det er ud fra en forestilling om, at indholdet i disse fag er af en karakter som gør det egnet til test. I forlængelse heraf ser dansklæreren en tilnærmelse mellem humaniora og naturvidenskab, som ikke er ønskelig fordi det vil gå ud over det almindelige formål med dansk. Ræsonnementet er, at fokus på faglige kompetencer er uforeneligt med almindelse. Det er interessant fordi som vi har redegjort for i afsnit 4.2 så er en af intentionerne med gymnasireformen netop at almindelsen skal fremmes via de faglige kompetencer. Endelig markeres den holdning, at den angelsaksiske testkultur er en trussel mod den skandinaviske tradition. Der er en modsætning mellem refleksion og problemorientering og test(kultur)

<b>4. Forskningsspørgsmål: Hvad er holdningen til læringsstilttest hos naturvidenskab/humaniora?</b>
--

### **Kondensering:**

Diskussionen om læringsstile bliver i lige så høj grad en diskussion om læringsmiljø:

*Det er overfyldt (folkeskolen) kompetencebegrebet er misbrugt og kommet hen et sted, hvor det ikke hører hjemme, det er interessant, men der er andet der er mere interessant, skab et godt læringsmiljø, det er fint man bruger elevens forskellige kompetencer, men jeg synes man har skudt lidt for højt i forhold til mange andre ting som er mere interessante for at skabe et godt læringsmiljø.*

*Det er utroligt vigtigt, at der kan opstå en faglig debat og et social miljø i en klasse, det er det vigtigste. Hvis du ikke kan skabe et reflekterende miljø, det får du først når eleverne er trygge ved hinanden.*

*Hvis vi skal tale om forskellige stile i undervisningen, jamen så handler det om at skabe et miljø som man er tryk ved, om der kommer noget faglig er mere ligegyldigt i starten, redskabet til indlæring er miljøet, fælles ansvar, så kan du skabe undervisning, om du laver foredrag eller hvad som helst, fordi du har skabt et fagligt skabt socialt miljø.*



*Hvis eleverne er opmærksomme på det med læringsstilen og man har testet dem og gjort opmærksom på dem og undervist dem i læringstile, den bevidstgørelse vil eleverne kunne bruge for at få alle eleverne med på en anden måde.*

*Du opnår ikke bare et læringsmiljø bare ved en formel....Det kan ikke bare bygge på en lærers erfaring.*

*Meget mere interessant at bruge midler og energi på at få det frem i stedet for 20 million på at få IKT ind eller en tvungen kultur.*

### **Fortolkning:**

Der ser ud som om informanterne anerkender, at eleverne kan have gavn af, at kende deres læringsstile, men der gives samtidigt udtryk for, at læringsmiljøet (i klassen) er det vigtigste middel i undervisningen. Det er forudsætningen for, at der kan opstå et reflekterende miljø. Det sidste citat bringer os samtidigt tilbage, der hvor vi begyndte med evalueringskulturen. Der bliver her sat en tyk streg under, at evalueringskultur og IKT-basering er noget sekundært i forhold til læringsmiljøet.

### **5.1.4 Opsamling**

Der er markant forskel på den måde de to fokusgrupper fungerer på. I den "jyske" gruppe kan meningsudvekslingen og meningsforhandlingen klart anskues ud fra en socialkonstruktivistisk diskurs. Det understreges blandt andet af de mange øh'er, de mange afbrydelser og talen i munden på hinanden samt ikke mindst de mange henvisninger til tidligere udtalelser i interviewet. Det fortæller samtidig noget om, at evalueringskulturen og synet på, hvordan evaluering kan eller skal foregå i høj grad er noget der forhandles og indgår i en løbende diskussion.

Den "færøske" gruppe fungerer markant anderledes, meningsforhandlingen er ikke nær så udpræget som i den jyske gruppe. Det virker som om, at præmissen nævnt i afsnit 4.4.1: "Fokusgruppen supplerer hinandens udtalelser, og på den måde sker der en form for verificering og nuancering af deltageres udsagn, da deltagerne på sin vis holdes i skak af de øvrige deltageres udtalelser" ikke helt holder her. Set i bakspejlet mangler der nok en modererende indsats her og det betyder, at validiteten af interviewet ikke er så stor som man kunne ønske sig.

## 5.2 Faglige test – analyse

Resultaterne af de faglige test analyseres såvel kvantitativt som kvalitativt. Indledningsvis gives en opsummerende beskrivelse af resultaterne af de to test hver for sig.

### 5.2.1 Biologi

Den egentlige test er som nævnt udført i to klasser med i alt 50 elever. 2 elever har gennemført testen uden at svare på nogle af spørgsmålene og en elev har i løbet af 5 min besvaret alle spørgsmål vel at mærke med 94 % rigtige svar. Vi har at udelade disse tre fra den videre resultatbehandling. Resultaterne af testen er vist i bilag 12. Som det fremgår af afsnit 4.2.3, er 6 af spørgsmålene af typen ”multiple correct” og pointgivningen i Quia fungerer på den måde, at der kun tildeles point ved korrekt svar dvs. hvis der kun svares rigtigt på 2 ud af i alt 3 rigtige så tildeler programmet 0 point. Point tildelingen for så vidt angår multiple-correct spørgsmål, er derfor sket manuelt efter, at eleverne har afleveret testbesvarelsen, sådan at eleverne får et antal point, der svarer til antallet af rigtige svar. For eksempel gælder for spørgsmål 8, at der tildeles 4 point for 4 rigtige svar, 3 point for 3 rigtige svar osv.

Question 8 (Multiple correct):		Average score: 1.5 out of 4 points
Hvad er den fysiologiske årsag til diarre (tynd afføring)? (Afkryds alle de svar du mener kan være rigtige)		% answered
<b>Possible answers</b>		
[A] Meget fiberrig kost		7%
[B] For meget vand i kosten		14%
• [C] Bakteriefektioner i tarmsystemet		61%
• [D] Der suges vand ind i tarmsystemet fra de omgivende væv/organer		14%
• [E] Virusinfektioner i tarmsystemet		46%
[F] For meget salt i maden		4%
• [G] Laktoseintolerance		29%
<i>blank answer</i>		21%

Tabel 5.1: Faglig test biologi spørgsmål 8. De sorte prikker til venstre angiver de rigtige svar.

Hvor godt eleverne har formået at besvare de enkelte spørgsmål er vist nedenfor i en samlet opgørelse over, hvor mange procent af eleverne der har fået maksimum point i de enkelte spørgsmål:

Spørgsmål	Type	Point	% elever med max.point
1	Multiple choice	1	71 %
2	Matching	3	44 %
3	Ordering	3	23 %
4	Pop-up	1	62 %
5	Matching	3	48 %

6	Multiple choice	1	37 %
7	Multiple correct	3	10 %
8	Multiple correct	4	6 %
9	Multiple choice	1	77 %
10	Multiple correct	2	29 %
11	Multiple correct	3	12 %
12	Multiple correct	2	25 %
13	Pop-up	1	38 %
14	Multiple correct	4	6 %
15	Multiple choice	1	33 %
<b>Gennemsnit</b>			<b>35 %</b>

Tabel 5.2: Biologielevernes besvarelse af de enkelte spørgsmål

Der viser sig at være en ganske stor spredning på, hvor mange elever, der får maksimum point i de forskellige spørgsmål. Af de 4 Multiple choice sp. må det konstateres at to har været ”svære” og to har været ”lette” at besvare. Til gengæld må alle multiple correct sp. betegnes som vanskelige idet under en 1/3 er i stand til at besvare dem fuldstændig korrekt. Ser vi på sp. 6, 8, 11 og 14 som repræsenterer de højeste taksonomiske niveauer er det karakteristisk, at kun en lille del af eleverne svarer helt rigtigt. For spørgsmål 6 vedkommende, er der tale om et multiple-choice spørgsmål (med tre svarmuligheder), hvor man må forvente, at nogle elever gætter svaret, så det kan være en del af forklaringen på, at eleverne klarer denne opgave noget bedre end de tre øvrige. Med hensyn til elevernes total score (i % af det maksimale antal point) er det sådan, at de to elever der har scoret højest har 82 %. De to elever, der har scoret lavest har scoret henholdsvis 9 og 3 %. Gennemsnits-scoren er på 53 %.

Sammenligner vi spørgsmål 1 og 9, der begge er multiple choice sp. er der i begge tilfælde tale om høje korrekte svarprocenter (se nedenfor). Sp. 1 er kategoriseret på højeste taksonomiske niveau, hvor eleven skal overføre viden fra teori til praksis. Hvorimod sp.9 er et rent paratviden spørgsmål. Hvis alle elever besvarer sp.1 som rent gæt vil man forvente, at 33 % af eleverne svarer rigtigt. Derfor må vi gå ud fra, at en stor del af eleverne har besvaret spørgsmålet ved at resonere sig frem til svaret. Dog skal det med i betragtning, at eleverne har adgang til hjælpemidler under testen. I sp. 9 ville man forvente, at ca. 25 % af eleverne svarer rigtigt, hvis der er tale om rene gæt, så her kan de mange rigtige besvarelser altså heller ikke forklares alene af gæt, men i og med, at der er tale om et rent paratviden spørgsmål antager vi, at en del elever har fundet svaret ved hjælp af hjælpemidlerne.

Question 1 (Multiple choice):

Average score: 0.712 out of 1 point

Hvilke af følgende svar passer bedst med nedenstående udsagn:  
 En analyse af en persons daglige kost (en kostanalyse) kan bruges til:  
 Possible answers

% answered

[A] At vurdere om personen taber sig eller tager på	10%
• [B] At vurdere om personens kostindtag er i overensstemmelse med den anbefalede næringsstoffordeling	71%
[C] At vurdere personens daglige energibehov	8%
<i>blank answer</i>	12%

---

**Question 9 (Multiple choice):** Average score: 0.769 out of 1 point  
Hvad er proteiner opbygget af?  
**Possible answers** **% answered**

• [A] aminosyrer	77%
[B] monosakkarider	0%
[C] nukleinsyrer	2%
[D] Fedtsyrer	2%
[E] kulhydrater	6%
<i>blank answer</i>	13%

Tabel 5.3: Biologifaglige test, spørgsmål 1 øverst og 9 nederst. Prikker = rigtigt svar

I Spørgsmål 15 derimod, er der tegn på, at en del af eleverne har gættet på et svar, idet kun en 1/3 har angivet det rigtige svar og resten af svarene fordeler sig nogenlunde ligeligt på de øvrige svar kategorier.

<b>Question 15 (Multiple choice):</b> <span style="float: right;">Average score: 0.327 out of 1 point</span> De muskelbevægelser, der får føden til at bevæge sig gennem fordøjelsessystemet kaldes: <b>Possible answers</b> <span style="float: right;"><b>% answered</b></span>	
• [A] Peristaltiske bevægelser	33%
[B] Nedadgående bevægelser	13%
[C] Koncentriske bevægelser	15%
[D] Cirkulære bevægelser	12%
[E] Længdegående bevægelser	17%
<i>blank answer</i>	10%

Tabel 5.4: Biologifaglig test, spørgsmål 15. Prik = rigtigt svar

### Opsamling:

Opsummerende kan vi konstatere, at spørgsmål af typen multiple correct er vanskelige for eleverne at besvare. Det er formentlig ikke alene denne type, der gør det svært, fordi vi kan se at også andre typer spørgsmål kan være vanskelige, men en kombination af flere faktorer herunder også det taksonomiske niveau, abstraktionsgrad, antallet af svaralternativer og deres relevans.

### 5.2.2 Musik

Den egentlige test er i musik ligeledes udført i to klasser med i alt 25 elever. Som det fremgår af afsnit 4.2.4 er der 20 spørgsmål i musiktesten, hvoraf de 13 er multiple choice, 3 fill-in, 2 matching og 1 ordering. I samtlige spørgsmål "belønnes" et rigtigt svar med 1 point, eleverne får dog ikke feedback om dette. Testen er ikke afprøvet i et pilotforsøg. Det har derfor ikke været muligt at justere på testspørgsmålene inden den egentlige test.

Nedenfor er vist, hvor godt eleverne har formået, at besvare de enkelte spørgsmål i en samlet opgørelse over, hvor mange procent af eleverne der har fået point i de enkelte spørgsmål (testresultaterne er i øvrigt vist bilag 13):

Spørgsmål	Type	Point	% elever med max.point
1	Multiple choice	1	16%
2	Fill-in	1	84%
3	Fill-in	1	48%
4	Fill-in	1	72%
5	Matching	1	0%
6	Matching	1	0%
7	Multiple choice	1	44%
8	Multiple choice	1	32%
9	Multiple choice	1	24%
10	Multiple choice	1	24%
11	Multiple choice	1	16%
12	Multiple choice	1	16%
13	Multiple choice	1	4%
14	Multiple choice	1	56%
15	Multiple choice	1	72%
16	Ordering	1	9%
17	Multiple choice	1	72%
18	Multiple choice	1	56%
19	Multiple choice	1	53%
20	Multiple choice	1	61%
<b>Gennemsnit</b>			<b>38%</b>

Tabel 5.5: Musikelevernes besvarelse af de enkelte testspørgsmål

Som det fremgår af tabellen, er der i gennemsnit 38 % af eleverne der får point i de 20 sp. Det dækker dog over meget stor spredning på spørgsmålene. I øvrigt er det bemærkelsesværdigt, at ingen elever har været i stand til at svare på de to ”matching” spørgsmål.

Ser man nærmere på elevernes besvarelse af spørgsmålene viser det sig, at alle elever faktisk matcher en eller flere begreber og navne i sp. 5, men som pointtildelingen fungerer belønnes man altså kun, hvis alle match er rigtige.

Sp.5	Begreb	Korrekt match	% elever som matcher korrekt
	Ragtime	Scott Joplin	54%
	New Orleans	Louis Armstrong	31%
	Blues	Robert Johnson	23%
	Swing	Duke Ellington	23%
	Be Bop	Charlie Parker	38%

Tabel 5.6: Musikelevernes besvarelse af spørgsmål 5

I sp. 6 derimod, er det ét match som ingen af eleverne er stand til at foretage rigtigt. Og sammenligner man i øvrigt de to spørgsmål fremgår det, sp.6 er noget sværere at besvare end sp.5.

Sp.6	Begreb	Korrekt match	% elever som matcher korrekt
	Ragtime	Synkoper	8%
	New Orleans	Two Beat	25%
	Blues	Shuffle	0%
	Swing	Big Band	58%
	Be Bop	Alteratíonir	25%

Tabel 5.7: Musikelevernes besvarelse af spørgsmål 6

I afsnit 4.2.4 er sp. 5 tillagt sværhedsniveau 5 og sp. 6 sværhedsniveau 4, så det var på forhånd ventet, at eleverne ville have svært ved at besvare spørgsmålene. Til gengæld er det så en smule uventet, at eleverne har svære ved at besvare sp.6 end sp. 5. Spørgsmål 10, 12 og 13 er kategoriseret til sværhedsniveau 2, alligevel ser vi, at eleverne har forholdsvis svært ved at besvare dem. Der er 5 svarmuligheder og som det fremgår af sp. 13 nedenfor er det meget få elever, der kender det rigtige svar. Så noget tyder på, at der ikke er overensstemmelse mellem forventet sværhedsniveau og reelt sværhedsniveau. Tallene viser i øvrigt, at ren gætning ikke synes, at forekomme i særlig stor udstrækning. En samlet vurdering af alle multiple choice spørgsmål indikerer, at ren gætning synes i det hele taget ikke at forekomme i særlig stor udstrækning.

13. Hvør av hesum svarmöguleikum eyðkennir bluesformin?

Possible answers	% answered
[A] AABA	80%
[B] AA	4%
• [C] AAB	4%
[D] ABA	4%
[E] ABAC	12%

Tabel 5.8: Musikelevernes besvarelse af spørgsmål 13. Prik = rigtigt svar.

Spørgsmål 12 viser et eksempel på, at det kan være svært at angive sandsynlige ”distractors”. Vi henviser i afsnit 3.4 til Rust & Golombok (1999), der anbefaler, at man frasorterer distractors, som ikke bruges. Men det har ikke været muligt her, idet testen som nævnt ikke er pilottestet.

12. Hvat fyri ljóðföri spældi Duke Ellington á?

Possible answers	% answered
[A] trompet	37%
• [B] piano	24%
[C] guitar	0%
[D] saxofon	28%

---

Tabel 5.9: Musikelevernes besvarelse af spørgsmål 12

Det er således et dilemma, at der gerne skal være mange svaralternativer, for at undgå gætning. På den anden side er det svært at konstruere mange relevante svaralternativer.

I spørgsmål 16 (se nedenfor) skal eleverne ordne begreber i rigtig rækkefølge, hvilket 2 ud af 25 elever er i stand til. De andre elever kan i princippet have placeret en eller flere rigtige begreber, uden at have fået point for det. Det skal endvidere med i betragtning, at spørgsmålet hører til sværhedsniveau 4, derfor er det forventeligt, at eleverne vil have svært ved at svare rigtigt.

---

16. Hvussu er tíðarraðfylgjan á hesum jazz-sjangrum (set tí elstu ovast)?

new orleans  
swing  
bebop  
cool  
modal

---

Tabel 5.10: Spørgsmål 16 i musiktesten.

### Opsamling

Selvom der er stor spredning på, hvor mange procent af eleverne, der svarer rigtigt på de enkelte spørgsmål, så er spredningen på eleverne ikke så stor. Den elev, der score mest opnår 60 %, den der score mindst opnår 20 % rigtige og den gennemsnitlige scoreprocent er 38. En forklaring på, at eleverne ikke differentieres mere kan være, at pointtildelingen i den ovennævnte opgavetype ikke differentierer efter, hvor mange rigtige match eleverne har. En anden forklaring er, at forholdsvis mange af spørgsmålene er af multiple choice typen som i sagens natur kun kan give 1 point alternativt 0 point.

Bortset fra sp. 10, 12 og 13, hvor der som omtalt ovenfor er uoverensstemmelse mellem forventet sværhedsniveau og reelt sværhedsniveau, tyder resultaterne på, at der er nogenlunde overensstemmelse mellem forventet og reelt sværhedsniveau.

### 5.3 Læringsstile – analyse

Man kan altid diskutere, hvor mange spørgsmål, der skal indgå i undersøgelse af læringsstile før, den kan være valid og reliabel. Vi mener, at med 12 henholdsvis 13 spørgsmål til hver kategori af læringsstilene, så må vi være dækket nogenlunde ind. Vi vil i det følgende analysere på elevernes foretrukne læringsstil med henblik på at undersøge om der er mønstre og sammenhænge i de forskellige præferencer. Forekomsten af en læringsstil hos en elev er i undersøgelsen sket ved pointtildeling, således at svarer en elev *enig* til et spørgsmål/udsagn så får eleven et point med hensyn til den pågældende læringsstil spørgsmålet hører til. Svarer eleven uenig, eller undlader at svare, får man ingen point. En læringsstil kan altså være helt fraværende eller relativt tilstede. Resultaterne af læringsstilsundersøgelserne er vist i bilag 14.

#### Biologielever

I afsnit 4.1.2 anføres det, at Dunn mener, der er dokumentation for at flertallet af akademiske studerende hører til den holistiske type. Det er en opfattelse, der deles af Lena Boström som fremfører, at flertallet af elever i slutningen af puberteten er globale/holistiske (Damberg 2006:351). Vores data støtter ikke dette, idet den analytiske læringsstil "score" flere point end den holistiske målt på gennemsnittet af alle elever:

	Holistisk	Analytisk
gennemsnit	6,62	8,47
standardafvigelse <sup>40</sup>	2,23	2,03

Tabel 5.11: Biologielevernes gennemsnitlige præference for holistisk og analytisk læringsstil.

På trods af den relativt store spredning, viser en signifikanstest<sup>41</sup>, at der er signifikant forskel på de to gennemsnitsværdier. Det kan dog ikke afvises, at en større stikprøve ville give et andet resultat. Nedenfor er eleverne sorteret efter holistisk præference med de 7 elever, der score højest til venstre og de 7 elever der score lavest til højre. Den holistiske præference er så korreleret til henholdsvis

<sup>40</sup> I alle de analyser, hvor vi har lavet gennemsnitsberegninger har vi samtidigt beregnet standardafvigelse. Denne parameter siger noget om, hvor spredte de enkelte værdier er i forhold til gennemsnittet. Et stort tal indikerer en stor spredning, et mindre tal en mindre spredning. Beregningerne er foretaget med den automatiske beregningsfunktion i regnearket Excel.

<sup>41</sup> I forbindelse med analysen af de kvantitative data, har vi i nogle tilfælde undersøgt om der er statistisk signifikant forskel på de målinger vi sammenligner. Dette er i alle tilfælde sket ved vha. t-test, der er en statistisk metode, der sammenligner to middelværdier eller gennemsnit. Testen tager udgangspunkt i en hypotese, der siger, at de to middelværdier er ens. Testen måler så, hvor sandsynligt det er. Det sker ved at beregne en signifikanssandsynlighed og i alle de test vi har lavet er denne 5%. Dvs. hvis testen beregner, at sandsynligheden er mindre end 5% så forkastes hypotesen ergo er middelværdierne/gennemsnittene forskellige. Er sandsynligheden dermed større end 5% så opretholdes hypotesen, dvs. så er middelværdierne/gennemsnittene ikke forskellige. Alle beregninger er foretaget med den automatiske beregningsfunktion i Excel.



auditiv, visuel og kinæstetisk/taktil præference. Resultaterne viser, at der er en korrelation mellem holistisk og visuel henholdsvis kinæstetisk/taktil præference, idet gennemsnittet for visuel og kinæstetisk/taktil præference i de to elevgrupper er testet signifikant forskellige, hvorimod der ikke er signifikant forskel på den auditive præference hos de to elevgrupper. Dette resultat støtter meget godt vores beskrivelse af den "holistisk lærende" i afsnit 4.1.2.

Elev Nr.	Holistisk	Auditiv	Visuel	Kinæstetisk	Elev Nr.	Holistisk	Auditiv	Visuel	Kinæstetisk
4	12	7	9	8	37	4	4	7	5
10	11	6	11	9	32	4	9	6	5
31	11	9	9	11	17	4	5	8	5
22	10	7	11	9	14	4	9	10	5
7	10	8	10	9	20	4	5	8	7
26	9	10	10	8	18	3	4	8	7
12	9	9	8	9	2	3	10	7	8
<b>Gennemsnit</b>		<b>8</b>	<b>9,71</b>	<b>9</b>			<b>6,57</b>	<b>7,71</b>	<b>6</b>
<b>Standardafvigelse</b>		<b>1,41</b>	<b>1,11</b>	<b>1</b>			<b>2,64</b>	<b>1,25</b>	<b>1,29</b>

Tabel 5.12: Korrelation mellem læringsstile biologielever

Laver vi den samme korrelation blot sorteret efter analytisk præference er der ingen signifikant forskel.

### Musikelever

Vi har foretaget de samme analyser for musikeleverne som for biologieleverne. Angående fordelingen mellem holistiske og analytiske præferencer viser den samme tendens sig som hos biologieleverne, baseret på sammenligning af de gennemsnitlige score, er der flere analytiske end holistiske og forskellen er også her signifikant.

	Holistisk	Analytisk
gennemsnit	7,09	8,26
standardafvigelse	2,78	1,81

Tabel 5.13: Musikelevernes gennemsnitlige præference for holistisk og analytisk læringsstil

Til gengæld er der ikke de samme korrelationer mellem holistisk og visuel henholdsvis kinæstetisk/taktil præference som findes hos biologieleverne. Der er ganske vist tegn på, at det samme gør sig gældende, men der er ikke tale om signifikant forskel. Stikprøven (antallet af musikelever) er dog betydeligt mindre en antallet af biologielever, derfor skal resultaterne tages med forbehold.

Elev Nr.	Holi-	Auditiv	Visuel	Kinæste-	Elev	Holi-	Auditiv	Visuel	Kinæstetisk
----------	-------	---------	--------	----------	------	-------	---------	--------	-------------

	stisk			tisk	Nr.	stisk			
6	11	8	10	10	8	4	10	11	8
9	11	7	8	4	12	4	10	6	7
14	11	7	8	8	21	4	6	9	5
1	10	8	9	8	23	4	7	6	8
5	10	5	7	7	4	3	8	3	5
11	10	9	10	9	10	3	7	8	8
<b>Gennemsnit</b>		<b>7,33</b>	<b>8,67</b>	<b>7,67</b>			<b>8,00</b>	<b>7,17</b>	<b>6,83</b>
<b>Standardafvigelse</b>		<b>1,37</b>	<b>1,21</b>	<b>2,07</b>			<b>1,67</b>	<b>2,79</b>	<b>1,47</b>

Tabel 5.14: Musikelevernes gennemsnitlige præference for holistisk og analytisk læringsstil

### Opsamling

Sammenfattende for læringsstilsundersøgelsen af de to elevgrupper kan vi konstatere, at vi med hensyn til analytisk og holistisk læringsstil får nogle andre resultater end de resultater vi kan sammenligne med fra litteraturen. Vi kan ikke give nogen plausibel forklaring på, hvorfor vi finder flere elever med analytisk præference end med holistisk præference, bortset fra, at en større stikprøve muligvis ville give et andet resultat. De to elevgrupper ser endvidere ud til at være ret ens med hensyn til læringsstilspræference. Der er i begge grupper tegn på, at den holistiske læringsstil følges af henholdsvis kineæstetisk/taktil og visuel præference i begge grupper. I biologigruppen er sammenhængen testet statistisk signifikant. Denne sammenhæng er i overensstemmelse med den beskrivelse af den "holistisk lærende" vi har beskrevet i afsnit 4.1.2.

### 5.4 Multiple intelligenser – analyse

Som redegjort for i afsnit 4.1.3 har af praktiske grunde været nødt til kun at stille 6 spørgsmål per intelligensform med den konsekvens, at det går ud over validitet og realibilitet. Fordelen ved at have flere spørgsmål i hver kategori er, at man bedre kan skelne de forskellige intelligensformer ad. Hvis der havde været f.eks. 12 spørgsmål per kategori ville man med større sikkerhed kunne vurdere, i hvilken grad en elev besidder en konkret intelligensform. I alle de sammenhænge, hvor vi bruger resultaterne af MI-undersøgelserne skal dette problem altså tages med i betragtning. Vi vil alligevel i det følgende foretage en analyse af de foreliggende data. Forekomsten af en intelligensform hos en elev er i undersøgelsen sket ved pointtildeling, således at svarer en elev *enig* til et spørgsmål/udsagn så får eleven et point i den intelligensform spørgsmålet hører til. Svarer eleven *uenig*, eller *undlader* at svare, får man ingen point. En intelligensform kan altså være helt fraværende eller relativt tilstede. Resultaterne af undersøgelserne er vist i bilag 14.

## Biologielever

Nedenfor er vist gennemsnitsværdien for de forskellige intelligensformer som de fremtræder hos biologieleverne. Det mest i øjenfaldende er den høje værdi for interpersonel intelligens. Vi betvivler dog, at værdien giver et retvisende billede af denne intelligensforms tilstedeværelse hos eleverne. Der er formentligt tale om, at eleverne tilpasser svarene efter, hvad de "tror" omgivelserne forventer af dem. Der gør sig derfor et sociokulturelt fænomen gældende her som betyder, at undersøgelsen er mindre valid for denne intelligensforms vedkommende. Et fænomen, som for så vidt også godt kan gøre sig gældende blandt svarene på de andre intelligensformer, men dog i mindre grad. For eksempel kan man forestille sig, at enkeltelever føler et pres fra omgivelserne, om at udvikle sig i en bestemt retning som gør, at de svarer på en måde som matcher denne forventning.

<b>MI intelligensform</b>	Lingvistisk-sproglig	Logisk-matematisk	Spatial	Kinæstetisk-kropslig	Interpersonel	Intrapersonel	Musikalsk
<b>Gennemsnit</b>	3,40	3,62	3,67	3,89	4,93	3,24	3,24
<b>Standardafvigelse</b>	1,34	1,47	1,46	1,39	0,81	1,42	1,82

Tabel 5.15: MI-form hos biologieleverne

I øvrigt er det bemærkelsesværdigt, at logisk-matematisk intelligens og lingvistisk-sproglig intelligens ligger midt i feltet set i lyset af, at det jo har heddet sig, at det almene gymnasium fortrinsvis har rekrutteret elever med disse intelligensformer (et forhold som vi har diskuteret i afsnit 4.2). På den anden side, kan der også gøre sig det forhold gældende, at spørgsmålene vedr. disse intelligensformer lyder mindre spændende, mere skoleagtigt, end spørgsmålene, der knytter sig til de øvrige intelligensformer. Derfor kan eleverne have en tilbøjelighed til i højere grad, at tilslutte sig udsagnene i spørgsmålene, der knytter sig til de øvrige intelligensformer.

## Musikelever

For musikeleverne gør der sig det samme forhold gældende, vedrørende interpersonel intelligens, som for biologieleverne og forklaringen er givetvis den samme. Men derudover er det bemærkelsesværdigt, at eleverne ligger meget højt i musikalsk intelligens.

<b>MI intelligensform</b>	Lingvistisk-sproglig	Logisk-matematisk	Spatial	Kinæstetisk-kropslig	Interpersonel	Intrapersonel	Musikalsk
<b>Gennemsnit</b>	3,57	3,83	3,87	4,48	5,13	3,91	5,30
<b>Standardafvigelse</b>	1,41	1,50	1,42	1,27	0,97	1,31	1,11

Tabel 5.17: MI-form hos musikeleverne

Umiddelbart er det måske ikke så bemærkelsesværdigt eftersom det drejer sig om musikelever, men der er formentlig to problemer, der gør sig gældende: 1) Som musikelev kan det opleves ”unaturligt”, hvis ikke man bekræfter de udsagn, der handler om musik, dvs. der er en forventning om, at man har de egenskaber udsagnene udtrykker. 2) En del af de spørgsmål, der vedrører musik, har karakter af at være for ”almindelige”, dvs. mange elever vil kunne sige ja til udsagnene uden, at være særlig ”musikalsk intelligente”. Spørgsmål 13 lyder for eksempel: ”Jeg nynner eller synger ofte for mig selv eller slår rytmer med mine hænder når jeg har en blyant eller pen i hånden”. Det kan godt være man siger ja til udsagnet, men det er næppe tilstrækkeligt til at kunne bruges som mål for musikalsk intelligens.

### **Opsamling**

For begge elevgrupper gælder, at MI undersøgelserne ikke er så valide som man kunne ønske sig. Især er der problemer med undersøgelsen af den interpersonelle intelligens, hvor elever generelt ligger meget højt. Et grundlæggende problem er, at antallet af spørgsmål per intelligensform er i underkanten af, hvad der er nødvendigt for at kunne adskille de forskellige intelligensformer. Det er bemærkelsesværdigt, at begge elevgrupper ligger lavt med hensyn til matematisk logisk intelligens og sproglig-lingvistisk intelligens, men forklaringen kan være, som i tilfældet med den interpersonelle intelligens, at der gør sig nogle sociokulturelle faktorer gældende forstået på den måde, at eleverne svarer efter, hvad omgivelserne forventer eller efter, hvordan de selv ønsker det skal være. Sammenligner vi musik- og biologielever er det iøjnefaldende, at musikeleverne gennemgående ligger højere på alle intelligensformer. Det har vi ikke nogen plausibel forklaring på.

### **5.5 holdningsundersøgelse analyse**

I det følgende analyser vi elevernes besvarelse af holdningsundersøgelsen (beskrevet i afsnit 4.3). Vi undersøger elevernes holdning til den test, de netop har gennemført og deres holdning til test generelt. Analysen tager udgangspunkt i den tematisering af spørgsmålene, vi har skitseret i afsnit 4.3 og den kombinerer beskrivende og diskuterende analyser af elevernes holdning til test i de to fag.

Resultaterne af holdningsundersøgelsen er vist i bilag 15 og 16. Som nævnt i afsnit 4.3 beregner Quia procentfordelingen på de enkelte svarkategorier og beregner en gennemsnitsværdi for alle svar. Gennemsnitsværdien er således et udtryk for, i hvor høj grad eleverne er enige eller uenige i et givet udsagn og en værdi på 3 vil være ensbetydende med at eleverne hverken er enige eller uenige i det pågældende udsagn.

### **Tema 1: Fagligt standpunkt og egen vurdering af faglige kompetencer**

For så vidt angår biologieleverne på det jyske gymnasium viser de spørgsmål, der har at gøre med elevernes opfattelse af testen i forhold til fagligt standpunkt og egen vurdering af faglige kompetencer (sp. 1,2,3,4,5,10) på den ende side, at et flertal af eleverne anerkender, at testen giver dem indblik i deres faglige kompetencer indenfor dette emne i faget (sp. 10) og det støttes på sin vis af svarene på sp. 3 (en test, som den jeg lige har gennemført, kan fortælle mig om jeg skal øge/justere min indsats i faget), hvor et flertal erklærer sig enig i dette. På den anden side peger svarene på sp. 1 og 2 i den anden retning. I modsætning hertil er et flertal af musikeleverne på Færøerne enige i de udsagn, der fremsættes på nær sp. 5. Med hensyn til sp. 5 er det er dog bemærkelsesværdigt, at det er et meget lille flertal (gennemsnit 3.08) på Færøerne, der erklærer sig uenige i, at ”en test, som den jeg lige har gennemført, giver mig mulighed for at vise, alt hvad jeg virkelig kan i faget”. Blandt biologieleverne, er der et større flertal, der er uenige, gennemsnittet er her 3.79. Det ser altså ud til, at brugbarheden af test vurderes mere positiv blandt musikeleverne end blandt biologieleverne. Der kan være forskellige forklaringer på, at eleverne i de to grupper svarer forskelligt på de ovennævnte spørgsmål. For det første er der tale om to forskellige faglige test og to forskellige fag. Det vil derfor ikke være underligt hvis eleverne har forskellig oplevelse af, hvad testen giver af viden om egne kundskaber, hvordan testen kan bruges formativt. For det andet er den kontekst, som testen indgår i givetvis også forskellig. De to elevgrupper medbringer jo en eller anden form for forståelse som blandt andet er grundlagt på den undervisning og læring, der er gået forud for testen. F.eks. er det en nærliggende tanke, at de jyske elever er påvirket af debatten om nationale test i folkeskolen (omtalt i afsnit 3.2) og på den måde har tilegnet sig en kritisk opfattelse af test som læringsværktøj. En tredje forklaring af mere teknisk karakter (og som gælder generelt i forhold til holdningsundersøgelsen) kan være det forhold, at spørgsmålene er administreret på dansk til de færøske musikelever. Det kan derfor ikke udelukkes, at der knytter sig nogle forståelsesvanskeligheder til musikelevernes besvarelse af spørgsmålene.

### **Tema 2: IKT kontra papir test**

Samlet set indikerer de fire spørgsmål, der omhandler IKT-baseringen (sp. 6, 8, 19, 21), at der er tilslutning til, at IKT-baseringen er at foretrække frem for papirbaseret. De fire spørgsmål afdækker forskellige aspekter af dette og derfor, kan der selvfølgelig være forskellige individuelle begrundelser for, at IKT-basering er at foretrække. Desuden skal det bemærkes, at sp. 21 ikke forholder sig til om IKT-basering er godt eller dårligt, men udelukkende om kvaliteten af den aktuelle

brugergrænseflade. Der er stort set ikke forskel på musik- og biologieleverne i disse spørgsmål. Begge elevgruppers besvarelse støtter altså antagelsen om, at IKT-baseringen er fremmede for elevernes deltagelse i faglige test.

### **Tema 3: Motivation**

Vurderet ud fra sp. 7 og 9, der handler om motivationsaspektet i test (det samme gør sp.6, men i forhold til Pc anvendelsen) ser det ikke ud til, at biologieleverne ser det som særligt motiverende at besvare en test i sig selv, og heller ikke ser noget særligt konkurrenceforhold i det. Musikeleverne derimod har en lidt anden opfattelse: I sp. 7, der vedrører konkurrencemomentet er gennemsnittet præcist 3, altså hverken enig eller uenig. Hvad angår sp. 9, er der faktisk et flertal af musikeleverne der mener, det er motiverende at deltage i en test. Forskellen i opfattelse kan igen skyldes de faktorer, der blev nævnt ovenfor i forbindelse med kompetence og standpunktstemaet. Især kan de nævnte forskelle i forforståelse være en væsentlig forklaring. Vi har endvidere i afsnit 4.3.1 problematiseret det forhold, at elevernes testresultat kan have indflydelse deres holdning til test. Det problem undersøges nærmere i afsnit 5.6.

### **Tema 4: Holdning til test, test positivt eller negativt**

Svarene på sp. 11, 16 og 17, der mere eller mindre direkte handler om elevernes holdning til test, giver ikke noget klart indtryk af biologielevernes holdning. Der er et ganske lille flertal, der ikke mener, det er positivt at få kontrolleret sine faglige kompetencer ved brug af test og som heller ikke vil have flere af denne slags faglige test. Til gengæld er der så et lidt større flertal, der heller ikke er interesseret i en spørgsmålstype, hvor de selv kan/skal formulere svarene (sp. 17). Hertil skal der dog bemærkes, at standardafvigelsen, på data, er noget større på sp. 17 end på sp. 11 og 16, dvs. der er større spredning på svarenes fordeling på svarkategorier i sp.17 end i de to første. Musikeleverne derimod er mere klar i deres tilkendegivelser. Der er et betydeligt flertal, som gerne vil have deres faglige kompetencer kontrolleret ved test og som gerne vil have flere test, samtidigt med, at et stort flertal afviser de hellere vil have testopgaver, hvor selv de kan/skal formulere svarene. Måden de to elevgrupper besvarer sp. 11 og 16 på, stemmer meget godt overens med de tidligere nævnte forskelle og forklaringerne er formentlig også de samme. Til gengæld er der er altså et flertal i begge elevgrupper, der afviser testopgaver, hvor de selv kan/skal formulere svarene. Her må vi imidlertid vedgå, at vi er i tvivl om, hvordan eleverne opfatter spørgsmålet. Forstår de det på samme måde som vi forstår det? Altså at de i den type opgaver, som vi sigter til, med spørgsmålet, skal skrive et "essay-agtigt" svar på problemløsningsbasis. Dvs. den type opgaver som har været fremherskende

indenfor den tyske og nordiske *didaktiktradition* (omtalt i afsnit 4.2.1). Eller forstår de det på den måde, at de skal skrive de udsagn, ord, begreber osv. der passer til de forskellige spørgsmål. Hvis de forstår det på den sidste måde, er der ikke noget at sige til, at de er afvisende overfor udsagnet for det er jo klart mere besværligt end at klikke et svar med musen. På den anden side er det heller ikke overraskende, hvis de afviser udsagnet ud fra den første forståelse af spørgsmålet. Som gymnasie-lærer ved vi godt, at mange elever har det svært med det skriftlige arbejde, og standard spørgsmålet til læreren, når en elev skal lave skriftlig opgave er altid: "Hvor meget skal det fylde?" og hvis man så siger minimum 2 sider, så er næste replik "nej – så meget kan jeg ikke skrive". Så alt i alt er det ikke uventet, at eleverne svarer som de gør, men der er selvfølgelig en markant forskel på, hvad begrundelserne for svarene er.

### **Tema 5: Spørgsmålenes karakter**

For så vidt angår spørgsmålenes karakter (sp. 12 og 13) kan man med en vis rimelighed konstatere, at biologieleverne, hverken betragter spørgsmålene som lette eller svære at forstå (gennemsnit = 3.06). Til gengæld er der et klart flertal, der er uenige i, at de fleste spørgsmål er lette at besvare. Blandt musikeleverne er der et flertal (gennemsnit 3.36), der er uenige i at spørgsmålene er svære at forstå og et endnu større flertal (gennemsnit 3.56), der er uenige i at spørgsmålene er lette at besvare. Besvarelsen af de to spørgsmål relaterer sig direkte til den faglige test og resultaterne af denne. Vi vender derfor tilbage til disse spørgsmål i det følgende afsnit (afsnit 5.6).

### **Tema 6: test som stressfaktor**

Et flertal af biologieleverne føler sig ikke under tidspres når de besvarer testen, fremgår det af svarerne på sp. 14. Til gengæld er der et flertal blandt eleverne, der erklærer sig enige i, at de bliver usikre når de skal besvare en test af denne type (sp.15). Der er dog ikke nødvendigvis nogen modsætning i svarene på de to spørgsmål. Et meget stort flertal af musikeleverne afviser, at de følte sig under tidspres (gennemsnit 4.48) dog skal det bemærkes, at der er en stor standardafvigelse (spredning) på svarene nemlig 1,45. Ligeledes er der et mindre flertal af musikeleverne, der afviser at de bliver usikre, når de skal besvare en test af denne type. Et flertal i begge elevgrupper afviser altså, de er under tidspres og det hænger da også ganske godt sammen med deres gennemførsel af testen, idet det viser sig, at kun få elever bruger mere en 25 minutter på at besvare testen.

### **Tema 7: Feedback**

Med hensyn til feedback peger resultaterne på, at den feedback eleverne får, ikke er tilstrækkelig (sp. 18) samtidigt med at eleverne gerne vil have lærerkommentarer (sp. 20) til deres testresultat. Og det gælder både musik- og biologielever. Musikeleverne er dog relativt set lidt mere uenige end biologieleverne i sp. 18 og relativt lidt mere enige end biologieleverne i sp. 20. I den forbindelse mindes om, at musikeleverne ikke har fået feedback under besvarelsen af testen, hvorimod biologieleverne har fået feedback i form af en markering af, om svaret er rigtigt eller forkert og hvor mange point, de har fået tildelt. Denne form for feedback vurderes altså ikke som tilstrækkelig af biologieleverne, hvilket også ud fra et formativt synspunkt virker forståeligt.

### Opsamling

Vi har på baggrund af tematiseringen givet en analyse af, elevernes holdning til test. På nogle områder er der overensstemmelse mellem musik og biologieleverne på andre områder er der ikke. IKT-baseringen vurderer begge elevgrupper positivt. Den feedback eleverne har fået vurderes som utilstrækkelig i begge grupper og begge grupper vil gerne have lærerkommentarer med til deres besvarelse. Elevgrupperne er også enige om, at spørgsmålene er lette at forstå men svære at besvare. Til gengæld er der forskel på elevgruppernes holdning til test. Musikeleverne er positivt stemt overfor test, det kan man ikke sige om biologieleverne. Brugbarheden af test i forbindelse vurdering af eget standpunkt vurderes også positivt af eleverne, hvorimod billedet er mere tvetydigt blandt biologieleverne, på den ene side vurderer de det som brugbart, på den anden side tyder noget på, at de ikke mener dette.

## 5.6 Korrelationsanalyser

### Korrelation mellem læringsstile og MI

Vi har analyseret på hvilken sammenhæng, der er mellem elevernes intelligensform og deres foretrukne læringsstil. Analysen er foretaget ved at gruppere eleverne efter, hvor mange point eleverne har fået med hensyn til kinæstetisk/kropslig intelligens og derefter korrelere grupperne med hensyn til kinæstetisk/taktil læringsstil.

Biologi		MI		Læringsstil	
Elev nr.	Kinæstetisk/kropslig	Kinæstetisk/taktil	Elev nr.	Kinæstetisk/kropslig	Kinæstetisk/taktil
2	6	8	13	2	7
6	6	11	18	2	7
7	6	9	32	2	5
10	6	9	34	2	8



23	6	9	37	2	5
27	6	10	31	1	11
5	5	8	30	0	10
<b>Gennemsnit</b>		<b>9,14</b>			<b>7,57</b>
<b>Standardafvigelse</b>		<b>1,07</b>			<b>2,56</b>

Tabel 5.18: Korrelation mellem MI og læringsstil

Når vi sammenligner gennemsnitsværdierne ser det ud som om, at elever med kinæstetisk/kropslig intelligens har tendens til præference for kinæstetisk/taktil læringsstil. På grund af den relativt store standardafvigelse (spredning) i elevgruppen med lav kinæstetisk/kropslig intelligens er forskellen dog ikke signifikant. Den samme tendens gør sig også gældende for musikeleverne, men også uden at være signifikant.

Tilsvarende analyser for spatial intelligens og visuel læringsstil viser samme tendens, med stadigvæk ikke signifikant.

Musik Elev Nr.	MI Spatial	Læringsstil Visuel	Elev Nr.	MI Spatial	Læringsstil Visuel
14	6	8	23	4	6
6	6	10	4	3	3
8	6	11	15	2	1
22	6	8	9	2	8
16	6	7	21	2	9
<b>Gennemsnit</b>		<b>8,8</b>			<b>5,4</b>
<b>Standardafvigelse</b>		<b>1,64</b>			<b>3,36</b>

Tabel: 5.19: Korrelation mellem MI og læringsstil

Der er således tegn på, at de relationer mellem intelligensformer og læringsstile, vi har beskrevet i afsnit 4.1.3 støttes af vores empiri om end der ikke er tale om en signifikant korrelation. Vi skal dog huske på, at de problemer, der gør sig gældende vedrørende validiteten af MI undersøgelsen, betyder at vores konklusion her kan anfægtes.

### Korrelation mellem MI og testscore

Vi har undersøgt, om der er korrelation mellem testscore og intelligensform. Det er gjort på den måde, at vi har korreleret elevernes "score" i MI undersøgelsen med deres score i den faglige test. Eleverne "sorteres" efter, hvor mange point de har fået indenfor de 7 forskellige intelligensformer. Derefter sammenlignes med, hvor mange point eleverne har fået, i gennemsnit, i test score (se nedenstående eksempel).

Elev nr.	Point matematisk-logisk intelligens	Point i faglig test (%)	Elev nr.	Point matematisk-logisk intelligens	Point i faglig test(%)
----------	-------------------------------------	-------------------------	----------	-------------------------------------	------------------------

6	6	73	10	2	39
20	6	76	14	2	67
23	6	82	15	2	42
32	6	67	16	2	48
<b>Gennemsnit</b>		<b>74,5</b>	36	2	55
<b>Standardafvigelse</b>		<b>6,24</b>	37	2	47
			41	2	58
2	5	64	<b>Gennemsnit</b>		<b>50,86</b>
11	5	61	<b>Standardafvigelse</b>		<b>9,75</b>
13	5	64			
17	5	52	22	1	52
25	5	58	34	1	58
26	5	55	31	0	39
29	5	42	<b>Gennemsnit</b>		<b>49,67</b>
38	5	70	<b>Standardafvigelse</b>		<b>9,71</b>
42	5	55			
45	5	61			
46	5	58			
<b>Gennemsnit</b>		<b>58,18</b>			
<b>Standardafvigelse</b>		<b>7,37</b>			

Tabel 5.20: Korrelation logisk-matematisk intelligens versus testscore i biologi

Som eksemplet viser, er der tegn på korrelation, mellem elevernes point antal i logisk-matematisk intelligens og antal point i den faglige test. Man kan udtrykke det på den måde, at jo flere matematisk-logiske kendetegn eleverne besidder jo flere korrekte svar afgiver de i den faglige test. Forskellen i testscore mellem de elever, der har scoret 6 point og de øvrige grupper er ydermere signifikant forskellig, hvilket er i overensstemmelse med den traditionelle opfattelse af, hvad det er, der måles ved intelligens test. I afsnit 4.1, hvor vi diskuterede Gardners teori nævnte vi blandt andet, at den endimensionelle intelligens (g-faktoren) har stort set udelukkende hentet empirisk støtte fra standardiserede ”papir og blyant” test, som næsten udelukkende basere sig på matematisk-logisk og sproglig intelligens. Hvis vi ser bort fra de tidligere omtalte validitets og reliabilitets problemer er ovenstående resultat et eksempel herpå.

Til gengæld er der ikke andre former for korrelation mellem testscore og intelligensform. De øvrige seks MI former viser ikke nogen form for korrelation med testscoren. Hos musikeleverne er der heller ingen korrelation mellem logisk-matematisk intelligens og testscore, hvilket formentlig hænger sammen med, at spredningen på elevernes testscore ikke er tilstrækkelig stor (den mindste score er 20 % og den højeste 60 %).

### Korrelation mellem læringsstile og testscore

Der kan ikke findes nogen korrelation mellem de læringsstile, vi har undersøgt og elevernes testscore. Dvs. vi kan ikke finde tegn på, at elever med en given præference for en læringsstil score tilsvarende højt, eller lavt i den faglige test, hverken hos biologier eller musikelever. Vi kan heller ikke finde korrelation mellem visuel læringsstil og testscore på grafik inkluderende spørgsmål. Dvs. der er ingen sammenhæng mellem elevernes score i de spørgsmål, der inkluderer grafik, i den faglige test og elevernes visuelle præference. Den visuelle præference skulle teoretisk set betyde, at grafiske elementer kunne være en hjælp i læringssammenhæng (som beskrevet i afsnit 4.1). At der i dette tilfælde ikke er denne sammenhæng kan forklares med, at en korrekt besvarelse af spørgsmålet afhænger af andre faktorer end grafikken, herunder måden spørgsmålet er formuleret på og måden, der skal svares på. Der er forskel på om grafikken er en forudsætning for besvarelse af spørgsmålet eller blot et supplement.

### Korrelation mellem testscore og elevernes holdning til test

Som nævnt tidligere får biologieleverne feedback med hensyn til om de svarer rigtigt eller forkert og hvor mange point de får tildelt. Problematikken omkring denne feedback i forhold til holdningen til test er også omtalt i afsnit 4.3. Vi har derfor undersøgt om der er nogen korrelation mellem elevernes score i den faglige test og deres holdning til test. Vi har gjort det ved at sammenholde testscoren for de 7 elever, der har scoret flest henholdsvis færrest point, med deres svar på spørgsmål 9, 11 og 16 i holdningsundersøgelsen.

Elev nr.	Score i faglig test	Svar på holdningsspørgsmål			sum
		sp.9	sp.11	sp.16	
20	82%	3	3	5	11
22	82%	2	5	3	10
9	76%	2	4	2	8
19	76%	3	4	4	11
6	73%	1	2	1	4
37	70%	4	2	2	8
10	67%	2	2	2	6
<b>Gennemsnit</b>					<b>8,29</b>
<b>Standardafvigelse</b>					<b>2,63</b>
39	39%	3	3	5	11
30	39%	3	2	3	8
42	36%	4	3	4	11
41	36%	5	5	4	14
32	30%	3	1	2	6
12	9%	4	5	5	14
26	3%	2	2	1	5
<b>Gennemsnit</b>					<b>9,86</b>
<b>Standardafvigelse</b>					<b>3,63</b>

Tabel 5.21: Biologielevernes holdning til test korreleret med testscore

Som det fremgår af tabellen er der en lille forskel i gennemsnitlig svar fordeling på de to grupper, men når man tager standardafvigelsen i betragtning, er det tvivlsomt om der er en signifikant forskel på de to grupper, hvilket betyder, at vi kan konstatere, at elevernes holdning til test ikke er korreleret med, hvor godt et resultat eleverne opnår i testen.

Musik eleverne har som tidligere nævnt ikke fået feedback under den faglige test, og har altså ikke haft kendskab til deres testscore da de besvarede holdningsspørgsmålene. Vi har alligevel på tilsvarende vis undersøgt om der er korrelation mellem testscore og holdning til test. Vi har her sammenlignet de 5 mest scorende med de 5 mindst scorende:

Elev nr.	Score i faglig test	Svar på holdningsspørgsmål			sum
		sp.9	sp.11	sp.16	
13	60%	1	1	2	4
11	55%	3	1	2	6
10	50%	4	4	4	12
16	50%	3	4	5	12
21	50%	4	2	4	10
<b>Gennemsnit</b>					<b>8,8</b>
<b>Standardafvigelse</b>					<b>3,63</b>
20	30%	4	3	3	10
3	25%	2	2	1	5
7	25%	4	3	2	9
9	25%	1	1	1	3
2	20%	2	2	2	6
<b>Gennemsnit</b>					<b>6,6</b>
<b>Standardafvigelse</b>					<b>2,88</b>

Tabel 5.22: Musikelevernes holdning til test korreleret med testscore

Vurderet ud fra gennemsnitsværdierne ser det ud som om, de elever der har fået færrest point i musiktesten også er dem, der er mest positive overfor test, men igen skal man være opmærksom på, at der er forholdsvis stor spredning på data. Det kan derfor ikke konkluderes, at der er signifikant forskel på de to grupper. Samlet set er der ikke noget, der tyder på, at elevens holdning til test er påvirket af, hvor godt de kan klare sig i test.

I Spørgsmål 12 i holdningsundersøgelsen bliver eleverne spurgt om testspørgsmålene er svære at forstå og i sp.13 bliver de spurgt om de er lette at besvare. Vi har efterfølgende undersøgt om der er korrelation mellem svarene på disse spørgsmål og elevernes testscore (se nedenfor):

---

**Svar på holdningsspørgsmål**

Elev nr.	Score i faglig test	sp.12	sp.13
20	82%	4	4
22	82%	2	3
9	76%	3	5
19	76%	3	3
6	73%	4	4
37	70%	5	3
10	67%	4	3
<b>Gennemsnit</b>		<b>3,57</b>	<b>3,57</b>
<b>Standardafvigelse</b>		<b>0,98</b>	<b>0,79</b>
39	39%	2	4
30	39%	4	4
42	36%	5	5
41	36%	2	4
32	30%	2	2
12	9%	2	5
26	3%	4	4
<b>Gennemsnit</b>		<b>3</b>	<b>4</b>
<b>Standardafvigelse</b>		<b>1,29</b>	<b>1,00</b>

Tabel 5.23: Korrelation mellem testscore og spørgsmålenes sværhedsgrad, biologielever

Et flertal af de højest scorende Biologielever er altså uenige i, at testspørgsmålene er svære at forstå, tilsvarende mener et flertal ikke spørgsmålene er lette at besvare. Når vi så sammenligner med den gruppe, der har scoret lavest er det bemærkelsesværdigt, at der i denne gruppe er færre som mener spørgsmålene er svære at forstå, men dog flere, der er uenige i at spørgsmålene var lette at besvare.

Ser vi på musikelevernes svar på samme spørgsmål, tegner der sig et billede af, at elevernes score i testen ikke er i overensstemmelse med, hvordan eleverne opfatter spørgsmålenes sværhedsgrad. De elever, der har scoret flest point har i gennemsnit fået lidt over 50 % point, alligevel mener et flertal blandt dem, at spørgsmålene var lette at besvare. Tendensen er endnu mere udbredt i den gruppe, der har fået færrest point. Her er der et relativt stort flertal, der mener spørgsmålene var lette at besvare. På trods af relativt stor spredning på svarene, er der noget, der tyder på at der er problemer med realibiliteten i disse spørgsmål. En forklaring kan muligvis være, at der ikke er tale om anonyme besvarelser og den kan som nævnt i afsnit 4.3 betyde, at validitet og realibilitet kommer til at lide skade. Den manglende anonymitet kan i dette tilfælde betyde, at nogle elever helst ikke vil give indtryk af, at de synes spørgsmålene er svære. Endelig skal man selvfølgelig huske, at fraværet af feedback også kan spille en rolle, da eleverne jo ikke ved om de har svaret rigtig eller forkert, ”kan de jo være i god tro”. Hvis de tror, de har svaret rigtig, er det jo på sin vis forståeligt nok, hvis de opfatter spørgsmålene som lette.

Musik Elev nr.	Score i faglig test	Svar på holdningsspørgsmål	
		sp.12	sp.13
13	60%	1	1
11	55%	3	1
10	50%	4	4
16	50%	3	4
21	50%	4	2
<b>Gennemsnit</b>		<b>3</b>	<b>2,4</b>
<b>Standardafvigelse</b>		<b>1,22</b>	<b>1,52</b>
20	30%	4	3
3	25%	2	2
7	25%	4	3
9	25%	1	1
2	20%	2	2
<b>Gennemsnit</b>		<b>2,6</b>	<b>2,2</b>
<b>Standardafvigelse</b>		<b>1,34</b>	<b>0,84</b>

Tabel 5.24: Korrelation mellem testscore og spørgsmålenes sværhedsgrad, musikelever

### Korrelation mellem auditiv læringsstil og auditive spørgsmål

I musik testen indgår, som beskrevet i afsnit 4.2.4, fire spørgsmål som inkluderer lyd. Eleverne skal lytte til et stykke musik, et instrument eller lign. og derpå svare på spørgsmålet. En analyse af disse spørgsmål med henblik på korrelation til den auditive læringsstil hos de elever, der svarer rigtigt på et eller flere af spørgsmålene, viser, at der ikke er nogen sammenhæng mellem graden af auditiv læringsstilspræference og antallet af rigtige svar på disse faglige spørgsmål. Det er altså ikke sådan, at elever med høj auditiv præference også har flere rigtige svar end elever med lavere auditiv præference.

### Opsamling

Vi har foretaget korrelationsanalyser med henblik på at undersøge hvilke sammenhænge, der findes mellem de kvantitative data, der indgår i vores empiri. Vi kan ikke finde nogen signifikant korrelation mellem læringsstile og MI. Blandt de læringsstile, vi har undersøgt er dog noget der tyder på en hvis sammenhæng mellem kinæstetisk/kropslig intelligens og kinæstetisk/taktil læringsstil samt for spatial intelligens og visuel læringsstil. Validitetsproblemer i forhold til MI-undersøgelsen gør dog at sammenhængen kan anfægtes.

Vi har fundet korrelation mellem biologielevernes testscore og deres logisk-matematiske intelligens, således at de elever, der ligger højest i logisk-matematisk intelligens også er dem, der score højest i den faglige test og omvendt. Derimod kan der ikke findes nogen korrelation mellem læringsstile og testscore.

Det er ikke noget der indikerer, at der er korrelation mellem elevernes holdning til test og deres test-score. Elever, der score lavt i den faglige test har altså ikke automatisk en negativ holdning til test og omvendt elever, der score højt i faglige test er ikke automatisk positive overfor faglige test.

Vi har undersøgt om, der er sammenhæng mellem, elevernes testresultater og deres vurdering af spørgsmålenes sværhedsgrad, for biologielevernes vedkommende viser det sig, eleverne i overvejende grad opfatter dem som vanskelige at besvare men lette at forstå. Musikeleverne derimod har en opfattelse, af spørgsmålene var lette at besvare uanset, hvor mange point de havde fået i den faglige test.

Musikelevernes besvarelse af faglige spørgsmål inkluderende lydoptagelser viser, at der ikke er nogen sammenhæng mellem graden af auditiv præference hos eleven og så antallet af rigtige svar på disse spørgsmål.

## 6. Diskussion og perspektiv

Vi fremlægger i dette kapitel en diskussion som er tematiseret i 4 overordnede afsnit. I det første afsnit diskuteres testdesign, multiple intelligenser og læringsstile ud fra en didaktisk synsvinkel. I det næste afsnit sætter vi fokus på faglige test og taksonomi og diskuterer, hvordan taksonomiske niveauer kan integreres i faglige test. I det tredje afsnit fokuserer vi på adaptive test med udgangspunkt i Vygotskys teori om 'den nærmeste udviklingszone'. Endelig vurderer vi i det sidste afsnit mulighederne for at integrere feedback i faglige test med henblik på formativ evaluering og læring.

### 6.1 Didaktisk testdesign, MI og Læringsstile

#### Multiple Intelligenser eller Læringsstile?

I et interview har Mogens Hansen<sup>42</sup> udtalt, at "I dag kender gud og hver mand Gardner, og der er jo fint. Men den kritiske refleksion har manglet i mange år. Der har været tale om en begejstringsbølge, som jeg finder alarmerende. Mange Intelligenser er blevet en fødselskage med for mange hurra-flag. Først rejste Gardner verden rundt, siden kom Daniel Goleman med sit begreb emotionel intelligens, og nu rejser Dunn & Dunn rundt med deres læringsstile. Tilsammen har de tre godt nok banket mange flag ned i kagen. Som jeg forstår Gardner, ville han egentligt helst selv være fri for følgeskabet af Dunn & Dunn" (Albrecht:26). Endvidere mener Mogens Hansen, at Gardner kun bør bruges i den fagligt didaktiske planlægning af undervisning, ikke til typeinddeling af eleverne. MI er først og fremmest en pædagogisk didaktisk teori og at "i stedet for at fremme 'talkloge' og 'ordkloge' børnetyper har vi brug for at se børn i helheder og skabe en undervisning, hvor alle intelligenser er i spil på én gang – ligesom i en musical med bevægelse, ord og musik" (ibid:27).

Grundlæggende er vi enige med Mogens Hansen i, at målet med operationaliseringen af Gardners teori ikke er at typeinddele børnene, men på den anden side mener vi, at et af midlerne til at kunne lave en fornuftig og relevant fagdidaktik er konkret kendskab til intelligensformernes fremtræden hos eleverne.

I modsætning til Gardner og Hansen mener Rita Dunn, at eleverne først og fremmest skal undervises efter deres primære læringsstilpræference, og at hvis en elev bliver undervist efter andre præferencer, kan det være ligefrem hindrende for læreprocessen. Men Mogens Hansen siger også i artiklen, at hvis de svage potentialer ikke bliver sat i spil, tørrer de ud og bliver til ingenting. Både Dunn

---

<sup>42</sup> Interview af Jakob Albrecht i artiklen "en historie om håb" link:  
<http://www.dpu.dk/everest/showdoc.asp?id=060828122025&type=doc>



og Hansen er dog nogenlunde enige i at menneskebørn fungerer altid i en helhed. Lorna M. Earl's følgende udsagn understøtter den holistiske tilgang, som også findes hos Dunn<sup>43</sup> at "When they know the purpose and direction of their learning and can see the connection of the learning intention of any particular task to the larger agenda, they are likely to be more motivated and task-oriented. Even more important, they can plan their own learning and see their own progress" (Earl 2003:80). Endvidere siger hun at "The goal is to make the next step in learning visible to the students so that they can use their minds to unravel ideas and see patterns" (ibid:80).

Det store spørgsmål er, hvordan man skal forstå henholdsvis MI og læringsstile. Er det to sider af samme sag, hvor det så bare er et spørgsmål om didaktisk tilgang? Den empiri vi har indsamlet viser tegn på, om end ikke signifikante, at der er korrelation mellem kinæstetisk/kropslig intelligens og kinæstetisk/taktil læringsstil, samt spatial intelligens og visuel læringsstil (se afsnit 5.6). Det er dog for spinkelt et grundlag, kun med baggrund i empirien, at sige noget afgørende om sammenhængen mellem MI og læringsstile.

Mogens Hansen fortolker Gardner på den måde, at det er nødvendigt at lave materiale og undervisning, der understøtter alle intelligenser, for at 'sikre' sig at alle får 'noget' differentiering på undervisningsniveau (materiale, undervisningsstil), mens Dunn i praksis handler om at differentiere på individniveau. Der er mange spørgsmål der trænger sig på, hvad er muligt i praksis? kan man gøre begge dele?

Hvordan kan vi bruge vores viden om elevernes præferencer? Vi mener, for det første, at det ikke kan foregå på enkeltlærer niveau, det skal op på team/klasse eller skoleniveau, således at elevens profil ikke kun handler om hvordan en enkelt lærer griber tingene an, men om hvordan skolen pædagogisk/didaktisk forholder sig til elevernes læringsstile og intelligenser. For det næste skal det tages seriøst, således at både undervisningsmaterialer, undervisningsmiljø og lærerkompetencer afspejler den viden som findes om elevernes præferencer og at der siden handles derefter. Det er store udfordringer for den enkelte skole, som ikke har taget fat i dette område førhen.

En frugtbar distinktion mellem tilgangene kunne være at bruge Gardner på et overordnet didaktisk niveau i planlægningen, mens man i den praktiske undervisning som lærer tager udgangspunkt, eller i baghovedet, har den enkelte elevs læringsstil in mente, således at der i læringssituationen kan sættes ind individuelt efter behov, f.eks. ved at uddybe og arbejde med faglige emner på den enkeltes præmisser. Der behøver ikke ligefrem at være modsætninger imellem positionerne.

Omvendt må vi konstatere, at hvis der blandt teoretikere ikke er enighed om, betydningen og rækkevidden af de teoretiske positioner, så bliver det ikke just nemmere på det operationelle niveau,

---

<sup>43</sup> se afsnit om holister/analytikere kapitel 4.1

altså blandt lærerne som skal bringe teorierne i anvendelse. Det illustreres godt af fokusgruppeinterviewet fra det jyske gymnasium, hvor informanterne forveksler teorierne (se afsnit 5.1). En interessant pointe som påpeges i det samme fokusgruppeinterview er, at fokus skal om ikke flyttes så i hvert fald deles mellem hvad lærerne kan bruge MI og læringsstile til og hvad eleverne kan bruge teorierne til. Det illustreres med følgende udsagn fra fokusgruppen om at læringsstile og MI ikke bare skal være en del af lærernes evalueringskultur, men i lige så høj grad en del af elevernes evalueringskultur. Det hedder med en omskrevet formulering fra den jyske fokusgruppe: ”Eleverne skal med kendskab til deres egne læringsstile/MI blive i stand til, at håndtere egne læreprocesser” (se afsnit 5.1.2).

Der er ingen tvivl om at den bedste mulighed for at få mere indsigt i disse problemstillinger er kontinuerlig teoretisk/forskningsmæssig forankring og praktisk afprøvning i vekselvirkning, hvilket kræver teoretisk indsigt hos praktikerne i samspil med den teoretiske udvikling på området.

### **Testdesign**

Vi har brugt testning i flere sammenhænge. Vi har brugt personsbaserede test/undersøgelser i to tilfælde og faglige test i to tilfælde. Testene bruges til at samle oplysninger om elevernes læringspræferencer og faglige niveau. Testene kan beskrives som diagnostiske i den forstand, at de bruges til at give et udgangspunkt for den videre planlægning af undervisningen. Eleverne har forskellige emotionelle tilstande, forskellige læringsstile og styrker, er på forskellige udviklingstrin og faglige stadier, har forskellige talenter og evner og forskellige behov.

Som omtalt i afsnit 5.6 viser vores empiri, at der er en sammenhæng mellem biologielevernes logisk-matematiske intelligens og de samme elevs testscore i den biologifaglige test. Denne sammenhæng findes dog ikke blandt musikeleverne. Alligevel vil vi tillade os at konkludere, at det er et tegn på at testresultater afspejler en bestemt intelligensform som også Gardner fremhæver (se afsnit 4.1.1). Det betyder, at man i sit testdesign bør overveje, hvordan man kan stille spørgsmål som inddrager andre intelligensformer end den logisk-matematiske, hvis eller når man vil lave fagdidaktisk undervisningsplanlægning med udgangspunkt i Gardners teori.

### **IKT-integration**

Traditionelle multiple choice test kritiseres for at være dekontekstualiserede. Kvale siger om denne form for test at “The multiple choice test is a central device. Knowledge tends to become decontextualized into atomistic data and formal rules of combination” (Kvale 1993:237). Vi vil imidlertid

hævde, at med IKT tilføres nogle nye muligheder til testopgaver, og vi har eksempler på, i de test vi har foretaget, at test kan kontekstualiseres. Eksempel: man kan forestille sig, at eleverne i en klasse har lavet en kostundersøgelse v.h.a et IKT-baseret kostanalyseprogram, hvorefter de besvare en test hvori følgende spørgsmål indgår:

Hvilke af følgende svar passer bedst med nedenstående udsagn:

”En analyse af en persons daglige kost (en kostanalyse) kan bruges til”:

svarmuligheder: 1) at vurdere om personen taber sig eller tager på, 2) At vurdere om personens kostindtag er i overensstemmelse med den anbefalede næringsstoffordeling, 3) At vurdere personens daglige energibehov.

Der er nogle perspektiver i IKT-baserede test som kun nævnes her i oversigtsform. Der kan benyttes interaktive metoder som tager udgangspunkt i registrering af museklik (fx på et kort, eller tryk på de 3 rigtige figurer), der kan benyttes drag and drop (placer hvert ikon i den rigtige rækkefølge), der kan benyttes hotspots, (som museklik eksemplet), interaktive responsformater, hvor feedback er afhængig af elevens svar. Desuden adaptive muligheder, multimedia og andet stimulusmateriale, som kan skabe en vis autenticitet (lyd, video, animation, simuleringer) og en virtuel kontekst. Derudover er der mulighed for automatisering af afvikling og efterbehandling, rapportering og dokumentation (tid og penge). Hvis vi sammenligner med udviklingen indenfor spilområdet, kan man forestille sig at simuleringer også bliver et redskab der vil blive benyttet indenfor testområdet i fremtiden.

Vores test er indholdsmæssigt kun delvis forskellige fra papirbaserede test, som ofte er begrænsede til den lavere del af de kognitive niveauer som at erindre, fortolke, klassificere, opsummere, sammenligne eller analysere ideer, begreber, principper og teorier. Ved at bruge andre digitale medier, i særdeleshed kommunikationssystemer/konferencer, online diskussioner og portefolier kan man opnå højere kognitive niveauer som evaluering, at producere ny indsigt hos eleven og give mulighed for refleksion over egen praksis. Men en IKT-integration gør grænserne mellem metoderne mere flydende. De nævnte metoder eller værktøjer og test behøves ikke at ses isolerede, men som midler/redskaber som alle kan bruges i undervisningssammenhæng i en reflekteret didaktisk planlægnings- og udførelsesproces.

Vi har i vores faglige test vist, at der kan stilles spørgsmål som omfatter højere taksonomiske niveauer end paratviden og vi har tilsvarende set, at når eleverne skal besvare spørgsmål på højere

taksonomisk niveau så har de vanskeligere ved at svare rigtigt (se f.eks. spørgsmål 8, 11 og 14 i bilag 7).

At undersøge paratviden er nødvendigt i undervisningen, det må bare ikke stå alene som evalueringskriterie. Det er et synspunkt der bakkes op af den 'naturvidenskabelige' fokusgruppe: "Nej det er ikke nok, det er også derfor de test og den type de udgør også kun er en lille del af min undervisning og min bedømmelse af eleverne det skulle meget nødtigt blive for dominerende, at jeg kun er optaget af, at de har en paratviden af en vis størrelse, men samtidig så vil jeg nok vove at påstå, at det er svært at have forståelse uden paratviden i et eller andet omfang, der er jeg nok lidt gammeldags på den måde – paratviden må der til."

Vi vil dog tilføje, at de forskellige 'fakulteter' formentlig ser forskellig på den sag. Den del af den færøske fokusgruppe, der repræsenterer humaniora virker ikke så begejstret for tanken, men de henviser jo på den anden side til naturvidenskab idet informanterne siger: "Faglige test er sikker brugbare i sprogfag og naturvidenskabelige fag, men jeg ikke kan forestille mig det i historie, dansk, færøsk, samfundsfag i særlig stor stil... Kunne bruges i fysiologien i idræt, hvor der er et bestemt fastlagt indhold, der ikke er til diskussion".

Der findes i øvrigt flere eksempler på at humaniora og samfundsfag arbejder med test som evalueringsmetode<sup>44</sup>.

Der er stor forskel på testopgaver som har ét rigtigt svar og opgaver som er mere komplekse. Den første type, som vi har kaldt objektive er dem med et begrænset antal valgmuligheder. Det er mange flere kriterier i spil når man taler om essays og andre produktformer. For at vurdere sådanne produkter kræves en analyse/vurderingsproces, hvor læreren deltager personligt. Hvis der er en klar målorientering er det nemmere at vurdere, dette gælder ikke kun i testsammenhæng.

Vi betragter de lavere taksonomiske vidensniveauer som forudsætninger for højere og mere abstrakte kognitive niveauer. I hvor stor grad det er muligt at arbejde med højere ordens viden i test er stadig et uafklaret spørgsmål. Men der er ingen tvivl om at hele fagets bredde bliver ikke dækket af traditionelle test.

---

<sup>44</sup> Se følgende eksempler fra nettet: <http://www.konsulenter.acu-arhus.dk/pt/MDC/EVALUERING%20MED%20IT%20kap%202.pdf>  
[http://www.unev.dk/files/Annelise\\_Grinsted\\_7.pdf](http://www.unev.dk/files/Annelise_Grinsted_7.pdf)

## 6.2 Faglige test og taksonomiske niveauer

I forbindelse med benyttelsen af Blooms reviderede taksonomi, kan man sætte spørgsmålstegn ved placeringen af det metakognitive niveau og hvordan det kan benyttes i testsammenhæng. Det opererer på et andet niveau end de andre kategorier på 'knowledge' akse (se kap. 4.2). I deres begrundelse for at inkludere metakognition siger de, at det 'reflects recent research on how students' knowledge about their own cognition and control of their own cognition play an important role in learning' (Anderson et al 2001:43). Det er altså gennem indflydelse fra socialkonstruktivistiske læringsteorier, at de har taget dette skridt.

De adskiller metakognition i 1) viden om kognition og 2) kontrol over kognitive processer (selvregulering): "This basic distinction between metacognitive knowledge and metacognitive control or self-regulating parallels the two dimensions in our Taxonomy Table. Accordingly, we have limited metacognitive knowledge to knowledge about cognition. The aspect of metacognition that involves metacognitive control and self-regulation reflects different types of cognitive processes and therefore fits into the cognitive process dimension, which is discussion" (ibid:43). I nærværende speciale bliver det for omfattende at komme i dybden med dette spørgsmål og forholdet til taksonomien, men man kunne forestille sig en inddragelse af metakognition i testformer. Man kunne lave test således, at der til hvert (fagligt) spørgsmål i en test også var en åben svarmulighed, hvor eleven kunne komme med reflekterende kommentarer. Dette kunne bruges som en form for feedback til læreren. Et eksempel på et metareflekterende elevsvar kunne være: i dette spørgsmål vidste jeg at shuffle og blues hørte sammen, men samtidig kunne shuffle jo også være en del af swingmusikken, men så brugte jeg udelukkelsesmetoden i det tilfælde, ellers synes jeg at spørgsmål hvor man skal matche svarmulighederne er en god udfordring.

Der mangler en dimension i Bloom's reviderede taksonomi, der siger noget om sværhedsniveau indenfor den enkelte kategori. Det er selve det enkelte læringsmåls sværhedsgrad. Man kan sagtens forestille sig at sværhedsgraden for en opgave på forståelsesniveau kan være sværere end en simpel analyseopgave. Desuden er der i praksis overlap mellem niveauerne i den kognitive dimension. Det vil med andre ord sige, at vi ikke kan bruge sværhedsgrad som parameter for skift i taksonomisk niveau, hvilket er problematisk! Disse problemer er nok ikke nye, oftest ser man taksonomien operationaliseret i opgaveformuleringer i gymnasiergi i en tre- eller firedeling som beskriv – analyser(forstå) – vurder. Det er også vanligt at dele op i paratviden, forståelse og højere ordensviden. Man kan tale om konceptuel viden på et bestemt taksonomisk niveau, hvor der kan være meget stor spredning i læringsmålets sværhedsgrad. Sværhedsgraden vil være afgørende for vægtning eller pointangivelse for det konkrete læringsobjekt, fx. kunne det være muligt at opdele spørgsmål i 5

sværhedsgrader (eller langt flere) med pointgivningen 1-5. Det er også nødvendigt at have en skala for hvor godt opgaven/svaret/præstationen er udført, hvis der er tale om udførelse af en opgave i ikke objektive test, eller når man skal vurdere en præstation. I disse tilfælde ville en mulighed være Dreyfus & Dreyfus hvor man opstiller 5 kategorier fra novice til ekspert (se uddybning i efterfølgende afsnit om adaptive test).

Taksonomien er, på trods af lidt uklarhed, et godt værktøj for læreren til at opstille og analysere læringsmål – undervisningsaktiviteter og elevvurdering. Disse tre didaktiske aspekter kan planlægges/tænkes ud fra taksonomien, man kan forestille sig at et læringsmål indebærer flere forskellige delaktiviteter. Den store risiko ved at planlægge læringsmål og faglige test i samme 'ramme' er, at det kan være med til at indsnævre de faglige mål således at de passer til testen og ikke omvendt. Vi mener at det er mest meningsfuldt at opstille læringsmål som udgangspunkt, hvorefter evalueringsformen tilpasses disse læringsmål. Man kan selvfølgelig forestille sig brede læringsmål, som kan deles op i nogle som kan testes og nogle som kræver en subjektiv vurdering.

Et problem, som ofte dukker op i forbindelse med debat om test, er udtrykket 'Teaching to the test' som er en risiko i meget test-orienterede miljøer. Dette kan også påvirke elevernes væremåde, som dette citat så tydelige beskriver: "Learners also activate the relevant situational, conditional, or cultural knowledge for solving a problem in a certain context. For example, they may know that the teacher uses only multiple-choice tests. Furthermore, they know that multiple-choice tests require only recognition of the correct answers, not actual recall of the information as in essay tests. This metacognitive knowledge might influence how they prepare for the test" (Anderson et al 2001:44).

### **Behaviorisme?**

Tanja Miller<sup>45</sup> siger om Blooms taksonomi at "Betragter vi taksonomien som en læringsteori, siger den kort fortalt, at man først må opbygge et solidt videnskabsmæssigt grundlag, for dernæst at forstå dette grundlag, således at man fx kan oversætte det til andre fremstillingsformer og fortolke det. Dernæst må man lære at anvende det i andre sammenhænge end deri, hvor det er lært – osv. indtil man har lært at foretage selvstændige vurderinger" (Miller:454). Senere siger hun at "Det vigtigste spørgsmål taksonomien og denne læringsteori stiller er: Hvordan er den enkeltes elevs forudsætninger. Der er ingen plads til de omstruktureringer af forhåndsviden og forforståelser andre læringsteorier opererer med, såsom konstruktivismen eller socialkonstruktionismen. Nej, læringen foregår trinvis og konfliktfri, hvis forudsætningerne er i orden. Den vigtigste opgave for læreren

---

<sup>45</sup> Bloom taksonomien s. 453-456 i kapitel 17 tanja miller Karaktergivning og afgangsprøver fra Introduktion til Undervisning – rammer, metoder, resultater (Frydenlund 2003)

bliver da følgelig dels at kende målene, være i stand til at diagnosticere forudsætningerne og endelig at kunne udarbejde en god undervisningsplan” (ibid:454). Vi er noget uenige med Tanja Miller, vi mener ikke det giver mening, at betragte Blooms taksonomi som en læringsteori. Taksonomierne siger efter vores opfattelse ikke noget om, hvilke læringsstrategier der er i spil, men derimod kun noget om, hvilke kognitive processer, der er i spil, når man lærer. Jens Dolin er inde på det samme, idet han konstaterer at “Taksonomien angiver (ideelt set) typen af intellektuel aktivitet som skal til for at opfylde pågældende trinmål, men da taksonomien ikke er baseret på en læringsteori, siger den ikke noget om indlæringsrækkefølge. Man kan godt forestille sig en didaktisk tilrettelæggelse som starter på et højere trin (fx af motivationsmæssige grunde) som indgang til arbejde med et lavere trin” (Dolin 2006:334).

I øvrigt er Blooms reviderede taksonomi påvirket af socialkonstruktivistisk læringsteori. Men man kan dog argumentere for at test indeholder et behavioristisk læringssyn, hvis test bruges til at bestemme hvad eleverne skal lære som isolerede bidder af viden, stramt ud fra læreplanen, hvor begreberne læres hierarkisk fra det simple til det mere komplekse. Vi ser feedback som det centrale bindeled mellem traditionel testning og formative testmetoder som kan fremme konstruktivistiske læringsprocesser. Dette vil vi forfølge i de næste afsnit.

### **6.3 Vygotsky og adaptive test**

Vi har i vores empiriske undersøgelser ikke haft mulighed for at foretage adaptive test. Vi har dog både i kapitel 2 og 3 diskuteret de adaptive IKT-baserede test i den danske folkeskole. Den indledende opmærksomhed på adaptive test og vores praktiske udførelse af ikke adaptive test har givet os ideer til og klarhed over, hvilke mangler vi ser ved de test vi har lavet. Det har særlig været den utilfredsstillende feedbackmulighed som har fået vores fokus. Det vil vi kigge nærmere på i næste afsnit.

Der er et vis fællesskab mellem dette afsnit og det næste, særlig gennem Vygotsky. Idet Vygotsky ikke har været teoretisk grundlag for vores empiri, gennemgår vi kort vores forståelse af hans teori også. Udgangspunktet for dette og det følgende afsnit er følgende: 1) I hvilken grad kan man inddrage Vygotskys teorier i testsammenhæng, særlig adaptive test? og det helt centrale 2) kan Vygotskys teorier være en afgørende faktor for læringspotentialer i test?

En faglig test, som den vi har gennemført, kan kun sige noget om elevens faktiske udviklingsniveau og ikke om det potentielle. Teorien om den nærmest udviklingszone handler netop om forholdet

mellem det faktiske og det potentielle udviklingsniveau. I Lindquist et al<sup>46</sup>, påvises netop denne problemstilling, hvor forfatteren nævner, at når man skal bestemme barnets udviklingsniveau, ansøger man ofte dette som statisk uden noget potentiale. Der gøres klart at Vygotsky var kritisk overfor testmetoder som kun bygger på det aktuelle niveau - slutresultatet, hvilket i særdeleshed var gældende for intelligens tests i hans levetid (Lindquist et al 2004: 286). Forfatteren nævner at den nærmeste udviklingszone ”aktualiserer spørgsmålet om, hvordan selve overgangen fra det ydre til det indre, internaliseringen<sup>47</sup>, foregår” (ibid:286). Internaliseringsprocessen er inderliggørelsen af de sociale og kulturelle former. Internaliseringsprocessen foregår gennem mediering<sup>48</sup>.

I Ivar Bråten et al<sup>49</sup> nævner forfatteren a + 1 princippet, hvor den gode undervisningssituation ligger i sværhedsgraden lidt over det udviklingsniveau, som børnene allerede har opnået (Bråten et al 2006:133). Her udfordres eleverne lige tilpas over det trivielle/banale uden at det bliver for svært. I lighed med Lindquist påpeger Bråten den Vygotsky-inspirerede tanke, at når først eleven har demonstreret sit aktuelle niveau, tilbyder læreren et samarbejde, som kan afdække det potentielle niveau. Vurderingsarbejdet, i vores tilfælde en test og lærerens viden om faglig progression/niveau, er så redskabet for planlægningen af den fremtidige undervisning og læringsmål. Dette kalder Bråten dynamisk vurdering (Bråten 2006:163). Den dynamiske vurdering adskiller sig fra traditionelle psykometriske tilgange ved at fokusere mere på at ‘beskrive individets ændring og modtagelighed for undervisning i løbet af problemløsningen end at fastslå individets relative placering’ (ibid:160). Der lægges vægt på vejledning, forklaringer og demonstrationer som hjælp til problemløsning (ibid:160). Han advarer om at en forståelse af udviklingszonen uden den sociohistoriske kontekst er mekanisk (ibid:164). Vi er således klar over at vores fortolkning af Vygotsky i dette tilfælde kan betrages på denne måde.

I den efterfølgende figur ser vi et konstrueret eksempel på en grafisk opstilling af en elevs resultat i en faglig test i matematik der er foretaget i en fiktiv klasse:

---

<sup>46</sup> ‘Vygotsky om læring og udviklingsvilkår’ Klim 2004

<sup>47</sup> internalisering: viden går fra det interpersonelle til det intrapersonelle

<sup>48</sup> Mediering betyder mægling, formidling, mellemkomst og betegner hvorledes den enkelte forbinder sig med omverdenen. (Gympæd:166)

<sup>49</sup> ‘Vygotsky i pædagogikken’ Frydenlund 2006



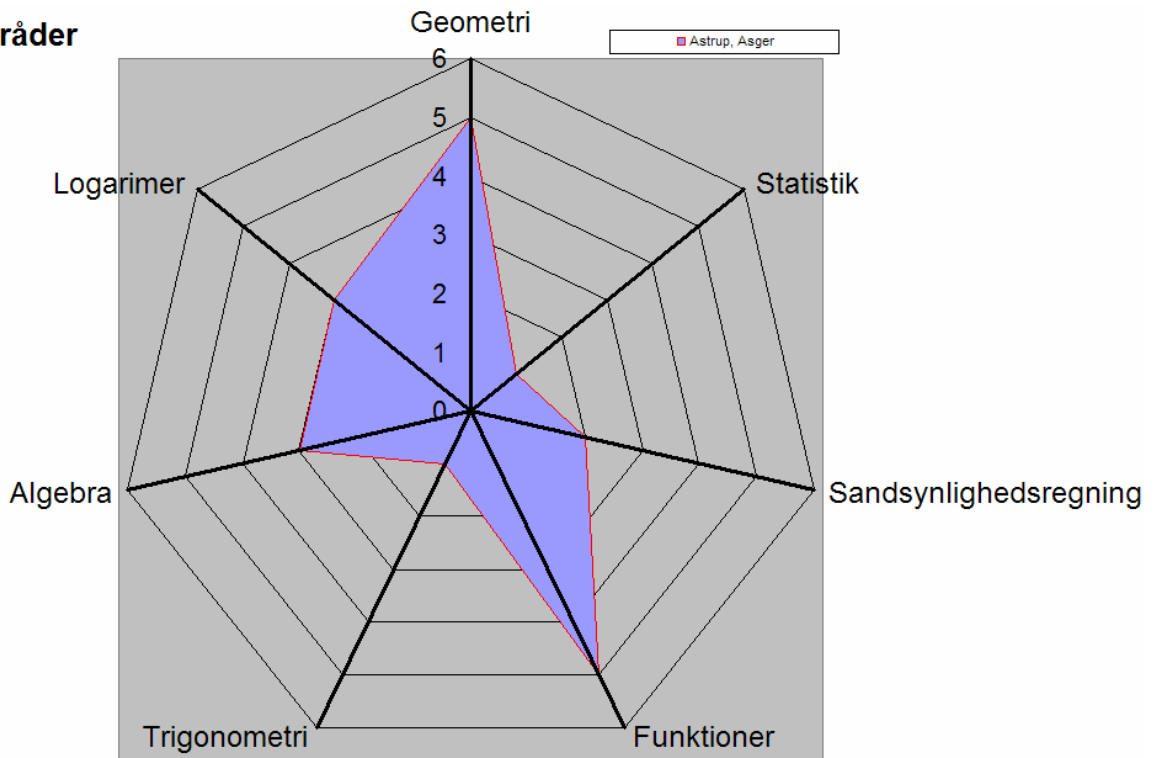
**Profilområder**


Fig. 6.3

Den viser de profilområder i faget som er blevet undersøgt, en skala fra 1-6 viser, hvilket niveau eleven har gennemført testen på (det kunne også være flere summative evalueringer i løbet af skoleåret som er opsamlet). Figuren viser altså elevens aktuelle niveau i de dele af fagområderne som kan testes, vi ser således bort fra test-problematikkens faglige dækningsgrad i eksemplet.

Hvis vi antager, at det tilfredsstillende niveau for en elev i klassen på det konkrete undervisningstidspunkt er 4, ser vi meget klart, hvor eleven har sine stærke og svage sider, i forhold til det ønskede faglige trinmål. Vi har en formodning om, at det er denne tankegang der ligger bag de nationale tests adaptive princip, uden at vi dog har fået det bekræftet. Hvad kan vi så bruge det til? Læreren har nu mulighed for at se hvordan den enkelte elev er stillet indenfor profilområderne. Her vil det så være muligt at se på egen undervisning og undersøge om tendenserne er fælles for alle elever, eller om der kun er enkelte elever som viser 'mangler' i de enkelte emneområder. Vi er klar over, at det er en viden som ikke nødvendigvis hænger sammen med test, men sagtens kan være en del af en lærers formative vurdering/dokumentation på anden vis.

I en undervisnings- og vurderingsproces som den ovenfor, trænger begrebet scaffolding sig særligt på. Det er blevet introduceret af Jerome Bruner og giver eleven en model eller struktur for, hvad der skal indlæres. Vi vil i det følgende beskæftige os mere indgående med de læringsmæssige muligheder som feedback og scaffolding kan give os i testsammenhæng.

## **6.4 *Formativ evaluering, feedback og læring***

Vi har som bekendt benyttet Quia (se afsnit 3.3) som testredskab, men hvilke feedbackmuligheder giver Quia? Feedbackmulighederne i Quia er begrænsede, idet der kun kan gives standardsvar til alle brugere. Der kan gives feedback på både forkerte og rigtige svar. Et eksempel på et svar er: ”Forkert, smør er federe end øl, husk at øl består mest af vand”. Dette svar indeholder operant betingning ”forkert” og info som uddyber hvorfor svaret er forkert. Men er et sådant automatisk feedback fra et testsystem dækkende? Kan et testsystem give tilfredsstillende feedback som fører eleven videre i sin læring? Det mener vi kun, at det kan i begrænset omfang, den menneskelige/personlige feedback er det centrale. Computeren har det positive, at den har mulighed for at give umiddelbar feedback, men vi mener ikke at den computergenererede feedback er tilfredsstillende for eleverne, i betydningen standardiserede tilbagemeldinger om rigtigt henholdsvis forkert, hvilket også understøttes af vores holdningsundersøgelse (se afsnit 5.5). På den anden side, er det den umiddelbare hurtige feedbacks force, at den foregår under opgaveløsningen altså i konteksten, i modsætning til face to face feedback, der ofte sker meget senere måske flere uger efter opgaveløsningen, hvor konteksten er en hel anden og eleven er måske allerede kommet videre eller har glemt sammenhængen.

Den autogenererede type feedback kan også kaldes mekanisk (Earl 2003:114). Hun siger, at feedback kan deles op i evaluerende og beskrivende feedback, hvor eksempler på evaluerende feedback er belønning og straf (fx. karakter) og at udtrykke anerkendelse eller misbilligelse. Mens eksempler på beskrivende feedback er at fortælle eleven, at de har ret eller uret, foruden at beskrive hvorfor et svar er korrekt. Man kan også fortælle elever, hvad de har opnået og hvad de ikke har opnået/præsteret, foruden at angive eller antyde en bedre måde at gøre noget på. Det optimale er, at få eleven til selv at foreslå måder hvorpå de kan blive bedre eller gøre fremskridt (ibid:91). Vores empiri problematiserer værdien af feedback yderligere: Holdningsundersøgelsen peger på, at eleverne gerne vil have feedback fra læreren, men omvendt betvivler lærerne i den jyske fokusgruppe værdien af feedback, idet fokusgruppen stiller spørgsmålstegn ved om eleverne bruger feedback konstruktivt (se afsnit 5.1.2). Her er altså et konkret eksempel på et skisma mellem eleverne opfattelse af feedback og lærernes opfattelse, af hvordan elever bruger feedback.

### **Metakognition og selvevaluering**

Metakognition kan beskrives som at lære at lære og omfatter evnen til at kunne kontrollere sin egen læreproces. Earl ser elevens rolle som værende ‘the critical connector between the assessment and learning process’ og ser dette som den regulerende proces i metakognition (Earl 2003:25). Hun si-

ger også om metakognition at “It occurs when students personally monitor what they are learning and use the feedback from this monitoring to make adjustments, adaptations, and even major changes in what they understand. Assessment as Learning is the ultimate goal, where students are their own best assessors” (ibid:25).

Hun har en tredelt opdeling af evaluering som er vist i figuren nedenfor:

Features of assessment of, for, and as Learning			
Approach	Purpose	Reference Points	Key Assessor
Assessment of Learning	Judgements about placement, promotion, credentials, etc.	Other students	Teacher
Assessment for Learning	Information for teachers' instructional decisions	External standards or expectations	Teacher
Assessment as Learning	Self-monitoring and self-correction or adjustment	Personal goals and external standards	Student

(Fig. 6.4 efter Earl 2003:26)

Her er den første type sammenlignelig med summativ evaluering, og den næste med formativ evaluering og den tredje type går et skridt videre og kan sammenlignes med selvevaluering: “In this reconfigured assessment environment, assessment would make up a large part of the school day, not in the form of separate tests, but as a seamless part of the learning process. And there would be test when the decisions to be made require identification of a few individuals or groups, or when a summative description is important for students and others as a milestone or rite of passage (ibid: 27-28). Hun har her nogle spændende principper om en meget høj grad af elevinddragelse, som umiddelbart kan virke som en enorm udfordring i undervisningssammenhæng. Et synspunkt hos Dolin som understreger dette er at “Det er ikke uproblematisk at arbejde med metakognition i undervisningen. Fokus skal flyttes fra det der skal læres, til hvordan det læres bedst. Et sådant perspektivskifte vil mange elever finde både besværligt og have svært ved at se nytten af. Dette gælder især elever som har en instrumentel tilgang til skolegangen, dvs. som vil lære det krævede fordi det kræves snarere end af interesse for emnet” (Dolin 2006:146).

### Vygotsky og feedback.

Earl beskriver feedback således ”Feedback can lead to increased effort or engagement, alternative strategies to understand the material, and the restructuring of understanding” (Earl 2003:89). Feedback kan betragtes som scaffolding, som igen peger i retning af det centrale i Vygotskys begreb om

den nærmeste udviklingszone<sup>50</sup>. Læring foregår, ifølge Vygotskys perspektiv, i en social kontekst, hvor eleven observeres og støttes af en mere kompetent person. I denne proces afdækkes læringsbehovet og den nødvendige hjælp ydes af den mere kompetente. I undervisningssammenhæng giver læreren, eller en anden person, feedback ud fra elevens tidligere præstationer og nuværende formåen. Læreren har kendskab til elevens niveau fra tidligere præstationer som f.eks. test. I kapitel 2 viste vi en figur som viser den formative evalueringsproces (kap. 2). Evalueringsprocessen kan også beskrives som en iterativ proces bestående af læringsmål – læringsaktiviteter - vurdering/testning af forståelse – feedback, som illustrerer den tydelige forskel mellem summativ og formativ evaluering. Earl formulerer det således “Taken together, curriculum teaching, learning, and assessment interact in an iterative and sometimes cyclical process” (Earl 2003:83).

Den formative evaluering består af scaffolding eller coaching, hvilket er fraværende i den summative evaluering. En virkelig brugbar feedback stiller krav til både formidler og elevens refleksive formåen. Hvis feedback skal give mening skal eleven kunne reflektere over sin egen forståelse, kunnen og præstationer, og kunne analysere eller metareflektere over det. Earl siger om den formative evaluering, at ”It helps teachers provide the feedback to scaffold next steps. And it depends on teachers’ diagnostic skills to make it work” (ibid:24). Vores empiriske materiale viser i den forbindelse tegn på, at eleverne tror på nytten af den faglige test (se afsnit 5.5). I de spørgsmål, der omhandler testen i forhold til fagligt standpunkt og egen vurdering af faglige kompetencer (sp. 1,2,3,4,5 og 10) er der en i overvejende grad positiv vurdering af testens brugbarhed (måske lidt tvetydigt hos biologieleverne). Hvis det er rigtigt medfører det jo alt andet lige en eller anden form for metakognition eller metarefleksion hos eleven (eleven reflekterer over, at testen kan være et nyttigt evalueringværktøj).

Efter et Vygotsky-inspireret socialkonstruktivistisk princip mener vi, at feedback skal indeholde en skriftlig eller mundtlig dialog mellem lærer og elev.

### **Synliggørelse af læringsmål.**

Vi har i det ovenstående argumenteret for, at summative test kun afdækker den aktuelle udviklingszone eller område, men ved hjælp af mediering, kan den mere erfarne lærer afdække elevens nærmeste udviklingszone, som er afstanden mellem de aktuelle og de potentielle udviklingsmuligheder. Hos Vygotsky opstilles ofte i forbindelse med mediering og internaliseringsprocessen forholdet mellem subjekt, artefakt og objekt (se Dolin 2006:166).

---

<sup>50</sup> Begrebet scaffolding er ikke udviklet af Vygotsky, men i tæt tilknytning til hans teori om den nære udviklingszone. (Dolin:169)

En analyserende opdelt oversigt over læringsmål er egnet som grundlag for feedback. Ved at bygge på kriterier, som er fastlagt i planlægningsfasen og at eleven kender progressionen før undervisningsforløbet starter, bliver det et redskab for eleven.

Her følger et eksempel på en holistisk eller evaluerende<sup>51</sup> oversigt over læringsmål:

Udførelse af opgave	Score
Elevens udførelse var meget sikker og overbevisende	4-5
Udførelsen var noget rutinepræget	3
Udførelsen var usikker og kun delvis sammenhængende	1-2
Udførelsen var usammenhængende	0

Fig. 6.5

Denne type opstillinger er ikke tilfredsstillende nok for eleven. Kan sammenlignes med denne feedback: "Du svarede rigtigt på 73% af spørgsmålene og får derfor karakteren 8". Eleven har mere brug for at få at vide hvordan han kan opnå bedre resultater, og dette skal synliggøres for eleven. En oversigt skal kunne bruges som grundlag under evalueringssamtaler eller som et selvevaluerings redskab.

I det efterfølgende eksempel, som kan kaldes analytisk efter Earls distinktion, har vi brugt Dreyfus & Dreyfus<sup>52</sup> som målestok for kvaliteten/sværhedsgraden af elevens præstation:

Læringsobjekt/område	Avanceret begynder <sup>53</sup>	Kompetent udøver	Kyndig udøver	Ekspert
Stemmebrug i klassisk jazz	Noget mekanisk og udtryksløs og ikke helt tonerent	Der synges rent bortset fra de sværeste passager, men der er ikke den store formning af fraserne	Der er en stilren frasering, og et dynamisk udtryk som ligger tæt op at originalen, uden at være efterabende	Udførelsen lever op til idealet for stemmebrug, og der er desuden tale om et personligt udtryk i fremførelsen
Jeg skal arbejde med:	Brug støtte på de lange toner i fraserne foruden en del vibrato, øv særlig passagerne i B-stykket			

Fig. 6.6

Dette er et konstrueret eksempel på en udspecificering af et forholdsvis snævert fagligt område, som bygger på et sang-teoretisk ideal, som der i undervisningen er skabt en fælles forståelse af. Der benyttes konkrete begreber som skal udtrykke hvilke mål og mangler der er i udførelsen og hvad eleven allerede har opnået. Sådanne oversigter kan også bruges som grundlag for udarbejdelsen af testopgaver (se afsnit 4 om faglige test). Det er således både et lærings- og et vurderingsværktøj og et grundlag for elevens konstruktion af læringen ved hjælp af scaffolding. Læringsmålbeskrivelser

<sup>51</sup> Earl skelner mellem evalative og descriptive feedback (s. 91)

<sup>52</sup> Efter Hermansens 2001:84

<sup>53</sup> De har et første niveau som vi har udeladt i eksemplet: nybegynder

er efter vores mening et fleksibelt værktøj i læringsmiljøet, som synliggør læringsmålet for læreren i planlægningsfasen og i evalueringsfasen, for eleven i undervisningssituationen og understøtter elevens refleksionsprocesser.

Et andet eksempel på en velegnet visuel fremstilling er figuren af elevens testprofil fra forrige afsnit. Dette er igen et eksempel på synliggørelse af læringsmål som en konkret artefakt i en scaffolding- eller vejledningsproces.

Det er som fremført i afsnittet om udviklingen af de faglige test omfattende at lave den slags klart definerede læringsplaner, men det er nødvendigt, at oversigten er nok udspecificeret så eleverne kan se en progression. Det nytter ikke, at der er fire niveauer, som skal opnås på fx. to år, så kan der blive motivationsproblemer, hvis eleven ikke selv fornemmer fremskridt.

## 7. Konklusion

I denne afhandling har vi haft et målrettet fokus på IKT-baseret evaluering og test, og bredere på elev- og læringsforudsætninger. De empiriske undersøgelser kan fra en overordnet betragtning ses som et projekt i projektet, hvor vi afprøver undersøgelsesformer som vi har valgt at kalde test. Dette er et bevidst valg, test som begreb kan kaldes kontroversielt, men er højaktuelt, også i Danmark og testene bruges bemærkelsesværdigt nok i mere eller mindre samme form, som dengang curriculum-forskningen startede i USA for over 50 år siden. Området har således ikke fulgt samme landvinding som den pædagogiske forskning i perioden. Det øgede fokus på standardisering og kravene om dokumentation og evaluering har også nået det danske skolesystem, og er en udvikling som ikke ser ud til at kunne standses. Vi har taget udgangspunkt i denne virkelighed og har stillet spørgsmålet om i hvilken grad IKT kunne være et egnet redskab i udviklingen af test, og endvidere undersøgt i hvilken grad der er et læringspotentiale i test.

Vores nære virkelighed er den daglige undervisning i Gymnasiet, hvor der er sket markante forandringer med en ny gymnasireform, som har rykket ved den eksisterende pædagogiske praksis. Det pædagogiske udgangspunkt i gymnasiet og folkeskolen synes noget forskellige, måske især pga. gymnasiefagernes forankring i de forskellige universitetsfakulteters videnskabsideal. Det er dog åbenlyst, at de pædagogiske teorier som operationaliseres i undervisningen i folkeskolen, også holder sit indtog i gymnasiet. Det er de samme elever som ind og udsluses i skolesystemet uanset aktuelt skoletrin.

Vores problemformulering fra indledningen lød således: **Hvordan kan IKT-baserede test inddrages i gymnasiet som et redskab der fremmer læring.** Vi vil nu komme med de afsluttende kommentarer til de hypoteser, som vi udkrystalliserede fra problemstillingen.

- **Elevers præstation i faglige test er betinget af deres læringsstile og multiple intelligensformer**

Gardners teori om multiple intelligenser og Dunn og Dunns teori om læringsstile, udgør en del af det teoretiske grundlag for vores empiri. Vi har vist, at der er en korrelation mellem elevernes testresultat i den biologifaglige test og den logisk-matematiske intelligensform. Denne hypotese er således verificeret med hensyn til den logisk-matematiske intelligens for biologi-eleverne, og denne sammenhæng er signifikant. Vedrørende læringsstile er der ikke

fundet nogen forbindelse mellem elevernes læringsstil og deres testresultat. Hvilket kan skyldes at testudformningen favoriserer logisk-matematisk intelligens.

Vi har i fremstillingen påpeget, at viden om elevernes foretrukne læringsstil og multiple intelligensform(er) kan være et pædagogisk redskab i undervisning og ved udvikling af test. Det betyder, at man som evaluator skal overveje, om man kan stille spørgsmålene på en sådan måde, at andre intelligensformer evt. kan tilgodeses, i hvert fald, hvis man vil operationalisere Gardner og Dunn & Dunn. Det kunne være et udgangspunkt for videre undersøgelse indenfor IKT-baseret testudvikling.

Vores empiri angående elevernes læringsstile viser, i modsætning til hvad Dunn & Dunn har fundet, at flertallet er af den analytiske type og mindretallet er af den holistiske type. Vi kan ikke komme med en plausibel forklaring på, hvorfor vores resultater peger i modsætning af deres resultater.

Vi har diskuteret ligheder og forskelle i mellem Gardners teori og Dunn og Dunns teori, uden at nå til en endelig afklaring. Vores empiri viser også tegn på, at intelligensformen kinæstetisk/kropslig intelligens korrelerer til den kinæstetisk/taktile læringsstil og spatial intelligens korrelerer tilsvarende til den visuelle læringsstil. Der er dog ikke tale om signifikante sammenhænge.

- **IKT-baserede test kan anvendes til at evaluere kundskaber på højere taksonomiske niveauer end paratviden**

Vi har foretaget testning på paratviden, forståelse og analyse indenfor snævre faglige grænser. Empirien kan således ikke sige os hvor grænsen går for hvad det er muligt at teste, fordi det er for lidt empirisk grundlag. Men vi har formodninger om, at fremtidig IKT udvikling vil gøre det muligt. Vi hævder på af baggrund vores overvejelser, gennem de test vi har udviklet og afprøvet, at test kan evaluere kundskaber på højere taksonomiske niveauer end paratviden. Vi kan dog ikke evaluere læring på metakognitivt niveau og metarefleksivt niveau altså niveauer, hvor eleverne skal forholde sig til deres egen læring med den testudformning som vi har benyttet, nemlig lukkede spørgsmål.



- **Feedback har betydning for testens nytteværdi i læringsammenhæng**

Denne hypotese er klart verificeret, både i empiri og i teoretisk fremstilling, og vi mener i fremstillingen at have afgørende dokumentation for at dette er tilfældet. Feedback er et godt grundlag og input for elevens refleksion. Formålet med testen afgør hvilken form for feedback er den mest hensigtsmæssige. Med de rapporteringsmuligheder, der følger med faglige IKT-baserede test har vi påvist, at IKT-baserede test giver gode muligheder for at dokumentere en evalueringsindsats. Dette er et formelt krav i gymnasireformen, men som nok så vigtigt også er en dokumentation i forhold til den enkelte elev. Derved bliver testen eller rettere testresultatet et læringsværktøj for den enkelte elev.

Vi argumenterer i den forbindelse for, at Vygotskys teori om den nærmeste udviklingszone kan komme i spil som læringsteoretisk grundlag for den formative evaluering. Nyttевærdien af dokumentationen står og falder dog med kvaliteten af den feedback, der gives på testen. De faglige test vi har afprøvet viser, at en standardiseret og instrumentel feedback ikke er tilstrækkelig. Eleverne ønsker en mere kvalificeret feedback.

- **IKT-basering påvirker elevernes oplevelse af at gennemføre test i positiv retning**

Hypotesen er verificeret i elevernes svar i holdningsundersøgelsen. Resultaterne af denne viser, at eleverne foretrækker IKT frem for papirbasering og de foretrækker også spørgsmål, hvori der indgår multimedievirkemidler. De virkemidler multimediet tilbyder med hensyn til lyd, illustrationer, billeder, animationer mv. gør, at der kan stilles en bredere vifte af spørgsmål og spørgsmål som tilgodeser en bredere vifte af læringsstile og intelligenser. Vi kan dog ikke med vores empiri vise, at disse virkemidler har en effekt på testresultaterne. Det skyldes givetvis, at spørgsmålenes sværhedsgrad afhænger af flere faktorer, herunder det taksonomiske niveau spørgsmålet sigter efter, formuleringen af spørgsmålet og svarmulighederne.

Der kan også være andre grunde til at eleverne svarer positivt på denne hypotese, fx at de synes, at de digitale test er nemmere at benytte end papirbaserede test, det kan være talen om generel computerfascination og holdningen til at alt er nemmere med computer. Mange elever er normalt ikke glade for at skrive længere passager, computeren er her en hjælp til dem, som har sværere ved at formulere sig, end dem der er sprogligt stærke. Vi mener at computermedieret testning indeholder et potentiale, som kan forbedre og kontekstualisere test, i

modsatning til traditionel testning. Vi har ikke haft innovative redskaber til rådighed til at udvikle og eksperimentere med computermedieret testdesign.

- **Testning og læring er i grundlæggende modstrid med hinanden, idet de baserer sig på forskellige paradigmer**

Denne hypotese er på sin vis falsificeret, idet det er muligt at lave test som har et formativt islæt. Fokusgruppeundersøgelsen viser modstridende forklaringer på nytteværdien af test og syn på test. Dette bunder måske i forskellen mellem amerikansk testtradition og dansk didaktiktradition, og mellem behaviorisme og socialkonstruktivisme.

Fokusgrupperne er skeptiske overfor om test kan eller skal være en del af evalueringskulturen. Fokusgruppen, der repræsenterer naturvidenskab er dog ikke helt så afvisende overfor test som den humanistiske fokusgruppe. Men det, der også træder frem her, er en diskurs, som fremhæver forskellene på angelsaksisk curriculumtradition og skandinavisk didaktiktradition. Man får det indtryk af fokusgrupperne, at på trods af at evalueringskulturen forekommer utydelig på de to gymnasier, så manifesterer der sig en socialkonstruktivistisk forankret diskurs i interviewene om, at en gymnasial evalueringskultur ikke skal forankres i curriculumtænkning.

Den endelige afrunding, som vi på falderebet vil uddrage fra afhandlingens undersøgelser, er at et skift fra summativ til formativ vurdering og fra et behaviouristisk til et socialkonstruktivistisk syn på læring er signifikant for at kvalificere højere ordens viden. Vi foretrækker at betragte disse positioner som eksisterende som et kontinuum, hvor det teoretiske udgangspunkt afhænger af formålet med de faglige intentioner, set i sammenhæng med praksis.

## 8. Bilagsoversigt

<u>Kapitel</u>	<u>Bilagsnr.</u>	<u>Emne</u>
Kap. 3	1	16 typer af prædefinerede spil og læringsaktiviteter i Quia
Kap. 4	2	Web materiale brugt som grundlag for MI og læringsstil undersøgelse
	3	Spørgsmål til undersøgelse af læringsstile
	4	Spørgsmål til undersøgelse af MI
	5	Blooms reviderede taksonomi
	6	Faglig test prototype biologi
	7	Faglig test endelig biologi
	8	Faglig test endelig musik
	9	Interviewguide fokusgrupper
	Kap.5	10
11		Fokusgruppeinterview det færøske gymnasium
12		Faglig test biologi elevbesvarelser
13		Faglig test musik elevbesvarelser
14		Elevbesvarelse af Læringsstile og MI former
15		Elevbesvarelse holdningsundersøgelse biologi
16		Elevbesvarelse holdningsundersøgelse musik

## 9. Litteraturliste

### **Albrecht, Jakob**

[http://www.dpu.dk/everest/Publications/Udgivelser/Asterisk/Asterisk%20nrx2E%2030%20september%202006/20060828122025/CurrentVersion/asterisk\\_30\\_s24\\_27.pdf](http://www.dpu.dk/everest/Publications/Udgivelser/Asterisk/Asterisk%20nrx2E%2030%20september%202006/20060828122025/CurrentVersion/asterisk_30_s24_27.pdf)

### **Anderson, Lorin (2001):**

A Taxonomy for Learning, Teaching, and Assessing.

### **Bates, Bill (2001):**

Supporting a range of learning styles using a taxonomy-based design framework approach.

<http://www.ascilite.org.au/conferences/melbourne01/pdf/papers/batesb.pdf>

### **Bendixen, Carsten (2005):**

Evaluering og læring. Kroghs forlag.

### **Bricheno, Pat and Younger, Mike (2004)**

Some unexpected results of a learning styles intervention?

<http://www-rba.educ.cam.ac.uk/Paper%20PB.pdf>

### **Bråten, Ivar (2006):**

Vygotsky i pædagogikken, Frydenlund

### **Christensen, Torben Spanget (2004):**

Integreret evaluering. Phd afhandling Odense. Syddansk Universitet.

<http://www.humaniora.sdu.dk/phd/dokumenter/filer/Afhandlinger-42.pdf>

### **Christensen, Torben Spanget (2006):**

Formativ evaluering. I Gymnasiepædagogik – en grundbog. Red. Damberg Erik, Jens Dolin og Gitte Holten Ingerslev. Hans Reitzels Forlag

### **Dahler-Larsen, Peter (2005):**

Evalueringskultur – et tiltrækkende begreb?

Evalueringskultur – forandringskultur, Unge pædagoger nr. 2/3 juni 2005

### **Dahler-Larsen, Peter (2006):**

Evalueringskultur: et begreb bliver til. - 1. udgave, 1. oplag. - Odense: Syddansk Universitetsforlag.

### **Damberg, Erik (2006):**

Gymnasiepædagogik - en grundbog red. af Erik Damberg, Jens Dolin og Gitte Holten Ingerslev.  
Hans Reitzels Forlag

**Dolin, Jens (2001):**

At lære fysik.

<http://pub.uvm.dk/2001/fysik/index.html>

**Dolin, Jens (2006):**

Dannelse, kompetence og faglighed. I Gymnasiepædagogik – en grundbog. Red. Damberg Erik, Jens Dolin og Gitte Holten Ingerslev. Hans Reitzels Forlag

**Dunn, Rita (2006):**

The Dunn and Dunn Learning-Style Model and Its Theoretical Cornerstone.

[http://www.learningstyles.net/index.php?option=com\\_docman&task=cat\\_view&gid=30&Itemid=73&lang=en](http://www.learningstyles.net/index.php?option=com_docman&task=cat_view&gid=30&Itemid=73&lang=en)

**Dunn, Rita (2003):**

Artikelsamling om læringsstile. Dafolo Forlag

**Earl, Lorna M. (2003):**

Assessment As Learning, Corwin Press

**Egelund, Niels (2003):**

Evaluering og internationale sammenligninger

Introduktion til undervisningen, kap.14, Finn Held (red.), Frydenlund.

**Egelund, Niels (2005):**

Evaluering – ja, men først og fremmest løbende intern evaluering i en dansk tradition.

Evalueringskultur – forandringskultur, Unge pædagoger nr. 2/3 juni 2005

**Fibæk Laursen, Per (2005):**

Individualiseringens pædagogik? I: perspektiver på de mange intelligenser. Hansen, Fibæk Laursen og Maj Nielsen 1. udg. Roskilde Universitetsforlag

**Fowler, F.J. (1993):**

Survey Research Methods.

Thousand Oaks: Sage Publications

**Gardner Howard (1993):**

Multiple Intelligences. The theory in practice. Basic books.

**Gardner, Howard (1983):**

Frames of mind. The Theory of Multiple Intelligences.

Heinemann London

**Gardner, Howard (1999):**

Den intelligente skole.

Gyldendal

**Gleerup, Jørgen (2006):**

Pædagogisk evaluering. I Gymnasiepædagogik – en grundbog. Red. Damberg Erik, Jens Dolin og Gitte Holten Ingerslev. Hans Reitzels Forlag

**Graaff, Erik de (2000):**

Assessment and educational development, Videncenter for Læreprocesser Aalborg Universitet.

**Halkier, Bente (2003):**

Fokusgrupper. Samfundslitteratur. Roskilde Universitetsforlag

**Hansen, Mogens (2005):**

De mange intelligenser – mangfoldighedens pædagogik. I: perspektiver på de mange intelligenser.

Hansen, Fibæk Laursen og Maj Nielsen 1. udg. Roskilde Universitetsforlag

**Haue, Harry (2004):**

Almendannelse for tiden: en ledetråd i dansk gymnasieundervisning. Odense Syddansk Universitetsforlag

**Haywood, H. Carl (1992):**

Interactive Assessment, Springer-Verlag.

**Hermansen, Mads (2001):**

Læringens univers, Forlaget Klim.

**Hermansen, Mads (2005):**

Læringens univers. Forlaget Klim

**Hiim Hilde og Else Hippe (1997):**

Læring gennem oplevelse og handling. En studiebog i didaktik. Gyldendal / Undervisning.

**Hildebrand Jensen, Annette; Kamilla Frimodt Madsen og Lilian Rohde (2005):**

Læringsstile - hvordan man kan ligge i blød sofa, gnaske gulerødder og samtidig lære noget nyt og svært.

AGORA nr. 6: <http://www.cvustork.dk/agoranr6.asp>

**Jackson, Charles (2002):**

Psykologisk testning, Psykologisk forlag

**Jacobsen, B., Schnack, K., Wahlgren, B. og Madsen, M. B. (1999):**

Videnskabsteori.

Kbh.: Gyldendal Uddannelse.

**Klafki, Wolfgang (2002):**

Dannelsesteori og Didaktik 2.udg. Forlaget KLIM.

**Krogh, Ellen (2006):**

Pædagogisk evaluering. I Gymnasiepædagogik – en grundbog. Red. Damberg Erik, Jens Dolin og Gitte Holten Ingerslev. Hans Reitzels Forlag

**Lave, Jean (1993):**

The practice of learning. I: Understanding practice – Perspectives on activity and context Edt. Seth Chaiklin & Jean Lave. Cambridge University Press.

**Lindqvist, Gunilla (2004):**

Vygotskij – Om læring som undervisningsvilkår, Klim

**Mayes, Terry:**

JISC e-Learning Models Desk Study

[http://www.jisc.ac.uk/uploaded\\_documents/Stage%202%20Learning%20Styles%20\(V%201\).pdf](http://www.jisc.ac.uk/uploaded_documents/Stage%202%20Learning%20Styles%20(V%201).pdf)

**Miller, Tanja (2003)**

Karaktergivning og afgangsprøver

Introduktion til undervisningen, kap.14, Finn Held (red.), Frydenlund.

**Miller, Tanja (2005):**

”Hvordan synes du selv, det går?”

Evalueringskultur – forandringskultur, Unge pædagoger nr. 2/3 juni 2005

**Miller, Tanja (2006):**

Summativ evaluering. I Gymnasiepædagogik – en grundbog. Red. Damberg Erik, Jens Dolin og Gitte Holten Ingerslev. Hans Reitzels Forlag.

**Mortimore, Peter; Maria David-Evans; Reijo Laukkanen; Jouni Valjarvi (2006):**

OECD-rapport om grundskolen i Danmark – 2004. Uddannelsesstyrelsens temahæfteserie nr. 5 – 2004.

**Olsen, Henning (2005):**

Fra spørgsmål til svar – konstruktion og kvalitetssikring af spørgeskemadata

Akademisk Forlag

**Pedersen, Kirsten Bransholm & Birgit Land (2001):**

Den kvalitative forskningsproces. I Pedersen og Nielsen (red.): kvalitativ metode – fra metateori til markarbejde, 1. Udg. Roskilde Universitetsforlag

**Rust, John and Susan Golombok (1999):**

Modern Psychometrics - The Science of Psychological Assessment. Second Edition  
Routledge.

**Schultz, Majken (2004):**

Kultur i organisationer – funktion eller symbol. Handelshøjskolens forlag

**Sørensen, Christen (2006):**

De nationale it-baserede obligatoriske test – frelser eller uhyre?, Syddansk Universitet.

**Undervisningsministeriet (2003):**

<http://www.uvm.dk/nyheder/gymaftale.doc>

**Undervisningsministeriet (2004):**

<http://www.retsinfo.dk/index/UND/AT001702.htm>

**Undervisningsministeriet (2006):**

Læreplan og vejledning til læreplanen kan ses på følgende adresse:

[http://us.uvm.dk/gymnasie/fagene/biologi/ny\\_stx.htm?menuid=15](http://us.uvm.dk/gymnasie/fagene/biologi/ny_stx.htm?menuid=15)

**Undervisningsministeriet (2007):**

<http://www.uvm.dk/evalueringskultur/fremme/index.htm>

**Williams, David (2006):**

Online assessment, measurement and evaluation, Science publishing.