

Working Paper Series
CSER WP No. 0010

Larm og læring: Klasserumsklimaets betydning for elevers læring

Bent Sortkær*

Maj 2013



Larm og læring:

Klasserumsklimaets betydning for elevers læring

Bent Sortkær*

This version: Maj 2013

Resumé:

Som en replik til de seneste års fokus på den urolige danske folkeskole, undersøger denne artikel, dels hvor meget uro der er i klasseværelserne rundt i Danmark, dels om de danske elever er særlige urolige i forhold til vores nordiske nabolande, og endelig hvad denne uro betyder for elevernes faglige udbytte. Med udgangspunkt i PISA 2003 og ved hjælp af deskriptiv statistik finder artiklen at over halvdelen af de danske 15-årige elever ofte føler sig forstyrret i timerne. Resultatet af en multilevel model viser at denne oplevede uro har en signifikant og betydelig sammenhæng med hvor godt den enkelte elev klarer sig fagligt. Ved nærmere analyse viser det sig desuden, at drenge og børn af forældre uden uddannelse bliver negativt påvirket i særlig grad af uro i klassen. Dette resultat kalder dog på yderligere forskning. Og endelig viser en sammenligning med Sverige og Norge, at de danske elever ikke er støjende i særlig grad.

Abstract:

As a remark to the ongoing discussion about the noisy Danish compulsory school this paper examines the disciplinary climate in the Danish classrooms. Using data from PISA 2003 and descriptive statistics this paper finds that more than half the 15 year old students often experience disorder in the classrooms. Results from a multi-level model show a significant relationship between the disciplinary climate in the classrooms and student achievement. Further analysis reveal that the negative effect of the disciplinary climate is larger for boys and students with low-educated parents. Further research is needed to elaborate this finding. Finally by comparison this paper finds that Denmark is not much different from Sweden and Norway in terms of the disciplinary climate and the relationship to student achievement.

* Institut for Uddannelse og Pædagogik, Aarhus Universitet, Niels Juels Gade 84, 8200 Aarhus N. E-mail: beso@dpu.dk

Total word count (including footnotes): 7,616, four tables.

Key words: cultural capital, educational success, causal effect, difference-in-difference, United States

Tak til CCIELO for faglig sparring og økonomisk støtte.

Antal ord (inklusive fodnoter): 17.073, 16 tabeller

Key words: disciplinært klima, læring, multilevel model, PISA 2003

1. Introduktion

Uro, larm og dårlig disciplin og det at drive skole har gennem tiderne gået hånd i hånd. Det vidner korporlige metoder som kindhæste og benyttelse af spanskrør om. Metoder som er blevet benyttet netop for at mindske omfanget af larm og uro og dermed øge chancen for læring. Forskningen viser da også entydigt, at den danske grundskole er præget af meget støj og mange forstyrrelser (Egelund & Hansen, 1997; Egelund, Nielsen & Randvid, 2011), og tal fra PISA 2003 placerer Danmark som nummer 30 ud af 41 deltagende lande på listen over de lande med det bedste disciplinære klima (Mejding (Red.), 2004).

Ligeledes viser forskningen en i Danmark negativ sammenhæng mellem hvor meget uro den enkelte elev oplever og dennes faglige standpunkt (Mejding (Red.), 2004; Hermansen (Red.), 2007). Dermed er forstyrrende adfærd ikke blot et problem for den enkelte elev, som dermed risikerer røde kinder eller en øm bagdel, men et problem for alle elever i klassen som derved forringer sine chancer for læring. Og endelig bliver det et problem for Danmark som et videnssamfund, da viden er den vare, der skal sikre os fortsat vækst og velstand i en globaliseret verden (Pedersen, 2011). I den sidste halvdel af 1900-tallet har antallet af ufaglærte jobs i Danmark været dalende. En udvikling der gradvis tager fart, og som i 2013 har resulteret i, at der kun er få ufaglærte jobs tilbage i Danmark. Samtidig med afviklingen af denne type jobs bliver der skabt masser af jobs, som kræver uddannelse. Det betyder at man i Danmark, som i resten af vesten, fokuserer på mere uddannelse, både hvad angår længden af uddannelserne, men også med fokus på hvor mange der får sig en uddannelse (Andersen & Breck, 2011).

Denne artikel vil med udgangspunkt i data fra PISA-undersøgelsen fra 2003 belyse sammenhængen mellem det disciplinære klima i klassen¹ og den enkelte elevs faglige standpunkt. Mens der er talrige studier der finder en signifikant sammenhæng mellem det disciplinære klima i klassen og eleveres faglige kunnen (Goh & Fraser, 1998; OECD, 2005a; Scheerens, 2005), er der få studier der beskriver, hvor stort omfang uroen og forstyrrelserne har (Egelund & Hansen, 1997; Egelund, Nielsen & Randvid, 2011), og om det disciplinære klima påvirker alle elever ens. Denne artikel vil bidrage med ved deskriptiv statistik, OLS og multilevel regression at sætte et særligt fokus på at undersøge, hvor stort et omfang forstyrrelserne og uroen i de danske udskolingsklasser antager, samt om drenge og piger og unge med forældre med forskellig uddannelsesbaggrund er påvirket i lige stor grad heraf.

Eleverne i PISA-undersøgelsen er på grund af stikprøvedesignet grupperet i skoler og ikke i klasser. Dette er problematisk da eleverne i undersøgelsen bliver spurgt til uro og forstyrrelser i den konkrete undervisning og målet for det disciplinære klima bliver derved et forhold ved den enkelte klasse. Tidligere analyser af PISA-data vælger enten at ignorere denne problematik, og behandle det disciplinære klima i klassen som en skolevariabel (Mejding (Red.), 2004; OECD, 2005a), eller at udføre analyserne på individniveau (Olsen, 2003; Hermansen (Red.), 2007). Jeg

¹ Begrebet 'det disciplinære klima' skal ikke forstås som kæft, trit og retning, men som en oversættelse af den engelske term 'disciplinary climate'. Begrebet vil i artiklen blive anvendt til at indfange fænomener som støj, uro og forstyrrelser, og skal tænkes som et kontinuum gående fra dårlig disciplin til god disciplin.

vil i denne artikel gennem analyse undersøge om et aggregeret skoleniveau for det disciplinære klima i klassen er et brugbart mål til at beskrive uroen i de enkelte klasser.

Endelig bidrager artiklen til tidligere forskning ved at give opmærksomhed til selve sammenhængen mellem det disciplinære klima og den enkelte elevs faglige standpunkt og denne sammenhængs retning. Analysens design muliggør ikke påvisning af sammenhængen som kausal. Men i modsætning til tidligere forskning vil det blive diskuteret, om sammenhængen mellem oplevet uro i klasseværelset og elevernes PISA-score kan være et udtryk for, at dygtige elever oplever eller laver mindre larm end de mindre dygtige elever i samme klasselokale. Diskussionen vil blive suppleret af analyser af data for på den måde empirisk at udforske sammenhængen og potentielt sandsynliggøre, at forstyrrende adfærd faktisk medfører mindre læring.

At undersøge klasserumsfaktoren uro er særlig interessant af to grunde; 1) Flere undersøgelser viser, at uro og forstyrrelser er de faktorer, foruden elevens socioøkonomiske baggrund, der forklarer den største del af den enkelte elevs faglige udbytte (Wang et al., 1993; Teodorovic, 2009); 2) Det disciplinære klima i klassen er i modsætning til fx social baggrund og køn noget, skolen og lærerne har mulighed for at ændre og forbedre. Derved bliver artiklens resultater relevante for den uddannelsespolitiske verden, men også for den enkelte lærer i grundskolen. Artiklen har dog ingen mulighed for at være handlingsanvisende i konkret forstand, i forhold til hvordan den enkelte skole og lærer kan mindske uroen i klassen. Til dette er yderligere forskning påkrævet.

Den resterende del af artiklen er opbygget som følger: Næste del gennemgår forskning der undersøger sammenhængen mellem det disciplinære klima og elevens faglige standpunkt, og diskuterer desuden validiteten af elevbesvarede undersøgelser. Anden del er en præsentation af data, metode og analysestrategi. Herefter præsenteres mine analyseresultater og endelig konkluderes artiklen.

2. Educational effectiveness

Det er sparsomt med forskning indenfor educational effectiveness der uddybende beskriver omfanget af uro og forstyrrelser i skoler. Der er dog især to danske undersøgelser der bør refereres: I 1997 undersøgte Egelund og Hansen (1997), på bestilling af Folketinget, omfanget af uro i folkeskolen. Rapporten var bestilt på baggrund af en bekymring om for megen uro i folkeskolen. Rapporten kortlægger fænomenet larm og uro i folkeskolen ved at undersøge, hvor ofte en række lærere oplever, at eleverne er urolige. Egelund og Hansen fandt frem til at 10 % af eleverne jævnligt forstyrrer undervisningen, hvoraf hver femte af disse elever næsten altid forstyrrer undervisningen for sig selv og for andre. 29 % af eleverne forstyrrer lejlighedsvis, og 61 % af eleverne forstyrrer aldrig undervisningen. Undersøgelsen dækker alle klassetrin fra 0. til 10. klasse og finder, at der bliver mere ro i undervisningen, jo højere klassetrin. Således er eleverne forholdsvis mere rolige på 9. klassetrin, som svarer til den alder deltagerne i PISA har, end på de yngre klassetrin.

Undersøgelsen finder desuden, at drengene er de mest urolige, eller rettere oftest er urolige, da det er frekvensen af forstyrrelser der undersøges. Undersøgelsen finder ikke nogen sammenhæng mellem skolens størrelse og uro, ej heller mellem klassekvotient og uro. Geografisk set er der mere uro, des længere skolen ligger mod øst i Danmark. Konklusionen af undersøgelsen bliver, at

køn er den faktor, der med en forklaringsgrad på 7,8 % forklarer den klart største del af uroen, når der testes for sammenhænge i en multivariabel regressionsanalyse.

10 år senere gennemfører Hermansen ((Red.), 2007) ligeledes en undersøgelse, der har som mål at undersøge omfanget af uro i undervisningen. Hvor Egelund og Hansen fik deres viden om uro ved alene at spørge lærerne, udmærker denne undersøgelse sig ved også at inddrage elevernes oplevelse. Nøgternt set kan det påstås, at det er elevernes oplevelse, der er den mest relevante. Hvis de oplever, at de bliver forstyrret, er det i bund og grund ikke så vigtigt, hvad læreren oplever (Nordahl, Mausehagen & Kostøl, 2009, s. 21).

Hermansen finder, at 78 % af eleverne svarer, at de bliver forstyrret af larm. Dette tal kan synes utrolig højt, især hvis det sammenlignes med de tal Egelund og Hansen fandt 10 år tidligere, men tallet dækker over både de elever der *tit* bliver forstyrret af støj, men også dem der kun *nogle gange* bliver forstyrret. Og hvor lærerne i undersøgelsen af Egelund og Hansen bliver bedt om kun at tænke på de seneste 14 dage, så præciseres tidshorisonten ikke i dette spørgeskema. Det synes at være svagheden ved rapporten af Hermansen, at målingen af uro derved ikke bliver nuanceret mere.

Hvor 78 % af eleverne oplever støj, angiver 53 % af lærerne, at eleverne bliver forstyrret af støj. De 53 % bliver i rapporten sammenlignet med elevernes svar, men denne sammenligning synes vanskelig, da svarkategorierne er forskellige og befinder sig på forskellige analyseniveauer. Lærerne svarer således ikke *tit* og *nogle gange*, men *ja, afgjort* og *ja, i nogen grad*, til hvorvidt eleverne bliver forstyrret af støj. Så hvor eleverne bedes forholde sig til *frekvensen* af uro på *individniveau*, bliver lærerne bedt om at vurdere *graden* af uro på *klasseniveau*².

Til spørgsmålet om hvorvidt det er en larmende klasse, svarer 69 % af eleverne og 22 % af lærerne, at det er de enige i. Stadig med vekslende ordlyd på svarmulighederne på spørgeskemaet. Hermansen (Red.) finder ved korrelationsanalyse en direkte moderat sammenhæng mellem uro i klassen og testscore ($\gamma = 0,28$), således scorer dem, der oplever mindst larm højest i matematiktesten (Hermansen (Red.), 2007, s. 111). Analysen foregår på individniveau, og tager således ikke i betragtning, at eleverne er grupperet i klasser og på skoler. Det er derfor vanskeligt at sige, om sammenhængen mellem larm og testscore blot er et billede på, at dygtige elever er sværere at forstyrre. Eller et billede på, at dygtige elever larmer mindre. Undersøgelsen inddrager ikke baggrundsfaktorer i analysen.

Af international forskning finder de to reviews 'Toward a knowledge base for school learning' (Wang et al., 1993) og 'Educational effectiveness' (Teodorovic, 2009), at jo tættere et givent forhold befinder sig på den konkrete undervisning, des større betydning har forholdet for den enkelte elevs læring. Således findes forhold på skoleniveau kun at have en lille eller ligefrem ubetydelig effekt på den enkelte elevs læring, hvorimod forhold på klasse- og individniveau findes at have en stor effekt. Det samme finder OECD-rapporten fra 2000, som analyserer data fra den første bølge af PISA og konkluderer, at faktorer, der direkte har med undervisningen at gøre, har større indflydelse på PISA-score end fx økonomiske eller organisatoriske forhold (OECD, 2005a). Således er der belæg for at sige, at hvis man vil effektivisere læringen, er det faktorer tæt på undervisningen,

² Link til elev- og lærerspørgeskemaet: <http://www.samfundslitteratur.dk/>

der må formodes at have den største effekt (Rangvid, 2003). Der er som følge heraf også en del forskning, som undersøger faktorer tæt på undervisningssituationen.

Meget forskning indenfor educational effectiveness har fokus på en række af forhold på én gang, hvor disciplinen i klasseværelset blot bliver én blandt mange faktorer, der undersøges. I et studie fra Canada undersøges hvilke forhold, der har betydning i forhold til, om unge får sig det, der svarer til en gymnasial uddannelse. Frempong, Ma og Mensah (2012) finder, at det disciplinære klima i klasseværelset har betydning, ligesom lærer-elev-forholdet og socioøkonomisk baggrund har betydning. Data er fra PISA 2000 suppleret med en national ungdomssurvey. Ligeledes finder et studie af 1512 tilfældigt udvalgte grundskoleelever i Singapore, at larm, uro, mobning og slåskamp i klassen har en signifikant sammenhæng med både læring og lyst til matematik (Goh & Fraser, 1998). Et observationsstudie foretaget i amerikanske grundskoler og kombineret med faglige test finder, at antallet af reprimander for dårlig opførsel har en negativ sammenhæng med faglig kunnen (Fisher et al., 1981).

To studier fra USA undersøger effekten af implementeringen af 'Positive Behaviour Support', som er et program med mange forskelligartede facetter (Luiselli, Putnam, Handler & Feinberg, 2005; Lassen, Steele & Sailor, 2006). Begge studier er eksempler på forskningsdesigns, hvor en udvalgt skole får en 'behandling', og effekter efterfølgende kan aflæses som forbedret disciplinært klima og faglige forbedringer. Målet for behandlingen er, at det disciplinære klima skal forbedres. Begge studier finder en betydelig nedgang i frekvensen af dårlig opførsel og en samtidig betydelig forbedring af elevernes faglige kunnen. Begge studier synes dog efter min mening at have svært ved at påvise en direkte sammenhæng mellem opførsel og læring, da 'Positive Behaviour Support'-programmet også indeholder tiltag som forbedring af instruktionen i undervisningen, forøgelse af klasserumsengagementet og støtte af positiv indsats i timerne. Tiltag som alle må formodes at kunne øge fagligheden, uafhængig af om uroen forbliver den samme.

En række rapporter har benyttet PISA-data til at undersøge effekten af støj og uro. I 2000 finder en OECD-rapport en signifikant sammenhæng mellem disciplinært klima på skoleniveau og PISA-score. Dette gælder hvis alle lande analyseres samlet. Hvis landene derimod testes hvert for sig, er der kun 8 lande, hvor sammenhængen er signifikant. Og Danmark er ikke et af dem (OECD, 2005a).

Nogle år senere analyserer en række finske forskere data fra PISA 2003 og finder, at støj og uro har signifikant sammenhæng med PISA-score i Finland, men ikke i OECD samlet set (Väljjarvi et al., 2007). Ud fra samme data, men med fokus på Danmark, finder Egelund, at kun én faktor på skoleniveau har signifikant sammenhæng med PISA-score; og det er det disciplinære klima i klasseværelset (Mejdning (Red.), 2004, s. 199).

Hvor ovenstående rapporter kun flygtigt analyserer effekten af støj og uro, udkom der i 2003 en rapport, hvor et helt kapitel er tilegnet at undersøge støj og uro i undervisningen og sammenhængen med PISA-score. Data er fra PISA 2000. Rapporten bærer titlen 'Northern Lights', og har til formål at sammenligne de nordiske lande internt, men også med resten af landene i OECD (Olsen, 2003).

PISA-data er på grund af samplingen konstrueret således, at der ikke umiddelbart kan udskilles noget klasseniveau, og da støjen og uroen opleves i klasseværelset, undersøger Olsen om det

giver nogen mening at benytte et disciplinindeks på skoleniveau. Denne undersøgelse mener jeg er både vigtig og nødvendig, inden skolegennemsnittet for disciplin benyttes i analyse. Jeg vil senere i artiklen gøre mig en lignende overvejelse, dog med et noget andet metodisk greb. Olsen udnytter, at både eleverne og skolens leder har angivet, hvordan de oplever disciplinen på skolen, og ved en triangulering, hvor elevernes og lederens svar sammenlignes, finder han, at der med rette kan argumenteres for, at der på den enkelte skole eksisterer en særlig kultur forskellig fra andre skoler, når det kommer til disciplinær opførsel. Herefter finder Olsen en sammenhæng mellem disciplin og PISA-score. Dette gælder både på individniveau og på det aggregerede disciplinindeks på skoleniveau, med størst funden effekt på skoleniveau. Sammenlignet med de andre nordiske lande har Island en smule bedre disciplinært klima end Danmark. Sverige og Finland er meget lig Danmark og Norge har en smule dårligere disciplinært klima. Alle nordiske lande ligger under OECD-landenes gennemsnit. Som jeg senere vil forklare uddybende, må sådanne internationale sammenligninger tolkes med forsigtighed.

Olsen finder i sin analyse tegn på, at forholdet mellem den enkelte elevs oplevelse af det disciplinære klima i klasseværelset og PISA-score er kurvelineært. Det vil i dette tilfælde sige, at dem der scorer lavt, har en tendens til at opleve ekstremer: Enten at de oplever støj hele tiden, eller også at de aldrig oplever støj, hvorimod de elever, der scorer højt i PISA-testen, har en tendens til at svare mere i midten på skalaen, hvilket angiver, at de kun oplever støj ind i mellem. Olsen finder, at elevbesvarelsen på spørgsmålet: ”Eleverne hører ikke efter, hvad læreren siger” har en kurvelineær sammenhæng med PISA-score i læsning i Norge. Analysen er desværre ikke særlig gennemsigtig i forhold til, hvor mange respondenter analysen bygger på og på hvilket analyseniveau. Det fremgår således ikke, om analysen er lavet på hele stikprøven, eller på den enkelte skole, og derved er det uklart, hvad man kan konkludere ud fra denne. Analysen synes umiddelbart at være lavet på baggrund af hele stikprøven og ikke på den enkelte skole, hvor sidstnævnte må være at foretrække, hvis der skal tages højde for, hvordan elever med forskelligt fagligt standpunkt oplever den samme støj forskelligt. Den Brok et al. (2004) påpeger ligeledes den mulige kurvelinearitet, men undersøger det ikke empirisk. Om der i PISA-data fra 2003 skulle være et kurvelineært forhold mellem larm og PISA-score, vil senere i artiklen blive undersøgt. Hvor denne rapport af Olsen udmærker sig ved sin grundighed, og især overvejelserne om skoleniveauet som analyseniveau synes interessante, så mangler den efter min mening at inddrage baggrundsfaktorer som socioøkonomisk baggrund og køn, da disse faktorer i anden forskning netop har vist sig at være interessante og relevante.

Det disciplinære klima i klassen synes ud fra ovenstående gennemgang af forskning inden for educational effectiveness at være en af de mest betydningsfulde faktorer, når det kommer til at skabe effektiv undervisning. Artiklen har nu gennemgået forskning, der anvender forskellige forskningsdesigns, forskellige analyseniveauer og forskellige metodiske tilgange. Alle studier finder i større eller mindre grad en sammenhæng mellem det disciplinære klima i klasseværelset og faglig kunnen. Et enkelt studie finder dog, at denne effekt ikke gælder for Danmark.

2.1 Elevrapporтерingers reliabilitet og validitet

Da artiklens mål er at undersøge sammenhænge mellem larm og læring, og målet for larm er udtaget fra elevspørgeskemaet, bliver det vigtigt at finde belæg for elevbesvarelsernes reliabilitet

og validitet. Dette afsnit undersøger forskning på området, om hvorvidt en elevs idiosynkratiske oplevelse kan bruges til at beskrive fænomenet larm på klasseniveau.

Artiklen benytter to tilgange til at få undersøgt reliabiliteten og validiteten af de elevbesvarede spørgsmål angående det disciplinære klima i klasseværelset; en teoretisk tilgang hvor forskning på området undersøges; og en empirisk tilgang hvor elevrapporteringens validitet og reliabilitet, gennem analyse af data fra PISA 2003, vil blive vurderet. Den teoretiske tilgang vil blive beskrevet nedenfor, mens den empiriske tilgang bliver præsenteret i artiklens analysedel.

Eleverapportering eller evaluering er hyppigt benyttet i undervisningsverdenen, da det er et nemt og billigt værktøj til at undersøge forhold vedrørende undervisningen. For at fastslå i hvilken grad man kan bruge elevvurderinger i forbindelse med evaluering af lærere, undersøgelse af forhold i klassen, forbedring af undervisningen og en række andre forhold i relation til klasseværelset og undervisningen, er der derfor også en del forskning på området, som undersøger validiteten og reliabiliteten af sådanne elevbesvarelser.

Hvor der er et væld af forskning på området, når det kommer til elever ældre end grundskolen (Marsh, 1987; Greenwald, 1997 – begge reviews; den Brok et al., 2004), er det straks mere sparsomt, når det kommer til undersøgelser med udgangspunkt i grundskolen. I 'The predictive validity of student evaluations in the identification of meritorious teachers' (Ebmeier, Jenkins & Crawford, 1991) testes validiteten af 6. til 12. klassetrins elevers vurdering af en række læreres kvalifikationer, ved at undersøge om eleverne vurderer lærernes kvalifikationer på samme måde, som en række eksperter vurderer de samme lærere. Studiet er designet således, at eleverne bliver bedt om at vurdere deres lærer ud fra en række parametre som; organisation af undervisningen; instruktioner; evaluering; samt om læreren formår at holde ro og orden. Samtidig med at eleverne vurderer lærerne, har en række undervisningsexperter vurderet de selvsamme lærere og placeret dem i en af følgende grupper: Særlige dygtige lærere og dygtige lærere. Undersøgelsen går så på at se, hvorvidt eleverne vurderer lærerne på samme måde som eksperterne. Konklusionen fra denne undersøgelse er, at i 93 % af tilfældene vurderer eleverne og eksperterne lærerne ens.

Jeg tillader mig dog at stille spørgsmål ved, om det er validiteten, der reelt undersøges, men i stedet reliabiliteten, idet designet kun kan undersøge om elever og eksperter måler ens og ikke om der måles rigtigt. Mere om denne komplikation senere i afsnittet.

Peterson & Stevens (1988) undersøger reliabiliteten i 7. til 12. klassetrins elevers evalueringer af deres lærere ud fra et longitudinalt design. De finder konsistens og stabilitet i disse vurderinger, også over en to-årig periode, og konkluderer samtidig, at eleverne formår at vurdere lærerne enkeltvis uden at blande dem sammen.

Af dansk forskning finder Hermansen ((Red.), 2007) en moderat positiv sammenhæng ($\gamma = 0,28$) mellem lærer- og elevopfattelsen af, om de bliver forstyrret af støj i klassen (s. 79), og ligeledes findes der en moderat positiv sammenhæng ($\gamma = 0,38$) mellem lærer- og elevopfattelsen af, om deres klasse er en larmende klasse. Denne undersøgelse har at gøre med elevbesvarelsernes reliabilitet. En interessant forskel mellem lærer- og elevperspektivet kan dog som tidligere skrevet observeres, idet 78 % procent af eleverne svarer, at de bliver forstyrret af støj, mens kun 53 % af lærerne mener, at eleverne bliver forstyrret af støj. Ligeledes mener 69 % af eleverne, at deres klasse er en larmende klasse, hvorimod kun 22 % af lærerne mener, at klassen er en larmende klasse. Forskelle som dels skyldes, at der i spørgeskemaerne til lærerne og eleverne bruges for-

skellig ordlyd, kan samtidig være med til at stille spørgsmålstejn ved, hvorvidt larm i klassen kan indfanges, uden at medtænke den som oplever det. Dette vil blive diskuteret sidst i dette afsnit.

'Northern lights on PISA' (Olsen, 2003) behandler data fra første runde af PISA fra 2000 og undersøger validiteten af elevbesvarelserne på larm og uro. Variabler meget lig dem der benyttes i denne artikel. Olsen finder ved hjælp af triangulering en vis enighed mellem elevernes og skoleledernes vurdering af det disciplinære klima på skolen og fortolker dette som, at de har at gøre med 'the same objective reality' (s.119). Dette på trods af at skolelederen ikke spørges til disciplinen blandt de 15-årige, men til det disciplinære klima på skolen som helhed. Olsen tænker som ovenstående forskning larm i klasseværelset som noget objektivt, og som derfor kan opleves ens af alle, og at der derfor findes én valid værdi for larm.

Men måske opleves larm ikke ens af alle. Det er i hvert fald den konklusion Kunter og Baumert (2007) uddrager af deres undersøgelse af korrelationer mellem elev- og lærervurderinger af forhold i forbindelse med matematikundervisningen. Data er fra den tyske del af PISA 2003, som adskiller sig fra de øvrige deltagende landes, idet der også gennemføres en udvidelse, hvori 288 matematiklærere og deres elever svarer på en række spørgsmål, udover dem resten af de deltagende lande svarer på i forbindelse med PISA 2003. Ud fra dette data undersøger Kunter og Baumert, om der er sammenhæng mellem, hvordan elever vurderer en række forhold i undervisningen, og hvordan deres respektive lærere vurderer de samme forhold. Altså i design meget lig den ovenfor refererede artikel af Ebmeier et al. (1991). Kunter og Baumert finder i deres undersøgelse, at det afhænger af, hvilket fænomen der vurderes, hvorvidt det vil blive vurderet ens. Classroom management bliver vurderet næsten ens af elever og lærere, hvorimod forskellige undervisningsmetoder bliver vurderet meget forskelligt. Forklaringen er, at selvom ordlyden synes næsten ens på hhv. elevernes og lærernes spørgeskemaer, så bliver de opfattet meget forskelligt. Des mere konkret et fænomen er, der skal vurderes af elever og lærer, des mere ens er de i deres vurderinger. I elevs og læreres vurderinger af larm og uro, finder artiklen en interfaktorkorrelation på $r = 0.64$ ($p < 0.05$) (s. 240).

Kunter og Baumert skriver, at den typiske forklaring på forskelle mellem elev- og lærerbesvarelser baseres på, at der findes et underliggende sandt latent fænomen, en objektiv sandhed, som de hver især ikke helt formår at beskrive ved deres oplevelser, hvilket også bl.a. Olsen ovenfor antager. Kunter og Baumert vil derimod hellere forklare forskelle ved, at uenigheden skyldes perspektiv-specifikke validiteter: At lærere og elever rent faktisk repræsenterer forskelle synsvinkler, og at de derfor vil opleve det samme fænomen forskelligt. Ved simple observérbare fænomener som uro, vil lærere og elever tendere til at opleve ens, mens de ved mindre konkrete forhold vil være mere forskellige (s. 241).

Denne nuancering af validitetsproblematikken betyder for denne artikel, at elevernes oplevede larm og uro med rette kan forstås som en perspektiv-specifik oplevelse af fænomenet. Og at denne oplevelse som udgangspunkt, ifølge Kunter og Baumert, kan betegnes som valid. Det betyder, at der i den enkelte klasse kan optræde forskellige oplevelser af det samme fænomen; at der på klasseniveau, blandt eleverne, kan være perspektiv-specifikke forskelle. En tese til afprøvning ved analyse vil således være, **at der i den samme klasse vil være forskellige rapporteringer af det**

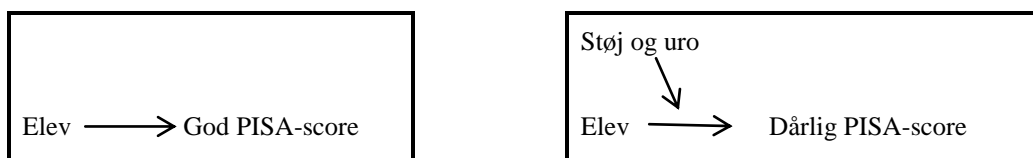
samme fænomen. Fx vil nogle elever svare, at der ofte er støj og uro, mens andre elever i samme klasse vil svare, at de aldrig eller næsten aldrig oplever støj og uro.

Samtidig med, at det forventes at finde en vis heterogenitet i besvarelsene, når det kommer til elever fra samme klasses oplevelse af larm, giver undersøgelserne af Peterson & Stevens (1988) og Ebmeier et al. (1991) en forventning om, at der på trods af denne idiosynkratiske forskellighed må formodes at optræde en vis grad af homogenitet på klasseniveau. De undersøger netop, om der på klasseniveau er en vis grad af enighed om det oplevede fænomen. Om elevbesvarelsene er konsistente. Det bliver således også en tese til afprøvning i artiklens analyse, **at der på klasseniveau kan findes en vis grad af homogenitet blandt elevernes oplevelse af støj og uro.** Graden af uro vil således variere fra elev til elev, men i den enkelte klasse må der forventes en vis grad af enighed.

Ovenstående forskning undersøger reliabiliteten og validiteten af elevbesvarelser med udgangspunkt i klasserummet. Derfor blev de to teser også formuleret med udgangspunkt i klasseniveauet. Men i den danske del af PISA 2003 er der ikke angivet noget klasseniveau, og tesoerne kan derfor ikke afprøves i ovenstående ordlyd, men må omformuleres så de får følgende ordlyd: **Der vil på den samme skole være forskellige oplevelser af det samme fænomen; og der vil på skoleniveau være en vis grad af homogenitet blandt elevernes oplevelse af støj og uro.** Jeg vil senere i artiklen diskutere, hvorvidt det giver mening at teste ovenstående to teser på skoleniveau i stedet for på klasserumsniveau. Videre vil jeg beskrive, hvorledes tesoerne er tænkt efterprøvet ved analyse, og hvordan resultaterne af disse analyser vil danne grundlag for en vurdering af, hvorvidt det disciplinære klima i klasseværelset kan indfanges på skoleniveau i analysen. Hvilket igen kan bestemme reliabiliteten og validiteten af elevbesvarelsene vedrørende det disciplinære klima i klasseværelset.

Målet med artiklen er som nævnt at undersøge, om der er sammenhæng mellem hvor meget støj og uro eleverne oplever i klasseværelset, og hvor godt de scorer i PISA-testen. Artiklens grundlæggende hypotese er dermed meget simpel og logisk: At støj og uro forstyrrer den enkelte elev, så denne ikke lærer så meget, som hvis der ikke havde været støj og uro, og dermed ikke får så god en PISA-score. Dette er illustreret i model 2.1.

Model 2.1. Model over artiklens grundlæggende hypotese



Eleven til venstre i modellen bliver ikke forstyrret og får en god PISA-score, mens eleven til højre bliver forstyrret og får dermed en mindre god PISA-score. Denne simple hypotese giver et forståelsesmæssigt simpelt, men solidt fundament for artiklen, men kræver at blive udfoldet og nuanceret.

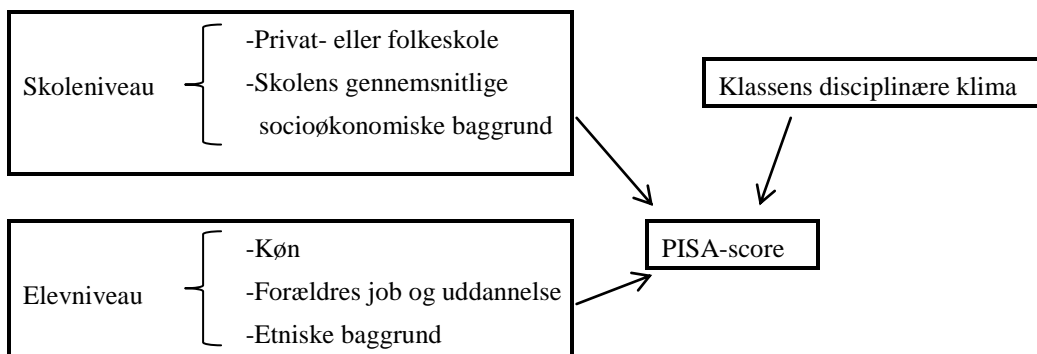
Hvor det i ovenstående model er et enten-eller-spørgsmål om der er larm og uro, vil det i artiklens analyse blive et spørgsmål om, hvor meget uro der er. Derved bliver det muligt at tænke sammenhængen mellem disciplin og PISA-score som en lineær sammenhæng, hvor PISA-scoren gradvist bliver dårligere des dårligere det disciplinære klima i klassen bliver.

PISA-scoren og et mål for det disciplinære klima i klasseværelset er således artiklens vigtigste analysegenstande, og det bliver derved omdrejningspunktet for hele analysen.

Foruden disciplinen i klasseværelset vil der i analysen også blive inddraget en række baggrundsfaktorer, som, anden forskning har påvist, kan forklare en del af variansen i PISA-score. Det drejer sig på individniveau om køn (Brophy, 1988; Rangvid, 2003; den Brok et al., 2004), forældres uddannelse og job (Rangvid, 2003; Bay, 2011; Frempong et al., 2012) og elevens etniske baggrund (Brophy, 1988; Rangvid, 2003; Bay, 2011; Frempong et al., 2012). På skoleniveau inddrages den gennemsnitlige socioøkonomiske baggrund for årgangen for således at kontrollere, om elevens kammerater har indflydelse på opnået PISA-score (Teodorovic, 2009; Legewie & DiPrete, 2012; Frempong et al., 2012) samt om skolen er privat eller offentlig drevet (den Brok et al. 2004; Bay, 2011; Frempong et al., 2012).

Med udgangspunkt i model 2.1 og med inddragelse af de ovenfor nævnte baggrundsfaktorer på hhv. elev- og skoleniveau kan der skitseres følgende model for analysen:

Model 2.2. *Analysemodel*



2.2 Overvejelser om kausalitet

Både hypotesen (model 2.1) og analysemodellen (model 2.2) har ensrettede pile, der antyder at støj og uro påvirker PISA-scoren, og ikke omvendt. Det skal endnu en gang understreges, at det valgte forskningsdesign ikke muliggør påvisning af kausale sammenhænge mellem det disciplinære klima i klasseværelset og PISA-score. Artiklens analyse er i den forstand udelukkende deskriptiv, og kan kun påvise statistiske sammenhænge. De ensrettede pile symboliserer således blot en sandsynlig effekt.

Sammenhængens retning mellem det disciplinære klima i klasseværelset og PISA-score synes oplagt, og det vil være et logisk udtryk for, at elever lærer mere, hvis der er ro i klasseværelset.

Men selvom sammenhængens retning umiddelbart synes oplagt, kan den ligeledes tænkes at være et udtryk for, at fagligt dygtige elever larmer mindre. Disse elever er så optaget af den spændende undervisning, at de ikke har tid og lyst til at larme. Og de fagligt svage elever finder undervisningen svær, kan ikke følge med, bliver derved lettere forstyrret, og de forstyrrer andre. En sammenhæng mellem disciplin og PISA-score kan endvidere være et tegn på, at fagligt dygtige elever ikke lader sig forstyrre, og at de derved ikke oplever så meget larm. Hvorimod fagligt svage elever nemt lader sig forstyrre og derved oplever meget larm. Og det er netop elevernes egen oplevelse af støj og uro, som konstruktet det disciplinære klima bygger på.

Det ville ud fra en række anerkendte læringsteorier være muligt at påstå, at for at læring kan ske optimalt, skal der være en tilpas udfordring i tilpasse omgivelser. Dermed sandsynliggøres en kausal sammenhæng mellem det disciplinære klima og PISA-score (Bruner, 1970; Vygotsky, 1982; Csikszentmihalyi, 2005). Men ingen af disse læringsteorier understøttes empirisk i en sådan grad, at en kausalsammenhæng kan påvises, så de kan kun bistå artiklen ved at sandsynliggøre sammenhængens tænkte retning. Således har jeg ingen forskning fundet indenfor educational effectiveness, der empirisk kan påvise en kausalsammenhæng mellem disciplin og læring. Ovenstående litteraturstudie finder nok sammenhænge, men ingen af disse kan betegnes kausale. Fisher et al. (1981) kommer med et empirisk bud på, hvordan sammenhængen mellem disciplinen i klasseværelset og elevens faglige kunnen kan opstå. I undersøgelsen måles den faktiske tid, hvor den enkelte elev arbejder uforstyrret og fagligt fordybet, og der findes en sammenhæng mellem denne faktiske læringstid og elevens faglige kunnen. Igen ingen kausaleffekt, men et empirisk bud på hvordan sammenhængen kan forstås.

Undervejs i artiklen vil sammenhængen mellem det disciplinære klima og PISA-score blive analyseret, og det vil blive diskuteret, hvorvidt det er sandsynligt, at sammenhængen er, som pilene i de to ovenstående modeller antyder: At disciplinen i klasseværelset har en effekt på den enkelte elevs læring.

3. Data og metode

Analysens data er hentet fra OECD's PISA-undersøgelse fra 2003. Undersøgelsen udmærker sig ved at have informationer om den enkelte elevs faglige standpunkt, samt informationer om hvordan den enkelte elev oplever det disciplinære klima i matematikundervisningen. Undersøgelsen består foruden den faglige test af et spørgeskema til hver elev, samt et til skolelederen.

Udvælgelsen af elever til deltagelse i PISA 2003 er foretaget over to etaper, en to-trins klyngeudvælgelse, hvor først skolen bliver udtrukket, og siden eleverne. For metodisk at tage højde for dette i estimeringen af standardfejl, benytter artiklens analyser hhv. Balanced Repeated Replication og multilevel modellering. Skolerne bliver udtrukket med en sandsynlighed svarende til det antal elever i PISA-alderen, der går på skolen. I stikprøvens anden etape udtrækkes op til 28 elever i den rette alder med lige stor sandsynlighed på hver af de udtrukne skoler (OECD, 2005b, s. 46). Således repræsenterer de enkelte klynger ikke den enkelte klasse, men den enkelte skole.

Stikprøvedesignet betyder for analysen, at der uden problemer kan undersøges for sammenhænge mellem den enkelte elevs PISA-score og dennes oplevelse af disciplin. Mens der ikke på samme måde kan undersøges for sammenhænge mellem den enkelte elevs PISA-score og den gennemsnitlige oplevede disciplin i *klassen*, men kun i forhold til hvordan disciplinen er blandt skolens 15-årige. Men disciplin er en klasserumsfaktor. Og heraf følger, at det ikke er vigtigt, hvordan disciplinen er i andre klasser. Det er udelukkende det disciplinære klima i den klasse, eleven sidder i, der kan have betydning for, hvor meget læring der sker, og hvor godt eleven klarer PISA-testen i matematik. Og ikke hvordan disciplinen er på skolen generelt set.

For at det alligevel giver mening at benytte den gennemsnitlige oplevelse af disciplin på skoleniveau som mål for disciplin i analyserne, må der i data findes en vis lighed mellem eleverne på den enkelte skole, som er forskellig fra andre skoler (Lüdkte, Trautwein, Kunter & Baumert, 2007, s. 217). For at undersøge dette udføres der i artiklens analyse en variansanalyse, som undersøger within-class og between-class korrelationer og herudfra beregnes intra-class korrelationen og om denne er signifikant forskellig fra nul. Inden variansanalysen udføres, vil elever, der ikke går i 9. klasse blive sorteret fra, for at sikre at det kun er variansen i det disciplinære klima på 9. årgang, der undersøges. Disse frasorterede elever opfylder PISAs krav om alder, men går af den ene eller anden grund enten i 8. eller 10. klasse. Da de er jævnt fordelt på mange skoler og samtidig udgør et relativt lille antal, vurderes de som unødvendigt støj, når det kommer til at benytte det disciplinære klima som klasserumsvariabel på skoleniveau. Det drejer sig om 487 elever ud af i alt 4218 elever. Ligeledes vil skoler, der herefter har færre end ti PISA-elever blive sorteret fra (82 elever). Dette for at sikre at gruppen af elever har en vis størrelse at lave variansanalyse ud fra, og at outliers ikke kan yde så stor indflydelse. Hvis analysen finder at der på hver skole er en vis enighed blandt elevernes oplevelse af det disciplinære klima i klassen, kan det tolkes som at hver skole har en særlig skolekultur når det kommer til det disciplinære klima. Det vil således også kunne bekræfte den tese der lød på, at der på hver skole ville være en vis grad af enighed i forhold til oplevet disciplinært klima. Tesen der lød på, at der på hver skole samtidig vil optræde en vis grad af forskellighed i forhold til, hvordan de forskellige elever oplever det disciplinære klima, vil blive undersøgt ved at se på spredningen af svar på de enkelte skoler, og derudfra vurdere om tesen kan af- eller bekræftes.

Hvis de to teser bliver bekræftet, betyder det, at klasserumsfaktoren disciplin kan benyttes som et aggregeret konstrukt på skoleniveau, og at dette mål for det disciplinære klima i klasseværelset er pålideligt og validt.

Hvor ovenfor beskrevne variansanalyse skal teste, om det giver mening at benytte skolen som analyseniveau, vil der i analysen også blive undersøgt, hvor præcist målet for disciplin på skoleniveau repræsenterer det disciplinære klima i den enkelte klasse. For at vurdere om disciplinen på skolen svarer til disciplinen i den enkelte klasse, vil det være nødvendigt i stikprøven at kunne udpege en række skoler, hvor det synes sandsynligt, at der på disse skoler kun er én 9. klasse. Disse ét-sporskoler kan så sammenlignes med den gruppe af skoler, hvor intet tyder på, at de kun skulle have én 9. klasse. Sammenligningen sker på baggrund af en variansanalyse af begge grupper af skoler. I en sådan analyse er min forventning, at intra-class korrelationen vil være større for den gruppe af skoler, som kun har ét spor, end i den gruppe af skoler som har flere spor, da eleverne trods alt går

i samme klasse, og det derfor er det samme klasserum, de har oplevet og rapporteret hændelser fra i elevspørgeskemaet. Men det bliver interessant at se, hvor stor en forskel der er i intra-class korrelationen for de to grupper af skoler. Hvis forskellen er lille, tyder det på, at disciplinen i de enkelte klasser på en skole med flere spor varierer meget lidt klasserne imellem. Hvis forskellen derimod er meget stor, betyder det, at den oplevede disciplin i de enkelte klasser på en skole med flere spor er meget forskellig klasserne imellem. Sidstnævnte vil betyde, at PISA-data bliver mindre velegnet til at undersøge sammenhængen mellem det disciplinære klima i klasseværelset og PISA-score, idet validiteten af konstruktet disciplin på skoleniveau ville svækkes.

For at udføre denne undersøgelse har jeg udvalgt 15 skoler, hvori de udvalgte PISA-elever højst sandsynligt går i én og samme klasse. Disse 15 skoler bliver brugt som kontrolgruppe til at sammenligne med de resterende skoler, hvori eleverne går på samme skole, men med stor sandsynlighed er fordelt over flere klasser. Metodisk bliver det en art triangulering, hvor variansen undersøges i forhold til oplevet larm på skoleniveau for de 15 udvalgte skoler, og den fundne varians sammenlignes så med variansen mellem de resterende 165 skoler.

At det lige skulle være 15 skoler, har ikke været noget mål, men et resultat af en udvælgelsesprocedure. For at finde frem til skoler som med overvejende sandsynlighed kun har ét spor på 9. årgang, har det været nødvendigt at definere et udvælgeskriterium. Først frasorteres som tidligere nævnt, elever, som ikke går i 9. klasse, og skoler der herefter har færre end ti PISA-elever. Videre stiller jeg som krav, at skolen ikke må have mere end gennemsnitligt 24 elever pr. årgang. Flere end 24 elever pr. årgang sandsynliggør, at skolen kan have to spor på alle eller udvalgte klassetrin; herunder 9. klassetrin. Folkeskolelovens kapitel 2, § 17 foreskriver et maksimalt elevtal pr. klasse på 28, dog med mulighed for dispensation, hvor loftet kan blive hævet til 30 elever pr. klasse³. I grundskolen betragtes klassestørrelser på eller i nærheden af 28 elever ikke som attraktive. Hverken af lærere, forældre eller af børn. Skoleledere må medtænke den umiddelbare økonomiske fordel ved store klasser, men betragter ikke store klasser som en populær løsning. Praksis i folkeskolen anno 2003 var derfor, at der var så få klasser som muligt med elevtal i nærheden af 28. Derfor vurderer jeg, at et gennemsnit på mere end 24 elever pr. årgang kan betyde, at skolen kan have mere end ét spor på udvalgte klassetrin.

I PISA 2003 kan der i skolelederspørgeskemaet findes oplysninger om, hvilke klassetrin den enkelte skole har, samt hvor mange elever der er i alt på skolen. I elevspørgeskemaet er der oplysninger om, på hvilket klassetrin den enkelte elev går, samt hvilken skole eleven går på. Ud fra disse oplysninger kan der udregnes et gennemsnitligt antal elever pr. årgang for den enkelte skole (se tabel 3.1).

Jeg har i tabellens sidste kolonne angivet hvor mange elever fra skolens 9. klassetrin, der deltog i PISA 2003. Det ses, at antal PISA-elever generelt set er en smule lavere end den udregnede gennemsnitlige klassekvotient på skolen. Dette kan sandsynligvis forklares med, at der generelt er færre elever i 9. klasse end i de yngre klasser, da flere vælger at tage 9. klasse på efterskole. Samtidig er der i PISA 2003 et vist frafald af elever, hvilket også kan være en forklarende årsag.

³ Link til folkeskoleloven 1993 (gældende i 2003): <http://inet.dpb.dpu.dk/ress/skolelove/folkeskole/30.06.1993.html>

Tabel 3.1. 15 skoler som med høj sandsynlighed kun har én 9. klasse.

SCHOOLID	Årgange på skolen	Antal årgange på skolen	Elever i alt på skolen	Gennemsnitligt antal elever pr. årgang	Antal PISA-elever fra 9. årgang
00111	5. – 9.	5	57	11,4	16
00148	8. – 10.	3	60	20	17
00181	7. – 10.	4	61	15,3	18
00056	9. – 11.	3	67	22	22
00173	1. – 10.	10	162	16,2	23
00091	1. – 8.*	8	163	20,4	13
00193	1. – 9.	9	170	18,9	13
00024	1. – 9.	9	183	20,3	12
00064	1. – 9.	9	200	22,2	14
00054	1. – 9.	9	203	22,6	13
00110	1. – 11.	11	223	20,3	18
00058	1. – 10.	10	236	23,6	10
00072	1. – 10.	10	243	24,3	17
00031	1. – 10.	10	246	24,6	26
00187	1. – 10.	10	247	24,7	24

Data: PISA 2003, skolelederspørgeskemaet og elevspørgeskemaet

* elevspørgeskemaet viser at eleverne går i 9. klasse

Ovenstående variansanalyser vil blive udført i artiklens analysedel, og her vil resultaterne ligeledes blive diskuteret. Hvis det viser sig, at variansen af skolegennemsnittene for disciplin ikke afviger alverden fra variansen af klassegennemsnittene for disciplin, vil artiklens analyser være valide analyser af disciplinen i klasseværelset. I så fald vil det imødekomme den kritik som blandt andre Hermansen er kommet med, hvor det kritiseres, at klasserummet er så svagt belyst i forskningen indenfor educational effectiveness (Hermansen (Red.), 2007, s. 24).

Til analysen er der udarbejdet to disciplinindekser. Det første er et indeks, der repræsenterer den enkelte elevs oplevelse af disciplinen i klassen. Det er konstrueret ud fra fem ordinale variabler uddraget fra elevernes svar på følgende udsagn: *Eleverne hører ikke efter, hvad læreren siger; der er støj og uro; læreren må vente i lang tid, før eleverne falder til ro; eleverne kan ikke arbejde ordentligt; og eleverne begynder først at arbejde længe efter, at timen er startet.* Svarene er givet på en likert-skala, hvorpå der kan svares: *I hver time; i de fleste timer; i nogle timer; aldrig eller næsten aldrig.* Svarene er kodet 1 til 4, hvor 4 svarer til 'aldrig eller næsten aldrig'. De kodede svar fra hver enkelt elev er lagt sammen, og tallet 5 er trukket fra. Således indikerer 0, at det disciplinære klima er elendigt, hvorimod 15 angiver, at der i matematiktimerne næsten aldrig er uro og forstyrrelser, altså et godt disciplinært klima. Indekset bevæger sig på individniveau, og det bliver

i analysen hovedsageligt brugt deskriptivt, men derudover også i undersøgelsen af mønstre i oplevelsen af disciplin på den enkelte skole i forhold til PISA-score, for at blive klogere på om dem der scorer højt i PISA eksempelvis oplever disciplinen anderledes end resten af eleverne i klassen.

Indekset benyttes desuden til at konstruere en aggregeret skala, hvor alle skolens PISA-elevs oplevelse af disciplin danner et mål for elevernes gennemsnitlige oplevelse af det disciplinære klima i klasseværelset på skoleniveau. Hvis den før beskrevne variansanalyse viser, at der på hver skole er en særlig disciplin forskellig fra andre skoler, vil dette disciplinindeks på skoleniveau blive benyttet som mål for det disciplinære klima på den enkelte skole, og det vil blive benyttet i artiklens analyser.

Tabel 3.2. *Deskriptiv statistik: Præsentation af analysens variable. Ikke vægtet.*

Variabel	N	Gennemsnit	SD	Min.	Max.
Individniveau:					
Disciplin	3439	8,92	3,21	0	15
PISA-score i matematik	3649	519,15	84,54	181,52	751,24
Køn (Pige=0; Dreng=1)	3649	0,49	0,50	0	1
Forældres højeste uddannelsesniveau (ISCED-skala)	3513	4,47	1,28	0	6
Forældres højeste jobstatus (ISEI-skala)	3562	49,31	15,31	16	90
Etnicitet (dummyer)					
Født i Danmark	3617	0,95	0,21	0	1
Taler dansk i hjemmet	3539	0,97	0,18	0	1
Skoleniveau:					
Disciplin	180	8,89	1,37	4,64	11,83
Foræld.høj.udd. (agg.)	180	4,47	0,49	2	5,77
Foræld. Status (agg.)	180	49,25	6,10	29,67	69,04
Skoletype (FS=0; Privatskole=1)	180	1,20	0,40	0	1

Data: PISA 2003

4. Resultater

4.1 Deskriptiv statistik

Stikprøvens 3649 danske elever er fordelt på 180 skoler med mellem 10 og 28 elever på hver og et gennemsnitligt antal PISA-elever på 20,3. Gennemsnitscoren for de danske elever i stikprøven er på 519,3 med en standardfejl på 2,98. Den lavest scorende elev har en PISA-score på 181,5 og den højest scorende på 751,2.

PISA-scoren i matematik er tilnærmelsesvis normalfordelt, og denne variabel vil derfor uden problemer kunne benyttes i artiklens senere analyser som afhængig variabel.

Hvis elevernes oplevelse af det disciplinære klima i klassen analyseres på individniveau, ses det i tabel 4.1, at cirka halvdelen af eleverne svarer 'i nogle timer' til hvert af spørgsmålene (mellem 47,2 % og 58,2 %). Herudover fordeler den restende del af besvarelserne sig på hhv. 'i de fleste timer' og 'aldrig eller næsten aldrig', som tilsammen står for cirka 40 % af besvarelserne. At de enkelte udsagn skulle ske 'i hver time', oplever færre end hver tiende PISA-elev, hvis der ses bort fra, om der er støj og uro, hvilket opleves af hele 17,1 % af eleverne i hver matematiktime. Der er således en klar tendens til, at svarene klumper sammen om svarmuligheden 'i nogle timer', og det er da også denne svarmulighed, der har den største frekvens efterfulgt af de tilstødende svarmuligheder hhv. 'i de fleste timer' og 'aldrig eller næsten aldrig'.

Det kunne se ud til at udsagnene 'Læreren må vente i lang tid, før eleverne falder til ro' og 'Eleverne begynder først at arbejde længe efter, at timen er startet' er to udsagn, der dækker over samme hændelse. Både hvad angår ordlyden og fordelingen af svar. Hvis dette er tilfældet, kunne der være god grund til at sortere det ene udsagn fra. Ved at undersøge svarene fra de enkelte elever, viser der sig imidlertid store forskelligheder i, hvad der bliver svaret til de to udsagn. Derfor vælger jeg at beholde begge variabler i konstruktet disciplin.

Tabel 4.1. *Fordelingen af elevernes besvarelser på spørgsmål 38 b, f, h, i og k fra elevspørgeskemaet: "Hvor tit sker følgende i matematiktimerne?"*

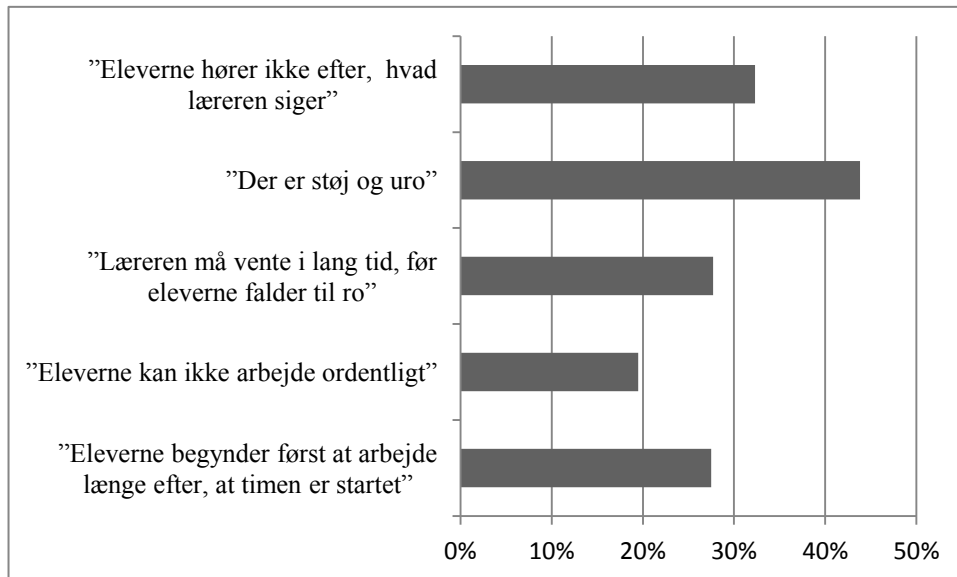
	I hver time	I de fleste timer	I nogle timer	Aldrig eller næsten aldrig	I alt:	N
"Eleverne hører ikke efter, hvad læreren siger"	8,7 %	23,6 %	58,2 %	9,6 %	100 %	3546
"Der er støj og uro"	17,1 %	26,7 %	47,2 %	9,0 %	100 %	3538
"Læreren må vente i lang tid, før eleverne falder til ro"*	8,6 %	19,1 %	49,5 %	22,8 %	100 %	3544
"Eleverne kan ikke arbejde ordentligt"	4,2 %	15,3 %	54,7 %	25,8 %	100 %	3528
"Eleverne begynder først at arbejde længe efter, at timen er startet"*	8,2 %	19,3 %	49,8 %	22,8 %	100 %	3546

Data: PISA 2003. *Disse to spørgsmål synes ens i forhold til tilslutning, men ved nærmere undersøgelse er fordelingerne i høj grad forskellige.

I nedenstående diagram (Figur 4.1) er der afbilledet, hvor stor en procentdel af eleverne der oplever hvert af de fem hændelse i hver time eller i de fleste timer. Diagrammet viser blandt andet, at 20 % af eleverne i den danske grundskole ofte oplever, at eleverne ikke kan arbejde ordentligt, mens 44 % ofte oplever, at der er støj og uro i matematikundervisningen. Videre fås det ved optælling, at 2.122 af eleverne oplever, at mindst én af hændelserne sker i de fleste eller i alle timer. Det vil sige, at 59 % af eleverne i den danske grundskole oplever, at der i de fleste eller i hver time er, hvad der kunne kaldes dårlig disciplin i matematikundervisningen.

Data viser således, at dårlig disciplin ikke er et sjældent fænomen i den danske grundskole. Tallene kan ikke direkte sammenlignes med tallene fra Hermansens forløbsundersøgelse ((red.), 2007) eller Egelunds og Hansens survey (1997), men peger sammen med disse tidligere undersøgelser i retning af, at larm, uro og forstyrrelser er hyppige fænomener i den danske grundskole.

Figur 4.1. Procentdel der oplever følgende hændelser i alle eller næsten alle matematiktimer.



Data: PISA 2003

4.2 Skolen som analyseniveau

Når data analyseres på skoleniveau gælder, at der på de enkelte skoler er en stor spredning i forhold til, hvor ofte de enkelte elever oplever dårlig disciplin. Til spørgsmålet om, hvor tit der opleves støj og uro i matematikundervisningen, er der på de fleste skoler både elever, der angiver, at det oplever de i hver time, mens der på samme skole er elever, der angiver, at det oplever de nærmest aldrig (Bilag 1). Således kan det konstateres, at larm og uro opleves forskelligt eleverne imellem. Dette er i overensstemmelse med den første af teserne fra litteraturstudiet, hvor det blev forventet, at elever oplever det samme fænomen individualspecifikt og derfor forskelligt til en vis grad.

Tabel 4.2. Variansanalyse af den enkelte elevs oplevelse af larm. Skolen som grupperingsvariabel.

	Hele stikprøven	'klasser'	'skoler'
Between-school varians	6058,63	496,38	5526,17
Within-school varians	29788,07	1785,72	28002,59
Total varians	35846,70	2282,11	33528,76
Intra-class correlation	0,169	0,218	0,197
p =	0,000	0,000	0,000

Data: PISA 2003

Samtidig viser resultatet af variansanalyse at skolen kan forklare 17 % af variansen i stikprøven (se tabel 4.2) og at forskellen er signifikant forskellig fra nul. Disse resultater er i overensstemmelse med den anden af teserne fra litteraturstudiet, der lød på, at der vil være en vis grad af lighed på den enkelte skole forskellig fra øvrige skoler, når det kommer til oplevet disciplinært klima. På trods af, at larm opleves forskelligt fra elev til elev, viser data, at der er en vis grad af overensstemmelse indenfor den enkelte skole.

Resultatet antyder, at hver skole har en form for særegen skolekultur, når det kommer til disciplinært klima, og at det derfor giver mening at benytte skolen som analyseniveau. At benytte det aggregerede mål for disciplin tilgodeser samtidig, at det er generelle informationer om læringsmiljøet, der er af interesse og ikke den enkelte elevs individuelle idiosynkrasier. At benytte det aggregerede mål for disciplin vil være i lighed med en række andre studier (Dowell & Neal, 1982; Kunter & Baumert, 2007; Teodorovic, 2009).

I spørgeskemaet til skolelederne bliver skolelederne spurgt, om elevernes indlæring bliver hæmmet af elever, der forstyrrer undervisningen. Spørgsmålet er generelt og gældende alle elever på hele skolen, og vedrører derfor ikke specifikt de 15-årige elever. Derfor vurderer jeg, at spørgsmålet kun i ringe grad kan indfange det disciplinære klima i skolens 9. klasser. I den tidligere omtalte artikel af Olsen (2003) vurderes det, at skolelederens vurdering kan være med til at vise, at hver enkelt skole har sin egen skolekultur. Ved korrelationsanalyse af skolelederens oplevelse af om indlæring bliver hæmmet af forstyrrende elever overfor skolegennemsnittet af elevernes oplevelse af det disciplinære klima får jeg en korrelationskoefficient på 0,21. Sammenhængens styrke er derfor ikke overvældende, men dette var heller ikke forventet, da skolelederens vurdering er for hele skolen. Men at der trods alt er en sammenhæng, og at denne er positiv, mener jeg, i tråd med Olsen, kan være med til at understrege, at der på hver skole optræder en særlig disciplinkultur.

Det disciplinære klima i klasseværelset er i sagens natur et klasserumsfænomen, og den enkelte elev i stikprøven angiver da også, hvorledes disciplinen opleves i netop denne elevs klasse. Men PISA 2003 er udvalgt ud fra skoler og eleverne er i data grupperet således. Dette afsnit vil undersøge, om skolegennemsnittet for oplevet disciplin siger noget præcist og brugbart om disciplinen i den enkelte elevs klasse. For det er nu en gang her, sammenhængen mellem det disciplinære klima og PISA-score kan tænkes at opstå.

Til at undersøge dette blev der i tidligere i artiklen udvalgt 15 skoler fra stikprøven, hvorfra de udtrukne PISA-elever med stor sandsynlighed stammer fra samme 9. klasse. Metoden er at sammenligne intra-class korrelationen fra denne gruppe af skoler med intra-class korrelationen fra de resterende skoler, hvor der med stor sandsynlighed er flere 9. klasser på hver skole.

Resultatet af denne sammenligning er angivet i tabel 4.2. Fra tabellen ses det, at intra-class korrelationen fra klasse-gruppen er en smule større end intra-class korrelationen for skolegruppen, hhv. 21,8 % og 19,7 %. Det kan tolkes heraf, at klasserne i klasse-gruppen er en smule mere homogene end skolerne i skole-gruppen, når det kommer til elevernes oplevede disciplin. Dette er i overensstemmelse med mine forventningen, og det er da heller ikke overraskende at elever fra én og samme klasse vil være mere ens i forhold til oplevet larm end elever, der blot går på samme skole. Den forholdsvis lille forskel mellem intra-class korrelationerne betyder, at det gi-

ver god mening at benytte klasserumsfaktoren disciplin på skoleniveau, på trods af stikprøvedesignet, hvor skolen og ikke klassen er grupperingsvariabel.

Disse resultater er dermed med til at validere skolegennemsnittet for oplevet disciplinært klima som inputvariabel i artiklens analyser, da de viser, at der på skoleniveau er et særligt disciplinært klima, som er meget lig det, man sandsynligvis ville kunne finde på klasseniveau på samme skole.

Samtidig viser resultatet af afsnittets første analyse, at denne særlige disciplinkultur på den enkelte skole er særegen og forskellig fra den, man finder på andre skoler.

Det aggregerede mål for det disciplinære klima på skoleniveau går teoretisk set fra 0 til 15, men reelt fra 4,64 til 11,83. Indekssets gennemsnit for de 180 skoler er på 8,89 og har en standardafvigelse på 1,37.

4.3 Bivariat analyse

Af tabel 4.3 ses det, at den gruppe af elever, der går på skoler med den dårligste oplevede disciplin, har en gennemsnitsscore i PISA-testen på 500,81, hvorimod gruppen af elever, som går på skoler med den bedst oplevede disciplin, har en gennemsnitlig PISA-score på 537,49. Videre ses det, at jo bedre disciplin der opleves på skolen, des bedre klarer den enkelte elev sig i PISA-testen, og analysen viser en tilnærmelsesvis lineær sammenhæng mellem disciplin og PISA-score. Med den undtagelse at der ingen større forskel er mellem den midterste gruppe elever og den gruppe elever, der har den næstbedste oplevede disciplin på skolen. Der er således en umiddelbar sammenhæng mellem disciplin på skoleniveau og den enkelte elevs PISA-score i matematik, når der ikke inddrages andre variabler.

Tabel 4.3. *Sammenhæng mellem disciplin på skoleniveau og PISA-score i matematik.*

	N	Gennemsnitlig PISA-score	St.fejl
0-7,99 (dårlig disciplin)	734	500,81	(5,30)
8-8,67	758	506,63	(6,71)
8,68-9,21	707	526,09	(6,36)
9,22-10,11	721	525,84	(6,02)
10,12-15 (god disciplin)	729	537,49	(5,39)

Data: PISA 2003

Tabel 4.4 viser resultatet af en tilsvarende analyse, hvis drengene og pigerne undersøges for sig. Drengene i stikprøven har en gennemsnitlig PISA-score på 528,64 med en standardfejl på 3,74, og pigerne en gennemsnitlig PISA-score på 510,44 med en standardfejl på 3,08.

I tabellen er det angivet hvordan de forskellige grupper klarer sig i forhold til den gennemsnitlige score for de to køn. Af gennemsnittene for de fem niveauer af disciplin ses, at der for både drenge og piger er en tilnærmelsesvis lineær sammenhæng. Des bedre disciplin i klasseværelset, jo bedre PISA-score i matematik.

Tabel 4.4. *Sammenhængen mellem køn, disciplin i klasseværelset og PISA-score.*

	Drenge				Piger			
	N	Score	(St.fejl)	Forskel*	N	Score	(St.fejl)	Forskel*
0-7,99 (dårlig disciplin)	354	506,93	(6,15)	-21,71	380	495,07	(6,22)	-15,37
8-8,67	375	519,01	(7,40)	-9,63	383	494,41	(8,26)	-16,03
8,68-9,21	358	534,19	(7,49)	5,55	349	517,87	(7,05)	7,43
9,22-10,11	342	532,01	(8,06)	3,37	379	520,31	(6,15)	9,87
10,12-15 (god disciplin)	348	551,36	(6,93)	22,72	381	525,01	(5,22)	14,57

Data: PISA 2003. * forskel i forhold til kønsgennemsnittet i PISA-score.

Hvis forskellene i forhold til det kønsspecifikke gennemsnit undersøges, kunne det ud fra denne tabel tyde på, at drengene bliver påvirket mere af ekstremerne, end det er gældende for pigerne. Forstået på den måde at drenge der går i en klasse med meget oplevet larm, scorer hele 21,7 point dårlige end gennemsnittet, hvorimod piger der går i klasser med meget larm, kun scorer 15,4 point dårligere end gennemsnittet. Ligeledes scorer drenge, der går i klasser med meget god disciplin, hele 22,7 point bedre end gennemsnittet, hvor pigerne kun scorer 14,6 point bedre.

Hvor drengene ser ud til at blive påvirket mest af ekstremerne og i mindre grad i de øvrige grupper, ser det ud til, at pigerne også er meget påvirket af disciplinen, hvis de går i klasser, der tilhører gruppen med den næstdårligste disciplin. Denne gruppe scorer for pigernes vedkommende 16,0 point dårligere end gennemsnittet.

Disse forskelle gør det interessant senere i analysen ved hjælp af regressioner at udforske denne tilsyneladende kønsspecifikke forskel yderligere. Hvis det forholder sig sådan, at drengene påvirkes mere af det disciplinære klima i klasseværelset end pigerne, bliver det uddannelsespolitisk interessant, da drengene i disse år ser ud til at have svært ved at begå sig i den danske grundskole.

En af de mest betydende faktorer, for hvordan en elev klarer sig fagligt i den danske grundskole, er dennes socioøkonomiske baggrund, og herunder forældrenes uddannelsesniveau (Hansen, 2003). Af tabel 4.5 ses det, at for alle de tre grupper af elever gælder, at der er en sammenhæng mellem disciplin og PISA-score, og at sammenhængen er tilnærmelsesvis lineær. Således er der en negativ sammenhæng mellem dårlig disciplin og opnået PISA-score for den enkelte elev uanset dennes forældres uddannelsesniveau.

Den første gruppe elever, hvis forældre højest har folkeskolens udvidede afgangseksamen, scorer i gennemsnit 468,0 i PISA-testen, med en standardfejl på 6,6. Gruppen er den mindste i antal af de tre konstruerede grupper med 354 elever og kaldes herefter gruppe 1. Den midterste gruppe, hvis forældre højest har en gymnasial eller erhvervsuddannelse, har en gennemsnitlig PISA-score på 521,3 med en standardfejl på 2,6. Denne gruppe kaldes herefter gruppe 2. Og den sidste gruppe, hvis forældre minimum har en mellemlang videregående uddannelse, har en gen-

nemsnitlig PISA-score på 549,1 med en standardfejl på 4,5. Og denne gruppe kaldes gruppe 3. Som i tabel 4.4 er der igen angivet forskel fra gruppegennemsnittet.

Tabel 4.5. *Sammenhængen mellem forældres højeste uddannelse, disciplin i klasseværelset og PISA-score. (Standardfejl i parentes).*

	Folkeskolen eller mindre			Gymnasial eller erhvervsudd.			Mellemlang eller mere		
	N	Score	Forskel*	N	Score	Forskel*	N	Score	Forskel*
0,00 -7,99	85	448,3 (13,0)	-19,7	487	508,8 (5,7)	-12,5	122	524,6 (10,1)	-24,5
8,00 -8,67	84	446,0 (13,7)	-22,0	488	512,5 (5,8)	-8,8	153	537,2 (8,9)	-11,9
8,68 -9,21	63	477,6 (13,6)	9,6	478	525,0 (5,8)	3,7	151	556,0 (9,7)	6,9
9,22-10,11	67	487,9 (17,4)	19,9	478	522,4 (5,6)	1,1	149	563,2 (6,8)	14,1
10,12-15	55	498,0 (14,9)	30,0	505	536,9 (4,9)	15,6	148	558,3 (11,0)	9,2

Data: PISA 2003. *forskul i forhold gennemsnittet for gruppen.

For gruppe 1 er den gennemsnitlige forskel i score, mellem dem der går i klasser med god disciplin, og dem der går i klasser med dårlig disciplin hele 49,7 point, hvor det samme tal for gruppe 2 og gruppe 3 er på hhv. 28,1 og 33,7 point. Samme mønster tegner sig i forskellen i gennemsnitlig PISA-score mellem de klasser, som ligger i grupperne med næstbedst og næstdårligst oplevet disciplin. Her er forskellen igen størst for den gruppe elever, hvis forældre højest har gennemført folkeskolen, hvor der er en forskel på 41,9 point, mens forskellen for gruppe 2 og gruppe 3 er hhv. 9,9 og 26,0 point. Disse resultater tyder på, at styrken af sammenhængen mellem disciplin og PISA-score er størst for de elever, hvis forældre højest har en folkeskoleeksamen, næsthøjest for elever hvis forældre minimum har en mellemlang videregående uddannelse, og mindst for gruppen hvis forældre enten har en gymnasial eller en erhvervsuddannelse.

Fra en uddannelsespolitisk vinkel vil det til alle tider være interessant at blive klogere på, hvordan vi hæver niveauet for dem, der har det sværest fagligt i den danske grundskole. Derfor er det en interessant opdagelse, at styrken af sammenhængen mellem disciplin og PISA-score er størst for den gruppe elever, hvis forældre højest har en folkeskoleeksamen. Sammenhængen mellem forældres højeste uddannelse og styrken af sammenhængen mellem disciplin og PISA-score vil blive undersøgt nærmere i næste kapitels OLS regressioner.

4.4 Kausalitet og sammenhængens retning

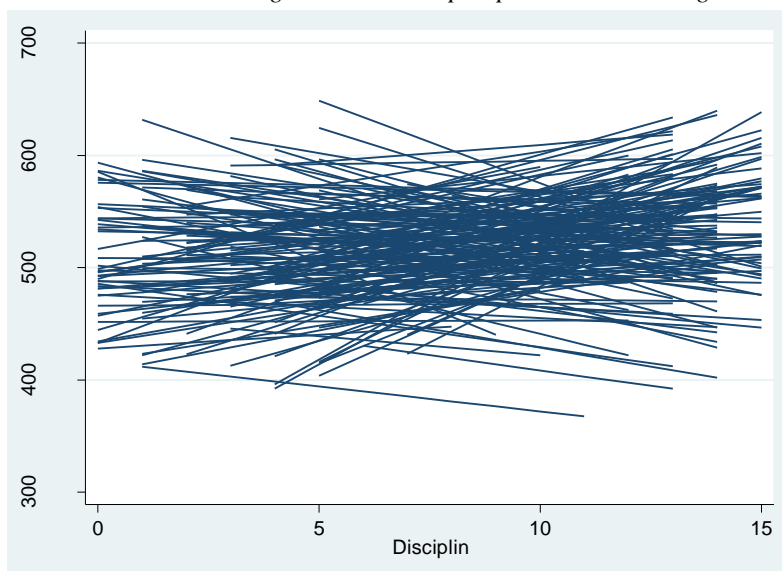
Som beskrevet i artiklens indledning er jeg bevidst om, at jeg med analysens design ikke kan påvise en kausalsammenhæng mellem det disciplinære klima i klasseværelset og PISA-score. Med dette afsnit vil jeg sandsynliggøre, at der kan være en sådan på spil.

Ved at benytte det aggregerede mål for disciplin sikrer jeg mig til en vis grad, at den fundne sammenhæng mellem disciplin og PISA-score ikke blot er et udtryk for, at de dygtige elever oplever mindre dårlig disciplin end de mindre dygtige elever, selvom de går i den samme klasse. Dette var en af bekymringerne fra afsnit to. Dette gælder dog ikke, hvis der ikke hersker en vis grad af

heterogenitet i forhold til PISA-score på de enkelte skoler. Hvis for eksempel alle dygtige elever går på samme skoler, vil skolegennemsnittet for oplevet disciplin være højt. Og det samme vil PISA-scoren for de enkelte elever. Nedenstående analyse vil i en vis grad imødekomme denne mulige bias ved for de enkelte skoler at undersøge sammenhængen mellem oplevet disciplin på individniveau og PISA-score. Jeg vil med andre ord undersøge, om der på de enkelte skoler er tegn på, at de dygtige elever skulle opleve mindre dårlig disciplin end de mindre dygtige. Eller om andre sammenhænge mellem disciplin og PISA-score skulle vise sig. Hvis der ikke findes nogen sammenhæng mellem disciplin og PISA-score på den enkelte skole, øger det sandsynligheden for, at PISA-scoren ikke er afgørende for, hvordan den enkelte elev oplever det disciplinære klima. Det er dermed med til at sandsynliggøre, at det er det disciplinære klima i klasseværelset, der yder indflydelse på PISA-score og ikke omvendt.

Andre studier har som tidligere nævnt fundet en kurvelineær sammenhæng mellem oplevet disciplin og faglig kunnen, dvs. at de dygtigste elever har en tendens til at vurdere det disciplinære klima til at være middel, hvor de mindre dygtige har en tendens til at vurdere det i ekstremerne, dvs. enten helt godt eller helt skidt (Olsen, 2003; den Brok et al., 2004). Jeg har for hver enkelt skole grafisk, ved hjælp af en scatterplot-graf, undersøgt sammenhængen mellem disciplin og PISA-score og samtidig beregnet korrelationskoefficienten for samme sammenhæng. Bilag 2 viser et tilfældigt udvalgt uddrag af denne undersøgelse. I undersøgelsen fandt jeg ingen tegn på, at der skulle være særlige tendenser i sammenhængen mellem disciplin og PISA-score for den enkelte skole. Således var der intet tegn på kurvelinearitet, eller at de dygtige elever skulle opleve det disciplinære klima anderledes end de mindre dygtige elever. Samtidig viser korrelationskoefficienterne heller ingen form for mønstre, idet sammenhængen skiftevis er positiv og negativ, samtidig med at den varierer i styrke fra næsten ingen sammenhæng til en stærk sammenhæng. Figur 4.2 viser tilpassede regressionslinjer for hver enkelt skole på sammenhængen mellem disciplin på individniveau og PISA-score. Heller ikke her er der tegn på noget mønster i sammenhængen når regressionslinjerne for de enkelte skoler sammenlignes.

Figur 4.2. *Tilpassede regressionslinjer for hver af stikprøvens skoler på sammenhængen mellem disciplin på individniveau og PISA-score.*



Hvor ovenstående analyse af sammenhængen mellem oplevet disciplin og PISA-score på den enkelte skole imødekommer den mulighed, at dygtige elever oplever mindre larm end de mindre dygtige elever, er det ikke på samme måde muligt empirisk at imødekomme den mulighed, at dygtige elever larmer mindre end mindre dygtige elever. Hvis dette er tilfældet, vil der i en klasse med mange dygtige elever simpelthen være mindre larm end i en klasse med mange mindre dygtige elever. Hvis dette er tilfældet, at dygtige elever larmer mindre, vil den fundne sammenhæng mellem det disciplinære klima i klasseværelset og PISA-score, nødvendigvis delvist skulle forklares herved. Data viser dog, at der både er skoler med god disciplin, hvor eleverne ikke klarer sig godt i PISA (fx skolerne med id-nummer 00025 og 00012), og der er skoler med dårlig disciplin, hvor eleverne alligevel klarer sig godt i PISA (fx skolerne med id-nummer 00175 og 00141). Så at dygtige elever larmer mindre, kan højst være en del af forklaringen på den fundne sammenhæng. Et norsk observationsstudie finder da heller ingen forskel i 'Klasseromsrelateret problematferd' mellem dygtige og mindre dygtige elever, hvilket er med til at sandsynliggøre at det er det disciplinære klima der har en effekt på den enkelte elevs læring (Nordahl et al., 2009, s. 16).

4.5 OLS regressionsanalyse

Tabel 4.6 viser resultatet af en OLS regression (model 1) som inddrager alle elever i stikprøven, og kontrollerer for en række baggrundsfaktorer. Det ses, at den tidligere fundne bivariate sammenhæng mellem disciplin og PISA-score, også gælder i denne regressionsmodel. Sammenhængen er signifikant, og har et parameterestimat på 7,3. Det betyder, at for hvert trin disciplinen i klassen opleves forbedret, stiger den enkelte elevs PISA-score med 7,3 point.

Af model 1 ses det yderligere, at drengene scorer 16,1 point højere end pigerne i PISA-testen, når der kontrolleres for de øvrige variabler. Forældres højeste uddannelse har ligeledes en signifikant sammenhæng med PISA-score, både den enkelte elevs, men også skolegennemsnittet for forældres højeste uddannelse har betydning. Forældres jobstatus har kun på individniveau en signifikant sammenhæng med PISA-score. Af de to kontrolvariabler for etnicitet har det kun betydning, om eleven er født i Danmark. Ligeledes ses det, at elever der går i folkeskolen, klarer sig bedre end de elever, der går i privatskole. Denne sammenhæng er ikke signifikant ($p=0.11$).

Model 1a og 1b tester om styrken af sammenhængen mellem disciplin og PISA-score er forskellig for drenge og piger. Som det blev fundet i forrige kapitels bivariate analyse, ses det ligeledes i denne analyse, at styrken er stærkere for drenge end for piger med parameterestimer på 8,3 for drengene og 6,3 for pigerne. Det vil sige, at drengene påvirkes mere af det disciplinære klima i klasseværelset, end det gælder for pigerne, når det kommer til at klare sig godt i matematik. Og som nævnt tidligere er denne forskel mellem kønnene politisk interessant. I de senere år har der været stor fokus på, at drengene får sværere og sværere ved at klare sig fagligt i folkeskolen i forhold til pigerne, og politikerne har svært ved at finde ud af, hvor der skal sættes ind. Men da disciplinen i klasseværelset, som nævnt indledningsvis i artiklen, er en faktor, som kan reguleres og forbedres, bliver dette forhold pludselig til et uddannelsespolitisk interessant indsatsområde, hvis man vil have drengene med.

Foruden denne forskel ses det i forhold til model 1, at det er de samme variabler, der fortsat er signifikante, dog med forskellig styrke. Forældres højeste uddannelse har mere end et dobbelt så stort parameterestimat hos pigerne, end det er gældende hos drengene, og omvendt har forældrenes jobstatus mest betydning hos drengene. Interessant er det også, at det har langt større betydning for drengene i forhold til PISA-score, om de er født i Danmark, end det har for pigerne. Denne forskel er et udtryk for, at tilflyttede piger klarer sig markant bedre end tilflyttede drenge.

Tabel 4.6. OLS regressioner.

	Model 1	Model 1a	Model 1b	Model 1c	Model 1d	Model 1e
	'OLS'	'Dreng'	'Pige'	'FS'	'Gym+erhv.'	'Mellemlang+'
	Est.	Est	Est.	Est.	Est.	Est.
Individniveau:						
Konstant	209,8**	242,5**	217,2**	194,3*	217,9**	188,8**
Dreng	16,1**	.	.	26,6**	17,7**	7,9
Foræld.Udd.	9,0**	5,3*	12,3**	.	.	.
Foræld.Jobstatus	0,9**	1,1**	0,8**	0,3	0,9**	1,5**
Født i Danmark	32,1**	42,2**	26,1**	-9,2	27,9**	55,8**
Taler dansk i hjemmet	7,2	0,5	16,9	-14,7	4,8	22,2
Skoleniveau:						
Disciplin	7,3**	8,3**	6,3**	12,3**	6,3**	7,2*
Foræld.Udd.(agg)	23,6**	22,9*	24,9*	34,4*	19,8*	24,1
Foræld.Job (agg)	0,2	0,2	0,1	-0,1	0,3	0,3
Privatskole	-10,1	-11,0	-9,2	-22,5	-8,1	-10,9
R ²	,144	,126	,146	,144	,092	,174
N =	3263	1565	1698	298	2293	672

Data: PISA 2003. * $p < 0.05$, ** $p < 0.01$.

I model 1c, model 1d og model 1e undersøges det, om styrken af sammenhængen mellem det disciplinære klima og PISA-score varierer, når det kommer til forældrenes uddannelsesniveau. Resultatet fra artiklens bivariante analyser bekræftes, idet sammenhængen synes stærkest for de elever, hvis forældre har en afgangsprøve fra folkeskolen som højeste uddannelsesniveau. Parameterestimatet for disciplin-variablen for denne gruppe elever er på 12,3 og signifikant, mens den for elever, hvis forældre har en gymnasial eller erhvervsuddannelse, har et estimat på 6,3, og elever, hvis forældre har en mellemlang videregående uddannelse eller mere, har et estimat på 7,2. I forrige kapitel fandt jeg, at den gennemsnitlige PISA-score for disse tre grupper af elever er på hhv. 468,0, 521,3 og 549,1 point. Det betyder, at man ved at forbedre det disciplinære klima i klasseværelset, kan forbedre alle elevers faglige kunnen, men samtidig være med til at mindske forskellen i faglig kunnen mellem elever med højtuddannede forældre og elever med lavtuddannede forældre.

Af tabel 4.6 ses det desuden ved en sammenligning mellem model 1c, 1d og 1e, at drengene klarer sig bedst i alle tre modeller, og at forskellen mellem drengene og pigerne bliver større, des mindre uddannelse forældrene har. Lidt frækt og uden teoretisk belæg kan man påstå, at i forhold

til matematik, er drengene som udgangspunkt meget bedre end pigerne, men at forældre til pigerne, som funktion af deres uddannelsesniveau, kan være med til at mindske denne forskel.

At være født i Danmark har størst betydning for den gruppe elever med de højest uddannede forældre, og mindst betydning for eleverne af lavtuddannede forældre, hvor sammenhængen med PISA-score ikke er signifikant. At sammenhængen for denne gruppe ikke er signifikant, kan delvist være et resultat af det lave antal elever i denne model, hvoraf samtidig meget få elever ikke er født i Danmark. Forældres højeste jobstatus betyder mere for opnået PISA-score, des højere uddannet forældrene er, mens skolegennemsnittet for forældres uddannelse har størst betydning for de elever, hvis forældre højest har gennemført grundskolen.

4.6 Multilevel regressionsanalyse

De følgende analyser tager udgangspunkt i en null-model (model 0 i tabel 4.7). Null-modellen benytter PISA-scoren som afhængig variabel og skolen som grupperingsvariabel. Derudover indeholder modellen ingen forklarende variabler på hverken individ- eller skoleniveau. Modellen tillader random intercept og indeholder et sådant for hver skole. Modellen finder det gennemsnitlige skolegennemsnit for PISA-scoren til at være 518,71 med en standardfejl på 2,70. Gennemsnitsscoren for skole j kan derved udregnes som $518,71 + \text{den skolespecifikke residual}$. Hvis en skole har en residual større end 0, har denne skole en gennemsnitlig PISA-score større end stikprøvegennemsnittet, og omvendt hvis en skole har en residual mindre end 0, har denne skole en gennemsnitlig PISA-score mindre end stikprøvegennemsnittet. Residualerne for de enkelte skoler vil blive undersøgt senere.

Null-modellen estimerer between-school variansen til 1020,35 og den totale varians til 7233,51, hvilket giver en intra-class korrelation på $(1020,35 / 7233,51 =)$ 14 %. Dermed kan 14 % af variansen i PISA-score forklares af forskelle skolerne imellem. En likelihood ratio test viser desuden at between-school variansen med overvejende sandsynlighed er forskellig fra nul ($p < 0,001$), og at det derved giver god mening fortsat at benytte multilevel regression.

Intra-class korrelationen for PISA-scoren på 14 % er en smule mindre end intra-class korrelationen for oplevet disciplin. Denne blev i kapitel fire fundet at være 17 %. Dermed er homogeniteten på den enkelte skole en smule større, hvad angår oplevet larm end opnået PISA-score, hvilket giver god mening, da der er grænser for hvor individspecifikt, man kan opleve disciplinen i samme klasse, hvorimod en elevs PISA-score ikke i samme grad er afhængig af de øvrige elever på skolens PISA-score.

Den fundne intra-class korrelation i forhold til PISA-score på 14 % er med til at beskrive en dansk grundskole, med stor faglig spredning på de enkelte skoler. Dette kan bidrage til den fortløbende kausalitetsdiskussion. Jeg skrev tidligere, at hvis skolerne i den danske grundskole var meget homogene, når det kom til fagligt standpunkt, ville sammenhængen mellem disciplinen i klasseværelset og PISA-score kunne dække over, at mindre dygtige elever oplever mere larm. Forstået på den måde, at med megen homogenitet, ville der være en række skoler med mange mindre dygtige elever, som alle ville opleve meget larm, og samtidig en række skoler med mange dygtige elever,

Tabel 4.7. *Multilevel regression.*

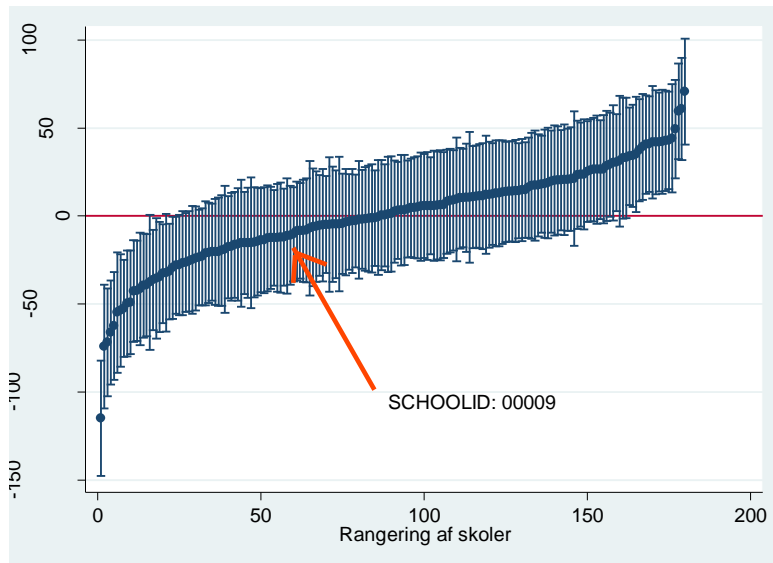
	Model 0 'Null-model'	Model 1 'OLS'	Model 2 'random intercept'
	Est.	Est.	Est.
Individniveau:			
Konstant/Intercept	518,71**	209,8**	204,5**
Dreng	-	16,1**	15,3**
Foræld.Udd.	-	9,0**	8,8**
Foræld.Jobstatus	-	0,9**	0,9**
Født i Danmark	-	32,1**	32,0**
Taler dansk i hjemmet	-	7,2	10,2
Skoleniveau:			
Disciplin	-	7,3**	7,4**
Foræld.Udd.(agg)	-	23,6**	22,4**
Foræld.Job (agg)	-	0,2	0,3
Privatskole	-	-10,1	-8,4§
Forklaret variation (%):			
Skoleniveau	-	-	71,0
Elevniveau	-	-	11,4
Total (R ²)	-	14,4	19,7
N =	3649	3263	3263

Data: PISA 2003. § $p < 0,1$, * $p < 0,05$, ** $p < 0,01$.

der ikke ville opleve så meget larm. Men med en intra-class korrelation på 14 % er det ikke tilfældet.

Figur 4.3 er genereret ud fra ovenstående null-model og viser et 'caterpillar plot' over rangerede skoleresidualer. De fede blå punkter er den enkelte skoles gennemsnitlige PISA-score med tilhørende 95 %-konfidensintervaller holdt op mod skoleresidualens rangering blandt alle skoler i stikprøvens residualer. Den røde linje repræsenterer den gennemsnitlige skoleresidual i stikprøven.

Figur 4.3. 'Caterpillar plot' over rangerede skoleresidualer.



Data: PISA 2003

De skoler, der er placeret over den røde linje, har et skolegennemsnit bedre end stikprøvegennemsnittet, mens de skoler, der er placeret under den røde linje, har et skolegennemsnit lavere end stikprøvegennemsnittet. Det er derved ud fra grafen muligt at identificere de skoler, som har en signifikant bedre eller signifikant dårligere gennemsnitsscore end den gennemsnitlige skole. Fx har skolen med id-nummer 00009 en estimeret residual på -12,23, hvilket giver denne skole en gennemsnitlig PISA-score på $518,71 - 12,23 = 506,48$. Skolen rangerer som nummer 56 fra bunden. Rangeringen skal i denne sammenhæng ikke tillægges nogen større betydning, da data kun er en stikprøve af alle skoler, og kun medtaget for at gøre grafen mere læselig. Det interessante bliver derved at se variationen i gennemsnitlig PISA-score mellem de enkelte skoler, som en illustration af intra-class korrelationen på 14 %, og hvorfor skoleniveauet bør medtænkes, når data behandles. Det skal i forhold til figuren bemærkes, at der for multilevel modeller gælder en shrinkage faktor, som bevirker, at residualerne for især højt og lavt præsterende skoler, estimeres mindre end de faktisk er (OECD, 2009, s. 213).

I PISA-undersøgelsen er elevernes faglige kunnen kun målt på ét tidspunkt. Det er derfor ikke muligt at fortolke skoleresidualerne som skoleeffekter, da dette ville kræve et mål for elevernes faglige forbedringer over tid. Rangeringen i figur 4.3 siger dermed ikke noget som helst om, hvilke skoler der er bedst til at lære deres elever matematik, blot hvor gode eleverne på skolen i gennemsnit er. Residualerne kan derfor kun fortælle, hvordan skolerne er forskellige, når det kommer til, hvor godt deres elever klarer sig i PISA-testen.

4.7 Random intercept model med fixed effects

I model 2 tilføjes baggrundsvariablerne fra den tidligere OLS regression til null-modellen som fixed effects. Resultatet fremgår af tabel 4.7. Variablen af mest interesse, disciplin på skoleniveau, viser sig med et estimat på 7,4, også i denne model at have stor betydning for, hvordan den enkelte elev klarer PISA-testen. Estimatet betyder, at hvis disciplinen i klassen bliver forbedret ét point, så vil den enkelte elev forbedre sig med 7,4 point i PISA-testen. Estimatet er meget lig det parameterestimat, der blev fundet i den tidligere OLS regression, som var på 7,3. OLS regressionen (model 1) er medtaget i tabel 4.7 for sammenligningens skyld. De øvrige kovariater har ligeledes parameterestimerer meget lig dem fra OLS regressionen. Maximum likelihood estimererne i denne model viser sig generelt set at være en smule lavere end parameterestimererne fra OLS regressionen.

Det er dog værd at bemærke, at estimatet for dummyen privatskole, nu er lige i nærheden af at være signifikant ($p=0,075$). I stikprøven er der 180 skoler, hvoraf kun 36 er privat drevne. Og det lille antal privatskoler kunne være en grund til at slække på 5 %-kravet til signifikansniveauet, hvilket i så fald vil betyde, at der er en signifikant forskel mellem privat- og folkeskoler, når det kommer til opnået PISA-score. Hvor folkeskoleeleverne er dem, der har den højeste score. Det skal bemærkes, at privatskole i datasættet dækker over både privat-, fri- og lilleskoler, som er tre typer skoler med meget varierende pædagogisk ståsted og elevoptag, hvilket efter min mening gør forskellen mindre interessant.

Ved at sammenligne variansen i model 2 med variansen i null-modellen, fås det, at knap 20 % af den totale varians af PISA-score kan forklares af de tilføjede variabler. Ligeledes fås det ved beregning, at de tilføjede variabler forklarer 71 % af forskellen mellem skoler og godt 11 % af forskellen på den enkelte skole (se tabel 4.7). Ved at benytte multilevel regression og tillade random intercept, forbedres modellens forklaringsgrad dermed betydeligt i forhold til at benytte standard OLS regression.

Som skrevet skelnes der i datasættet ikke mellem de forskellige typer af privatskoler, der tilbyder skolegang for de 15-årige. Denne analyse er lavet med det in mente. På trods af denne mangel i data, mener jeg stadig, at det er interessant at se, om folkeskolerne og privatskolerne i Danmark er forskellige, når det kommer til sammenhængen mellem det disciplinære klima i klasseværelset og PISA-score. I tabel 4.8 fremgår resultatet af tre random intercept modeller. Den første er en gengivelse af model 2 fra tabel 4.7, som er medtaget til sammenligning. De to næste modeller er estimeret ud fra samme variabler, men med en separat model for hhv. folkeskoler og privatskoler. Der indgår 138 folkeskoler og 36 privatskoler i beregningerne.

Det ses, at parameterestimatet for disciplin er på 5,6 for folkeskolen og 11,7 for privatskolerne. Begge signifikante. Dette betyder, at eleverne i den danske folkeskole ikke påvirkes i samme grad af disciplinen i klasseværelset, som det er tilfældet i på privatskolerne. Da privatskole dækker

Tabel 4.8. *Multilevel regression. Privat- og folkeskole.*

	Model 2 'random intercept'	Model 'Folkeskole'	Model 'privatskole'
	Est.	Est.	Est.
Individniveau:			
Konstant/Intercept	204,5**	221,3**	99,1§
Dreng	15,3**	15,1**	17,0**
Foræld.Udd.	8,8**	8,5**	10,1**
Foræld.Jobstatus	0,9**	1,0**	0,6**
Født i Danmark	32,0**	35,8**	19,0
Taler dansk i hjemmet	10,2	5,8	37,5
Skoleniveau:			
Disciplin	7,4**	5,6**	11,7**
Foræld.Udd.(agg)	22,4**	20,3**	31,5§
Foræld.Job (agg)	0,3	0,2	0,3
Privatskole	-8,4§	-	-
Forklaret variation (%):			
Skoleniveau	71,0	79,8	42,3
Elevniveau	11,4	12,6	7,4
Total (R ²)	19,7	21,9	12,2
N =	3263	2620	643

Data: PISA 2003. § $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

over vidt forskellige typer af skoler, bliver forskellen mellem folkeskolen og privatskoler mindre interessant.

Et parameterestimat på 5,6 for den danske folkeskole er 1,8 point lavere end det blev fundet for hele stikprøven i model 2. Da uddannelsespolitikken i Danmark laves med folkeskolen for øje, er dette resultat interessant. En delkonklusion for artiklen vil således blive, at sammenhængen mellem det disciplinære klima i klasseværelset og PISA-score er på 7,4 PISA-point pr. forbedring på disciplinindekset for hele stikprøven, men at sammenhængen i folkeskolen kun er på 5,6.

4.8 Danmark er et larmende land?

Resultaterne fra nedenstående transnationale sammenligninger må tages med et vist forbehold. På trods af PISA-undersøgelsens design (OECD, 2005b), bør man ifølge Erik Jørgen Hansen og Bjarne Hjorth Andersen stadig være påpasselig med sådanne sammenligninger, da sociale fænomener kan spille forskellige roller ”i forskellige samfund, der af historiske, geografiske og demografiske årsager har befolkninger, hvis livsbetingelser, behov og værdier er mere eller mindre forskellige.” (Hansen & Andersen, 2000, s. 234). Niels Egelund skriver i rapporten ’PISA 2003 – Danske unge i en international sammenligning’ om netop disciplin-spørgsmålene: ”Der må naturligvis tages forbehold over for, hvad der i forskellige lande opleves som passende adfærd, ligesom sprogbrugen i spørgsmålene kan give forskellige nuancer. Derfor bør analyser inden for hvert lands egne rammer tillægges størst relevans” (Mejding (Red.), 2004, s. 199).

På trods af, men med udgangspunkt i problematikken omkring internationale sammenligninger, mener jeg, at det er muligt at sammenligne resultaterne fra Danmark med vores to nordiske nabolande, Sverige og Norge. Jan Mejding skriver da også videre i den før nævnte PISA-rapport, at det på grund af kulturelle forskelle landene imellem giver bedst mening at lave sammenligninger med de øvrige nordiske lande (Mejding (Red.), 2004, s. 8). Danmark, Sverige og Norge ligner hinanden tilpas meget kulturelt set til, at en sammenligning mellem netop disse tre lande vil give mening.

For at få et så ens sammenligningsgrundlag som muligt, anvendes samme procedure for sortering af data, som blev anvendt på de danske tal. Derfor frasorteres elever, som ikke går i 9. klasse i Sverige eller i 10. klasse i Norge. Herefter frasorteres skoler med færre end 10 elever i de to lande. I den danske stikprøve blev der frasorteret 13 %, mens tallene for Sverige og Norge er på hhv. 5 og 3 %. Det giver stikprøvestørrelser på hhv. 3649, 4411 og 3954 for Danmark, Sverige og Norge.

Tabel 4.9 viser, for hvert af de tre lande, hvor ofte den enkelte elev oplever en række udsagn. Overordnet set ser de tre landes elevers svar rimelig ens ud. Svarene fra alle lande ser ud til at klumpe sammen omkring svarmuligheden ’I nogle timer’, og der er ikke umiddelbart noget land, der skiller sig ud. De danske elever oplever støj og uro en smule oftere end Norge, som igen oplever det en smule oftere end Sverige. For de andre udsagn placerer Danmark sig rimeligvis som de to nabolande, hvis ikke en smule bedre.

Tabel 4.9. *Fordelingen af danske, svenske og norske elevers besvarelser på spørgsmål 38 b, f, h, i og k fra elevspørgeskemaet: ”Hvor tit sker følgende i matematiktimerne?”*

	I hver time	I de fleste timer	I nogle timer	Aldrig eller næsten aldrig	N
”Eleverne hører ikke efter, hvad læreren siger”					
DK	8,7 %	23,6 %	58,2 %	9,6 %	3546
SWE	6,4 %	19,4 %	58,5 %	15,7 %	4343

	NOR	7,2 %	27,1 %	55,4 %	10,3 %	3801
”Der er støj og uro”	DK	17,1 %	26,7 %	47,2 %	9,0 %	3538
	SWE	13,1 %	22,6 %	51,8 %	12,5 %	4333
	NOR	14,8 %	26,9 %	49,4 %	9,0 %	3788
”Læreren må vente i lang tid, før eleverne falder til ro”	DK	8,6 %	19,1 %	49,5 %	22,8 %	3544
	SWE	11,7 %	20,7 %	48,5 %	19,1 %	4353
	NOR	11,7 %	24,7 %	48,3 %	15,4 %	3775
”Eleverne kan ikke arbejde ordentligt”	DK	4,2 %	15,3 %	54,7 %	25,8 %	3528
	SWE	4,7 %	15,3 %	55,5 %	24,6 %	4335
	NOR	6,1 %	22,5 %	58,5 %	12,9 %	3771
”Eleverne begynder først at arbejde længe efter, at timen er startet”	DK	8,2 %	19,3 %	49,8 %	22,8 %	3546
	SWE	9,5 %	19,2 %	50,7 %	20,6 %	3266
	NOR	10,1 %	26,5 %	47,6 %	15,8 %	3793

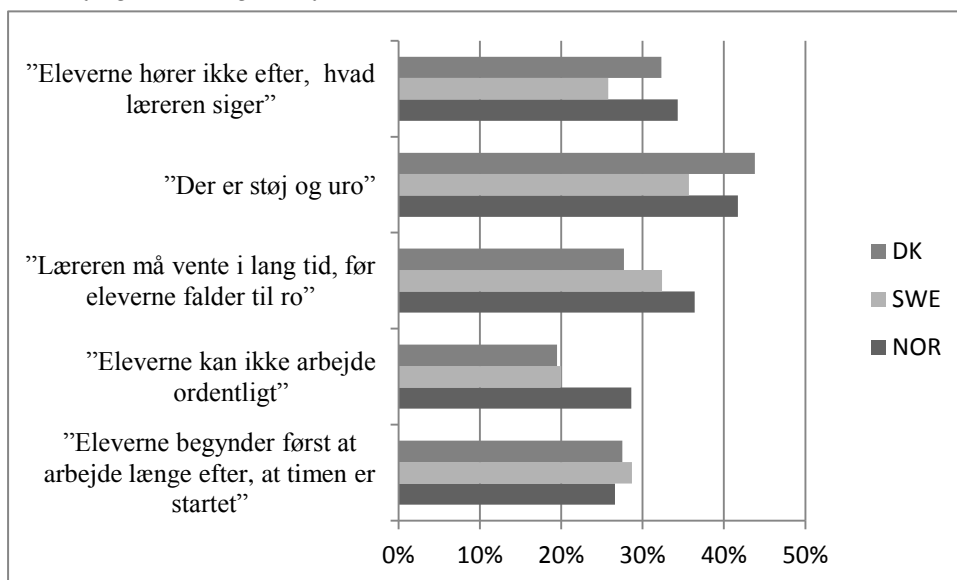
Data: PISA 2003

Figur 4.4 viser, hvor mange elever i de enkelte lande, der oplever de enkelte udsagn enten i hver time eller i de fleste timer. Af figuren ses det, at kun i forhold til oplevet støj og uro placerer Danmark sig dårligst. Ved de andre fire udsagn placerer Danmark sig to gange mellem Sverige og Norge og to gange som det bedst placerede land.

Når det kommer til hvor stor en del af eleverne fra de enkelte lande, der oplever mindst ét af udsagnene i hver time eller i de fleste timer, fås det, at det gælder for 59 % af de danske elever, mens det gælder for 57 % af de svenske elever og 65 % af de norske elever.

Der kan ud fra klassegennemsnittene af oplevet larm udregnes et vægtet landsgennemsnit for de tre stikprøver. Klassegennemsnittene bevæger sig som nævnt på en skala fra 0-15, hvor 15 svarer til, at eleverne aldrig eller sjældent oplever nogen af hændelserne. Landsgennemsnittet for Danmark er på 8,84 med en standardafvigelse på 1,41, hvor det for Sverige er på 8,98 med en standardafvigelse på 1,27 og for Norge på 8,41 med en standardafvigelse på 1,16. Disse tal viser, at det disciplinære klima i klasseværelserne i Danmark er en smule dårligere end i Sverige, men bedre end i Norge, men at forskellene ikke er signifikante.

Figur 4.4. Søjlediagram over hvor stor en procentdel i Danmark, Sverige og Norge der oplever følgende udsagn i de fleste eller i hver matematiktime.



Data: PISA 2003

Ovenstående deskriptive analyser finder en smule variation landene imellem. I grove træk kun nuancer, men ved at se på nuancerne tegner der sig et billede af, at Sverige har den bedste disciplin, tæt efterfulgt af Danmark, og Norge med den dårligste disciplin af nabolandene. Således er der ingen tegn på, at Danmark skulle være særligt, når det kommer til graden af dårlig disciplin i landets skoler. De tre landes indbyrdes placering er helt i overensstemmelse med de konklusioner Olsen fandt ud fra data fra PISA 2000 (Olsen, 2003). Nedenfor sammenlignes styrken af sammenhængen mellem det disciplinære klima i klasseværelset og PISA-score i matematik for de tre lande, Danmark, Sverige og Norge.

Tabel 4.10 viser resultatet af multilevel regressionerne for hvert af landene. Variablen af størst interesse, disciplinen i klasseværelset, ses at have en signifikant sammenhæng med PISA-scoren i alle tre lande. Styrken af sammenhængen er størst for Danmark med et maximum likelihood estimat på 7,4, hvor det for Norge er på 7,2 og for Sverige på 5,9. Det vil sige, at for hvert trin disciplinen i klasseværelset forbedres på skalaen fra 0-15, forbedrer den enkelte elev i klassen sig med mellem hhv. 7,4, 7,2 og 5,9 point i PISA-testen.

Tabel 4.10. *Multilevel regressioner for Danmark, Sverige og Norge.*

	Danmark	Sverige	Norge
	Est.	Est.	Est.
Individniveau:			
Konstant/Intercept	204,5**	235,3**	197,3**
Dreng	15,3**	7,4**	7,4**
Foræld.Udd.	8,8**	2,0§	4,6**
Foræld.Jobstatus	0,9**	1,4**	1,5**
Født i DK/SWE/NOR	32,0**	43,8**	18,6*
Taler dansk/svensk/norsk i hjemmet	10,2	23,5**	16,4*
Skoleniveau:			
Disciplin	7,4**	5,9**	7,2**
Foræld.Udd.(agg)	22,4**	16,9**	11,7
Foræld.Job (agg)	0,3	0,1	0,3
Privatskole	-8,4§	-9,0	24,7
Forklaret variation (%):			
Skoleniveau	71,0	79,5	59,5
Elevniveau	11,4	13,3	11,7
Total (R ²)	19,7	19,7	14,2
Intra-class korrelation	14,1	9,8	7,8
N =	3263	3205	3542

Data: PISA 2003. § $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Styrken af sammenhængen mellem disciplin i klasseværelset og PISA-score er dermed en smule større for Danmark, end det er gældende i vores nabolande. Resultaterne af de øvrige variabler i tabel 4.10 vil ikke blive kommenteret uddybende.

5. Konklusion

Artiklens mål var at undersøge, hvordan det disciplinære klima er i den danske grundskoles 9. klasser, og om der i Danmark er en sammenhæng mellem dette klima, og hvor gode eleverne er til matematik. Her er hvad jeg fandt:

- **Eleverne i den danske grundskole oplever ofte, at der er et dårligt disciplinært klima i klasseværelset.** Fx oplever knap 44 % af de 15-årige elever, at der altid eller næsten altid er støj og uro i matematiktimerne, 32 % at der ikke bliver hørt efter, hvad læreren siger, og 19 % at man ikke kan arbejde ordentligt.
- **Der er i Danmark en signifikant sammenhæng mellem det disciplinære klima i klasseværelset, og hvordan eleverne klarer sig i PISA-undersøgelsens matematiktest.** En multilevel regressionsmodel, der kontrollerer for køn, socioøkonomisk baggrund og etnicitet, finder, at ved en forbedring af det disciplinære klima i klassen med ét point på det kon-

struerede disciplinindeks (0-15) ville eleverne i klassen score 7,4 point højere i PISA-testen.

- **Den fundne sammenhæng mellem det disciplinære klima i klasseværelset og PISA-score er større for de 15-årige drenge end for de 15-årige piger.** For hver stigning på det konstruerede disciplinindeks ville drengene forbedre sig 8,3 point i PISA-testen, mens pigerne ville forbedre sig 6,3 point. Begge sammenhænge er signifikante.
- **Det disciplinære klima i klassen betyder mest for de elever, hvis forældre har folkeskolens afgangsprøve som højeste uddannelsesniveau.** For hver stigning på det konstruerede disciplinindeks ville denne gruppe elever forbedre sig med 12,3 point i PISA-testen.
- **Sammenhængen mellem det disciplinære klima i klassen og PISA-score falder 1,8 point i styrke hvis folkeskolen undersøges for sig.** Således er sammenhængen større for de privatskoler end for folkeskolen.
- **Det disciplinære klima i de danske klasseværelser ligner til forveksling det disciplinære klima i de svenske og norske klasseværelser.** I en sammenligning med Sverige og Norge fandt jeg, at det disciplinære klima i klasseværelserne i Danmark opleves marginalt dårligere end i Sverige og en smule bedre end i Norge.
- **I Sverige og Norge er der som i Danmark en signifikant sammenhæng mellem det disciplinære klima i klasseværelset og PISA-score.** Hvor sammenhængens styrke i Danmark er på 7,4 PISA-point pr. forbedring på det konstruerede disciplinindeks, er den i Sverige på 5,9 og i Norge på 7,2. Sammenhængen er dermed ikke særlig for den danske grundskole, om end en smule højere end i Sverige. Sammenligninger mellem lande skal dog ifølge blandt andre Hansen og Andersen tolkes med forsigtighed (Hansen & Andersen, 2000).
- **Tidligere forskning viser, at det disciplinære klima i klassen er et af de undervisningsrelaterede forhold, der har den største effekt på elevernes faglige kunnen.** Dette er med til at understrege vigtigheden af forskning, der har det disciplinære klima i klasseværelset som interesse.

Jeg baserer ovenstående konklusioner på analyse af data fra PISA 2003. For at gøre resultaterne så valide som muligt, har det været nødvendigt med en række metodiske overvejelser og en grundig vurdering og analyse af data.

5.1 Disciplinen i klassen målt ved skolegennemsnit

Jeg fandt, at forskningen peger på, at et aggregeret mål for det disciplinære klima i klassen baseret på hver elevs oplevelse heraf er pålideligt og validt. Men data fra PISA 2003 har i Danmark kun et skoleniveau og et elevniveau. Og da oplevet disciplinært klima er en klasserumsfaktor, har det været nødvendigt at diskutere og analysere, om et skolegennemsnit af elevernes oplevelse af det disciplinære klima i klasseværelset kan bruges i analysen.

Ved at benytte variansanalyse fandt jeg, at man ud fra data kan vise, at hver skole har sin egen kultur, når det kommer til det disciplinære klima i klasseværelserne, og at det derved giver mening at benytte skolegennemsnittet for oplevet larm i analysen, på trods af at det er en klasserumsfaktor.

For yderligere at validere skolegennemsnittet af oplevet larm udvalgte jeg 15 skoler i stikprøven, som med stor sandsynlighed kun har én 9. klasse. Denne gruppe skoler sammenlignede jeg med de resterende skoler i stikprøven, hvorpå der med stor sandsynlighed er mere end én 9. klasse. Jeg fandt, at variansen i elevernes oplevelse af det disciplinære klima er en smule mindre blandt de udvalgte 'klasser' end blandt de resterende 'skoler'. Den fundne forskel mellem de to grupper fortolkede jeg som, at skolegennemsnittet af oplevet larm ikke til fulde, men i overvejende grad indfanger klassegennemsnittet af oplevet larm. Dette er med til yderligere at validere skolegennemsnittet for oplevet larm som mål for det disciplinære klima i klasseværelset.

5.2 Sammenhængens kausalitet

Jeg er med artiklens forskningsdesign ikke i stand til at påvise sammenhængen mellem det disciplinære klima i klasseværelset og PISA-score som værende kausal. En ambition har dog været at sandsynliggøre sammenhængens retning, og dermed vise at forstyrrelser i undervisningen *medfører*, at eleverne lærer mindre. Og ikke omvendt. Jeg er ud fra data kommet frem til følgende:

- Ved at benytte skolegennemsnittet for oplevet disciplin undgår jeg, at sammenhængen blot er et resultat af, at mindre dygtige elever oplever mere larm. Dette argument gælder dog ikke, hvis der på skolen kun går mindre dygtige elever, som derved alle vil opleve mere larm. Med en intra-class korrelation i forhold til PISA-score på 14 % er der dog intet, der tyder på, at en sådan grad af faglig homogenitet skulle være at finde i den danske grundskole.
- Hvis skolerne undersøges enkeltvis, findes der i data intet, der tyder på, at det oplevede disciplinære klima skulle være afhængigt af elevens PISA-score. Der er således ingen tegn på, at eksempelvis dygtige elever skulle opleve disciplinen i klasseværelset anderledes end de mindre dygtige.
- Det har med data ikke været muligt, at undersøge om sammenhængen mellem det disciplinære klima i klasseværelset og PISA-score delvist dækker over, at dygtige elever lærer mindre end de mindre dygtige elever. Muligheden er langt fra utænkelig. Jeg fandt dog i data eksempler på skoler med god disciplin, som ikke scorede godt i PISA-testen, og ligeledes skoler med dårlig disciplin der scorede godt i PISA-testen. Et norsk studie finder dog heller ingen sammenhæng mellem problemadfærd og fagligt standpunkt. Ovenstående forklaring kan dog sagtens tænkes at være en medforklarende årsag på sammenhængen.
- Ved at inddrage de baggrundsfaktorer som anden forskning finder at have stor indflydelse på faglig formåen, sikrer jeg, at disciplinen i klasseværelset ikke dækker over en eller flere af disse baggrundsfaktorer, og dermed ikke blot er en spuriøs sammenhæng.

Jeg vil ud fra ovenstående konkludere, at den fundne sammenhæng mellem det disciplinære miljø i klasseværelset og PISA-score kan tolkes således, at forstyrrelser i undervisningen *medfører*, at eleverne lærer mindre. Parameterestimatet på 7,4 kan ud fra ovenstående tolkes som en øvre grænse for sammenhængens størrelse.

5.3 Diskussion

Data er fra PISA 2003, og ovenstående konklusioner siger dermed kun noget om de 15-årige i den danske grundskole anno 2003. Der kan således være sket ændringer i det disciplinære klima i klasseværelserne siden da. I 2003 oplevede knap 44 % af eleverne, at der i de fleste eller alle matema-

tiktimer var støj og uro, mens det tilsvarende tal fra dansktimerne i 2009 er 35 %. Tilsvarende var der i 2003 knap 28 % af eleverne, der oplevede, at det i de fleste eller alle timer tog lang tid, inden eleverne faldt til ro, mens tallet fra 2009 er 22 %. Tallene fra 2009 viser sig dog reelt at være en smule underestimerede, da elever med indvandrebaggrund oplever det disciplinære klima i klasseværelset signifikant bedre end elever uden indvandrebaggrund, og stikprøven i 2009 netop har en overrepræsentation af elever med indvandrebaggrund (Egelund et al., 2011, s. 106).

Ud fra en antagelse om at det disciplinære klima i matematiktimerne og dansktimerne er sammenligneligt, og med forbehold for stikprøvebiasen i 2009-data, ser det ud til, at der frem til 2009 er sket en lille forbedring i forhold til det disciplinære klima i klasseværelserne i den danske grundskole. Jeg har ikke mulighed ud fra eksisterende data at belyse, om der frem mod 2013 skulle være sket yderligere forbedringer.

Ligeså vel som der kan være sket ændringer i det disciplinære klima i klasseværelserne siden 2003, kan der også være sket ændringer i det disciplinære klimas sammenhæng med PISA-score.

Jeg fandt i analysen en sammenhæng mellem det disciplinære klima i klasseværelset og PISA-score på 7,4 PISA-point pr. forbedring på det konstruerede disciplinindeks. Jeg fandt samtidig, at sammenhængen var tilnærmelsesvis lineær, således at en forbedring fra dårlig til middel disciplin har den samme effekt på PISA-score, som en forbedring fra middel til god disciplin har. Derfor mener jeg, at der er grund til at tro, at på trods af ovenstående tegn på et forbedret disciplinært klima i klasseværelserne er den fundne sammenhæng med PISA-score stadig gældende i 2013.

Artiklens mål er ikke at ændre praksis i den danske grundskole, ej heller forsøge at være handlingsanvisende i forhold til undervisning. Artiklens mål har udelukkende været undersøgende.

Alligevel er jeg bevidst om, at artiklens resultater kan få betydning for den danske grundskole. Resultaterne fra artiklen kalder i den grad på, at det disciplinære klima i klasseværelset bliver forbedret i den danske grundskole. Men lige så klare som konklusionerne af denne artikel fremstår, lige så mangelfuld er artiklen, når det kommer til, at foreslå hvordan man kan ændre det disciplinære klima i klasseværelset.

Jeg er som uddannet og tidligere praktiserende folkeskolelærer bevidst om, at selvom en undersøgelse som denne kan konkludere, at der er en sammenhæng mellem disciplinen i klasseværelset, og hvor meget eleverne lærer, så er det disciplinære klima ikke bare noget, der ændres fra den ene dag til den anden. Det ville i det hele taget være en forkert tilgang at betragte det disciplinære klima som et isoleret fænomen, der kan elimineres uden samtidig at ændre andre forhold i og omkring skolen. Jeg mener, at det er nødvendigt at betragte det disciplinære klima som et resultat af interaktioner mellem mennesker. Derfor får læreren og skolelederen en stor rolle heri, men også hver enkelt elev og deres forældre har en medvirkende rolle. I fremtidige studier ville det derfor være interessant at undersøge hvilke forhold i og omkring skolen, der kan være medvirkende til hhv. at forbedre eller forringe det disciplinære klima i klasseværelserne.

Litteratur

- Andersen, L & Breck, J. (2011, 16. november). *370.000 ufaglærte job er forsvundet på 20 år*. Politiken. Retrieved from politiken.dk.
- Bay, H. (2011). Udviklingen i matematik målt i de danske PISA rapporter. *Nationaløkonomisk Tidsskrift*, 149(2011), 126–135.
- den Brok, P., Brekelmans, M., & Wubbels, T. (2004). Interpersonal Teacher Behaviour and Student Outcomes. *School Effectiveness and School Improvement*, 15(3-4), 407–442.
- Brophy, J. (1988). Research Linking Teacher Behavior to Student Achievement: Potential Implications for Instruction of Chapter 1 Students. *Educational Psychologist*, 23(3), 235.
- Bruner, J. S. (1970). *Om å lære*. Oslo: Dreyer.
- Csikszentmihalyi, M. (2005). *Optimaloplevelsens psykologi*. Virum: Dansk psykologisk forlag.
- Dowell, D., & Neal, J. (1982). A selective review of the validity of student ratings of teachings. *The Journal of Higher Education*, 53(1), 51–62.
- Ebmeier, H., Jenkins, R., & Crawford, G. (1991). The predictive validity of student evaluations in the identification of meritorious teachers. *Journal of Personnel Evaluation in Education*, 4(4), 341–357.
- Egelund, N. & Hansen, K. F. (1997). *Urolige børn i folkeskolens almindelige klasser – en kvantitativ og kvalitativ undersøgelse af urolige elever i folkeskolens almindelige klasser*. København: Undervisningsministeriets forlag.
- Egelund, N., Nielsen, C. P., & Rangvid, B. C. (2011). *PISA Etnisk 2009 – Etniske og danske unges resultater i PISA 2009*. København: AKF.
- Fisher, W. C., Berliner, D. C., Filby, N. N., Marliave, R., Cahen, L. S., & Dishaw, M. M. (1981). Teaching behaviors, academic learning time, and student achievement: An overview. *The Journal of classroom interaction*, 17(1), 2–15.
- Frempong, G., Ma, X., & Mensah, J. (2012). Access to Postsecondary Education: Can Schools Compensate for Socioeconomic Disadvantage? *Higher Education*, 63(1), 19–32.
- Goh, S., & Fraser, B. (1998). Teacher interpersonal behaviour, classroom environment and student outcomes in primary mathematics in Singapore. *Learning Environments Research*, (January), 199–229.
- Greenwald, a G. (1997). Validity concerns and usefulness of student ratings of instruction. *The American psychologist*, 52(11), 1182–6.
- Hansen, E. J. & Andersen, B. H. (2000) *Et sociologisk værktøj - Introduktion til den kvantitative metode*. København: Hans Reitzels Forlag.
- Hansen, E. J. (2003). *Uddannelsessystemerne i sociologisk perspektiv*. Kbh.: HR
- Hermansen, M. (Red.). (2007). *Skolens gode og onde cirkler – en empirisk forløbsundersøgelse i 4. og 5. klasse af elevstandpunkt i holdninger til uro, disciplin og læring*. Frederiksbjerg: Forlaget Samfundslitteratur.

- Kunter, M., & Baumert, J. (2007). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, 9(3), 231–251.
- Lassen, S., Steele, M., & Sailor, W. (2006). The relationship of school • wide Positive Behavior Support to academic achievement in an urban middle school. *Psychology in the Schools*, 43(6).
- Legewie, J., & DiPrete, T. a. (2012). School Context and the Gender Gap in Educational Achievement. *American Sociological Review*, 77(3), 463–485.
- Luiselli, J. K., Putnam, R. F., Handler, M. W., & Feinberg, A. B. (2005). Whole-school positive behaviour support: effects on student discipline problems and academic performance. *Educational Psychology*, 25(2-3), 183–198.
- Lüdtke, O., Trautwein, U., Kunter, M., & Baumert, J. (2007). Reliability and agreement of student ratings of the classroom environment: A reanalysis of TIMSS data. *Learning Environments Research*, 9(3), 215–230.
- Marsh, H. W. (1987). Students' evaluations of University teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11(3), 253–388.
- Mejding, J. (Red.). (2004). *PISA 2003 – Danske unge i en international sammenligning*. København: Danmarks Pædagogiske Universitets Forlag.
- Nordahl, T., Mausestagen, S., & Kostøl, A. (2009). *Skoler med liten og stor forekomst av atferdsproblemer. En kvantitativ og kvalitativ analyse av forskjeller og likheter mellom skolene*. Elverum.
- OECD. (2005a). School factors related to quality and equity – results from PISA 2000. OECD Publishing.
- OECD (2005b). *PISA 2003 Technical Report*. OECD Publishing.
- OECD (2009). *PISA Data analysis manual. SAS – second edition*. OECD Publishing.
- OECD (2012). *PISA 2009 Technical Report*. OECD Publishing.
- Olsen, R. V. (2003). Student and teacher behavior. I S. Lie, P. Linnakylä, & A. Roe, *Northern lights on PISA. ... in the Nordic countries in PISA 2000*. OECD Publishing.
- Pedersen, O. K. (2011). *Konkurrencestaten*. København: Gyldendal Akademisk.
- Peterson, K. D., & Stevens, D. (1988). Student reports for school teacher evaluation. *Journal of Personnel Evaluation in Education*, 2(1), 19–31.
- Rangvid, B. (2003). Educational Peer Effects - Quantile Regression Evidence from Denmark with PISA2000 data. *Do Schools Matter?*, (45).
- Scheerens, J. (2005). *Review of school and instructional effectiveness research*. Ikke-publiceret artikel skrevet i forbindelse med 'the EFA Global Monitoring Report 2005'.
- Teodorovic, J. (2009). Educational effectiveness: Key findings. *Zbornik Instituta za pedagoska istrazivanja*, 41(2), 297–314.

- Vygotsky, L. S. (1982). *Om barnets psykiske udvikling: en artikelsamling*. Kbh: Nyt Nordisk Forlag.
- Väljörvi, J., Kupari, P., Linnakylä, P., Reinikainen, P., Sulkunen, S., Törnroos, J., & Arffman, I. (2007). *The Finnish success in Pisa-and some reasons behind it: Pisa 2003*. 2.
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). *Toward a Knowledge Base for School Learning*. *Review of Educational Research* (Vol. 63, pp. 249–294).

Bilag 1 - Hvor tit oplever den enkelte elev at der er larm og uro. Grupperet efter skole.

Skole id: 00004	Frekvens	Procent
Hver time	1	4,4
De fleste timer	7	30,4
Nogle timer	14	60,9
Aldrig eller næsten aldrig	1	4,4
Total	23	100

Skole id: 00103	Frekvens	Procent
Hver time	3	13,0
De fleste timer	5	21,7
Nogle timer	12	52,2
Aldrig eller næsten aldrig	3	13,0
Total	23	100

Skole id: 00092	Frekvens	Procent
Hver time	2	9,5
De fleste timer	3	14,3
Nogle timer	13	61,9
Aldrig eller næsten aldrig	3	13,3
Total	21	100

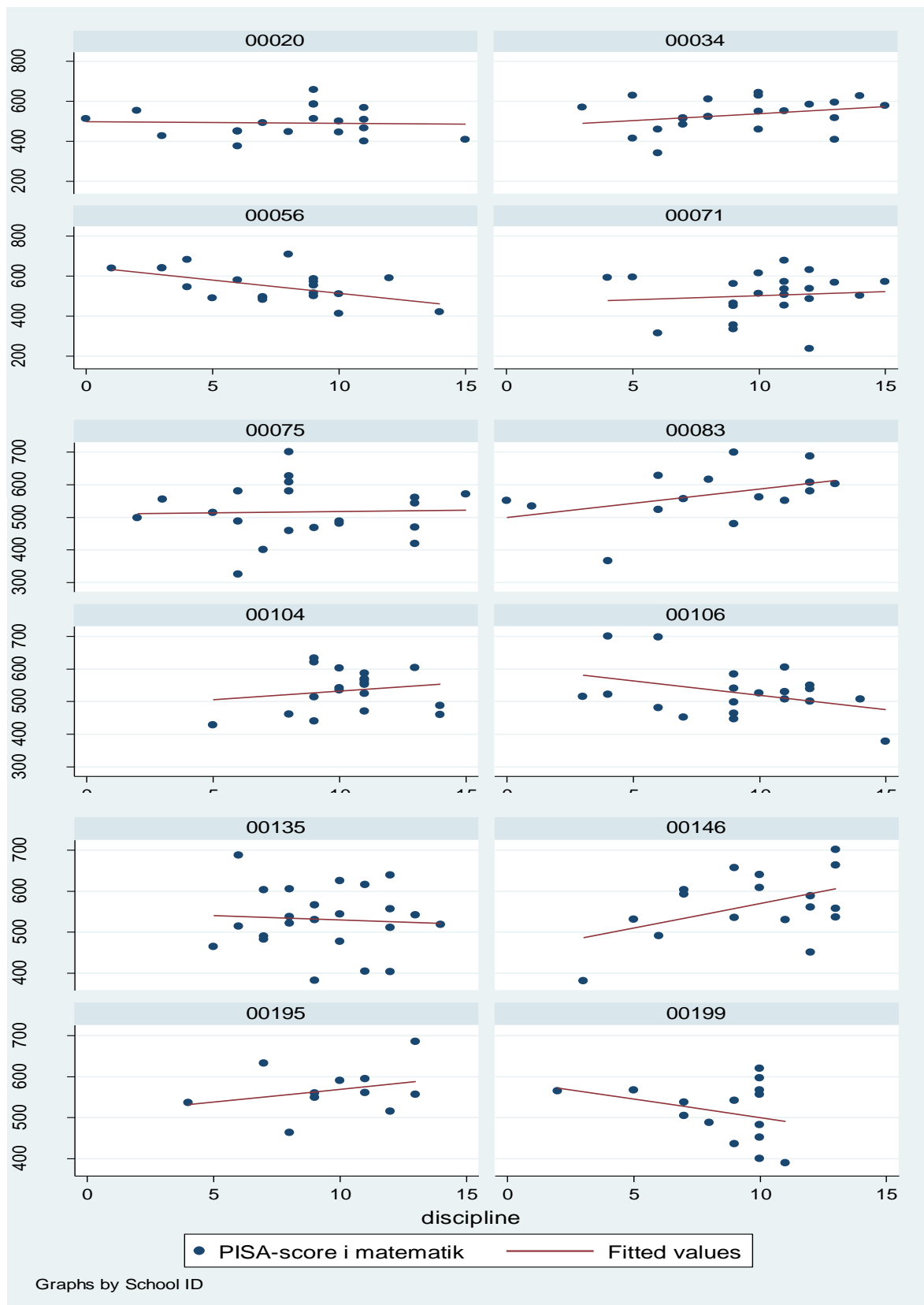
Skole id: 00053	Frekvens	Procent
Hver time	11	45,8
De fleste timer	5	20,8
Nogle timer	8	33,3
Aldrig eller næsten aldrig	0	0
Total	24	100

Skole id: 00202	Frekvens	Procent
Hver time	4	15,4
De fleste timer	11	42,3
Nogle timer	11	42,3
Aldrig eller næsten aldrig	0	0
Total	26	100

Skole id: 00182	Frekvens	Procent
Hver time	2	9,5
De fleste timer	1	4,8
Nogle timer	13	61,9
Aldrig eller næsten aldrig	5	23,8
Total	21	100

Bilag 2 – Sammenhæng mellem oplevet disciplin på individniveau og PISA-score på de enkelte skoler

Graferne viser sammenhængen mellem oplevet disciplin på individniveau og PISA-score på tilfældigt udvalgte skoler.



Tabel til bilag 2. *Korrelationskoefficient for sammenhængen mellem den enkelte elevs oplevelse af det disciplinære klima i klasseværelset og PISA-score. For en række tilfældigt udvalgte skoler.*

Skoleid	Korrelationskoefficient
00020	-0,036
00034	0,277
00056	-0,521
00071	0,101
00075	0,034
00083	0,435
00104	0,176
00106	-0,381
00135	-0,068
00146	0,466
00195	0,287
00199	-0,311

