

Helpt competitief beoordelen mannelijke studenten weer op de rails? Een overzicht van experimenteel bewijs

Sander Onderstal

In de laatste decennia presteren mannen gemiddeld slechter dan vrouwen in alle lagen van het onderwijs. In dit artikel onderzoek ik of een verhoogde competitieve druk op tentamens mannelijke studenten in het hoger onderwijs weer op de rails kan helpen. De literatuur geeft geen eenduidig antwoord op deze vraag. Verschillende empirische onderzoeken laten zien dat mannen relatief goed presteren op competitieve tentamens. In een veldexperiment op de Universiteit van Amsterdam wordt echter alleen binnen een kleine subgroep een geslachtseffect gevonden.

1 Inleiding

De afgelopen decennia lijkt er een onoverbrugbare kloof te zijn ontstaan tussen de prestaties van mannen en vrouwen in het onderwijs. De gemiddelde man presenteert systematisch slechter dan de gemiddelde vrouw in alle lagen van het onderwijs (zie bijvoorbeeld Salvi del Pero en Bytchkova 2013; Guiso et al. 2008). Populair-wetenschappelijke boeken over dit fenomeen gaan als warme broodjes over de toonbank, gepusht door deprimerende titels als *Boys Adrift* (Sax 2007), *The War Against Boys* (Hoff Sommers 2000) en *The End of Men* (Rosin 2012). De literatuur biedt een waaier aan verklaringen voor de geslachtskloof, variërend van biologisch bepaalde verschillen in leerstijlen tussen de geslachten tot onderwijssystemen die beter geënt zijn op vrouwen dan op mannen. In dit artikel zoom ik in op een andere mogelijke verklaring: Een gebrek aan prikkels voor mannen om hun best te doen op tentamens.

Meer specifiek geeft dit artikel een overzicht van de literatuur over hoe de seksen verschillend reageren op concurrentie tijdens toetsmomenten in het hoger onderwijs. Experimenten in andere contexten laten zien dat mannen beter gaan presenteren naarmate de omgeving competitiever wordt, terwijl de prestaties van vrouwen niet afhangen van de mate van concurrentie.¹ Zo maten Gneezy en Rustichini (2004) hoe hard 9- en 10-jaar oude jongens en meisjes renden over een parcours van 40 meter. Alle kinderen legden eerst het parcours alleen af. In een twee-

¹ Zie Gneezy et al. (2003) voor een vroege bijdrage en Niederle and Vesterlund (2011) en Croson en Gneezy (2009) voor overzichten van de literatuur

de ronde liepen sommige kinderen in paren en anderen nog een keer alleen. Gneezy en Rustichini observeerden dat jongens in de tweede ronde snellere tijden neerzetten als ze met iemand anders renden, terwijl het voor de meisjes niet uitmaakte of ze alleen liepen of tegen een ander. Als de resultaten van dit experiment kunnen worden geëxtrapoleerd naar het hoger onderwijs dan zouden mannen hun achterstand op vrouwen in het onderwijs kunnen inlopen als toetsmomenten competitiever worden omdat jongens dan harder hun best gaan doen terwijl vrouwen even hard blijven werken.

In de komende paragrafen passeren geslachtseffecten van competitiviteit op drie verschillende toetsmomenten de revue. Ik vat daarin inzichten uit de empirische literatuur samen, waarbij ik zoveel mogelijk gebruik maak van gerandomiseerde veldexperimenten. In paragraaf 2 concentreer ik me op selectie aan de poort door universiteiten. Voor zover ik weet, zijn er geen veldexperimenten gedraaid op dit gebied zodat ik me baseer op resultaten van niet-experimentele veldstudies. Paragraaf 3 bevat een samenvatting van studies waarbij onderzoekers concurrentie op kleine tussentoetsen varieerden. In paragraaf 4 beschouw ik een veldexperiment waarin mijn collega's en ik een competitieve en niet-competitieve beoordeling van een tentamen met elkaar vergeleken. Paragraaf 5 geeft een korte conclusie.

2 Selectie aan de poort

Wereldwijd selecteren veel universiteiten studenten op basis van toelatingsexamens. Onderzoekers kunnen gebruik maken van variatie tussen verschillende universiteiten, tussen verschillende tentamens of over de tijd om te meten in hoeverre genderverschillen in tentamenresultaten afhangen van de competitiviteit van het toelatingsexamen. In deze paragraaf bespreek ik achtereenvolgens een studie van Jurajda en Munich (2011) naar toelatingsexamens op Tsjechische faculteiten en het werk van Ors et al. (2013) die toelating tot een populaire Parijse businessschool bestuderen. Morin (2015) analyseert het effect van strengere selectie aan de poort op studieprestaties aan de universiteit van Toronto.

Jurajda en Munich (2011) onderzoeken geslachtsverschillen in toelatingsexamens op universiteiten in Tsjechië. Ze maken daarbij gebruik van resultaten van een landelijke toets die middelbare-schoolstudenten in 1999 hebben afgelegd. Volgens de auteurs speelden de scores op de landelijke toets geen enkele rol in de selectie van studenten door de universiteiten. Van alle studenten die deelnamen aan de toets observeren de auteurs bij welke universiteiten de studenten zich aanmeldden alsmede of ze er al dan niet werden toegelaten.

De auteurs identificeren genderverschillen op de volgende wijze. Voor elke faculteit nemen ze als maat voor competitiviteit de fractie van aangemelde studenten die werden toegelaten. De competitiviteit varieerde aanzienlijk tussen de 116 faculteiten, met een gemiddelde van 0,31 en een standaardafwijking van 0,21. Vervolgens vergelijken de auteurs toelatingkanssen voor mannen en vrouwen met verge-

lijkbare resultaten op de landelijke toets en checken ze in hoeverre de verschillen verklaard kunnen worden door de competitiviteit van de faculteit.

De auteurs observeren dat vrouwen slechter presteren dan mannen in de 25% meest competitieve faculteiten en vinden geen duidelijke geslachtsverschillen voor minder competitieve faculteiten. Dit resultaat blijkt robuust als de analyse wordt beperkt tot faculteiten die het meest gebalanceerd zijn in termen van geslacht (de faculteiten met de 25% hoogste fracties mannen of vrouwen worden niet meegenomen). Dezelfde conclusie geldt voor analyses met alleen de studenten die als beste of als slechtste uit de bus kwamen in de landelijke toets. Omdat scholieren zich typisch aanmeldden voor verschillende faculteiten konden de auteurs ook binnen individuen reacties op competitiviteit identificeren. Ook dan blijken vrouwen onder hoge competitieve druk relatief slechter te presteren dan mannen.

Ors et al. (2013) doen een vergelijkbare veldstudie naar het effect van competitiviteit op tentamens op geslachtsverschillen. Ze bestuderen examens voor toelating tot een Masterprogramma in Management aan de Ecole des Hautes Etudes Commerciales (HEC) in Parijs, de meest prestigieuze businessschool van Frankrijk. HEC selecteert bijzonder streng: Slechts 10-15% van de kandidaten wordt toegelaten tot het programma. De auteurs vergelijken de prestaties van studenten op het toelatingsexamen met hun resultaten voor een landelijk examen voor middelbare scholieren en de eerstejaarsvakken van de opleiding. De auteurs beargumenteren dat zowel het middelbare-schoolexamen als de eerstejaarstentamens veel minder competitief zijn dan het toelatingsexamen: De slagingspercentages op de middelbare school zijn 89% en slechts 1% van de eerstejaarsstudenten haalt het eerste jaar niet. Volgens de onderzoekers is het daarom mogelijk te analyseren hoe geslachtsverschillen afhangen van de mate van concurrentie op de tentamens.

Net als Jurajda en Munich (2011) vinden Ors et al. een aanzienlijke effect van competitiviteit op de prestatieverschillen tussen mannen en vrouwen. Waar binnen de cohorten van HEC-kandidaten vrouwen significant beter presteren dan mannen op het eindexamen doen ze het significant slechter dan mannen op het toelatingsexamen. Bovendien observeren de auteurs dat in de groep van de studenten die wordt toegelaten, vrouwen significant beter presteren dan mannen in het Masterprogramma op de niet-kwantitatieve vakken. Er wordt geen statistisch significant genderverschil waargenomen in prestaties op de kwantitatieve vakken. Al met al laten de resultaten zien dat mannen relatief goed presteren in een omgeving met een hoge competitieve druk waar relatieve prestaties ten opzichte van andere participanten essentieel zijn.

Morin (2015) maakt gebruik van een dubbel cohort middelbare scholieren in de Canadese provincie Ontario om de effecten van een toename van concurrentie tussen studenten aan de universiteit van Toronto te meten. Een nieuwe opzet van het middelbareschoolsysteem in Ontario rond de millenniumwisseling verkortte de middelbareschoolduur van vijf naar vier jaar. Als gevolg hiervan studeerden in juni 2003 twee cohorten middelbare scholieren af. Omdat de capaciteit van de universiteiten lang niet navenant toenam, groeide de concurrentie om plaatsing op de uni-

versiteiten aanzienlijk. Daardoor nam volgens de auteur de gemiddelde kwaliteit van de toegelaten studenten toe ten opzichte van andere jaren. Het onderwijsreglement van de universiteit van Toronto schrijft bovendien voor dat docenten alle tentamens relatief beoordelen volgens een vaste, belvormige, verdeling. Vanwege de toegenomen kwaliteit van het cohort zouden studenten dus beter moeten presteren dan andere cohorten om dezelfde cijfers te behalen. Deze toename in concurrentie op de tentamens maakt het mogelijk om geslachtsverschillen te identificeren wat betreft de gevoeligheid voor competitiviteit.

Morin's resultaten resoneren met die van de eerder besproken onderzoeken van Jurajda en Munich (2011) en Ors et al. (2013). Hij observeert dat de prestaties van mannen in een competitieve omgeving ongeveer één procentpunt meer verbeteren dan die van vrouwen. Dit effect, hoewel klein, blijkt consistent over de gehele studieduur van vier jaar. Hij vindt vergelijkbare effecten op het aantal behaalde studiepunten, met name voor relatief zwakke studenten, en op de fractie studenten die geen studieovertraging oplopen.

Hoewel alle drie de studies vinden dat competitiviteit op tentamens correleert met relatief betere prestaties van mannen ten opzichte van vrouwen, mag niet automatisch geconcludeerd worden dat mannen baat hebben bij toenemende competitiviteit op een tentamen. Het geobserveerde effect kan namelijk voortkomen uit twee andere kanalen. Ten eerste is de keuze van een student om al dan niet deel te nemen aan een competitief examen mogelijk endogeen. Zelfselectie zou zo de correlatie tussen competitiviteit en de genderverschillen verklaren, bijvoorbeeld als relatief veel getalenteerde vrouwen afhaken bij het vooruitzicht op competitieve examens. Het is niet helder of de auteurs voldoende kunnen corrigeren voor studentkarakteristieken om zelfselectie uit te sluiten. Een tweede mogelijk kanaal is dat de beoordelaars van de tentamens discrimineren bij het corrigeren van de tentamens zodat de onderzoekers niet kunnen uitsluiten dat de beoordeling van de tentamens plaatsvindt op andere karakteristieken dan de kwaliteit van studenten. Een gerandomiseerd experiment kan een zelfselectieverklaring uitsluiten omdat de toewijzing van studenten aan een al dan niet competitief examen exogeen is. Beoordelen op zuiver objectieve maatstaven kan het kanaal van discriminatie bij de beoordeling dichtgooien. De onderzoeken die ik in de volgende paragrafen bespreek, voldoen geheel of gedeeltelijk aan beide voorwaarden.

3 Deeltoetsen

In deze paragraaf bespreek ik twee onderzoeken naar hoe de mate van competitiviteit uitwerkt op de prestaties van mannelijke en vrouwelijke studenten op kleine deeltoetsen die werden afgenomen op Italiaanse universiteiten.

Bigoni et al. (2011) bestuderen de prestaties van studenten in een econometrie-cursus in een bachelor-programma in Management aan de Universiteit van Bologna. Studenten moesten voor de cursus vijf tussentoetsen maken waarmee ze een gelimiteerd aantal bonuspunten konden verdienen voor het eindtentamen. Elke

tussentoets bestond uit vijf meerkeuzevragen zodat de beoordeling ervan niet kon worden vervuild door geslachtsdiscriminatie. De onderzoekers verdeelden de 131 deelnemers aan het experiment willekeurig over drie groepen: een controlegroep en twee treatmentgroepen. In de controlegroep werden alle vijf de tussentoetsen ‘absoluut’ becijferd, in de zin dat het cijfer van een student onafhankelijk was van de resultaten van andere studenten. In de twee treatments werden de eerste en de laatste tussentoets ook absoluut becijferd. De drie tussenliggende toetsen werden anders beoordeeld. Elke student in deze treatments werd gekoppeld aan een willekeurige andere student in hetzelfde treatment. In het competitieve treatment kreeg een student twee extra punten voor haar tussentoets als zij het strikt beter deed dan haar partner. In het tweede, coöperatieve, treatment ontving een student één extra punt als haar partner minstens een bepaalde score behaalde op de tussentoets. In de controlegroep kregen alle studenten één punt extra voor de tussentoetsen in de tussenliggende weken om ze niet te benadelen ten opzichte van de studenten in de treatments.

De onderzoekers voorspellen op basis van een speltheoretisch model dat de studenten zich (zwak) meer zullen inspannen in het competitieve treatment dan in de controlegroep die op haar beurt weer (zwak) beter zal presteren dan de coöperatieve treatmentgroep. De data zijn consistent met het voorspelde patroon. Belangrijker: in de competitieve treatmentgroep presteren mannelijke studenten statistisch significant beter dan de mannen in de controlegroep en het coöperatieve treatment. De effecten zijn ook economisch significant: de geschatte gemiddelde toename in inspanning voor mannen is in de orde van 40%. Voor vrouwen worden geen significante effecten waargenomen. De resultaten suggereren dat competitiviteit op toetsen mannen kan aanzetten beter te presteren om zo de achterstand met hun vrouwelijke medestudenten goed te maken.

De Paola et al. (2015) passen een soortgelijke competitieve beoordeling toe als Bigoni et al. (2011) op een tussentoets van een economietentamen waaraan 720 bachelorstudenten deelnamen. De onderzoekers deelden de studenten in in paren van gelijke ingeschatte bekwaamheid. Een student kreeg 5 bonuspunten als zij beter presteerde dan haar partner en slechts 2 als ze het slechter deed op de tussentoets. In tegenstelling tot Bigoni et al. laten De Paola et al. studenten zelf kiezen of ze meedoen aan de competitieve toets. Op deze wijze konden de auteurs onderzoeken of er sprake is van zelfselectie en in hoeverre die geslachtsafhankelijk is, een zorg in de veldstudies die ik in paragraaf 2 besprak. Overigens was het kostbaar niet deel te nemen aan de tussentoets: Studenten die niet meededen kregen geen bonuspunt en waren dus slechter af dan de ‘verliezers’ op de tussentoets.

De auteurs vinden dat vrouwen en mannen gelijkwaardig presteerden op zowel de tussentoets als het (niet-competitieve) eindtentamen. Ze observeren evenmin geslachtsverschillen in de keuze om deel te nemen aan de competitieve deelttoets. De resultaten lijken er dus op te wijzen dat meer concurrentie op het tentamen niet altijd in het voordeel uitpakt voor mannen.

4 Tentamens

Samen met Eszter Czibor, Randolph Sloof en Mirjam van Praag heb ik in de herfst van 2013 een veldexperiment gedraaid waarin we twee manieren om tentamens te becijferen met elkaar vergeleken (Czibor et al. 2014). Het experiment was onderdeel van het vak Economie van Markten en Organisaties dat verplicht is voor verschillende tracks in het BSc-programma Economie en Bedrijfskunde van de Universiteit van Amsterdam. We verdeelden ruim 500 studenten op basis van gestratificeerde randomisatie willekeurig over twee groepen. Alle studenten moesten dezelfde tussentoets maken die bestond uit 30 louter meerkeuzevragen. De wijze van beoordelen verschilde echter tussen de twee groepen. Eén groep werd beoordeeld op een absolute schaal (*absolute grading*) en de andere op een relatieve schaal (*relative grading*; *grading on a curve*).² Onderstaande tabel geeft aan hoe we scores op het tentamen vertaalden in het tentamencijfer.

Absolute schaal (score)	Cijfer	Relatieve schaal (beste)
29 – 30	10	0% – 1%
27 – 28	9	1% – 5%
24 – 26	8	5% – 15%
22 – 23	7	15% – 36%
19 – 21	6	36% – 64%
17 – 18	5	64% – 85%
14 – 16	4	85% – 95%
12 – 13	3	95% – 99%
0 – 11	2	99% – 100%

In het Nederlandse hoger onderwijs is het gebruikelijk een absolute schaal te gebruiken. De empirische literatuur die ik tot nu toe besproken heb, lijkt de hypothese te rechtvaardigen dat mannen de huidige achterstand ten opzichte van vrouwen kunnen verkleinen als studenten worden beoordeeld op een relatieve schaal in plaats van op een absolute schaal: In tegenstelling tot een absolute schaal is een relatieve schaal competitief in de zin dat een student beter moet presteren dan een gegeven fractie van de studenten om een bepaald cijfer te halen.

Ons experiment verschilt op ten minste vier manieren van bovenstaande empirische studies. Ten eerste maken wij, net als Bigoni et al. (2011), gebruik van een gerandomiseerd veldexperiment. Hiermee kunnen we zelfselectie in de al dan niet competitieve omgeving uitsluiten. In de tweede plaats volgen we Bigoni et al. (2011) door louter meerkeuzevragen te gebruiken. Hiermee sluiten we uit dat geslachtsdiscriminatie de beoordeling van de tentamens kan beïnvloeden. Ten derde is het belang van het tentamen relatief groot: 50% van het eindcijfer werd bepaald door het resultaat op de tussentoets. Hiermee staat er voor studenten meer op het

² Voor het eindtentamen draaiden we de beoordelingswijze van de groepen om.

spel dan in de tussentoetsen bij Bigoni et al. (2011) en De Paola et al. (2015). Ook gebruiken we een andere manier van competitief becijferen dan Bigoni et al. (2011) en De Paola et al. (2015) die de beste student in elk willekeurig samengesteld koppel bonuspunten geven. We hebben daar goede redenen voor. Voor zover we weten, is het in de onderwijspraktijk niet gebruikelijk om bonuspunten te geven binnen willekeurig samengestelde koppels. Bovendien hangt het cijfer van een student veel af van de medestudent waaraan zij gekoppeld wordt. Onze beoordeling op relatieve schaal is minder gevoelig voor het toeval omdat een student met al haar medestudenten wordt vergeleken. Bovendien wordt zij veel toegepast in de praktijk, bijvoorbeeld op universiteiten in Angelsaksische landen.

Op geaggregeerd niveau laten onze resultaten noch voor mannen noch voor vrouwen duidelijke verschillen zien in termen van tentamenvoorbereiding of -resultaten onder de twee manieren van beoordelen. Deze bevindingen zijn in strijd met de meeste resultaten uit de boven besproken literatuur waar mannen relatief beter presteren naarmate tentamens competitiever zijn. Een mogelijke verklaring is de zesjescultuur die op Nederlandse universiteiten lijkt te heersen: Studenten willen hun diploma met minimale inspanning halen en voelen zich niet geprikkeld hoge cijfers te scoren.³ Studenten zijn dus weinig geïnteresseerd in hun cijfers zolang ze hun vakken maar halen. Onze data ondersteunen deze verklaring: als we de analyse beperken tot een subgroep van studenten ‘op het randje’ van zakken of slagen dan nemen we wel geslachtsverschillen waar. Mannelijke studenten presteren significant beter onder relatieve dan onder absolute becijfering, terwijl we voor vrouwen geen significant effect waarnemen.

5 Samenvatting en conclusie

In dit artikel heb ik de vraag gesteld of mannen en vrouwen anders reageren op toenemende concurrentie op tentamens. De literatuur geeft geen eenduidig antwoord. Verschillende niet-experimentele empirische onderzoeken laten zien dat mannen relatief goed presteren ten opzichte van vrouwen op competitieve toelatingstentamens op Tsjechische faculteiten (Jurajda en Munich, 2011), een prestigieuze Parijse businessschool (Ors et al., 2013) en de universiteit van Toronto (Morin, 2015). In experimenten met tussentoetsen op Italiaanse universiteiten vinden Bigoni et al. (2011) dat mannen relatief goed presteren op competitieve tussentoetsen terwijl De Paola et al. (2015) geen geslachtseffecten waarnemen. Mijn co-auteurs en ik vinden alleen voor een subgroep een duidelijk geslachtseffect op een tentamen op een Nederlandse universiteit.

Deze resultaten hebben consequenties voor de vraag hoe mannelijke studenten geholpen kunnen worden hun achterstand op vrouwen in het hoger onderwijs in te ha-

³ Zie bijvoorbeeld Leuven et al. (2010) en Inspectie van het Onderwijs (2014).

len.⁴ Overstappen van een absolute schaal op een relatieve schaal lijkt niet de heilige graal om mannelijk studenten weer op de rails te krijgen, althans niet in de zescultuur zoals die lijkt te heersen in het Nederlandse hoger onderwijs.⁵ Vervolgonderzoek moet uitwijzen in hoeverre onze bevindingen inderdaad specifiek zijn voor Nederland. Ons experiment zou bijvoorbeeld kunnen worden herhaald in een omgeving waar studenten aantoonbaar meer gemotiveerd zijn dan in Nederland.

Auteur

Sander Onderstal (e-mail: Onderstal@uva.nl) is universitair hoofddocent aan de Universiteit van Amsterdam.

Literatuur

- Bigoni, M., M. Fort, M. Nardotto en T. Reggiani, 2011, Teams or tournaments? A field experiment on cooperation and competition among university students, IZA Discussion Paper 5844.
- Croson, R. en U. Gneezy, 2009, Gender Differences in Preferences, *Journal of Economic Literature*, vol. 47(2): 448-74.
- Czibor, E., Onderstal, S., Sloof, R., en Van Praag, M., 2014, Does relative grading help male students? Evidence from a field experiment in the classroom. IZA Discussion Papers 8429.
- De Paola, M., F. Gioia en V. Scoppa, 2015, Are Females Scared of Competing with Males? Results from a Field Experiment, *Economics of Education Review*, te verschijnen.
- Gneezy, U. en A. Rustichini, 2004, Gender and Competition at a Young Age, *American Economic Review*, vol. 94 (2): 377-81.
- Gneezy, U., M. Niederle en A. Rustichini, 2003, Performance In Competitive Environments: Gender Differences, *The Quarterly Journal of Economics*, vol. 118(3): 1049-1074.
- Guiso, L., F. Monte, P. Sapienza en L. Zingales, 2008, Culture, Gender, and Math, *Science*, vol. 320(5880): 1164-65.
- Hoff Sommers, C., 2000, *The War against Boys*, Simon & Schuster.
- Inspectie van het Onderwijs, 2014, De staat van het onderwijs. Onderwijsverslag 2012/2013. Inspectie van het Onderwijs, April.

⁴ Ik ga hier niet in op de vraag of het wenselijk is dat de geslachtskloof in het onderwijs wordt gedicht. In de arbeidsmarkt lijken mannen nog steeds een grote voorsprong te hebben op vrouwen zowel wat salaris als wat carrièrekansen betreft.

⁵ Daarnaast kleven er nadelen aan het gebruik van een relatieve schaal. Zo zal in een klas waarin alle studenten de stof beheersen een deel zakken. Ook geeft een relatieve schaal studenten een prikkel om samen te spannen tegen de docent en weinig tijd in het vak te steken. Tot slot hebben studenten een prikkel om hun medestudenten niet met de stof te helpen of hen zelfs te saboteren om zo hun eigen concurrentiepositie te verstevigen. Deze bezwaren zijn overigens stuk voor stuk relevanter in kleine groepen dan in grote.

- Jurajda, S. en D. Munich, 2011, Gender Gap in Performance under Competitive Pressure: Admissions to Czech Universities, *American Economic Review*, vol. 101(3): 514-18.
- Leuven, E., H. Oosterbeek en B. van der Klaauw, 2010 The Effect of Financial Rewards on Students' Achievement: Evidence from a Randomized Experiment, *Journal of the European Economic Association*, vol. 8(6): 1243-65.
- Morin, L.P., 2015, Do men and women respond differently to competition? Evidence from a major education reform. *Journal of Labor Economics*, vol. 33(2): 443-91.
- Niederle, M. en L. Vesterlund, 2011, Gender and competition, *Annual Review of Economics*, vol. 3: 601-30.
- Ors, E., F. Palomino en E. Peyrache, 2013, Performance Gender Gap: Does Competition Matter?, *Journal of Labor Economics*, vol. 31(3): 443-99.
- Rosin, H., 2012, *The End of Men: And the Rise of Women*, Riverhead Books, Penguin Group.
- Salvi del Pero, A. en A. Bytchkova, 2013, *A Bird's Eye View of Gender Differences in Education in OECD Countries*, OECD Social, Employment and Migration Working Papers.
- Sax, L., 2007, *Boys Adrift: The Five Factors Driving the Growing Epidemic of Unmotivated Boys and Underachieving Young Men*, Basic Books.