

Opzet, resultaten en implicaties van het PISA-toetsprogramma op het gebied van natuurwetenschappen

Harrie Eijkelfhof

Freudenthal Instituut voor Didactiek van Wiskunde en Natuurwetenschappen, Universiteit Utrecht

Inleiding

In 1997 zijn door de OECD de voorbereidingen gestart van het PISA-programma: het Programme for International Student Assessment (zie: <http://www.pisa.oecd.org>). PISA richt zich op het meten van resultaten van onderwijssystemen op basis van de prestaties van leerlingen op de voor de economie belangrijke gebieden van lezen, wiskunde en natuurwetenschappen. Anders dan in het door de International Association for the Evaluation of Educational Achievement (IEA) al veel langer georganiseerde TIMSS-programma (zie: <http://timss.bc.edu>) gaat het hierbij niet om de prestaties op de grootste gemene deler van de bestaande onderwijscurricula, maar om de resultaten van leerlingen met het oog op hun toekomstig functioneren in de samenleving. Vandaar dat gekozen is voor de populatie van 15-jarigen, de leeftijd waarop in veel landen de leerplichtige leeftijd eindigt. Aan PISA doen alle 34 OECD-landen mee, maar inmiddels hebben circa 30 andere landen besloten om ook mee te doen. Om enige diepgang te bereiken en overbelasting van leerlingen en scholen te voorkomen worden de testen om de drie jaar afgenomen, waarbij steeds een van de kennisgebieden centraal staat. In 2000 lag de nadruk op lezen, in 2003 op wiskunde, in 2006 op natuurwetenschappen, in 2009 weer op lezen, enzovoort.

Onderliggende raamwerken

De vragen die in de PISA-testen aan bod komen zijn gebaseerd op raamwerken voor elk van de drie kennisgebieden. Voor de natuurwetenschappen is dat het Scientific Literacy Framework (OECD, 2009). Centraal in dit raamwerk staan drie competenties: (1) het identificeren van natuurwetenschappelijke issues, (2) het verklaren van verschijnselen en (3) het hanteren van evidentie-redeneringen. Onder (1) vallen bijvoorbeeld het herkennen van typisch natuurwetenschappelijke vraagstukken aan een aantal kenmerken. Bij (2) gaat het om het beschrijven, verklaren en voorspellen van verschijnselen. Tot (3) behoren het interpreteren van evidentie, het identificeren van aannames en het reflecteren op maatschappelijke implicaties.

In deze drie competenties komen een aantal componenten samen: (a) natuurwetenschappelijke kennis gebruiken, (b) de attitude ten opzichte van natuurwetenschappen en (c) het kunnen toepassen van redeneringen in levensreële situaties waarin natuurwetenschappelijke kennis een rol speelt. Bij (a) wordt onderscheid gemaakt tussen kennis *van* en kennis *over* natuurwetenschappen. Kennis *van* natuurwetenschappen betreft inhoud op de gebieden natuur- en scheikunde, biologie en aarde & ruimte. Bij kennis *over* natuurwetenschappen gaat het om inzicht in het functioneren van natuurwetenschappen, dat wil zeggen procedurele en epistemische aspecten zoals doelen, methoden, interpreteren van data, onzekerheden en soorten verklaringen – aspecten die in Nederland centraal staan in het schoolvak Algemene Natuurwetenschappen (ANW). Tot attitudes (b) worden gerekend de interesse in natuurwetenschappen en het onderwijs hierin, en meningen over het belang van deze disciplines voor de samenleving. Bij (c) gaat het om contexten op het gebied van gezondheid, natuurlijke hulpbronnen, milieu, risico's en veiligheid, en grenzen van kennisontwikkeling op het gebied van natuurwetenschappen en technologie. De contexten betreffen de persoonlijke omgeving, de regionale en landelijke situatie, en wereldwijde contexten.

Procedure

Circa vier jaar voor de eigenlijke afname van de toetsen beginnen de voorbereidingen, in volgorde: het gunnen door de OECD van deelprojecten, het herzien van de raamwerken, het ontwerpen en inhoudelijk en cultureel beoordelen van vragen, het vertalen van de vragen, het samenstellen van 13 verschillende vragenboekjes (zodat leerlingen niet alle vragen hoeven te beantwoorden maar wel alle vragen even vaak aan bod komen), het selecteren van scholen en 15-jarige leerlingen (geen klassen) en het benoemen van schoolcoördinatoren. De afname vindt in maart plaats: er is twee uur uitgetrokken voor de cognitieve test en een half uur voor een leerlingenvragenlijst over persoonlijke achtergrond, studiegewoontes, attitudes en motivatie. Door een van de schoolleiders wordt een vragenlijst ingevuld over bijvoorbeeld samenstelling van de leerlingenpopulatie en de kenmerken van de leeromgeving.

Daarna worden de data verzameld door scoring van de antwoorden en worden publicaties voorbereid. De resultaten worden publiek in december van het daaropvolgende jaar. Zo kunnen we de PISA-resultaten van 2012 (met wiskunde als hoofdgebied) verwachten in december 2013. De publicaties omvatten een serie internationale delen (bijvoorbeeld zes over de 2009-resultaten met betrekking tot kennis, socio-economische status, attitude en motivatie, school- en overheidsbeleid, meerjarige trends, digitale vaardigheden) (OECD, 2010) en een nationaal rapport (de Knecht-van

Eekelen, Gille & van Rijn, 2007; Gille et al., 2010). Daarnaast worden alle data openbaar gemaakt voor gebruik door wetenschappers. Vervolgens verschijnen, naast korte berichten in de media, tal van artikelen in wetenschappelijke tijdschriften over deelaspecten.

De vragen

De meeste PISA-vragen worden in een context gesteld, bijvoorbeeld aan de hand van krantenberichten of een kort beschreven situatie. Een deel van de vragen is van het type ja/nee of meerkeuze. Andere vragen zijn open van karakter, waarbij redeneringen gevraagd worden. Het scoren van de meerkeuzevragen is eenvoudig, bij de open vragen is dat lastiger. Hiertoe zijn uitgebreide scoringsvoorschriften geformuleerd, die precies aangeven hoe antwoorden moeten worden beoordeeld. De meeste vragen blijven geheim om in de toekomst weer te kunnen worden gebruikt. Enkele vragen worden vrijgegeven en zijn te vinden in de rapporten (OECD, 2007; OECD, 2010; de Knecht-van Eekelen, 2007) en op de PISA-website (zie <http://www.pisa.oecd.org>).

Enkele resultaten

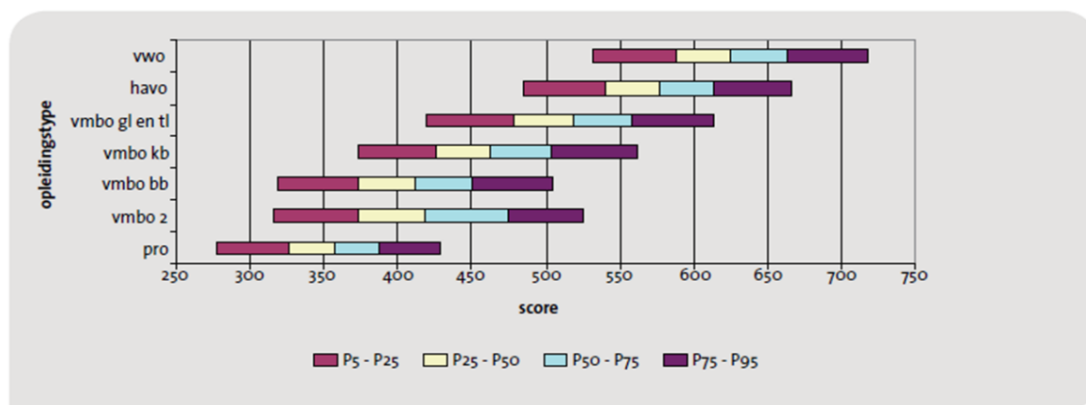
De resultaten worden op veel verschillende wijzen gepresenteerd, bijvoorbeeld naar landengemiddelde voor de kennisvragen, naar soort competenties, naar kennisgebieden, in verschillen tussen jongens en meisjes en naar bereikte niveaus. Hieronder enkele voorbeelden.

	Shanghai China	Finland	Singapore	Japan	Nederland	Duitsland	België	Verenigde Staten	Italië	Indonesië
Natuurwetenschappen	575 (1)	554 (2)	542 (4)	539 (5)	522 (11)	520 (13)	507 (21)	502 (23)	489 (35)	383 (54)
Wiskunde	600 (1)	541 (6)	562 (2)	529 (9)	526 (11)	513 (16)	515 (14)	487 (31)	483 (35)	371 (61)
Lezen	556 (1)	536 (3)	526 (5)	520 (8)	508 (10)	497 (20)	506 (11)	500 (17)	486 (29)	402 (57)

Figuur 1. Cognitieve landenresultaten PISA 2009 voor elk van de kennisgebieden, tussen haakjes de plaats in de ranglijst, $N = 65$ (bron: OECD, 2010).

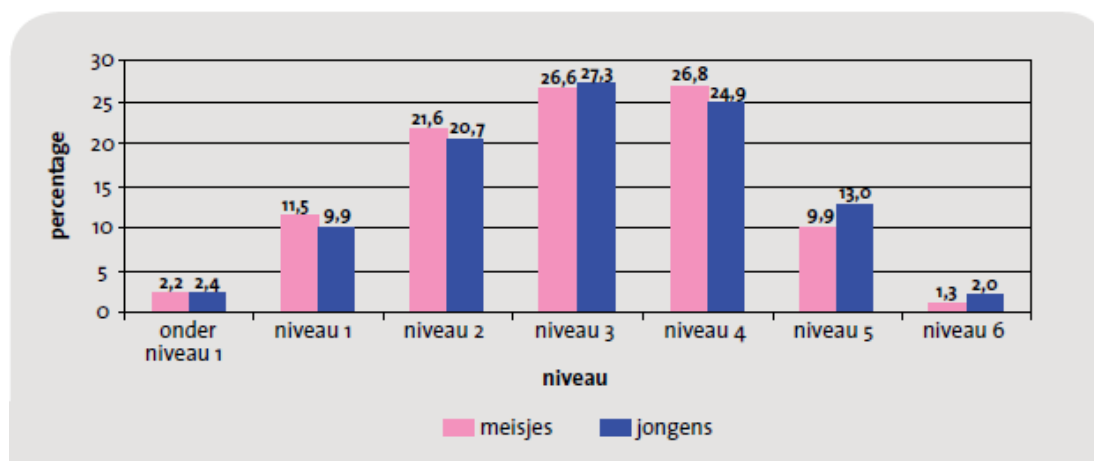
Uit figuur 1 wordt duidelijk dat Nederlandse leerlingen het op elk van de drie terreinen redelijk goed doen in vergelijking met de ons omringende landen, al blijven we ver achter bij een aantal Oost-Aziatische landen. Opvallend is dat de positie op de ranglijst voor elk van de drie gebieden voor de meeste landen vrijwel gelijk is. De VS scoort relatief laag bij wiskunde. Overigens moet worden opgemerkt dat de posities niet altijd statistisch significant verschillen. Zo kun je ook stellen dat Nederland op het gebied van natuurwetenschappen op een gedeelde 7^e plaats staat met Nieuw Zeeland, Canada, Estland, Australië, Chinese Taipei, Duitsland, Liechtenstein, Zwitserland, Verenigd Koninkrijk en Slovenië.

Figuur 2 laat zien dat de cognitieve resultaten in Nederland gemiddeld genomen overeenkomen met wat men zou verwachten naar schoolniveau, al is er flink wat overlap: zo doet een kwart van de vmbo-TL-leerlingen het een stuk beter dan een deel van de havo- en vwo-leerlingen.



Figuur 2. Cognitieve resultaten PISA 2006 Natuurwetenschappen voor Nederland naar schoolsoort en percentielen (bron: de Knecht-van Eekelen, Gille & van Rijn, 2007).

Uit figuur 3 blijkt dat de jongens het in Nederland iets beter deden dan de meisjes. De aantallen op het hoogste en laagste niveau zijn klein.



Figuur 3. Cognitieve resultaten PISA 2006 Natuurwetenschappen voor Nederland naar niveaus voor jongens en meisjes (bron: CITO, 2007). Voor de scores bij de verschillende niveaus geldt: niveau 6: > 669; niveau 5: 607-669; niveau 4: 545-607; niveau 3: 482-545; niveau 2: 420-482; niveau 1: 358-420; niveau <1: <358.

Bezien we in figuur 4 de percentages op de twee hoogste niveaus, dan zien we dat Nederlandse toppers het een stuk minder doen dan in Shanghai, waarbij het lijkt (vergelijk met figuur 3) alsof de gemiddelde populatie 15-jarigen in Shanghai zich op het Nederlandse vwo-niveau bevindt. Overigens zien we in dit opzicht weinig verschillen met onze beide buurlanden.

	Shanghai China	Finland	Singapore	Japan	Nederland	Duitsland	België	Verenigde Staten	Italië	Indonesië
Natuurwetenschappen	24,3	18,7	19,9	17,0	12,7	12,8	10,1	9,2	5,8	0,0
Wiskunde	50,4	21,6	35,6	20,9	19,8	17,8	20,4	9,9	9,0	0,1
Lezen	19,4	14,5	15,7	13,4	9,8	7,6	11,2	9,9	5,8	0,0

Figuur 4. Cognitieve resultaten PISA 2009 voor elk van de kennisgebieden: percentages leerlingen op de twee hoogste niveaus (5 en 6) (bron: OECD, 2010).

Nederlandse leerlingen scoren cognitief significant hoger dan het OECD-gemiddelde, maar dat is anders waar het attitude betreft (zie figuur 5). Nederlandse leerlingen lijken gemiddeld minder interesse te hebben in het leren op het gebied van de natuurwetenschappen, maar ze zien wel het belang van dit kennisgebied in voor de samenleving.

	Nederland	OECD
Plezier in het leren over NW	46	64
Graag lezen over NW	41	50
Leuk om NW vraagstukken op te lossen	33	43
Leuk om nieuwe NW kennis te verwerven	56	67
Geïnteresseerd in het leren van NW	46	63
NW en technologie van belang voor de economie	80	80
NW is erg relevant voor mij	46	57
Mogelijkheden om later NW kennis te gebruiken	58	59
NW en technologie zijn belangrijk voor de samenleving	84	75

Figuur 5. Attitude-resultaten PISA 2006 Natuurwetenschappen (NW) voor Nederland in vergelijking met het OECD-gemiddelde, in percentages positieve houding (bron: OECD, 2007).

Interpretatie van de resultaten

Het is opvallend hoe de resultaten van PISA verschillend worden geïnterpreteerd. Sommigen concluderen dat het Nederlandse onderwijs het vrij goed doet, al is natuurlijk verbetering mogelijk. Anderen trekken de conclusie dat het Nederlandse onderwijs onder de maat blijft, ofwel omdat we dalen op de ranglijst, ofwel omdat we niet tot de top-5 in de wereld behoren – onze politieke ambitie. In dit licht moeten enkele kanttekeningen worden geplaatst bij de resultaten:

- 1 PISA meet niet de resultaten van ons onderwijs naar de maatstaf van alle doelen die we voor ogen hebben in het onderwijs. PISA richt zich op scientific literacy voor iedereen, niet op voorbereiding op een exacte vervolgstudie. Bovendien is PISA niet afgestemd op de curricula in de diverse landen maar op doelen die wenselijk worden geacht met het oog op het latere leven.
- 2 Het is niet altijd duidelijk of de PISA-resultaten vooral iets zeggen over het onderwijs of ook kunnen worden toegeschreven aan de aandacht voor wetenschap in de media, met name omdat de PISA-vragen weinig schools zijn en veel te maken hebben met de context van het dagelijks leven.
- 3 Het was oorspronkelijk niet de bedoeling veel waarde te hechten aan de 'league tables'. Het was veel meer de bedoeling een gemeenschappelijke maatstaf te hanteren om knelpunten te signaleren en binnenlandse ontwikkelingen te monitoren. In de media en in politieke kringen hebben de ranglijstjes meer aandacht gekregen dan beoogd.
- 4 Zoals hierboven al gemeld zijn de posities op de ranglijsten statistisch lang niet altijd significant. Groepen landen vergelijken is meer verantwoord dan afzonderlijke landen.
- 5 De positie op de ranglijst is niet alleen afhankelijk van nationale inspanningen, maar ook van het al dan niet meedoen van andere landen en van de progressie die andere landen maken. Het gemiddelde is steeds 500.
- 6 Het is niet duidelijk met welke motivatie leerlingen participeren in de PISA-studie. In sommige landen staat onderwijs in hoog aanzien en wordt overdag en 's avonds hard gewerkt om goede resultaten te behalen. Bovendien is in sommige landen het collectieve gevoel belangrijker dan het individuele. Het is denkbaar dat in dit soort landen de vragen met grotere inzet worden beantwoord.
- 7 Bij de interpretatie van de attituderesultaten kunnen we ons afvragen of hier de cultuur geen invloed heeft op de manier waarop de vragen worden beantwoord. Zo is bekend dat in sommige landen een kritische houding niet wordt gewaardeerd, wat kan leiden tot veel positieve antwoorden. In Nederland is wellicht het tegenovergestelde het geval: jongeren zijn vaak kritisch en steken hun mening niet onder stoelen of banken.

Dit alles neemt niet weg dat de PISA-resultaten elk land wel een spiegel voorhouden: voldoet het onderwijs aan de verwachtingen die we hebben ten aanzien van algemene vorming, moeten we de verwachtingen of het onderwijsbeleid bijstellen, kunnen we leren van andere landen of willen we bepaalde kenmerken van de onderwijscultuur elders juist niet overnemen?

Beleidsimplicaties

Ook in andere landen krijgen de PISA-resultaten aandacht. In het Verenigd Koninkrijk verschenen zorgelijke berichten in de media (Gardner, 2010; Mortimore, 2008). Ook op Europees niveau wordt aandacht geschonken aan de PISA-scores (Figazollo, 2009; Grek, 2009). Daarnaast zien we de reactie van bezorgde ministers op tegenvallende resultaten door meer te investeren in curriculum-herziening (Duitsland) of veel nadruk te leggen op toetsing (Denemarken en Duitsland) (Dolin & Krogh, 2010; Kauertz, Neumann & Fischer, 2010). Politieke implicaties hebben de PISA-resultaten ook in bijvoorbeeld Nieuw Zeeland (Baker & Jones, 2005), Japan (Takayama, 2008) en de Verenigde Staten (Anderson, Milford & Ross, 2010).

Een van de meest opvallende gevolgen van het PISA-programma is dat veel belangstelling is ontstaan om over de grens te kijken. Enerzijds is dit zichtbaar door analyses van resultaten van landen die cultureel en economisch weinig van elkaar verschillen (Basl, 2011; Kordes et al., 2010; Nentwig et al., 2009). Anderzijds zijn landen als Finland al jaren erg in trek bij politici en beleidsmedewerkers om de geheimen achter de goede resultaten van Finse leerlingen te ontdekken (Lavonen & Laaksonen, 2009; Sahlberg, 2011).

Een andere trend is om onderwijsstelsels op veel factoren met elkaar te vergelijken: TIMSS- en PISA-scores maar ook investeringen, rendementen van voortgezet en hoger onderwijs, aantal schooljaren, leraar-leerlingratio en salarissen in het onderwijs. Recent heeft Pearson een rapport gepubliceerd waarin 40 landen met elkaar zijn vergeleken (Pearson, 2012). Als toplanden komen Zuid-Korea en Finland uit de bus. Interessant is dat beide landen volgens het rapport in veel opzichten van elkaar verschillen. Zuid-Korea is sterk op toetsing en reproductie gericht, kent een star curriculum, leerlingen bezoeken massaal privé huiswerkklassen in de avond (crammer schools (hagwans)), de klassen zijn groot en de salarissen van docenten zijn relatief hoog (1,64 maal het gemiddelde bruto loon). Finland: leerlingen maken een late start in onderwijs, het aantal lessen per jaar is relatief gering, huiswerk is niet gebruikelijk, er wordt weinig frontaal lesgegeven, het onderwijs is gericht op toepassen van kennis, de klassen zijn klein en de salarissen van docenten zijn relatief laag (0,87 maal het gemiddelde bruto loon). Maar er zijn ook overeenkomsten: het belang van de lerarenopleiding wordt hoog geacht en goed onderwijs staat in hoog aanzien, wordt min of meer als een morele verplichting gezien.

Wat betekent dit voor Nederland? In ons land is de overheidsreactie in het Akkieplan Beter Presteren (MOCW, 2011) om meer aandacht te geven aan talentontwikkeling en afspraken te maken met scholen om beter te gaan scoren op de PISA-rondes van 2015 en volgende, zelfs gekwantificeerd:

“Ik wil dat Nederland met leesvaardigheid, rekenen en ‘science’ de komende PISA-metingen (in 2015 en 2018) systematisch vooruitgang boekt. Daarbij gaat het om de *gemiddelde* score, en om het verhogen van met name de scores op de *hoogste* vaardigheidsniveaus.

- Bij *wiskunde* wordt gestreefd naar een vooruitgang van 5 punten per meting van 526 in 2009 naar 536 in 2015, en naar 541 in 2018,
 - bij *lezen* een vooruitgang met 4 punten per meting van 508 in 2009 naar 516 in 2015, en naar 520 in 2018,
 - bij *science* een verhoging van 2 punten per meting: van 522 in 2009 naar 526 in 2015 en naar 528 in 2018.”
- (p. 2)

Het is de vraag of dergelijke precieze benchmarks realistisch zijn omdat de resultaten niet alleen afhangen van de scores van Nederlandse leerlingen, maar ook van die van leerlingen in andere landen. Bovendien blijken kleine verschillen in scores tussen landen met vergelijkbare resultaten als Nederland (in PISA 2006 voor science scores 522-532 en in PISA 2009 voor science scores 522-529) statistisch niet significant (OECD, 2007, pp 56-57; 2010, p. 152).

Wanneer de Nederlandse overheid de PISA-resultaten serieus wil nemen, dan is het wel raadzaam goed notie te nemen van de overeenkomsten en verschillen tussen de raamwerken van PISA en de curricula voor de onderbouw van het voortgezet onderwijs. Een dergelijke vergelijking is voor science gemaakt aan de hand van het raamwerk 2006 en de resultaten van PISA 2009 (Eijkelhof, Kordes & Savelsbergh, 2013; Kordes et al., 2010; Kuiper et al., 2010). Daaruit blijkt een verwantschap wat betreft de aandacht voor contexten in het onderwijs. Ook wordt duidelijk dat vmbo-leerlingen met name lager scoren bij de open vragen. In dit soort vragen worden redeneringen gevraagd, en daarbij is noodzakelijkerwijs ook enige lees- en schrijfvaardigheid vereist.

De PISA-raamwerken zijn echter ook in ontwikkeling. Zo zijn er aanwijzingen dat in het Science Literacy Framework van 2015 meer aandacht zal zijn voor epistemische kenmerken van kennis in de natuurwetenschappen (de Jong, 2012), hetgeen verwantschap vertoont met de ontwikkeling van nieuwe science standards in de Verenigde Staten (NRC, 2012). Dit soort aspecten dreigt in het Nederlandse onderwijs juist minder aandacht te krijgen door verzwakking van de positie van het schoolvak ANW.

Wanneer de OECD-landen de nieuwe raamwerken onderschrijven, verdient het aanbeveling om bijvoorbeeld in EU-verband nader te analyseren hoe het onderwijs beter afgestemd zou kunnen worden op wat kennelijk als kennis, vaardigheden en attitudes belangrijk wordt gevonden voor het leven en werken in de toekomstige samenleving. Dat vereist waarschijnlijk geen grote stelselwijzigingen, maar wel focus en uitwisseling van good practices. Een mooie taak voor vakdidactici.

Literatuur

- Anderson, J.O., Milford, T., & Ross, S.P. (2010). An opportunity to better understand schooling: The growing presence of PISA in the Americas. *International Journal of Science and Mathematics Education*, 8(3), 453-473.
- Baker, R., & Jones, A. (2005). How can international studies such as the International Mathematics and Science Study and the Programme for International Student Assessment be used to inform practice, policy and future research in science education in New Zealand? *International Journal of Science Education*, 27(2), 145-157.
- Basl, J. (2011). Effect of school on interest in natural sciences: A comparison of the Czech Republic, Germany, Finland, and Norway based on PISA 2006. *International Journal of Science Education*, 33(1), 145-157.
- Eijkelhof, H.M.C., Kordes, J.H., & Savelsbergh, E.R. (2013). Implications of PISA outcomes for science curriculum reform in the Netherlands. In: Prenzel, M., Kobarg, M., Schöps, K., & Rönnebeck, S. (Eds), *Research in the context of the Programme for International Student Assessment* (pp. 7-21). Berlin/Heidelberg: Springer Verlag.
- Dolin, J., & Krogh, L.B. (2010). The relevance and consequences of PISA science in a Danish context. *International Journal of Science and Mathematics Education*, 8(3), 565-592.
- Figazollo, L. (2009). *Impact of PISA 2006 on the education policy debate*. Brussels: Education International.
- Gardner, R. (2010). *British schools slump in global league table*. The Independent, December 10.
<http://www.independent.co.uk/news/education/education-news/british-schools-slump-in-global-league-table-2154048.html> (retrieved January 9, 2013)
- Gille, E., Loijens, C., Noijons, J., & Zwitter, R. (2010). *Resultaten Pisa-2009. Praktische kennis en vaardigheden van 15-jarigen*. Arnhem: Cito.
http://www.cito.nl/onderzoek%20en%20wetenschap/onderzoek/deelname_onderzoek/pisa/resultaten.aspx (retrieved January 11, 2013)
- Grek, S. (2009). Governing by numbers: the PISA ‘effect’ in Europe. *Journal of Educational Policy*, 24(1), 23-37.
- Jong, J.H.A.L. de (2012). Framework for PISA 2015. What 15-year-olds should be able to do. Paper presented at the 4th Annual Conference of Educational Research Center, March 24-25, 2012, Broumana, Lebanon.
<http://www.educationalrc.org/conference2012/PDF/Prof%20John%20de%20Jong.pdf> (retrieved January 11, 2013)
- Kauertz, A., Neumann, K., & Fischer, H.E. (2010). From PISA to educational standards: The impact of large-scale assessments on science education in Germany. *International Journal of Science and Mathematics Education*, 8(3), 545-563.
- Knecht-van Eekelen, A. de, Gille, E., & Rijn, P. van (2007). *Resultaten Pisa-2006. Praktische kennis en vaardigheden van 15-jarigen*. Arnhem: Cito.
http://www.cito.nl/onderzoek%20en%20wetenschap/onderzoek/deelname_onderzoek/pisa/resultaten.aspx

- (retrieved January 11, 2013)
- Kordes, J., Smeets, P., Wouda, J., Marsman, M., Groen, M. van, Eijkelhof, H., & Savelsbergh, E. (2010). *Nederlandse 15-jarigen en de natuurwetenschappen. Hun kennis, vaardigheden en visie volgens PISA*. Arnhem: Cito.
http://www.cito.nl/onderzoek%20en%20wetenschap/onderzoek/deelname_onderzoek/pisa/resultaten.aspx (retrieved January 11, 2013)
- Kuiper, W., Hoeven, M. van der, Folmer, E., Graft, M. van, & Akker, J. van den (2010). *Leerplankundige analyse van PISA-trends*. Enschede: SLO.
<http://www.slo.nl/downloads/2010/leerplankundige-analyse-van-pisa-trends.pdf/> (retrieved January 11, 2013)
- Lavonen, J., & Laaksonen, S. (2009). Context of teaching and learning school science in Finland: Reflections on PISA 2006 results. *Journal of Research in Science Teaching*, 46(8), 922-944.
- MOCW (2011). *Actieplan 'Beter Presteren' vo: opbrengstgericht en ambitieus. Het beste uit leerlingen halen*. Kamerstuk 23 mei 2011. Den Haag: Ministerie van Onderwijs Cultuur en Wetenschappen.
<http://www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2011/05/23/actieplan-vo-beter-presteren.html> (retrieved January 11, 2013)
- Mortimore, P. (2008). *A league table to worry us all*. The Guardian, January 8.
<http://www.guardian.co.uk/education/2008/jan/08/schools.uk1> (retrieved January 9, 2013)
- Nentwig, P., Rönnebeck, S., Schöps, K., Rumann, S., & Carstensen, C. (2009). Performance and levels of contextualization in a selection of OECD countries in PISA 2006. *Journal of Research in Science Teaching*, 46(8), 897-908.
- NRC (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington D.C.: National Academies Press.
http://www.nap.edu/catalog.php?record_id=13165 (retrieved January 11, 2013)
- OECD (2007). *PISA 2006. Science competencies for tomorrow's world. Volume 1 – Analysis*. Paris: OECD.
- OECD (2009). *PISA 2009 assessment framework. Key competences in reading, mathematics and science*. Paris: OECD.
- OECD (2010). *PISA 2009 results: What students know and can do. Student performance in reading, mathematics and science. Volume 1*. Paris: OECD.
- Pearson (2012). *The learning curve. Lessons in country performance in education*. London: Pearson.
<http://www.thelearningcurve.pearson.com> (retrieved January 11, 2013)
- Sahlberg, P. (2011). *Finnish lessons. What can the world learn from educational change in Finland?* New York: Teachers College Press.
- Takayama, K. (2008). The politics of international league tables: PISA in Japan's achievement crisis debate. *Comparative Education*, 44(4), 387-407.