

Grote kwesties

U hebt er uiteraard weinig van gemerkt, neem ik aan. Maar de voorbije twee jaar leefde ik in een toestand die men wellicht het meest adequaat omschrijft als geestelijke ontreddeering. De nood aan enige opfrissing van hun normale verblijfplaats noopte mij ertoe om het merendeel van mijn boeken naar veiliger oorden te verkassen. Tijdelijk uiteraard. Maar vooreerst duurde 'opfrissen' langer dan voorzien. Daarna ontbrak de tijd om het hele boeltje terug te brengen en vooral ook de tijd om de talrijke dozen uit te pakken en hun inhoud terug een ordentelijke plaats te geven. Vandaar dus: twee jaar lang een ongelukkig gevoel bij het vullen van Boekenplanken of het schrijven over Grote Kwesties omwille van geen mogelijkheid om terug te vallen op mijn vertrouwde bibliotheek voor inspiratie of achtergrondinformatie. Nu wil ik eerlijkheidshalve wel toegeven dat het gevoel van ontreddeering vooral toesloeg tijdens het uitpakken van de dozen. Niet dat ik boeken terugvond waarvan ik niet meer wist dat ik ze ooit kocht. (Moeilijk te geloven volgens sommigen van mijn kennissen, maar toch waar. Trouwens, ik kocht in de loop van die twee jaar kennelijk maar een boek of drie dat nog ergens in een doos zat – allemaal romans – en ik beschouw dat eerder als een vorm van 'revealed consistency of preferences', een bewijs van consistentie in mijn koopgedrag, dan als een falen van mijn geheugen.) Maar de confrontatie met wat in de voorbije periode buiten mijn blikveld bleef, leidde toch tot een soort hernieuwde kennismaking via af en toe bladeren in publicaties die je lang terug hebt gelezen. Net zoals je bij beroepsmatige afspraken met goede vrienden die je een tijd uit het oog verloor eerst even moet bijpraten voor je tot de orde van de dag kan over gaan.

Ik vertel u dit alles omdat deze ervaring er mij nog meer van overtuigd heeft dat wij – en als u wil mag u dit 'wij' interpreteren als een koninklijk meervoud, zodat u zelf niet bent geïmpliceerd – veel te

weinig artikels en boeken herlezen. Zeker als het gaat om boeken die in de intellectuele geschiedenis van een bepaald domein of die van jezelf een belangrijke rol speelden. Verplichten is niet mijn ding, maar op dit punt ben ik bereid eventueel een uitzondering te maken. Bijscholen of levenslang leren hoeft niet noodzakelijk in te houden dat men zichzelf enkel beperkt tot nieuwe dingen. Ook herhaaldelijk terug bekijken en herinterpreteren van het historisch erfgoed hoort daartoe. Vandaar: een pleidooi voor grasduinen in het verleden en een Grote Kwesties met een portie nostalgie als vertrekpunt.

Over onderwijs als mobiliteitskanaal toen en nu

Neem nu de klassieker van Christopher Jencks over ongelijkheid – C. Jencks, *Inequality. A Reassessment of the Effect of Family and Schooling in America*, New York, Basic Books. Ik vermoed dat iedereen die een beetje begaan is met ongelijkheid – wetenschappelijk of beleidsmatig – nog wel van het bestaan van dit boek afweet. Maar leest men het ook nog? En heeft men weet van het debat daar rond? Ik vrees van niet. Zelfs voor ernstig werkende lieden in mijn omgeving blijkt dat het geval. Wetenschappelijke werken gelijken in dat opzicht vrij sterk op werknemers. Eens een bepaalde leeftijd gepasseerd begint men hun nut in twijfel te trekken.

Ik kocht en las *Inequality* vermoedelijk in 1973 of 1974, kort na afstuderen. Of ik het daarna nog een tweede keer heb doorgenomen, betwijfel ik. Geconsulteerd misschien, maar zeker niet helemaal en in detail terug doorgewerkt. Bij het doorbladeren een paar weken geleden realiseerde ik me dat ik alleszins een aantal opmerkelijke details vergeeten was. Bijvoorbeeld dat het boek door de auteur

gepresenteerd werd als een antwoord op het in 1966 verschenen Coleman-report en als een heranalyse van de gegevens waarop dat rapport was gebaseerd. Of dat het een uitgebreide appendix bevat die de techniek van pad-modellen uitlegt, iets waar ik in mijn opleiding nauwelijks zicht op had gekregen, en een opmerkelijke en zeer genuanceerde appendix over de erfelijkheid van IQ-scores. Ik herinnerde mij trouwens ook niet meer dat het boek niet het werk van één persoon was. Christopher Jencks prijkt weliswaar als auteur op het kaft en voor de interpretatie die aan de resultaten wordt gegeven tekent hij alleen verantwoordelijk, maar de binnenbladzijde vermeldt nog een aantal medeauteurs die het gros van de analyses hebben uitgevoerd – waaronder Mary-Jo Bane en Herb Gintis, namen die mij toen wellicht niet zijn opgevallen maar vandaag in de literatuur over ongelijkheid op eigen kracht het nodige respect opwekken.

Inequality verscheen in 1972 in de Verenigde Staten en kreeg snel grote weerklank ook buiten de VS – een Engelse uitgave verscheen in 1973. Het boek kreeg erg veel weerklank binnen en buiten de academische wereld. Het beroerde de publieke discussie veel meer dan, bijvoorbeeld, het boek dat een andere socioloog (Boudon) een jaar later (in 1974) zou publiceren, ook al poneerde dit boek een gelijkaardige stelling als Jencks en veroorzaakte het omwille van de modelmatige aanpak wél degelijk ook de nodige deining binnen de beroepsgroep. Waarom *Inequality* aanleiding was tot zoveel controverse en discussie is gemakkelijk aan te geven.

De hoofdstukken behandelen achtereenvolgens de ongelijkheid binnen en tussen de scholen, de ongelijkheid inzake verworven cognitieve vaardigheden, de ongelijkheid op het vlak van non-cognitieve karakteristieken, de ongelijkheid in onderwijsresultaten, de ongelijkheid inzake beroepsstatus, de ongelijkheid inzake inkomen en de ongelijkheid inzake jobsatisfactie. Jencks gebruikte het uitgebreid empirisch materiaal dat in het boek werd verwerkt om een belangrijke overtuiging van een groot deel van de politieke en maatschappelijke commentatoren naar de schroothoop te verwijzen. De feiten wijzen uit, schrijft hij, dat het streven naar grotere gelijkheid van kansen, in het bijzonder van onderwijskansen, niet leidt tot meer sociale mobiliteit en gro-

tere gelijkheid op volwassen leeftijd. “We have seen that educational opportunities, cognitive skills, educational credentials, occupational status, income and job satisfaction are all unequally distributed. We have not, however, been very successful in explaining most of these inequalities. The association between one variety of inequality and another is usually weak, which means that equalizing one thing is unlikely to have much effect on the degree of inequality in other areas.” (p. 253)

Of deze interpretatie volledig strookt met de resultaten van de analyses – waarvoor de meeste commentatoren erg lovende woorden gebruikten – en of deze interpretatie op zich wel methodologisch koesjer was, werd in de vele symposia en de talloze besprekingen vaak betwijfeld. Het is zelfs de vraag of de medeauteurs die op de eerste binnenbladzijde van het boek prijken het allemaal en volledig eens waren met Jencks' interpretatie. Maar over één zaak is geen twijfel mogelijk. In verschillende opzichten vormt het boek werkelijk een ‘discursief evenement’. James Coleman karakteriseerde *Inequality* in de eerste zin van zijn bespreking als “an example of a new activity” (*American Journal of Sociology*, vol. 78, 1973, p. 1523). Zowel omwille van de poging om op basis van een groot aantal bestaande gegevensbronnen via secundaire analyses een ‘macro-sociaal’ probleem te behandelen als omwille van de poging om op deze analyses de nodige beleidsadviezen te enten. Thomas Pettigrew begon zijn bespreking (*idem*, p. 1527) met de opmerking dat er feitelijk drie verschillende ‘Jencks reports’ bestonden: het eerste bestaat uit ingenieuze, ambitieuze en interessante empirische analyses, het tweede is de provocatieve, goed geschreven maar erg discutabele en dikwijls erg ver van de eigenlijke resultaten verwijderde interpretatie die Jencks aan deze analyses geeft en het derde is de verzameling van vulgariseerde versies in de media, versies die slechts een vage relatie hebben met de tekst zelf en virtueel niks met de feitelijke resultaten te maken hebben. Meerdere commentatoren vragen zich hardop af waarom Jencks enkele honderden bladzijden aan empirische analyses nodig heeft om in het laatste niet-empirische en niet-analytische hoofdstuk uit te komen bij een stelling die met het voorgaande nauwelijks te maken heeft; namelijk dat een onrechtstreekse aanpak (via de school) niet werkt om de inkomensongelijkheid aan te pakken en dat daarom een directe aanpak nodig is.

Een belangrijke reden voor de heisa die het boek ontketende was uiteraard dat het tegen de haren in streek van wie 'onderwijs' beschouwde als het kanaal bij uitstek voor de verbetering van de sociale mobiliteit en het bannen van de ongelijkheid. Zeker ook omdat men de lieden die deze boodschap brachten moeilijk een conservatief imago kon toedichten. De beleidsvoorstellen voor een rechtstreekse ingreep in de inkomensvorming in het slothoofdstuk typeren zij zelf trouwens als 'socialistisch'. Maar wat mij vooral treft bij het doorbladeren van het boek en het herlezen van een aantal passages is de erg brutale manier waarop Jencks tewerk gaat – Pettigrew spreekt in zijn bespreking niet voor niets over een tekst met “a negative, debunking, almost nihilistic character” (*idem*, p. 1528).

De aanhef van hoofdstuk 2 (over “ongelijkheid in de scholen”) illustreert perfect het procédé dat Jencks ook in de andere hoofdstukken gebruikt.

In de eerste paragraaf typeert hij twee manieren waarop men volgens hem scholen kan bekijken. Sommigen beoordelen de kwaliteit van een school op basis van de langetermijneffecten. Anderen zien scholen als gemeenschappen waarvan ze de kwaliteit eerder beoordelen op basis van het plezier of de interesse waarmee leraren en leerlingen eraan deelnemen. In de tweede paragraaf beschrijft hij in detail wat volgt uit een beoordeling van de school op haar langetermijneffecten. “If you judge schools according to their long-term effects, if you believe these effects are substantial, and if you are an egalitarian, you are likely to feel that everyone should get the same kind of schooling, whether they want it or not. Egalitarians of this persuasion often argue that students who want to drop out of high school should be encouraged or even coerced into returning, because otherwise they will have little chance of earning a good living. They also argue that students who do not want to attend college should be persuaded to do so, for similar reasons. They fight for systems of school finance that provide equal resources in every school, because they believe this is the only way to make the alumni of different schools equal. They demand an end to segregation because they think that this is a crucial step in eliminating the advantage of ‘haves’ over ‘have-nots’, and they oppose both elementary school tracking and distinctive high school curriculums on the

grounds that these arrangements doom certain students to subordinate roles in adult life.” In de derde paragraaf volgt de anticlimax. “The evidence discussed in this book has convinced us, and may convince some readers, that such arguments are misguided. Chapters 3 through 8 argue that differences between schools have rather trivial long-term effects, and that eliminating differences between schools would do almost nothing to make adults more equal. Even eliminating differences in the amount of schooling people would get do relatively little to make adults more equal. If this is true, schools ought to be judged largely by their short-term effects.” (pp. 16-17)

Te lang citaat? Oké, mijn excuses. Maar het fragment illustreert erg goed waarom *Inequality* ook vandaag nog best het lezen waard is. Al was het maar om te beseffen dat de visies waarvan Jencks meent de onhoudbaarheid empirisch te kunnen aantonen ook vandaag nog steeds een belangrijk onderdeel vormen in het debat over de rol en de plaats van het onderwijs. Of om te evalueren in welke mate we dertig jaar later over meer empirisch materiaal en wetenschappelijke kennis beschikbaar hebben om de conclusies van *Inequality* te nuanceren, te weerleggen, aan te vullen of te bevestigen. Bijvoorbeeld om ons ervan te vergewissen of we nu echt (al dan niet) veel meer weten over de ongelijke verdeling van niet-cognitieve karakteristieken en de invloed ervan op arbeidsmarktuitkomsten dan de vier schamele bladzijden die Jencks en zijn medeauteurs hebben kunnen bijeenrapen – “a confession of ignorance... We ... believe that the noncognitive effects of schooling are likely to be more important than the cognitive effects. But we do not know what these noncognitive effects are likely to be” (p. 134).

In “La Psychanalyse: mère et chienne” vertelt Elizabeth Roudinesco dat de Duitse theatermaker Bert Brecht ooit het plan opvatte om een toneelhuis te bouwen in de vorm van een Panopticon. In dit toneelhuis wou Brecht de grote rechtszaken uit de geschiedenis laten opvoeren om zo de toeschouwers individueel te confronteren met de beoordeling van goed en kwaad en de gevoeligheid hiervoor bij de bevolking aan te scherpen. Zelf heb ik het voorstel nooit in het werk van Brecht kunnen terugvinden, maar dat belet niet dat het als inspiratie kan dienen voor over hoe we met ‘discursieve evenementen’

zoals *Inequality* zouden kunnen en misschien moeten omgaan.

Laat ons een procédé zoeken om dit soort boeken collectief te herlezen en het debat dat er toen over werd gevoerd herdoen met inbreng van alle kennis die we sindsdien hebben verzameld. Waar wachten wij op? Misschien leren we er ook beleidsmatig wel meer uit dan uit veel enkel op nieuwe en recente gegevens gebaseerd onderzoek.

Survey, administratieve data of nog iets anders: wat zal het wezen?

Sociale wetenschappers die empirisch materiaal gebruiken voor onderzoeksdoeleinden – en als men beleidsrelevant wil werken dan is dit in de meeste gevallen een sine qua non – vallen meestal terug op gegevens die voortkomen uit surveys. Deze werkwijze is niet zonder problemen. Toevalssteekproeven zijn, zoals de naam zegt, onderhevig aan toeval en representeren de doelgroep dus niet adequaat. Ze kunnen ook onderhevig zijn aan structurele afwijkingen als men niet kan beschikken over een correcte en helemaal volledige lijst van de populatie waarover het onderzoek handelt. Men moet rekening houden met non-response, algemeen of per vraag. En, tenslotte, surveys zijn erg duur, zeker als men mikt op een hoge kwaliteit van het interviewproces en gebruik maakt van face-to-face bevraging. Daarom zoeken onderzoekers naar alternatieve pistes. Ofwel door over te stappen op minder dure vormen van enquëtering, zoals post- of telefonische enquêtes, of een mix daarvan. Ofwel door gebruik te maken van administratieve bestanden, eventueel in combinatie met survey-gegevens.

Deze laatste optie lijkt erg aantrekkelijk omwille van de lage kost en vooral ook omdat administratieve bestanden in principe toegang geven tot populatiegegevens en dus grote aantallen respondenten. Op internationale congressen heb ik al meermaals meegemaakt hoe aanwezigen jaloers toekijken als een van hun (meestal Scandinavische) collega's analyses presenteert op basis van tien-, zometert honderd-duizenden respondenten. Maar ook aan administratieve bestanden zitten potentiële nadelen vast. Een van de belangrijkste is wellicht dat administratieve bestanden meer nog dan surveys onderhevig zijn aan wetgeving op de bescherming van de private le-

venssfeer, zeker als men een koppeling met zelf verzamelde persoonsgegevens zou beogen. Terecht uiteraard. Maar, het zal wel voor iedereen duidelijk zijn, neem ik aan, de rationaliteit van waaruit beschermers van de private levenssfeer redeneren, spoort niet altijd met de vereisten van kwalitatief wetenschappelijk onderzoek. Meer klaarheid in welke van de alternatieven nu betere resultaten oplevert dan andere en welke afwegingen men moet in rekening brengen is dus meer dan wenselijk. Dat een dergelijke zorg duidelijk aanwezig is in de onderzoeksgemeenschap, blijkt uit de steeds grotere regelmaat waarmee de vaktijdschriften bijdragen op dit terrein bevatten.

Recente goede voorbeelden zijn: Annette JÄCKLE, Caroline ROBERTS, Peter LYNN, *Assessing the Effect of Data Collection Mode on Measurement*, Institute for Social and Economic Research (ISER Working Paper 2008-08), University of Essex, 2008, 23 p. en Julia LANE, *Optimizing the Use of Microdata: An Overview of the Issues*, in: *Journal of Official Statistics*, vol. 23 (2007) n° 3, pp. 299-317. De eerste publicatie rapporteert over een experiment dat in het raam van de European Social Survey werd opgezet om na te gaan in welke mate verschillende interviewtechnieken verschillen in resultaten opleveren. De tweede publicatie formuleert een voorstel dat moet toelaten om een balans te zoeken tussen de privacy-vereisten én de eisen betreffende nuttigheid die men kan stellen aan (administratief) verzamelde gegevens.

Een aparte rubriek in dit domein betreft het zogeheten 'valideringsonderzoek', al zijn de mogelijkheden om dit soort analyses uit te voeren niet echt dik gezaaid. Hierbij koppelt men voor een groep respondenten gegevens die men verkrijgt via survey-onderzoek aan gelijkaardige gegevens uit administratieve bestanden. De bedoeling van deze koppeling is om na te gaan in welke mate men de geconstateerde verschillen kan toeschrijven aan de verschillende soorten onzuiverheden die kunnen voorkomen bij gegevens uit survey-onderzoek. Kenmerkend voor deze bijdragen is dat ze uitgaan van de veronderstelling dat de administratieve gegevens correct zijn en dat men de geconstateerde afwijkingen tussen deze bestanden en de gegevens voortgebracht door survey-onderzoek enkel dient toe te schrijven aan meetfouten die optreden tijdens de verzameling van survey-gegevens. Dit uitgangspunt

impliceert dat administratieve gegevens de werkelijkheid weergeven zoals ze is en dat bij hun constructie of bij de koppeling van gegevens uit meerdere bestanden geen enkele fout wordt gemaakt. Wie ooit met administratieve gegevens werkte, weet dat dit een heroïsche veronderstelling is.

In een erg technische artikel onderzoeken Arie Kapteyn en Jelmer Y. Ypma – Arie Kapteyn, Jelmer & Y. Ypma, Measurement Error and Misclassification: A Comparison of Survey and Administrative Data, in: *Journal of Labor Economics*, vol. 25 (2007) n° 3, pp. 513-551 – wat de gevolgen zijn als men deze veronderstelling loslaat. Veronderstel een keer, zo is hun uitgangspunt, dat er een ‘echte’ werkelijkheid is en dat zowel de administratieve bestanden als de survey-gegevens deze realiteit weergeven met een zekere foutenmarge. Wat dan? Minstens rijst dan de vraag welke soort gegevens onder welke omstandigheden dan de voorkeur geniet en of men hierbij rekening moet houden met de methode die men wil gebruiken om een model te schatten. In het eerste gedeelte van hun tekst geven ze een antwoord op deze vragen op basis van een formeel model dat toelaat om verschillende soorten meetfouten te onderscheiden. In een tweede deel illustreren ze hun werkwijze door de survey-gegevens verzameld voor het Zweedse deel van het Europese SHARE-project te confronteren met de administratieve gegevens, waaraan men ze wil koppelen. De conclusie van het artikel is duidelijk. Verlaat men de veronderstelling dat administratieve gegevens de ‘echte werkelijkheid’ volledig correct weergeven dan gaan een aantal conclusies van de bestaande literatuur over dit onderwerp niet helemaal meer op. Meer zelfs, welke schattingsmethode men gebruikt blijkt wel degelijk van belang en de gevolgen verschillen naargelang men de gegevens hanteert als afhankelijke dan wel als onafhankelijke variabelen. Nog belangrijker is echter een opmerking die Kapteyn en Ypma als conclusie neerschrijven. Er zijn uiteraard erg goede redenen om administratieve gegevens te gebruiken: steekproefgrootte, kostenplaatje, kwaliteit van de gegevens, en dergelijke meer. Maar tenzij deze volledig foutloos zijn, vormen ze zeker geen wondermiddel. En men mag hierbij niet vergeten dat meetfouten in deze redenering niet enkel betrekking hebben op mismatches in de koppeling van gegevens met een bepaalde respondent, maar ook – en veel belangrijker want nagenoeg onvermijdelijk – meetfouten die een gevolg

zijn van discrepanties tussen de conventies op basis waarvan de administratieve gegevens worden geconstrueerd en de theoretische concepten die men hanteert. “Our illustrative calculations... suggest that biases resulting from using administrative data as right-hand variables may be very substantial. As always, one has to be careful in modeling the sources and nature of errors and take into account when investigating hypotheses of substantive interest.” (p. 536)

Een goed voorbeeld van de meerwaarde van het combineren van administratieve en survey-gegevens, maar tegelijk ook van de ernstige beperkingen van administratieve gegevens als men aan de empirische vaststellingen ook duidelijke conclusies wil verbinden is: Per Johansson, The importance of employer contacts: Evidence based on selection on observables and internal replication, in: *Labour Economics*, vol. 15 (2008) pp. 350-369. Het uitgangspunt van dit artikel is een studie van de effecten van SWIT, een specifiek opleidingsprogramma georganiseerd door de Federation of Swedish Industries en de Federation of IT-companies. De bedoeling van dit programma was het wegwerken van de bottleneck op deze specifieke arbeidsmarkt door het verhogen van de contacten tussen werkzoekenden en werkgevers via het voorzien van zogeheten ‘gastbedrijven’. Zowel in de bestaande opleidingen (AMV) als in SWIT was werkloosheid en inschrijving bij de publieke arbeidsvoorzieningsdienst een voorwaarde tot deelname. Kwaliteit van de opleidingen was vergelijkbaar want ze werd verzorgd door dezelfde instellingen. Maar een belangrijk verschil was “within the Swit organization, a project leader was responsible for the complete labor market training process (e.g. the selection of individuals to training, advising on labor market training courses, providing a host company, etc.), whereas these tasks were shared by many different employees within the AMV.” Tijdens de opleiding konden de deelnemers werkervaring opdoen bij meerdere werkgevers. Een studie gebaseerd op administratieve bestanden toonde aan dat SWIT in vergelijking met de traditionele programma’s (AVM) een stijging van ongeveer 20% werkzaamheidsgraad teweeg bracht. De vraag die Johansson zich stelt is of de veronderstelling dat de administratieve bestanden alle relevante informatie bevatten over de factoren die de participatie aan het programma en het vinden van een job bepalen.

Onder meer omdat de betrokken organisaties (SWIT en AMV) verschillende cijfers rapporteerden over het aantal SWIT-deelnemers, de timing van hun deelname en hun tewerkstellingsgraad werd beslist een telefonische survey te organiseren bij participanten aan beide programma's. Het zijn deze gegevens die Johansson gebruikt om de administratieve data aan te vullen. Wat blijkt? Uit de surveygegevens blijkt dat in het geval van de traditionele programma's (AMV) werkervaring een belangrijke factor is die de kansen op een job na de opleiding verhoogt. Voor SWIT is dat niet het geval. Bovendien is het effect van werkervaring voor AMV ongeveer even groot als het geschatte effect van SWIT. Een mogelijke verklaring voor deze vaststelling is dat 'werkervaring in AMV' en 'gastbedrijven in SWIT' functionele equivalenten zijn oftewel twee verschillende manieren om contacten met werkgevers te organiseren. Een aantal bijkomende gegevens laten toe om deze hypothese te testen. "All in all, this suggests that the effect of the Swit stems from increased employer contacts and not from differences in screening on (for us) unobserved characteristics or from organisational differences between the Swit and the AMV." Een belangrijke aanduiding dus betreffende het mechanisme dat de positieve resultaten van dit arbeidsmarktprogramma voortbrengt, maar een conclusie die men niet kan maken op basis van uitsluitend administratieve gegevens.

Over job-interviews en alloctonen

Dat de arbeidsmarkt alloctonen slechter behandelt dan hun autochtone collega's, zelfs als ze voor andere karakteristieken erg op elkaar gelijken, weten we. Maar waarom dat het geval is, is nog altijd niet erg duidelijk. Ronduit discriminerend, om niet te zeggen racistisch gedrag, waarover men soms berichten vindt in de pers is uiteraard een mogelijkheid. Maar een voor de hand liggende opmerking die rationeel ingestelde lieden met enig economisch inzicht daarbij maken, is waarom werkgevers zich in hemelsnaam van de voordelen van betere werknemers zouden beroven omwille van vooroordelen. Accidenteel kan een of andere werkgever zich wellicht een dergelijke dure smaak permitteren. Maar *alle* werkgevers *altijd*? Onderzoek dat meer licht laat schijnen op de mechanismen die hierbij een rol spelen is dan ook belang-

rijk. En dat de resultaten van dergelijk onderzoek niet altijd in dezelfde richting wijzen, soms zelfs erg moeilijk vergelijkbaar zijn, maar in ieder geval nopen tot genuanceerd denken, bewijzen de twee volgende artikels: David H. Autor, David Scarborough, Does Job Testing Harm Minority Workers? Evidence from Retail Establishments, in: *Quarterly Journal of Economics*, February 2008, pp. 219-277 en Sarah Campbell, Celia Roberts, Migration, ethnicity and competing discourses in the job interview: synthesizing the institutional and personal, in: *Discourse and Society*, vol. 18 (2007) pp. 243-271.

De analyse van Autor en Scarborough vertrekt van de vaststelling dat de technologische mogelijkheden werkgevers er meer en meer toe brengen om allerlei formele methoden in te zetten om bij aanwerving hun toekomstige werknemers te screenen. Van een dergelijke ontwikkeling mag men verwachten, zeggen ze, dat ze de onzekerheid over de 'echte' potentiële productiviteit van de kandidaat vermindert en dus de productiviteit in het algemeen ten goede komt. Maar dikwijls vreest men dat deze winst in efficiëntie gepaard gaat met minder gelijkheid omdat men weet dat risicogroepen – en zeker leden van etnische minderheden – bij dergelijke screenings minder goed presteren.

De centrale vraag die de auteurs trachten te beantwoorden is of deze vrees gegrond is. Zij doen dit in een dubbele beweging. Eerst construeren ze een formeel model dat meer inzicht moet geven in het eventuele samenspel van de twee mechanismen die de beoordeling van werknemers beïnvloeden. Het eerste mechanisme betreft de verbetering van de precisie van de screening. Het tweede mechanisme typeren Autor en Scarborough als 'changing beliefs', dit wil zeggen het binnenbrengen van informatie die systematisch afwijkt van de beoordeling gemaakt op basis van gewone interviews, waarbij deze systematische afwijking een gevolg kan zijn van ofwel vertekening door de (nieuwe) formele methode ofwel vertekening door de (vroegere) informele methode ofwel een (verschillende) vertekening door beide methoden.

De theoretische conclusie die uit deze analyse volgt, is duidelijk: "... if job tests are relatively unbiased, they do not pose an equality-efficiency trade-off; if job tests are bias-reducing, they pose an equality trade-off if and only if interviews are

minority-favoring; if job tests are bias-enhancing, they may pose an equality-efficiency trade-off – or they may simply reduce equality and efficiency simultaneously.” (p. 245) Kortom, op theoretische gronden kan men besluiten, aldus Autor en Scarborough, dat er geen intrinsiek probleem van trade-off bestaat. Efficiëntere selectie kan wel degelijk leiden tot meer ongelijkheid maar niet per definitie en dus moet de empirie uitsluitsel brengen.

Daartoe confronteren de auteurs de resultaten van deze formele analyse met de gegevens van een bedrijf dat eind vorige eeuw zijn informele werkwijze – spontane sollicitaties inwachten, bij vacatures iemand uitnodigen voor een interview en beslissen – verving door vormen van meer formele screening. Sollicitaties dienden vanaf dan te gebeuren via een soort ‘kiosk’ en vereisten de invulling van een geautomatiseerde vragenlijst. Deze vragenlijst verzamelde persoonsgegevens en gegevens over vroegere werkervaring, maar bevatte ook een persoonlijkheidstest van ongeveer 100 items, gebaseerd op het ‘big five-model’. Eens de sollicitatieformulieren ingevuld worden ze automatisch verwerkt en krijgt de manager een zogeheten ‘aanwervingsrapport’ met de resultaten van de test en een advies betreffende de eventuele aanwerving. Daarna verloopt de aanwervingsprocedure nagenoeg zoals in de vorige periode. Sollicitanten krijgen een job interview en in het beste geval een jobaanbieding. Leidt deze nieuwe procedure nu tot een meer accurate beoordeling van de sollicitanten en gaat dit gepaard met een kost in de vorm van een kleinere kans op een aanwerving voor etnische minderheden?

Het onderzoek van Campbell en Roberts, verbonden aan het Department of Education and Professional Studies van het befaamde King’s College van de University of London, is van een heel andere soort. Het onderzoek zelf is besteld door het Britse Department for Work and Pensions en wie de uitgebreide rapporten – ‘Talk on Trial’ (Research Report 344) 2006, 220 p. en ‘Talking Like a Manager’ (Research Report 510) 2008, 160 p. – wil consulteren kan terecht op de website van deze overheidsadministratie. Beleidsrelevant dus, maar wel op basis van zogeheten ‘kwalitatief materiaal’. (Voor een ander voorbeeld van hun methode kan ik verwijzen naar: Celia ROBERTS, Sarah CAMPBELL, *Fitting Stories into boxes: rhetorical and textual constraints on candidates’ performances in British*

job interviews, in: *Journal of Applied Linguistics*, vol. 2 (2005) n° 1, pp. 45-73.)

Sarah Campbell en Celia Roberts zijn socio-linguïsten. Hun empirisch materiaal bestaat uit een beperkt aantal op tape vastgelegde interviews met kandidaten voor relatief laagbetaalde jobs (receptionist, winkelbediende, ...) en van feedback sessies met de interviewers. Zij bekijken de interviews niet enkel op de inhoud; het gaat dus niet om een traditionele inhoudsanalyse. In wat men ‘discourse analysis’ noemt bekijkt men vooral ook een aantal vormkarakteristieken. Zo kijken ze naar de stijl van antwoorden, naar het aantal keren dat interviewers vragen herformuleren, naar de ‘omfloerste’ manier waarop geïnterviewden eventueel bepaalde moeilijke punten ter sprake brengen of omzeilen, naar de aarzelingen in het gesprek en dergelijke meer.

De aanleiding tot het werk van Campbell en Roberts is de ongelijke behandeling die etnische minderheden in Groot-Brittannië krijgen bij aanwervingen en de veronderstelling dat een op het eerste gezicht ‘neutraal’ selectie-instrument – het job-interview – ten aanzien van hen toch discriminerend werkt. De vraag is dan waarom? De onderzoeksters vertrekken van de stelling dat in moderne organisaties die een onderdeel vormen van de nieuwe economie erg veel nadruk ligt op het gebruik van taal en communicatie – met andere woorden op ‘discursieve vaardigheden’. Want, zo argumenteren Campbell en Roberts, typisch voor de moderne visie van organisaties is dat zij van werknemers vereisen dat zij in staat zijn om hun persoonlijke verhalen te (re)construeren in de bredere institutionele context. Dit is het geval omdat organisaties er niet enkel naar streven om individualisme de overhand te laten halen op de traditionele Fordistische structuren, maar tegelijk ook naar een grotere standaardisering van de waarden binnen de organisatie en een toenemende ‘accountability’. “On the one hand, the ‘empowerment’ and ‘making visible’ of individual employees is pursued through strategies of latent hierarchies, devolved decision-making power and emphasized flexibility. On the other hand, accountability is realized via techniques of surveillance, digitization and the standardization of procedures, measures and competences. Through this dual process, individual characteristics have become equated with institutional categories through the notion of ‘skills’ or ‘competences’, which are ‘measured’ in interviews

by the production of a written record and marks from one to ten.” (p. 246) Het interview tijdens het selectieproces speelt daarom een erg belangrijke en specifieke rol. Het moet toelaten aan de interviewers om na te gaan hoe en in welke mate de kandidaten erin slagen om een aanvaardbare identiteit te produceren waarin persoonlijke en institutionele aspecten naadloos in elkaar overvloeien. Dat men sinds de jaren 1990 ernaar tendeeft om bij selecties niet enkel meer CV's te bekijken, maar om het interview te structureren op basis van een competentiedenkader maakt dat zaken als institutionele prioriteiten en attitudes van de kandidaat expliciet ter sprake komen. Vooral de in dergelijke interviews veel voorkomende vraag om de eigen zwakke punten aan te geven en te beschrijven hoe zij zullen trachten zichzelf te verbeteren om zichzelf beter in lijn te brengen met de waarden van de organisatie demonstreert hoe de vereiste van persoonlijke investering in de organisatie (best) beantwoord wordt met een welbepaalde vorm van zelf-presentatie, gebruik makend van persoonlijke elementen in overeenstemming met de waarden van de organisatie: “... this slippage between the personal and the bureaucratic, the public and private domains in new corporate discourses means that the interviewee's ability to resolve these tensions by convincingly synthesizing the personal and institutional is crucial to their success.”

In hun artikel bespreken Campbell and Roberts in detail transcripts van interviews met verschillende kandidaten en de manier waarop de interviewers hierop reageerden. Zij maken hierbij een onderscheid tussen drie groepen van kandidaten: 'witte' Britse kandidaten, Britse kandidaten van allochtone afkomst en kandidaten (BA) die in het buitenland werden geboren en voor wie Engels niet hun moedertaal is. Zij constateren niet alleen dat de laatste groep de selectieprocedure duidelijk minder succesvol afrondt – tussen de eerste twee groepen is er geen verschil – maar ook dat de interviews met deze groep duidelijk meer 'problemen' opleverden. “Twice as many misalignments, and more than three times as many reformulations of questions occurred in interviews with BA candidates compared with their white and minority ethnic counterparts.” (p. 245) Leden van deze groep slagen er kennelijk niet in om een aanvaardbare synthese te brengen van de persoonlijke en de institutionele elementen in hun verhaal: “... it is particularly problematic for those

whom the dominant culture defines as 'other', whose extra-occupational life-world is viewed as foreign by the interviewer, and who are potentially less practised in the kinds of institutional and synthetic personal discourse that are privileged by the interview.” (p. 266) Vandaar dat de interviewers hun gedrag of uitleg veel meer typeren met negatieve termen als 'onbetrouwbaar', 'aangeleerde codewoorden', 'onprofessioneel', 'inconsistent' of 'opervlakkig'.

Kortom, zo concluderen Campbell en Roberts, de boodschap van dit alles is duidelijk. Potentiële werknemers die in hun dagelijkse praktijk (bijvoorbeeld omwille van hun werkomstandigheden) minder de gelegenheid krijgen om een 'synthetisch zelf' te produceren of die in omstandigheden leven die hen beletten om dit te doen (bijvoorbeeld omdat ze afgescheiden leven van de dominante organisatiecultuur) zijn tijdens selectieprocedures duidelijk gehandicapt als men de interviews als alleenzaligmakend beschouwt en dit zelfs als het gaat om jobs waarin 'discursieve vaardigheden' van ondergeschikt belang of zelfs helemaal niet belangrijk zijn.

En tot slot: iets over voorspellingen van Olympisch eremetaal

Mogelijk vindt u het niet echt een Grote Kwestie, maar vele mensen zijn erdoor gepassioneerd. En onze Vlaamse minister van Sport heeft er zich zorniet kwaad dan toch bezorgd over getoond: het aantal Olympische Medailles dat door onze atleten uit Peking is meegebracht. Veel te weinig in vergelijking met wat het zou moeten zijn geweest, zegen veel commentatoren – en om het mes in de wonde nog een keer rond te draaien werd daarbij dikwijls de vergelijking gemaakt met onze Noorderburen.

Laat me vooreerst bekennen dat ik graag een ouderwets en romantisch beeld blijf koesteren waarin sportactiviteiten als spektakel – bij zelf mee doen gelden andere criteria, denk ik – hun aantrekkelijkheid hoofdzakelijk verkrijgen door het onvoorspelbare van de uitkomst (inbegrepen het onverwacht boven hun eigen kunnen uitgroeien van sommige atleten) en slechts in ondergeschikte mate door de mate van technische perfectie waarmee handelin-

gen worden uitgevoerd (al zal het verband tussen beiden wel van sporttak tot sporttak verschillen). Aan de andere kant begrijp ik ook wel dat de confrontatie van de reëel behaalde uitslagen met de voorspelde resultaten van groot belang kan zijn voor mensen die het sportbeleid moeten bepalen en uitvoeren. Kwestie van niet met de natte vinger tewerk te gaan maar op een of ander manier aan 'evidence-based policy' te doen en de centen in te zetten op plaatsen waar het nodig is. Vandaar en omwille van pure wetenschappelijke nieuwsgierigheid is er een vraag die mij niet meer loslaat: op welke voorspelling baseren de commentatoren zich en waarop rusten deze voorspellingen?

Een beperkte zoektocht – 'Olympic Medals' intikken in Google – leverde een aantal verrassende resultaten op. Eerst en vooral, het aantal academici (van onder andere de universiteiten van Zürich, Groningen, Utrecht, Tokyo, Madrid, Colorado) en consultants (PriceWaterhouse Coopers, bijvoorbeeld) dat met deze materie bezig houdt, is groter dan ik dacht – en het aantal papers met voorspellingen dus ook. Zelfs Gary Becker, berucht Chicago econoom en ooit laureaat van de Nobelprijswinnaar economie, heeft aan deze kwestie op zijn blog een bijdrage gewijd – zij het om, zoals men kan verwachten, de superioriteit van het Amerikaanse marktmodel te bepleiten. Maar vooral: ik vond in eerste instantie geen spoor van een voorspelling die voor België het aantal medailles noemde dat men in de Vlaamse media hanteerde om de prestaties van de Belgen te beoordelen.

De meeste voorspellingen geven slechts een lijst van de eerste dertig landen van de ranglijst. In een voorspelling van Nederlandse onderzoekers – G. Kuper, E. Sterken, De winnaars van Beijing, in *Economisch-Statistische Berichten*, 25 juli 2008 – wordt deze dertigste plaats gedeeld door Zwitserland en Thailand met zeven medailles (2 gouden, 2 zilveren, 3 bronzen). Een andere voorspelling – A.W. Bernard, *Going for the Gold: Who Will Win the 2008 Olympic Games in Beijing*, Dartmouth (Tuck School of Business), July 2008 – sluit de lijst af met een aantal landen waarvoor men zes medailles verwacht (Tsjechië, Argentinië, Kazachstan, Zuid-Afrika en Iran). Een consultant – John Hawksworth, *Economic Briefing Paper: Modelling Olympic Performance*, PriceWaterhouseCoopers, June 2008 – publiceerde een lijst met Mexico op de dertigste

plaats met acht medailles (ex aequo met Iran, Zuid-Afrika en Indonesië). In geen van deze lijsten komt België voor, wat impliceert dat hun voorspelling voor ons land in elk geval lager ligt dan de elf medailles die we op basis van het bevolkingsaantal volgens De Morgen (28-08-2008) of de acht medailles die we op basis van rijkdom, welvaart en bevolking volgens Het Nieuwsblad (26-08-2008) hadden moeten behalen.

Gelukkig is er ook nog de website van prof. Daniel K.N. Johnson (Colorado College) – auteur, samen met Ayfer Alfi, van een van de meest geciteerde recente artikels over dit onderwerp *A Tale of Two Seasons: Participation and Medal Counts at the Summer and winter Olymnic Games*, in: *Social Science Quarterly*, vol. 85 (2004) n° 4, pp. 974-993. Op deze website staat een lijst waarin België wel voorkomt. Op de 23ste plaats – samen met Polen, Oostenrijk en Zwitserland – en een voorspelling van elf medailles (maar geen enkele gouden). Goed, denk je dan. Minstens één van onze sportjournalisten haalt zijn informatie dan toch uit een betrouwbare wetenschappelijke bron. En je begrijpt ook waarom onze Vlaamse Minister van Sport een twintigste plaats op de internationale ranglijst ten laatste in 2016 haalbaar acht.

Maar wat dan te denken van de positieve evaluatie van de Belgische prestaties door deze Amerikaanse prof zelf. Op 24 augustus 2008 – dus voor de geciteerde krantenartikels – was er een update op zijn website beschikbaar (<http://faculty1.coloradocollege.edu/~djohnson/Olympic.html>). "Who won?", vraagt hij zich af. Niet China of de US, zegt Johnson, hoewel ze zoals verwacht het hoogste aantal medailles behaalden. "After all, resources matter. The true 'winning nations' should be the ones that did amazing things with the resources given to them. That is, who outperformed expectations by the widest margin?"

Maar wat betekent 'beter doen dan verwacht'? Johnson heeft hiervoor drie criteria: het percentage waarmee het aantal reëel behaalde medailles het aantal voorspelde medailles overtreft, medailles halen als ze niet voorspeld werden, en gouden medailles halen die niet voorspeld werden. En ja, u hebt het wellicht al geraden, op basis van dit laatste criterium – en het is duidelijk dat hij aan dat criterium de voorkeur geeft – is volgens Johnson België

één van de grote winnaars – samen met Kenia, Canada, Denemarken en Zimbabwe.

Wat moeten we hier nu allemaal van denken? Vormen dit soort voorspellingen een goede basis voor een ‘evidence-based policy’ op het gebied van de sport? Ik weet het niet. Maar mijn korte toeristische uitstap in dit domein – en dan nog enkel via Google – maakt alleszins een aantal zaken duidelijk in verband met vragen die u zich wellicht ook stelt. Waarom wijkt de voorspelling van Johnson voor België zo sterk af van deze van andere academici? En wanneer mag men dit soort voorspellingen als succesvol beschouwen?

In tegenstelling tot wat de Vlaamse media suggereren gebruiken de meeste voorspellers modellen die iets meer factoren in rekening brengen dan bevolkingsgrootte en rijkdom (ofwel BNP per hoofd). Sinds decennia weet men dat ‘autoritaire regimes’ en ‘gastnaties’ door de band meer medailles winnen dan men op basis van de twee andere factoren zou verwachten. Ook het aandeel van sport in de overheidsbegroting speelt een rol. Johnson en Ali voegen daar in hun analyse van de medailles behaald in vorige Olympische Spelen zelfs de factor klimaat aan toe. Het voorbeeld van Indië, groot land en toch weinig medailles hoewel andere arme landen wel succesvol zijn, doet een aantal onderzoekers wijzen op culturele elementen – “Indiërs zijn enkel geïnteresseerd in cricket en da’s geen Olympische sport”. Een voorbeeld dat aangeeft dat men best ook rekening houdt met een weging naar sporttakken. Vandaar de mogelijkheid dat voorspellingen erg kunnen verschillen. Maar ook een meer technische factor speelt een rol. Het gros van de medailles wordt weggeplukt door een erg be-

perkt aantal landen – China, de VS, in het verleden de Sovjet-Unie – en bovendien is het totaal aantal beschikbare medailles beperkt. Twee elementen, met name ‘outliers’ die de resultaten sterk bepalen en een vorm van ‘censurering’, kunnen doen twijfelen aan de adequaatheid of zelfs de bruikbaarheid van de meest courante schattingsmethoden.

Dat men dergelijke voorspellingen dus best met enige argwaan bekijkt – en ondanks het zware wetenschappelijke alaaam dat ervoor wordt gebruikt niet als zaligmakend moet beschouwen, illustreert het volgende voorval.

Op een website – Recursivity (<http://recursed.blogspot.com>) – waarop men wel eens meer dergelijke problemen van het wetenschapsbedrijf ten berde brengt, stelt men op een erg sceptische toon de vraag of de voorspellingen van Johnson – de man was in de aanloop tot de Peking Spelen niet uit de US-media weg te branden – succesvol waren. Jazeker, antwoordt de professor in hoogst eigen persoon, want mijn voorspelling heeft een correlatie van .92 met de reëel behaalde resultaten. Goed en wel, zegt iemand anders, maar als men het aantal medailles behaald in Peking vergelijkt met het aantal medailles behaald in Athene dan vindt men een correlatie van .94. Kortom, veronderstellen dat landen het in Peking even goed (of slecht) zouden doen als in Athene presteert zelfs iets beter dan een voorspelling op basis van het model van Johnson en Ali. Toch iets om over na te denken, nietwaar?

Walter Van Trier