Leren redeneren over statistische verdelingen; een ontwikkelingsonderzoek

Arthur Bakker en Koeno P.E. Gravemeijer
Freudenthal Instituut, Universiteit Utrecht

Samenvatting
In dit ontwikkelingsonderzoek is onderzocht hoe brugklasleerlingen kunnen leren redeneren over statistische verdelingen. Het onderzoek bouwt voort op ontwikkelingsonderzoek van Cobb, Gravemeijer en collega's in de Verenigde Staten met computerprogramma's die wij vertaald en aangepast hebben (Cobb, 1999; Cobb, Gravemeijer, Bowers, & Doorman, 1997; Cobb, McClain, & Gravemeijer, in druk). De ontwikkeling van het redeneren over verdeling is geanalyseerd in drie fasen in overeenstemming met de gebruikte representaties: staafgrafieken, dotplots en door leerlingen ontworpen grafieken. Vier typen activiteiten kwamen tijdens dit onderzoek naar voren als zinvol bij het leren redeneren over verdeling: verzinnen van data bij bepaalde situaties, 'groeiende steekproeven', grafieken laten ontwerpen en voorspellende 'wat-als-vragen'.

1. Problemen in het statistiekonderwijs

Het doel van het beschreven onderzoek is onder andere om didactische oplossingen te ontwikkelen voor twee problemen binnen het statistiekonderwijs. Het eerste probleem is dat leerlingen in de basisschool en basisvorming geneigd zijn een dataverzameling als een collectie losse waarden te zien in plaats van als een geheel dat weer eigenschappen kan hebben (Hancock, Kaput, & Goldsmith, 1992; Konold & Higgins, 2002). Het onderliggende probleem is dat de meeste jonge leerlingen een meetwaarde als 1,52 m niet zien als een mogelijke waarde van een variabele (lengte) maar als een eigenschap van bijvoorbeeld een persoon. Wat leerlingen moeten leren is om de meetwaarde los te zien van het object of de persoon die gemeten wordt en om deze waarde te zien als een van de mogelijke waarden van een variabele. Vervolgens moeten leerlingen de verzameling meetwaarden als een geheel met een vast patroon leren zien, met andere woorden als een verdeling. 'Verdeling' is namelijk een conceptuele structuur waarbij de dataverzameling als een geheel gezien kan worden in plaats van als een verzameling losse elementen (Cobb, 1999; Gravemeijer, 1999a, b; Petrosino, Lehrer, & Schauble, in druk). We zeggen ook wel dat verdeling voor leerlingen een object (Sfard, 1991) zou moeten worden, dat weer bepaalde eigenschappen heeft.

Het tweede probleem is de cognitieve sprong die leerlingen in de tweede fase moeten maken van beschrijvende statistiek naar meer formele statistiek, als ze bijvoorbeeld leren over de normale verdeling. Ze leren dat de normale verdeling gezien kan worden als de limiet van histogrammen met steeds kleinere klassenbreedtes. Vervolgens leren ze onder andere hoe ze kunnen berekenen hoeveel procent van de mannen langer is dan 1,90 m als het gemiddelde en standaarddeviatie gegeven zijn. Veel leerlingen leren zulke sommen als een procedure en begrijpen niet goed waar ze mee bezig zijn.

Deze twee problemen vormden de reden om te onderzoeken hoe leerlingen in de basisvorming zonder statistische voorkennis (behalve het gemiddelde) op een voor hen betekenisvolle manier een informeel begrip van verdeling kunnen ontwikkelen. De te ontwikkelen kennis zou een goede basis moeten vormen voor de meer formele kennis die ze later leren.

2. Verdeling

Onbegrip

De roepstem van zijn vrouw, reeds op de trap:
'Wo bleibst du denn?' (Het eten stond al klaar)
Hè, altijd weer dat oponthoud, verdomme…

'Ich komme,' sprak hij wrevelig tot haar
Die binnentrad. 'Sei ruhig schon – ich komme'
En zij, geïntrigeerd: 'Was machst du doch?'

Het was zijn sedertdien beroemde Kromme
Die Carl F. Gauss liet zien; u kent die toch?
Een ware mijlpaal in de wetenschap

'Ganz reizend,' zei ze zonder veel mimiek
Ach ja, wat wist zij ook van statistiek?

Drs. P (drs. P & Kool, 2000, p. 105)[1]

Net als voor de vrouw van Gauss is een grafiek van een verdeling voor leerlingen in de basisvorming moeilijk te interpreteren. Aan de hand van Tabel 1 illustreren we deze complexiteit van het begrip 'verdeling' en schetsen we de globale conceptuele ontwikkeling die we nastreefden in dit onderzoek. Het uitgangspunt is dat leerlingen data als losse waarden zien en het einddoel is dat zij verdeling als een object ontwikkelen waarmee ze data kunnen modelleren. Daarbij zijn verschillende verdelingsaspecten van belang: centrum, spreiding, dichtheid en scheefheid.

Tabel 1. Structuur van het begrip verdeling.

<table>
<tr><td colspan="4" align="center">verdeling
(conceptuele eenheid)</td></tr>
<tr><td colspan="4" align="center">plaats en vorm
(globale informele aspecten)</td></tr>
<tr><td align="center">centrum</td><td align="center">spreiding</td><td align="center">dichtheid</td><td align="center">scheefheid</td></tr>
<tr><td colspan="4" align="center">data
(individuele waarden)</td></tr>
</table>

Om het redeneren over deze verdelingsaspecten te bevorderen is het nodig situaties te creëren waarin leerlingen gestimuleerd worden te kijken waar de meeste waarden zitten, of ze verspreid zijn of niet, waar de data dicht op elkaar zitten, of er uitschieters zijn, en waar de meerderheid van de data zit. Daarbij zijn grafische representaties noodzakelijk. Voor leerlingen zonder veel statistische voorkennis biedt het redeneren over de *vorm* van een verdeling een ingang. Die vorm heeft namelijk niet alleen een visueel, maar ook een conceptueel aspect. De vorm wordt namelijk beïnvloed door allerlei statistische eigenschappen van de verdeling: een hoge piek wordt veroorzaakt door een hoge frequentie van een bepaalde klasse; een lange staart aan één kant duidt op een scheve verdeling; een langgerekte vorm op een grote spreiding. Op grond van het onderzoek dat eerder uitgevoerd was in Nashville door Cobb, Gravemeijer en collega's, verwachtten we dat leerlingen informele woorden zouden gebruiken om dichtheid, spreiding en vorm te beschrijven (op elkaar, verspreid, heuvel). Zo hoopten we aan de hand van de vorm de achterliggende wiskunde bespreekbaar te maken zonder veel technische termen te gebruiken. Door het bespreken van de vorm van een verdeling hoopten we ook dat leerlingen op een andere manier zouden leren kijken naar een dataverzameling, namelijk als naar een geheel dat bepaalde kenmerken heeft. Dit betekent dat Tabel 1 op twee manieren gelezen kan worden, vanuit een opwaarts perspectief of vanuit een neerwaarts perspectief.

Het opwaartse perspectief is typerend voor nieuwelingen in de statistiek. Leerlingen zien data als losse waarden waarmee ze bijvoorbeeld het gemiddelde of de kwartielen kunnen berekenen. Dit betekent nog niet dat ze het gemiddelde als een centrummaat of een groepsrepresentant zien (Mokros & Russell, 1995; Konold & Higgins, 2002). Leerlingen moeten eerst enig begrip van verdeling hebben voordat ze kunnen beoordelen wanneer een gemiddelde of mediaan een zinvolle centrummaat is (Zawojewski & Shaughnessy, 2000). In andere woorden, leerlingen moeten ook het omgekeerde, neerwaartse perspectief ontwikkelen. Ze moeten centrum, spreiding en scheefheid als eigenschappen van een verdeling leren zien en ze moeten leren dat centrum gemeten kan worden met bijvoorbeeld het gemiddelde en de mediaan, en dat spreiding gemeten kan worden met spreidingsbreedte, kwartielafstand en standaarddeviatie, enzovoort. In de meeste schoolboeken worden dergelijke formele maten aangeleerd zonder een goede intuïtieve voorbereiding. Op grond van de domeinspecifieke theorie van het realistisch wiskundeonderwijs achten wij het zinvoller dat leerlingen eerst een notie van verdelingsaspecten ontwikkelen voordat ze statistische maten leren.

Experts combineren het opwaartse en neerwaartse perspectief. Voor hen is het gemiddelde zowel een berekening op empirische data als een eigenschap van een theoretische verdeling. Het opwaartse perspectief leidt tot een discrete frequentieverdeling, terwijl in het neerwaartse perspectief continue kansverdelingen (of een informeel voorstadium daarvan) gebruikt worden om de werkelijkheid te modelleren.

Welke grafiektypen zijn nu geschikt om te gebruiken met leerlingen zonder statistische voorkennis? Het is belangrijk om grafiektypen te kiezen waarin de verschillende verdelingsaspecten aan de vorm te zien zijn voor leerlingen. Conventionele grafieken als histogram en boxplot blijken moeilijk toegankelijk te zijn (Friel, Curcio, & Bright, 2001). Daarom is in de software en in het hypothetisch leertraject gekozen voor wat in het Engels 'dot plots' heten (Wilkinson, 1999). In een dotplot wordt iedere meetwaarde gerepresenteerd door een bolletje op de juiste plaats op de horizontale as. Leerlingen spreken van 'bolletjesgrafieken' of 'stippendiagrammen' (Figuur 5-7). Opdat leerlingen dit type grafiek goed interpreteren is het noodzakelijk dat ze inzien dat de bolletjes staan voor waarden van een variabele. Omdat dit inzicht niet voor alle leerlingen vanzelfsprekend is, worden ter voorbereiding de data gerepresenteerd met lijnstukken van 0 tot de juiste waarde op de horizontale as (Figuur 1). We hebben aanwijzingen dat dergelijke lijnstukken voor leerlingen inzichtelijker zijn dan stippen op een as. Horizontale staven representeren op een natuurlijke manier variabelen als remafstand, levensduur van batterijen en spanwijdte van vogels.

Deze twee grafiektypen, de staafgrafiek waarbij iedere staaf een meetwaarde voorstelt en het dotplot, zijn gebruikt in twee computertools die voor het Amerikaanse experiment ontwikkeld zijn: in Minitool 1 worden data met horizontale staven weergegeven en Minitool 2 als gestapelde bolletjes.[2] Er is gekozen voor statistieksoftware omdat leerlingen met software makkelijker en sneller data kunnen exploreren: grote hoeveelheden data zijn nauwelijks met de hand te verwerken. Bovendien is software dynamisch en interactief, in tegenstelling tot het papier. Hiermee is de gebruikte software gemotiveerd, is het leren redeneren over verdeling als einddoel beargumenteerd en is een schets gegeven van het beoogde leerproces.

3. Ontwikkelingsonderzoek

Om de hoofdvraag van dit artikel te kunnen beantwoorden, hadden we een nieuwe leergang nodig en om die leergang te kunnen maken moesten we onderzoek doen. We maakten daarom gebruik van de methode van ontwikkelingsonderzoek (Edelson, 2002; Freudenthal, 1991; Gravemeijer, 1994; van den Akker, 1999). Ontwikkelingsonderzoek, zoals wij het uitvoeren, heeft tot doel leergangen (toetsen, software, nascholing, …) en lokale onderwijstheorieën te ontwikkelen die fungeren als inspiratiebron voor anderen, zoals schrijvers van schoolboeken, onderzoekers en softwareontwikkelaars. Soms kunnen op basis van domeinspecifieke onderwijstheorieën ook domein-onafhankelijke onderwijstheorieën ontwikkeld of aangescherpt worden (Lijnse, in druk). Ontwikkelingsonderzoek wordt ook gebruikt om domeinoverstijgende onderzoeksthema's te onderzoeken, zoals het proces van symboliseren en modelleren in het wiskundeonderwijs (Cobb, Gravemeijer, Yackel, McClain, & Whitenack, 1997; zie ook Symbolen en Modellen*,* themanummer van TDβ*,* 1999).

Binnen een ontwikkelingsonderzoek kunnen we drie fasen onderscheiden die een cyclus vormen: een ontwerpfase, klassenexperimenten en een retrospectieve analyse. In de door ons gekozen opzet wordt vooraf op basis van bestaande of voorlopige lokale onderwijstheorieën een zogenaamd hypothetisch leertraject (HLT) geformuleerd. Daarin wordt geanticipeerd op de mentale activiteiten van leerlingen bij bepaalde opdrachten en wordt doordacht hoe die vervolgens gebruikt kunnen worden om dichter bij het einddoel te komen. Daarmee heeft het HLT (Simon, 1995) dezelfde functie als een 'scenario' (Klaassen, 1995). Het HLT wordt tijdens de klassenexperimenten steeds aangepast waar nodig, waarbij de empirische en theoretische redenen voor die aanpassingen gedocumenteerd worden. Op basis van de retrospectieve analyse worden het HLT en eventueel de lokale instructietheorie herzien, zodat ze tijdens een volgende ronde gebruikt kunnen worden.

De beschrijving van het onderzoeksresultaat en van het proces dat tot dit resultaat heeft geleid, moet bruikbaar en toegankelijk zijn voor buitenstaanders. Geïnteresseerden kunnen dan de achterliggende ideeën en overwegingen begrijpen en in andere situaties toepassen. Freudenthal formuleerde dit als volgt:

> *Developmental research means: experiencing the cyclic process of development and research so consciously, and reporting on it so candidly that it justifies itself, and that this experience can be transmitted to others to become like their own experience* (1991, p. 161).

De zo ontwikkelde prototypische leergangen en de daarbij horende lokale onderwijstheorieën worden tegelijkertijd en in samenhang met elkaar ontwikkeld, en wel zo dat ze tegelijkertijd theoretisch en empirisch onderbouwd worden. Ontwikkelingsonderzoek heeft daardoor tegelijkertijd een exploratief en een evaluatief karakter. Oude ideeën uit het HLT worden getoetst terwijl nieuwe gegenereerd en getoetst worden. Dit type onderzoek verschilt dus aanzienlijk van de aanpak waarbij een kant-en-klaar ontwerp in een experimentele en controleconditie wordt vergeleken en in een volgende ronde eventueel wordt aangepast. In dit artikel worden op basis van ontwikkelingsonderzoek vier nieuwe onderwijsactiviteiten gepresenteerd, waarbij we met 'nieuw' bedoelen nieuw ten opzichte van het vorige HLT en de lokale onderwijstheorie voor aanvankelijk statistiekonderwijs.

4. Onderzoeksmethode

Na exploratieve interviews met 26 willekeurig gekozen brugklasleerlingen op een Amsterdamse mavo-havo-vwo-school en vier losse lessen in een havoklas hebben we klassenexperimenten uitgevoerd in twee vwo- en twee havobrugklassen op een scholengemeenschap in Breukelen. In deze vier klassen waren de verzamelde gegevens audio-opnamen, leerlingenwerk, observaties en eindtoetsen. In de laatste twee experimenten zijn ook voortoetsen afgenomen en zijn video-opnamen gemaakt. De voortoetsen waren bedoeld om te kijken hoe goed leerlingen het gemiddelde konden berekenen (redelijk) en of ze misschien al wisten wat wij ze wilden leren (dat was niet het geval). De gebruikte software is op enkele punten aangepast. Voor het ontwerp van het lesmateriaal verwijzen we verder naar Bakker (in druk).

De protocollen van het laatste experiment zijn ingevoerd in het programma MEPA (multiple episode protocol analysis) van Erkens (2001). De uitspraken

Tabel 2. Overzicht van het onderzoek en de dataverzameling

<table>
<tr><th>Leerlingen (brugklas)</th><th>Type experiment</th><th>Dataverzameling</th><th>Aantal lessen</th><th>Niveau</th></tr>
<tr><td>26 leerlingen</td><td>interviews (20 minuten voor twee leerlingen)</td><td>audio-opnamen</td><td>-</td><td>mavo, havo, vwo</td></tr>
<tr><td>klas A (25)</td><td>vier lessen</td><td rowspan="3">leerlingwerk, eindtoets, observaties, mini-interviews, audio</td><td>4</td><td>havo</td></tr>
<tr><td>klas F (27)</td><td>eerste experiment</td><td>12</td><td>vwo</td></tr>
<tr><td>klas E (28)</td><td>tweede experiment</td><td>15</td><td>vwo</td></tr>
<tr><td>klas C (23)</td><td>derde experiment</td><td rowspan="2">idem plus voortest en video</td><td>12</td><td>havo</td></tr>
<tr><td>klas B (23)</td><td>vierde experiment</td><td>12</td><td>havo</td></tr>
<tr><td>twaalf klassen</td><td>implementatie op school</td><td>e-mailverslagen van twee docenten, observatieverslagen van bezoeken</td><td>144</td><td>havo en vwo</td></tr>
</table>

zijn gecodeerd naar opgave, gebruikte woorden en de statistische begrippen waar het om ging. Dit maakte het mogelijk alle fragmenten terug te zoeken die informatie konden geven over de betekenis van bepaalde begrippen en grafieken voor leerlingen. Dit was nuttig tijdens de retrospectieve analyse van dit experiment om gerezen vermoedens te kunnen toetsen.

Tijdens het eerste jaar is gewerkt met een zeer ervaren docente die ook ervaring heeft als schoolboekauteur. Hierdoor konden we ons vooral op het leerproces van de leerlingen in relatie tot het hypothetisch leertraject richten. In de jaren erna is de leergang gebruikt door twee onervaren docenten in twaalf andere brugklassen. Een van de auteurs heeft deze docenten ingewerkt, enkele lessen bijgewoond en verslagen van de docenten verzameld.

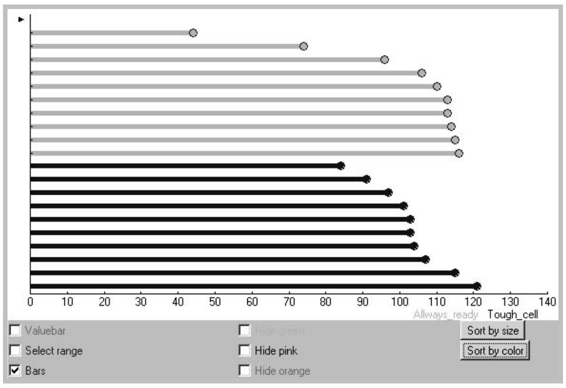
Een belangrijk hulpmiddel om uit te vinden wat begrippen en grafieken voor leerlingen betekenden was het houden van mini-interviews. De mini-interviews, die in de klas werden afgenomen, varieerden in lengte tussen ongeveer twintig seconden en vier minuten. Doordat wij veel reflectievragen stelden beïnvloedden deze interviews uiteraard het leerproces van de leerlingen. Toch was de validiteit van het onderzoek onzes inziens niet in gevaar omdat we wilden weten hoe brugklasleerlingen kunnen leren redeneren over verdeling, niet of de ontwikkelde leergang geschikt was voor willekeurige andere klassen in dezelfde tijd.

Voor de retrospectieve analyses van het laatste experiment gebruikten we een methode die lijkt op een methode van Cobb en Whitenack (1996) en op de ‘constant comparative method’ van Glaser en Strauss (Strauss & Corbin, 1998). We lazen de uitgewerkte protocollen, bekeken de videobanden en formuleerden vermoedens over het leerproces in de verschillende episodes. Deze vermoedens werden getoetst aan andere episodes van de protocollen en de videobanden, maar ook aan de rest van het datamateriaal (observaties, leerlingmateriaal, toetsen). Deze werkwijze werd herhaald op hetzelfde materiaal. Interessante protocolfragmenten, ook die in dit artikel opgenomen zijn, zijn besproken met collega’s (‘peer examination’) om onze interpretaties te staven. Een voorbeeld van een hypothese die gegenereerd, getest en bevestigd is, is dat leerlingen geneigd zijn om data in drie groepen in te delen, namelijk in ‘laag’, ‘gemiddeld’ en ‘hoog’.

Op basis van de gebruikte grafieken is de analyse onderverdeeld in drie fasen. In de eerste fase werkten leerlingen met grafieken waarin data door horizontale staven gerepresenteerd werden (Minitool 1, Figuur 1). In de tweede fase (les 5-12) werkten de leerlingen vooral met dotplots (Minitool 2, Figuur 3). In de derde fase (les 11-15) redeneerden leerlingen over grafieken die ze zelf bedacht hadden (de voorbeelden komen uit de tweede brugklas).

5. Fase 1: data gerepresenteerd door staven

In de tweede les moesten de leerlingen een rapport schrijven aan de consumentenbond over de kwaliteit van twee batterijmerken. Het doel van deze activiteit was om leerlingen over verdelingsaspecten te laten redeneren, zoals gemiddelde, spreiding en de meerderheid, in een representatievorm die betekenisvol was voor leerlingen. Met verschillende opties in Minitool konden ze de data sorteren op grootte en splitsen per merk. Ze gebruikten de verticale waardelijn om waarden af te lezen en later soms om visueel het gemiddelde te schatten.



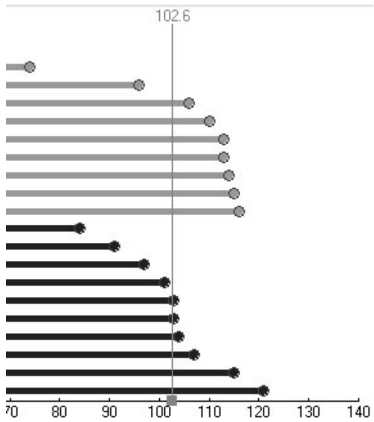
Figuur 1. Minitool 1. Batterij-dataverzameling van twee merken.

Tijdens de eerste les in klas A was er geen uitgebreide klassendiscussie over hoe de data totstandgekomen waren. Veel leerlingen keken naar de hoogste waarde: “Ik zou die [batterij] van 121 uur nemen”. Ze hadden niet door dat het om een test of een steekproef ging en dat een eenmaal geteste batterij waardeloos is. Dit illustreert dat leerlingen naar individuele waarden neigen te kijken (paragraaf 2) en niet aan steekproeven denken. Daarom hebben we in de volgende klassen meer tijd besteed aan de vraag waar de data vandaan komen. Cobb (1999) noemt dit “talking through the process of data creation”. Hij spreekt van “data creation” omdat voor de meeste vragen data gecreëerd moeten worden; ze zijn er niet zonder doel, geschiedenis of meetactiviteit (Lehrer & Romberg, 1996; Roth, 1996). Voor het beantwoorden van een statistische vraag moeten leerlingen eerst nadenken over de variabele die gemeten moet worden. We hoopten dat bij een betere doordenking van dit proces van data creëren de leerlingen een notie van variabele, bijvoorbeeld van levensduur, zouden ontwikkelen.

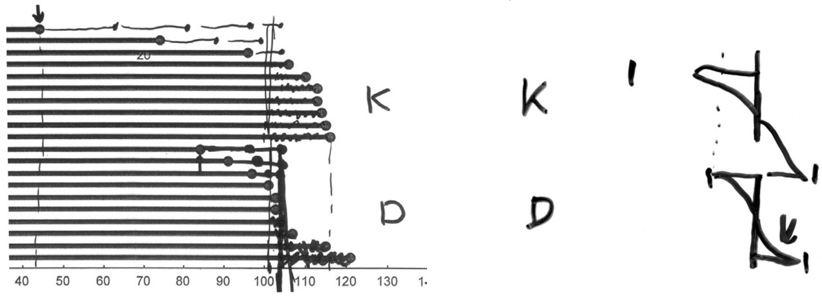
Ook realiseerden we ons dat ‘verdeling’ als complexe begrippenstructuur ook weer onderdeel is van een grotere begrippenstructuur. Verdeling is namelijk niet los te zien van bijvoorbeeld variatie, steekproef en populatie (Watson, 2002). Verdeling is een conceptueel instrument waarmee patronen in variatie beschreven kunnen worden. Zonder een notie van steekproef of populatie is niet duidelijk waar de gecreëerde data voor staan. Daarom hebben wij er bij nader inzien voor gekozen om al deze statistische begrippen tegelijk en informeel te behandelen op een coherente manier. Daarmee wijkt onze benadering af van wat er in de schoolboeken staat, die vooral centrummaten en grafieken behandelen.

Al bij het batterijenprobleem redeneerden leerlingen in alle klassen over aspecten van verdelingen: merk K heeft uitschieters maar je hebt meer kans op een goede. Merk D is betrouwbaarder omdat je weet dat ze meer dan 80 uur meegaan. Betrouwbaarheid bleek een bruikbaar aanknopingspunt te zijn om het over spreiding te hebben. Tot zover lijken onze observaties op die tijdens het experiment in Nashville waarin leerlingen over ‘consistency’ spraken (zie bijvoorbeeld Cobb (1999) en Sfard (2000)). In het vervolg van deze paragraaf besteden we aandacht aan twee punten waarop onze experimenten in fase 1 afweken van dat in Nashville: visueel gemiddelden schatten en zelf data verzinnen. Het visueel schatten van gemiddelden was in Nashville wel de bedoeling maar was niet gelukt. Het tweede, zelf data verzinnen, was wel een nieuwe activiteit in het hypothetisch leertraject.

Tijdens de exploratieve interviews waren er kinderen die spontaan gemiddelden schatten van temperaturen in een staafgrafiek. Ze zeiden een stukje van juli aan januari te geven en van augustus aan februari, enz. Omdat dit visuele schatten met staven inzichtelijker is dan alleen een rekenprocedure (Strauss & Bichler, 1988), hebben we activiteiten ontworpen die kwalitatief en visueel redeneren over het gemiddelde konden uitlokken (Bakker, in druk). Minitool 1 ondersteunt deze strategie met de beweegbare verticale waardelijn. Leerlingen in alle klassen legden uit dat ze iets van de langere staven afhakten en dat aan de kortere staven gaven (Figuur 2 en 3). In verschillende klassen konden leerlingen uitleggen waarom dit mocht: het totaal blijft hetzelfde en het gemiddelde is het totaal gedeeld door het aantal staven. Toen leerlingen zeiden dat merk D beter was omdat het gemiddelde hoger was, was het gemiddelde voor hen meer dan de uitkomst van een berekening: ze gebruikten het om iets over het merk als geheel te zeggen. Dit betekent een eerste stap bij de ontwikkeling van het ‘neerwaartse perspectief’ waarbij het informele verdelingsaspect ‘gemiddelde’ gebruikt worden om iets over de



Figuur 2. Gemiddelde schatten met verticale waardelijn.

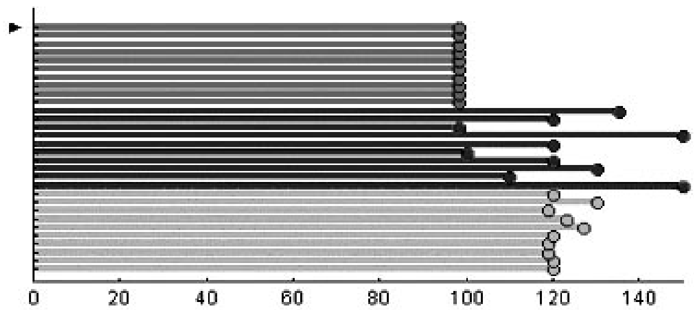


Figuur 3. Schetsjes op transparanten tijdens klassendiscussies over het schatten van gemiddelden. Het gemiddelde van data van D is iets hoger dan dat van de data van K.

data te zeggen.

Eerste voorbeeld van een nieuwe activiteit.
Tijdens de eerste drie lessen waren er nuttige discussies over de betrouwbaarheid van batterijen in relatie tot de staafgrafieken geweest, maar we zochten een manier waarop we konden onderzoeken wat leerlingen nu precies onder betrouwbaarheid en spreiding verstonden. Tijdens het tweede experiment was gebleken dat het zinvol was om leerlingen hun eigen grafieken te laten ontwerpen bij hypothetische situaties (zie paragraaf 7). Daarom besloten we bij de laatste twee experimenten de originele batterij-vraag om te draaien: we vroegen kinderen tijdens de vierde les om hun eigen data te verzinnen die hoorden bij de uitspraken “merk A is slecht maar betrouwbaar”, “merk B is goed maar onbetrouwbaar” en “merk C is goed en redelijk betrouwbaar”. Bijna alle leerlingen maakten in Minitool 1 iets wat op Figuur 4 leek, waaruit blijkt dat ze binnen deze context inzicht hadden verworven in het gemiddelde (hoe goed het merk is) versus de betrouwbaarheid (spreiding). Ook waren ze actief betrokken bij deze opgave.

Bij de retrospectieve analyse hebben we geconcludeerd dat het streven naar een dergelijke heen-en-weerbeweging tussen begrippen en grafieken een bruikbare heuristiek is bij het ontwerpen van lesmateriaal voor statistiek en wel om verschillende theoretische redenen.



Figuur 4. Door twee leerlingen bedachte data behorend bij “merk A is slecht maar betrouwbaar”, “merk B is goed maar onbetrouwbaar” en “merk C is goed en redelijk betrouwbaar”.

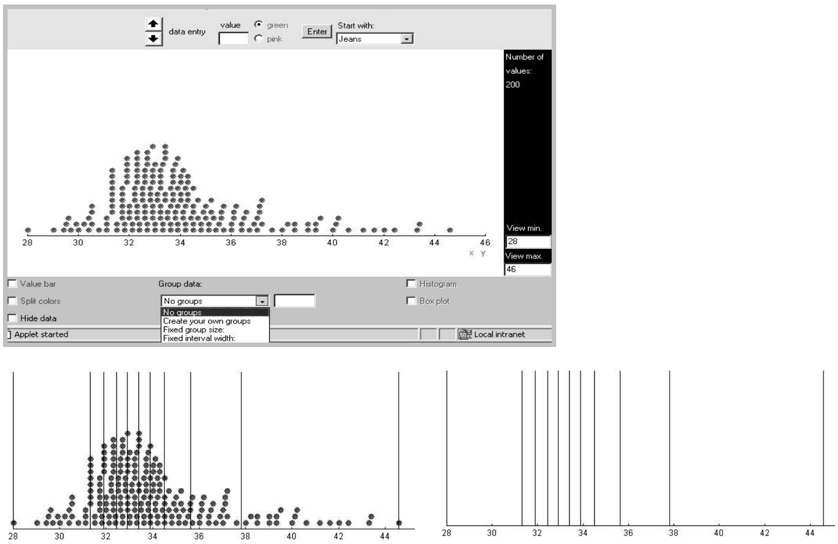
1. Omdat leerlingen met grafieken ideeën kunnen uitdrukken die ze niet met woorden kunnen uitdrukken (Lemke, in druk), kunnen leraren en onderzoekers beter toetsen wat leerlingen begrijpen als ze hun eigen data en grafieken verzinnen.
2. Als leerlingen denken aan eigenschappen als “goed maar onbetrouwbaar”, dan weerhoudt het gebrek aan data ze ervan om puur individuele data te construeren. Op deze manier gebruiken we bewust de cognitieve beperkingen: het is onmogelijk zich erg veel losse punten voor te stellen. Er is dus een cognitieve noodzaak voor een begripsstructuur waarmee data samengevat kunnen worden. Met de genoemde activiteit creëren we de noodzaak voor zo’n begrip, wat moet leiden tot de ontwikkeling van het begrip ‘verdeling’, omdat daarmee een collectie data met bepaalde eigenschappen beschreven kan worden. Met een dergelijke ‘omgekeerde’ opgave creëren we

dus een cognitieve impuls om een begrip te vormen dat eenheid brengt in de veelheid.
3. In de meeste schoolboeken moeten leerlingen kant-en-klare grafieken interpreteren (Friel et al., 2001) en als ze grafieken moeten maken, is meestal voorgeschreven welke en hoe. De Lange en collega's (1993) en Meira (1995) bevelen aan om leerlingen ook hun eigen grafieken te laten ontwerpen, omdat dat bij hen tot betekenisvol redeneren leidde. De vrijheid die leerlingen hebben om eigen symboliseringen te maken zorgt ervoor, naar we aannemen, dat die voor hen betekenis en een bepaalde functie hebben.
4. Het belang van het heen en weer gaan tussen data en grafieken (of verschillende grafieken) wordt onderschreven door theorieën over symboliseren (Steinbring, 1997; Yackel, 2000). Onder 'symboliseren' wordt hier het proces verstaan van symbolen maken, interpreteren, gebruiken en verbeteren, dus als een pars pro toto voor 'symbolen maken'. Steinbring onderscheidt bij leerlinginteractie over symbolen referentiesystemen en symboolsystemen. Leerlingen interpreteren een onbekend symboolsysteem in relatie tot een referentiesysteem dat ze wel kennen. Voor het leren van een nieuw symboolsysteem is het belangrijk dat leerlingen heen en terug redeneren tussen de twee systemen. Zo kan de staafgrafiek van Minitool 1 een symboolsysteem zijn als kinderen bekend zijn met getallen in tabellen. Als Minitool 1 eenmaal bekend is, kan dat weer als referentiesysteem voor een nieuw symboolsysteem dienen, zoals de dotplot van Minitool 2.

6. Fase 2: bolletjes vervangen staven

Minitool 1 is een onconventionele grafieksoort die voor leerlingen eenvoudig is; Minitool 2 is geavanceerder en komt meer in de richting van een conventionele grafieksoort waarin verdelingen meer lijken op kansverdelingen. Zoals vermeld in de vorige paragraaf, kan Minitool 1 als referentiesysteem gezien worden voor Minitool 2, het nieuwe symboolsysteem. De overgang werd als volgt gemaakt. Bij het oplossen van statistische problemen met Minitool 1 redeneerden de leerlingen met de eindpunten van de staven. In de herziene versie van Minitool 1 kunnen de staven dan ook verborgen worden zodat alleen de eindpunten overblijven. Sommige leerlingen vonden dit overzichtelijker, zeker als het aantal data groter werd. We krijgen de bolletjesgrafiek van Minitool 2 door de eindpunten in Minitool 1 in gedachten te laten vallen richting de juiste plek op de as. Als er al andere bolletjes liggen, stapelen ze op. Daarbij vallen de bolletjes niet opzij, wat open plekken in de grafiek verklaart (Figuur 5). Om de relatie tussen de twee Minitools inzichtelijker te maken is er bij de laatste twee experimenten voor gekozen om leerlingen dezelfde dataverzameling in beide representaties te laten vergelijken. Voorbeelden van vragen die we stelden waren: "in welke grafiek vind je dat het gemiddelde makkelijker te vinden is?" en "in welke grafiek kun je makkelijker de spreiding van de data zien?" Tijdens de retrospectieve analyse van de gecodeerde protocollen was het algemene beeld dat leerlingen het gemiddelde makkelijker konden vinden in Minitool 1 en de spreiding makkelijker zagen in Minitool 2, maar daar waren ook uitzonderingen op. Dit duidt erop dat met verschillende representaties verschillende verdelingsaspecten ontwikkeld kunnen worden.

Minitool 2 biedt meer mogelijkheden om data te organiseren dan Minitool 1. Behalve sorteren op grootte en kleur (subgroep) kunnen leerlingen ook data in eigen groepen indelen, in twee even grote groepen (voor de

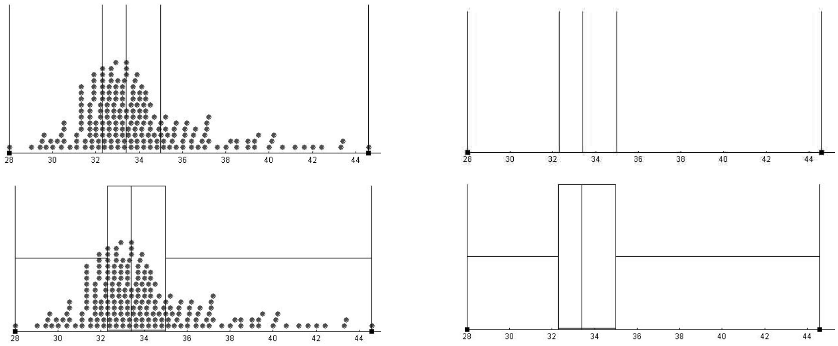


Figuur 5. Minitool 2 met spijkerbroekendata (taille in inches). Boven: het volledige venster met alleen data; onder links: vaste groepsgrootte (20 data punten per groep); rechts: ‘verberg data’.

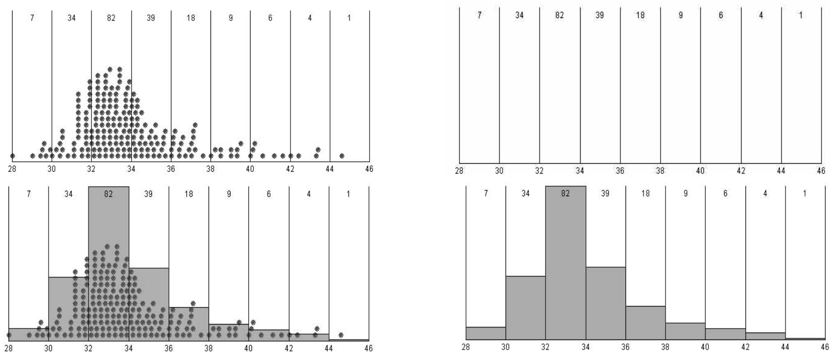
mediaan), in vier even grote groepen (voor het boxplot), vaste intervalbreedte (voor het histogram) en vaste groepsgrootte. De laatste optie bleek nuttig om redeneringen over dichtheid uit te lokken.

Een van de statistische problemen waaraan de klassen werkten met Minitool 2 was het zogenaamde spijkerbroekenprobleem. De opdracht was om een spijkerbroekenfabrikant advies te geven over het percentage broeken dat gemaakt moest worden van elke maat. De beschikbare dataverzameling gaf taillematen in inches van 200 mannen. Deze opdracht, die meestal tijdens de negende les gedaan werd, was bedoeld om de aandacht op de hele verdeling te richten met een grafische representatie die dichter bij de conventionele representatie van verdelingen kwam. We hoopten dat leerlingen een histogram zouden maken en het nut zouden inzien van die grafische representatie om een verdeling te beschrijven.

Veel leerlingen gebruikten in eerste instantie de vier even grote groepen om hun conclusie te onderbouwen dat “je veel [spijkerbroeken] moet maken van [maat] 34-36 en minder van 44-46”. Om exactere antwoorden te krijgen stelden we meestal een sceptische vraag: “Als de fabrikant jou inhuurt voor 1000 gulden, denk je dat hij tevreden zou zijn met dit advies?” De meesten kozen dan uiteindelijk de vaste-intervaloptie en maakten een tabel met percentages per maat. Tijdens de retrospectieve analyse concludeerden we dat de keuze voor vier even grote groepen bij de geïnterviewde leerlingen voortkwam uit de associatie met percentages: met vier even grote groepen zijn alle groepen 25%. Mogelijk sloot deze optie aan bij de gewoonte van leerlingen om groepen in laag (laagste 25%), gemiddeld (middelste 50%) en hoog (hoogste 25%) onder te verdelen.



Figuur 6. Vier even grote groepen met en zonder data. De boxplotoptie is toegevoegd in de revisie van Minitool 2.



Figuur 7. Vaste intervalbreedte met en zonder data. Histogramoptie is later toegevoegd.

Tweede voorbeeld van een nieuwe activiteit

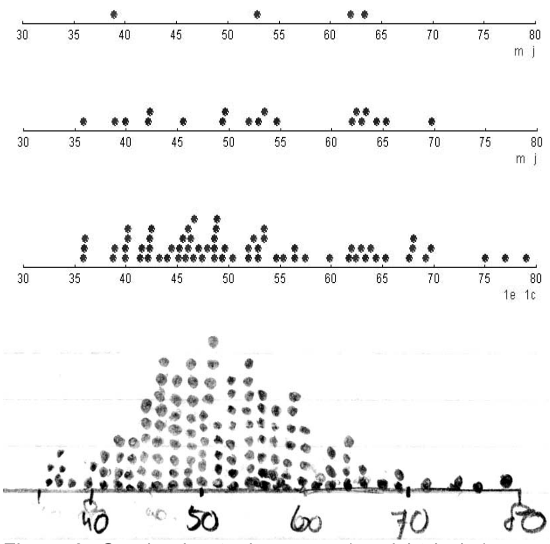
Ons einddoel was, als gezegd, dat leerlingen op een coherente en betekenisvolle manier over verdeling leerden redeneren in samenhang met het begrip steekproef en geschikte grafische representaties. In het voorgaande zijn aanzetten tot de ontwikkeling van verdelingsaspecten te zien en is ook in toenemende mate aandacht besteed aan de vraag hoe we aan data komen. Toch zochten we een activiteit waarbij de notie van steekproef explicieter in samenhang met verdeling en grafieken ontwikkeld kon worden. Door uit te gaan van de ideeën die leerlingen zelf hadden, werd ons een dergelijk idee in de schoot geworpen: groeiende steekproeven. Leerlingen bleken in eerste instantie steeds maar kleine steekproeven te willen nemen om iets te onderzoeken. De vraag naar een betrouwbaardere manier riep vanzelf om meer data.

Voor de achtergrond van deze onderwijsactiviteit gaan we terug naar de derde les, waarin leerlingen hadden nagedacht over het volgende probleem:

Deze luchtballon kan het gewicht van acht volwassenen dragen. Hoeveel brugklasleerlingen kunnen er veilig mee als je alleen op het gewicht let?

Deze vraag was bedoeld om ze te laten nadenken over variatie van gewichten, steekproeven en de representativiteit van het gemiddelde. Een veel voorkomende oplossing was om het gemiddelde gewicht van een brugklasser en een volwassene te schatten. Daarna gebruikten sommigen de verhouding daarvan om aan het aantal brugklassers te komen, maar de meesten berekenden eerst het toegestane maximumgewicht en deelden dat door het geschatte gemiddelde brugklasgewicht. De antwoorden varieerden van 10 tot 16 leerlingen per ballon.

Om leerlingen beter te laten nadenken over variatie en steekproeven vroegen we in een volgende les naar een betrouwbaardere manier om een antwoord op deze vraag te krijgen. Een leerling stelde voor om twee jongens en twee meisjes te wegen en het gemiddelde te bepalen. Uiteindelijk stelde iemand voor de hele klas te wegen en zo geschiedde. In de tweede brugklas werd ook ieders lengte gemeten.



Figuur 8. Groeiende steekproeven (gewicht in kg) en een voorspelling van een grote steekproef.

In een volgende les lieten we de gewichtgegevens van twee jongens en twee meisjes zien in Minitool 2 (Figuur 8) en vroegen we de klas wat ze verwachtten te zien als we de data van de rest van de klas toevoegden. Sommigen dachten dat het gemiddelde preciezer zou worden, maar omdat we hen over het geheel wilden laten redeneren, vroegen we explicieter naar de vorm van de grafiek. De een dacht dat de spreidingsbreedte groter zou worden omdat er meer uitschieters kwamen en de ander dat de grafiek hoger zou worden omdat er meer kinderen bijkwamen. Nadat we hun de data hadden getoond, vroegen we hun de data van drie klassen te voorspellen, lieten die zien en vroegen om een voorspelling van zes klassen. Op deze manier werden in dit laatste experiment extreme waarden, spreiding en vorm onderwerpen van discussie. Wij concludeerden bij de retrospectieve analyse dat ook het idee van de groeiende steekproeven bruikbaar was bij het leren redeneren over verdelingen op een betekenisvolle manier. Tijdens vervolg-experimenten in een tweede klas en twee groepen 8 werd dit bevestigd.

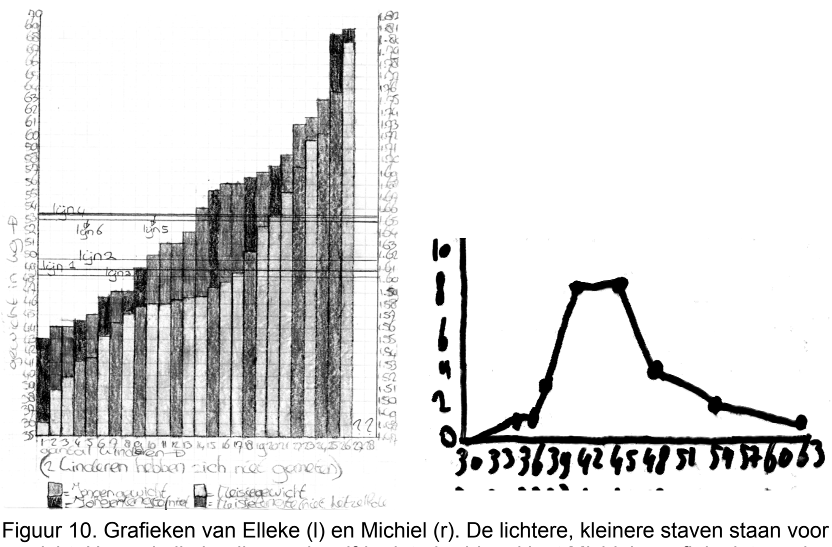
7. Fase 3: van data via ‘bult’ naar verdeling

Hoewel de leerlingen in de eerste twee experimenten in de tweede fase begonnen te redeneren over ‘meerderheid’ en de ‘meeste’, sprak niemand over de vorm van de grafiek. We hadden gehoopt dat ze over heuvels zouden praten omdat leerlingen in het Amerikaanse experiment dat deden (Cobb, 1999). Daar leidde het praten over heuvels (“if you look at the graphs and look at them like hills, …”, p. 19) ertoe dat leerlingen een dataverzameling als een verdeling gingen zien[3]. We zochten daarom een andere manier om leerlingen over grafieken als een geheel te laten redeneren.

Derde voorbeeld van een activiteit die nieuw was binnen het leertraject

Tijdens een onderzoek van De Lange en collega’s (1993) was gebleken dat leerlingen veel statistisch inzicht tentoonspreidden als ze zelf grafieken mochten maken. Omdat ook eigen constructies binnen de domeinspecifieke theorie van het realistisch wiskundeonderwijs een belangrijke rol spelen (Treffers, 1987), besloten we om leerlingen hun eigen grafieken te laten ontwerpen bij hun eigen data. We namen aan dat die grafieken dan iets voor hen zouden betekenen en een bepaalde functie zouden hebben – meer dan wanneer ze een grafiek voorgeschoteld krijgen of wanneer ze een bepaald type grafiek volgens een vaste procedure moeten maken.

Als vervolg op de ballonvraag lieten we de leerlingen een grafiek maken waarmee de ballonvaarder kon bepalen hoeveel brugklasleerlingen mee



Figuur 10. Grafieken van Elleke (l) en Michiel (r). De lichtere, kleinere staven staan voor gewicht. Hoewel alle leerlingen dezelfde data hadden, klopt Michiels grafiek niet precies met die van Elleke; Michiels grafiek is meer een ruwe schets.

mochten. Dit leidde tot verschillende grafieken; sommige leken op Minitool 1 en 2, andere waren nieuw (bijvoorbeeld een puntenwolk van gewicht tegen lengte). De docente koos ervoor om de meeste discussietijd te besteden aan de grafieken van Michiel en Elleke (Figuur 10), omdat die een verwantschap vertoonden met de Minitools en omdat in die van Michiel duidelijk een heuvel te zien was. Ze besprak de puntenwolk niet omdat die vermoedelijk te complex was voor veel leerlingen (Minitool 3 biedt een puntenwolk-representatie, maar daar zijn we niet aan toegekomen).

Michiels grafiek is interessant omdat de bespreking ervan leidde tot redeneren over bulten, dus over de vorm van verdelingen.

Michiel:	*Ja kijk, je hebt ongeveer gemiddeld genomen dan, uhm hoeveel kinderen zoveel wogen, en daar heb ik dan een stipje gezet.*
Lerares:	*Ja.*
Michiel:	*En dan heb ik links [y-as] het aantal kinderen. Er is één kind die ongeveer 35 weegt, en er is er eentje die 36 weegt, en twee die 38 wegen. En dan heb ik/*
Lerares:	*Heb je gewoon stippen gezet.*

Hij legde verder uit waar de stippen voor stonden. De stip boven 48 geeft aan dat er ongeveer vier leerlingen met een gewicht rond de 48 kg waren. Nadat enkele andere grafieken besproken waren, waaronder die van Elleke, vroeg de lerares het volgende.

Lerares:	*Wat zie je daar heel goed aan [aan Michiels grafiek]?*
Laila:	*Nou, dat het gemiddelde, dat de meeste kinderen in de klas, uhm, nou tussen de 39 en de, nou 48 zitten.*
Lerares:	*Ja, hier kan je in één klap zien welk gewicht de meeste kinderen hier in de klas zo'n beetje hebben, wat daar zo'n beetje de grootste groep is. Gewoon doordat je hier zo'n bult ziet. De bult die ben je hier [in Ellekes grafiek] een beetje kwijt.*

De lerares gebruikte het woord 'bult' om het over de vorm van de verdeling te kunnen hebben. Met haar opmerking dat de bult verdwenen was, lokte zij een cognitief conflict uit waardoor de leerlingen over de wiskundige achtergrond van de bult moesten nadenken. Nadia legde uit waar de bult in Ellekes grafiek was.

Nadia:	*Het verschil, ze staan van klein naar groot, dus de bult, dat is daar waar de dinges, waar de staafjes het dichtst bij elkaar liggen.*
Lerares:	*Wat bedoel je: waar de staafjes het dichtst bij elkaar liggen?*
Nadia:	*Het verschil, uiteinde, verschilt het niet zo met de volgende.*

Dit gebied, waar de grafiek recht liep, stond dus voor de meerderheid van de gegevens. Eva voegde het volgende toe.

Eva:	*Nou, als je goed kijkt, dan zie je bijna, in het midden zit bijna eentje recht en eh, ja die/*
Lerares:	*En dat is wat jij [Nadia] ook zei van eh, daar zitten ze vlak bij elkaar en hier zitten ze zeg maar op een kluitje, wat de*

	lengtes betreft of wat de gewichten betreft/
Eva:	*En dat is dan ook die bult.*

Uit dergelijke episodes bleek dat de bult niet alleen maar een visuele betekenis had. Het gebruik van het woord 'bult' was wel visueel gemotiveerd, maar de situatie was zodanig dat leerlingen zelf een wiskundige betekenis moesten construeren van iets wat aanvankelijk nog geen wiskundige betekenis had. Verschillende leerlingen konden goed uitleggen waar de bult volgens hen voor stond, namelijk de meerderheid van de leerlingen of de 'gemiddelde' leerlingen.

Vierde voorbeeld van een nieuw type activiteiten

Het eerste voorbeeld van een nieuwe activiteit dat we gaven over het omdraaien van de batterijopgave was onder andere gemotiveerd door het vierde en laatste voorbeeld dat we hier behandelen. Omdat we wilden dat leerlingen met verdelingen als objecten leren redeneren, zochten we naar situaties waarin ze niet meer over de data konden nadenken maar met de verdeling als geheel moesten redeneren. Zoals Sfard (1991) schrijft, moeten we situaties creëren waarin leerlingen het wiskundig object als object nodig hebben, bijvoorbeeld door ze er operaties mee uit te laten voeren. Een voorbeeld uit een ander wiskundig domein kan dit mogelijk verduidelijken. Om te stimuleren dat leerlingen functies als objecten zien (en niet puur als rekenprocedure), kan het helpen ze functies bij elkaar op te laten tellen of functies met een bepaalde factor te laten vermenigvuldigen. Naar aanleiding hiervan zochten wij informele operaties die leerlingen op verdelingen zouden kunnen uitvoeren. We dachten aan het verschuiven en uitvergroten van verdelingen en we deden dat in hypothetische situaties om te voorkomen dat leerlingen op individuele getallen zouden letten. Voorbeelden van zulke voorspellingsvragen, die we 'wat-als-vragen' noemen, volgen in de volgende fragmenten uit de dertiende les. Eerst een voorbeeld van het verschuiven van de verdeling:

Lerares:	*Wat gebeurt er als je alle tweede klassen zou wegen [in plaats van alle eerste klassen]?*
Luuk:	*Ik denk ook ongeveer hetzelfde, alleen andere maat, andere getallen.*
Guyonne:	*De bult meer naar rechts.*
Lerares:	*Wat zou dat voor die boxplots betekenen?*
Michiel:	*Schuift ook naar rechts. Die bult in het midden is eigenlijk gewoon de boxplot, die schuift meer naar rechts.*

Luuk denkt mogelijk nog aan losse getallen; Guyonne heeft het over de bult als enkelvoud die als geheel meer naar rechts verschuift; Michiel combineert zijn antwoord met een type grafiek dat de klas geleerd heeft (vermoedelijk gebruikte Michiel het woord 'boxplot' voor de box in het midden in plaats van voor de hele grafiek).

De volgende voorspellingsvraag gaat over wat er gebeurt als we een grotere steekproef zouden nemen.

Onderzoeker:	*Als je bijvoorbeeld niet alle brugklassen van deze school maar alle brugklassen in Utrecht zou nemen, hoe zou de grafiek veranderen of verandert die niet?*

Elleke: *Dan zouden er misschien iets meer aan de linkerkant en iets meer aan de rechterkant bijkomen. Dan zou die bult weer iets breder worden, denk ik.*
Onderzoeker: *Is iemand het hier niet mee eens?*
Michiel: *Ja, als er meer kinderen zijn, dan wordt het gemiddelde, dus het meeste, dat wordt ook meer. Dus de bult blijft gewoon hetzelfde.*
Albertine: *Ik denk dat alleen het aantal kinderen meer wordt en de bult hetzelfde blijft.*

Kennelijk zien enkele leerlingen een belangrijk aspect van verdelingen dat we hoopten te ontwikkelen: de globale constante vorm ongeacht de steekproefgrootte. Elleke is in staat om een begrip als uitschieter op een zinvolle wijze te verbinden met een vormaspect (breedte van de bult). Interessant aan de opmerkingen van Michiel en Albertine is dat deze leerlingen bij meer data toch aan dezelfde vorm blijven denken. Kennelijk denken ze niet aan absolute frequenties, want dan zou de bult veel hoger worden. De opmerking dat de bult hetzelfde blijft, duidt eerder op een intuïtief begrip van relatieve frequentie. We zijn er niet aan toegekomen om expliciet onderscheid te maken tussen de verdeling van een steekproef en die van de populatie, maar met de vraag over veel grotere steekproeven is daar wel een aanzet toe gemaakt.
We zagen in bovenstaande voorbeelden van redeneren over steekproeven een bevestiging van onze vermoedens over het nut van opereren met bulten. Het is nuttig om wat-als-vragen te stellen die niet op beschikbare data betrekking hebben, maar met voorspellingen van andere situaties te maken hebben. Hiervoor is een cognitieve verklaring denkbaar die vergelijkbaar is met het tweede argument in paragraaf 5: bij voorspellingen is het niet goed mogelijk om aan individuele data te denken. Een voorspellingsvraag waarbij met verdelingen als objecten geopereerd moet worden, kan het denken aan een geheel uitlokken. Er vindt dus een soort omkering plaats: de leerlingen redeneren niet meer over een veelheid van bestaande data, maar redeneren andersom met een intuïtieve notie van verdeling over hypothetische data. Dit betekent overigens niet dat verdeling voor alle leerlingen een object geworden was dat weer bepaalde eigenschappen kon hebben. We hebben maar een paar voorbeelden kunnen vinden waarin leerlingen over verdelingen zeggen dat ze scheef of normaal zijn.

8. Conclusies

De centrale vraag in dit artikel was hoe brugklassers op een betekenisvolle manier kunnen leren redeneren over verdeling. In drie fasen analyseerden we hoe opdrachten met en zonder computer dit leerproces ondersteunden. We vatten de drie fasen samen en noemen daarbij enkele didactische inzichten en vier nieuwe activiteiten die in dit ontwikkelingsonderzoek empirisch en theoretisch gerechtvaardigd werden.

Bij het oplossen van statistische problemen met Minitool 1 gebruikten leerlingen woorden als meerderheid, uitschieters, betrouwbaarheid en ‘verspreid’. Ze leerden visueel het gemiddelde te schatten. Ze bedachten dataverzamelingen die pasten bij eigenschappen van batterijtypen, zoals “goed maar onbetrouwbaar”. We beargumenteerden (voorbeeld 1) dat dit soort verzinopgaven kunnen helpen bij de overgang van data zien als losse waarden naar een dataverzameling als een geheel. Ook raakten we in toenemende mate ervan overtuigd dat er aandacht aan steekproeven besteed moet worden en dat daardoor makkelijker bepaalde verdelingsaspecten ontwikkeld kunnen worden. Dit didactisch inzicht is verwerkt in de gereviseerde leertrajecten.

Bij het werken aan opgaven met Minitool 2 redeneerden leerlingen over verdelingsaspecten als frequentie, klassen, spreiding, kwartielen, mediaan en dichtheid. Een veelbelovende activiteit (voorbeeld 2) bleek de ‘groeiende steekproef’ te zijn omdat bij die activiteit verdelingsaspecten met begrip van steekproeven en grafieken verbonden worden.

In de derde fase begonnen leerlingen over de vorm van verdelingen te redeneren in relatie tot statistische begrippen als meerderheid, uitschieters en steekproefgrootte in hypothetische situaties. Dit is in het laatste experiment ook bereikt, maar de voorbeelden komen uit het tweede experiment. We betoogden (voorbeeld 3) dat leerlingen de kans moeten krijgen om hun eigen grafieken te ontwerpen. Bovendien leidde het bespreken van verschillende grafieken tot het redeneren over ‘bulten’. Verder beargumenteerden we (voorbeeld 4) dat ‘wat-als-vragen’, voorspellingen over de vorm en plaats van verdelingen in hypothetische situaties, nuttig waren om leerlingen het beoogde ‘neerwaartse’ modelleerperspectief (paragraaf 3) op data te laten ontwikkelen. Opvallend daarbij is dat de redeneerkwaliteit tijdens lessen zonder computer, vooral als leerlingen zelf grafieken en data moesten verzinnen, hoger was dan tijdens de computerlessen. Toch zijn er aanwijzingen (de grafieken van Michiel en Elleke bijvoorbeeld) dat het exploreren met de computerprogramma’s een goede basis was voor de activiteiten zonder computer. We benadrukken hier dat de leerlingen in de eerste twee experimenten waarschijnlijk naar het vwo zouden gaan, en dat de citaten in fase 3 van de leerlingen kwamen die goed waren in wiskunde. Andere leerlingen zouden waarschijnlijk meer tijd nodig hebben of ouder moeten zijn om op een vergelijkbare manier over verdeling te leren denken. Het einddoel dat leerlingen een verdeling als een object moesten gaan zien, bleek voor een groot deel van de leerlingen te hoog gegrepen. Uit de toetsen blijkt verder dat de meeste leerlingen grafieken juist konden interpreteren, zelf grafieken bij een verhaaltje met informele en statistische termen konden verzinnen en nog niet veel wisten van steekproeven, behalve dat “een steekproef een deel van het geheel is”. In vervolgonderzoek willen we dus meer aandacht besteden aan steekproeven in relatie tot verdeling. We concluderen dat het op zijn minst mogelijk is om brugklassers over verdeling te leren redeneren op de manier die in dit artikel beschreven is.

Dankbetuigingen

Wij danken Mieke Abels, Maarten Jasper en Mirjam Jansens voor het lesgeven tijdens dit project. Het onderzoek is gefinancierd door NWO onder nummer 575-36-003B, dat onderdeel is van het aandachtsgebied ‘Wiskunde en ICT’. Onze dank gaat ook uit naar Jantien Smit voor haar correctie en naar de referenten en redacteur voor hun uitgebreide commentaar.

Noten

1. Hoewel de grafiek van de normale verdeling wel eens de kromme van Gauss (1777-1855) genoemd wordt, was deze verdeling al eerder bekend bij De Moivre (1667-

1754) en Laplace (1749-1827). Onafhankelijk van elkaar hebben Galton (1822-1911) en anderen haar de ‘normale verdeling’ genoemd (Stigler, 1999).
2. De Minitools kunnen als applets bekeken worden op http://www.fi.uu.nl/wisweb/ en http://peabody.vanderbilt.edu/depts/tandl/mted/. Van de eerste site kunnen ze ook als applicaties gedownload worden.
3. Een mogelijke verklaring voor het uitblijven van het redeneren over vorm is dat onze experimenten uit 12 of 15 lessen bestonden terwijl dat van Cobb, Gravemeijer en collega’s uit 34 lessen bestond en het redeneren over heuvels pas in de laatste fase daarvan voorkwam.

English summary

Learning to reason about statistical distributions; an example of design research

In this design research we investigated how seventh-grade students could learn to reason about statistical distributions. This research builded upon design research of Cobb, Gravemeijer, and colleagues in Nashville, USA, with computer tools that we have translated and revised (Cobb, 1999; Cobb, Gravemeijer, Bowers, & Doorman, 1997; Cobb, McClain, & Gravemeijer, in press). The development of students’ reasoning about distribution is analyzed in three stages according to the representations used: value bar graphs, dot plots, and student invented graphs. Four new instructional activities came to the fore as useful for learning to reason about distribution: invention of data according to some situation, ‘growing samples’, invention of graphs, and what-if questions about hypothetical situations.

Literatuur

Bakker, A. (in druk). The Early History of Statistics and Implications for Education. *Journal of Statistics Education*.

Cobb, P. (1999). Individual and Collective Mathematical Development: The Case of Statistical Data Analysis. *Mathematical Thinking and Learning, 1*, 5-43.

Cobb, P., Gravemeijer, K. P. E., Bowers, J., & Doorman, M. (1997). Statistical Minitools (applets en applicaties). Nashville & Utrecht: Vanderbilt University, TN & Freudenthal Instituut, Universiteit Utrecht.

Cobb, P., Gravemeijer, K. P. E., Yackel, E., McClain, K., & Whitenack, J. (1997). Mathematizing and Symbolizing: The Emergence of Chains of Signification in One First-Grade Classroom. In D. Kirshner & J. A. Whitson (Eds.), *Situated cognition theory: Social, semiotic, and neurological perspectives* (pp. 151-233). Hillsdale, NJ: Erlbaum.

Cobb, P., McClain, K., & Gravemeijer, K. P. E. (in druk). Learning about statistical covariation. *Cognition and Instruction*.

Cobb, P., & Whitenack, J. W. (1996). A Method for Conducting Longitudinal Analyses of Classroom Videorecordings and Transcripts. *Educational Studies in Mathematics, 30*, 213-228.

de Lange, J., Burrill, G., Romberg, T., & van Reeuwijk, M. (1993). *Learning and Testing Mathematics in Context: The case: Data Visualization.* Madison, WI: University of Wisconsin, National Center for Research in Mathematical Sciences Education.

Doorman, M., & Gravemeijer, K. P. E. (1999). Modelleren als organiserende activiteit in het wiskundeonderwijs. *Tijdschrift voor Didactiek der Bèta-wetenschappen, 16*, 38-55.

drs. P & Kool, M. J. H. (2000). *Wis- en natuurlyriek; met chemisch supplement.* Amsterdam: Nijgh & Ditmar.
Edelson, D. C. (2002). Design Research: What We Learn When We Engage in Design. *Journal of the Learning Sciences, 11*, 105-121.
Erkens, G. (2001). MEPA: Multiple Episode Protocol Analysis (Versie 4.6.2). Utrecht: Utrecht University.
Freudenthal, H. (1991). *Revisiting Mathematics Education: China Lectures*. Dordrecht: Kluwer Academic Publishers.
Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making Sense of Graphs: Critical Factors Influencing Comprehension and Instructional Implications. *Journal for Research in Mathematics Education, 32*, 124-158.
Gravemeijer, K. P. E. (1994). *Developing Realistic Mathematics Education*. Utrecht: CDβ Press.
Gravemeijer, K. P. E. (1999a*). ICT in dienst van het heruitvinden van statistische concepten en representaties.* Paper gepresenteerd op Onderwijsresearchdagen (ORD) 1999, Nijmegen.
Gravemeijer, K. P. E. (1999b). *An Instructional Sequence of Analysing Univariate Data Sets*. Paper gepresenteerd op de AERA 1999, Montreal, Canada.
Hancock, C., Kaput, J. J., & Goldsmith, L. T. (1992). Authentic Enquiry with Data: Critical Barriers to Classroom Implementation. *Educational Psychologist, 27*, 337-364.
Klaassen, C. W. J. M. (1995). *A problem-posing approach to teaching the topic of radioactivity*. Utrecht: CDβ Press.
Konold, C., & Higgins, T. (2002). Working with Data. In S. J. Russell & D. Schifter & V. Bastable (Eds.), *Developing Mathematical Ideas: Collecting, Representing, and Analyzing Data*. Parsippany, NJ: Dale Seymour Publications.
Konold, C., Robinson, A., Khalil, K., Pollatsek, A., Well, A. D., Wing, R., & Mayr, S. (2002). Students' Use of Modal Clumps to Summarize Data. In B. Phillips (Ed.), *Developing a Statistically Literate Society. Proceedings of the International Conference on Teaching Statistics* [CD-ROM], Cape Town, South-Africa, July 7-12, 2002.
Lehrer, R., & Romberg, T. (1996). Exploring children's data modeling. *Cognition and Instruction, 14*, 69-108.
Lemke, J. L. (in druk). Mathematics in The Middle: Measure, Picture, Gesture, Sign, and Word. In M. Anderson & A. Saenz-Ludlow & S. Zellweger & V. V. Cifarelli (Eds.), *Educational Perspectives on Mathematics as Semiosis: From Thinking to Interpreting to Knowing.* Ottawa: Legas Publishing.
Lijnse, P. (in druk). Op weg naar een didactische structuur van de natuurkunde? De ontwikkeling van didactische structuren volgens een probleemstellende benadering. *TDβ*.
Meira, L. (1995). Microevolution of mathematical representations in children's activity. *Cognition and Instruction, 13*, 269-313.
Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education, 26*(1), 20-39.
Petrosino, A. J., Lehrer, R., & Schauble, L. (in druk). Structuring Error and Experimental Variation as Distribution in the Fourth Grade. *Mathematical Thinking and Learning*.

Roth, W. M. (1996). Where is the context in contextual word problems? Mathematical practices and products in grade 8 students' answers to story problems. *Cognition and Instruction, 14*, 487-527.
Sfard, A. (1991). On the dual nature of mathematical conceptions: reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*, 1-36.
Sfard, A. (2000). Steering (Dis)Course Between Metaphors and Rigor: Using Focal Analysis to Investigate an Emergence of Mathematical Objects. *Journal for Research in Mathematics Education, 31*, 296-327.
Simon, M. A. (1995). Reconstructing Mathematics Pedagogy from a Constructivist Perspective. *Journal for Research in Mathematics Education, 26*, 114-145.
Steinbring, H. (1997). Epistemological investigation of classroom interaction in elementary mathematics teaching. *Educational Studies in Mathematics, 32*, 49-92.
Stigler, S. M. (1999). *Statistics on the Table. The History of Statistical Concepts and Methods*. Cambridge, MA: Harvard University Press.
Strauss, A., & Corbin, J. (1998). *Basics of Qualitative Research. Techniques and Procedures for Developing Grounded Theory*. Thousand Oaks, CA: SAGE.
Strauss, S., & Bichler, E. (1988). The Development of Children's Concepts of the Arithmetic Mean. *Journal for Research in Mathematics Education, 19*, 64-80.
Treffers, A. (1987). Three dimensions: a model of goal and theory description in mathematics instruction - The Wiskobas project, Dordrecht, the Netherlands: Kluwer.
Van den Akker, J. (1999). Principles and methods of development research. In J. van den Akker &. Branch, R. M. & K. Gustafson & N. Nieveen & T. Plomp (Eds.), *Design approaches and tools in education and training* (pp. 1-14). Boston: Kluwer.
Symbolen en Modellen. (1999). Themanummer *TDβ*, *16(1).*
Watson, J. M. (2002). Creating Cognitive Conflict in a Controlled Research Setting: Sampling. In B. Phillips (Ed.), *Developing a Statistically Literate Society; Proceedings of the Sixth International Conference of Teaching Statistics* [CD-ROM]. Cape Town, South Africa.
Wilkinson, L. (1999). Dot Plots. *The American Statistician, 53*, 276-281.
Yackel, E. (2000). Introduction: Perspectives on Semiotics and Instructional Design. In: P. Cobb & E. Yackel & K. McClain (Eds.), *Symbolizing and Communicating in Mathematics Classrooms* (pp. 1-13). Mahwah, NJ: Erlbaum.
Zawojewski, J. S., & Shaughnessy, J. M. (2000). Mean and Median: Are They Really so Easy? *Mathematics Teaching in the Middle School, 5*, 436-440.