

Redeneren

over steekproeven op de lerarenopleiding basisonderwijs

Vanwege het toenemende belang van data in onze samenleving, is het belangrijk dat burgers leren geldige inferenties te maken op basis van data. In de voorstellen voor een nieuw rekenen-wiskundecurriculum wordt dan ook voorgesteld om basisschoolleerlingen al kennis te laten maken met redeneren over steekproeven.

Dat zou kunnen door hen informele statistische inferenties (ISI) te leren maken. ISI houdt in dat er een generalisatie gemaakt wordt die gebaseerd is op data, maar die tegelijkertijd verder reikt dan deze data en die gekenmerkt wordt door onzekerheid. (Toekomstige) leraren moeten zelf voldoende ISI-kennis hebben. In dit artikel beschrijf ik de resultaten van mijn promotieonderzoek, waarvoor ik onderzocht heb hoe de ISI-kennis van studenten aan de lerarenopleiding basisonderwijs (pabo) bevorderd kan worden. De resultaten laten zien dat een meerderheid van de studenten die deelnamen aan een korte lessenserie over ISI kernelementen van ISI begrepen, maar dat als studenten zelf lesgeven over ISI op de basisschool ze moeite hadden uit te leggen waardoor het maken van een inferentie mogelijk is. Op basis van deze resultaten doe ik een voorstel voor een ISI-lessenserie voor de pabo.

Ultimately, teachers must be able to know and use mathematics in practice, not merely do well in courses or answer pedagogically contextualized questions in interviews.

Ball, Lubienski, and Mewborn (2001, p. 451)

Arjen de Vetten, ICLON,
Universiteit Leiden

De Vetten, A.J. (2019). Redeneren over steekproeven op de lerarenopleiding basisonderwijs. *Volgens Bartjens – ontwikkeling en onderzoek*, 39(2), 49-61

Inleiding

Inferenties in het dagelijks leven

‘Nederland zegt massaal ja tegen rekeningrijden’, stond er op 29 januari 2019 in het Nederlands Dagblad. Deze uitspraak, die pretendeert iets te zeggen over alle Nederlanders, is gebaseerd op een steekproef van 3263 volwassen Nederlanders. Het is een voorbeeld van een inferentie, of generalisatie, gemaakt over een populatie op basis van een steekproef. Data en statistiek doortrekken steeds meer onze maatschappij (Gravemeijer, Stephan, Julie, Lin, & Ohtani, 2017) en inferenties op basis van steekproeven komen dus ook steeds vaker voor in het dagelijks leven. Het is dus belangrijk dat burgers de geldigheid van inferentiële uitspraken kunnen beoordelen. Ook op de werkvloer blijkt dat inferentieel redeneren (het proces om een conclusie te trekken op basis van data of logica) steeds belangrijker wordt, mede doordat deze vaardigheid niet gemakkelijk geautomatiseerd kan worden (Liu & Grusky, 2013). Omdat inferentieel redeneren zo’n belangrijke maatschappelijke vaardigheid is, kan het raadzaam zijn kinderen al op de basisschool met deze vaardigheid in aanraking te laten komen. Het voorstel van het ontwikkelteam rekenen-wiskunde van curriculum.nu onderkent dit ook door statistiek een prominentere plek in het curriculum van de basisschool te geven. Ze suggereren dat leerlingen op de basisschool al hun eerste ervaringen moeten opdoen met het trekken van steekproeven en het doen van uitspraken op basis van deze steekproeven (Ontwikkelteam Rekenen & Wiskunde, 2019).

Informele statistische inferenties

Het maken van inferenties zou op de basisschool de vorm kunnen krijgen van informele statistische inferenties (ISI). ISI kan als volgt gedefinieerd worden: ‘een generalisatie die gebaseerd is op data, maar tegelijkertijd verder reikt dan deze data en die gekenmerkt wordt door onzekerheid’ (Ben-Zvi, Bakker, & Makar, 2015, p. 293). Een goede informele statistische inferentie is gebaseerd op drie componenten (Makar & Rubin, 2009):

1. Data als bewijs: De inferentie is gebaseerd op de data, en niet op eigen ervaring of overtuiging.
2. Generalisatie voorbij de data: De inferentie gaat verder dan een beschrijving van de steekproefgegevens door een probabilistische claim te maken die verder reikt dan de data.
3. Onzekerheidstaal: In de inferentie moet rekening gehouden worden met het feit dat uitkomsten van steekproeven variëren en dat de gebruikte steekproef meer of minder representatief is. Voor een correct gebruik van onzekerheidstaal is het daarom belangrijk dat de oorzaken van deze onzekerheid begrepen worden. Daarom onderscheiden we de volgende subcomponenten (De Vetten, Schoonenboom, Keijzer, & Van Oers, in druk):
 - a. Steekproefvariabiliteit: De inferentie is gebaseerd op begrip van steekproefvariabiliteit, wat inhoudt dat de uitkomsten van steekproeven variëren, maar onder bepaalde omstandigheden vergelijkbaar zijn en dat daarom een steekproef gebruikt kan worden voor een inferentie.
 - b. Methode van steekproeftrekking: De inferentie bevat een bespreking van de methode van steekproeftrekking en de implicaties daarvan voor de representativiteit van de steekproef.
 - c. Steekproefgrootte: De inferentie bevat een bespreking van de steekproefgrootte en de implicaties daarvan voor de representativiteit van de steekproef.
 - d. Onzekerheid: De inferentie bevat een beschrijving van de implicaties van de gebruikte methode van steekproeftrekking en steekproefgrootte op de waarschijnlijkheid dat de steekproef een betrouwbare voorspelling geeft over de hele populatie.

Probleemstelling

Als kinderen op de basisschool in aanraking moeten komen met ISI, dan moeten hun leraren voldoende kennis hebben van ISI. Onderzoek laat echter zien dat dit vaak niet het geval is (Groth & Meletiou-Mavrotheris, 2018). Er is echter nog weinig bekend hoe de ISI-kennis van studenten van de lerarenopleiding basisonderwijs (pabostudenten) bevorderd kan worden. In mijn promotieonderzoek heb ik daarom verkend hoe interventies de ISI-kennis van pabostudenten kunnen stimuleren (De Vetten, 2018). In dit artikel presenteer ik de belangrijkste bevindingen, waarin ik uiteindelijk uitkom bij een lessenserie voor ISI-onderwijs aan pabostudenten. Ik zal in het bijzonder stil staan bij welke kennis studenten laten zien als ze lesgeven over ISI op de basisschool.

De uitgevoerde interventies zijn op zo’n manier opgezet dat ze gemakkelijk vertaald kunnen worden naar curricula voor reken-wiskundeonderwijs op de lerarenopleiding basisonderwijs (pabo).

In de eerste plaats hebben we korte lessenseries ontworpen. In tweede plaats zijn de activiteiten getest in reguliere pabo-klassen, dus niet met vrijwillige, extra gemotiveerde of extra getalenteerde studenten. In de derde plaats vereisen de activiteiten weinig organisatie.

Vooronderzoek

De eerste twee studies van mijn proefschrift hebben in kaart gebracht wat de kennis over ISI van beginnende pabostudenten is, zodat de volgende studies konden voortbouwen op deze kennis (De Vetten, Schoonenboom, Keijzer, & Van Oers, 2019b; De Vetten, et al., in druk). De eerste studie was een vragenlijststudie onder 722 eerstejaars pabostudenten van zeven lerarenopleidingen basisonderwijs. De gebruikte vragenlijst bestond uit vijf taken, elk een combinatie van open vragen en correct-incorrect stellingen, die tezamen een beeld geven van de ISI-voorkennis van beginnende pabostudenten. Voor de tweede studie werd het redeneren van drie klassen eerstejaars pabostudenten tijdens een les over ISI onderzocht. In deze les deden de pabostudenten op basis van steekproeven van mannelijke en vrouwelijke pabostudenten een uitspraak over het verschil in de rekenattitude van mannelijke en vrouwelijke pabostudenten in het algemeen.

De opvallendste resultaten van de eerste twee studies waren de volgende. Ten eerste baseerden de meeste studenten hun conclusies op de data, in plaats van op eigen ideeën, hoewel in de vragenlijststudie aanzienlijke minderheden zich baseerden op een eigen mening of ervaring. Ten tweede beschreven veel studenten op een correcte manier de steekproefverdelingen. Het bleef echter vaak bij een beschrijving van de data; de meest studenten maakten geen expliciete generalisatie naar de populatie, waardoor ze ook geen aandacht besteedden aan steekproefgrootte en onzekerheid. Ten derde begrepen veel studenten niet waardoor het mogelijk is om op basis van een steekproef een uitspraak te doen over de populatie. Dit is in de inferentiële statistiek een cruciaal punt: De getrokken steekproef levert een bepaald resultaat op, maar een andere steekproef kan een ander resultaat opleveren. Of steekproeven uit dezelfde populatie tot vergelijkbare resultaten opleveren hangt vervolgens af van de kenmerken van deze steekproeven. Als de gebruikte steekproef voldoende groot is en op een correcte, bij de populatie passende manier geselecteerd is, zullen soortgelijke steekproeven een vergelijkbaar resultaat opleveren, waardoor inferentie op basis van een steekproef mogelijk is. Als studenten dit niet begrijpen, kunnen ze terughoudend zijn om de uitkomst te generaliseren naar de hele populatie.

Interventieonderzoek op de lerarenopleiding basisonderwijs

Straks bespreek ik uitgebreid welke ISI-kennis drie pabostudenten laten zien tijdens het lesgeven over ISI op de basisschool. Deze drie studenten namen deel aan een interventie op de pabo die als doel had de ISI-kennis van de pabostudenten te bevorderen (De Vetten, Schoonenboom, Keijzer, & Van Oers, 2018). Ik licht kort deze interventie en de belangrijkste uitkomsten toe.

De interventie vond plaats op Hogeschool iPabo in een klas van 21 tweedejaars pabostudenten, waarbij ik de docent was. In totaal duurde de interventie vijf (delen van) lessen, waarbij in het eerste deel de nadruk lag op de vakinhoudelijke kennis van de studenten en in het tweede deel op de voorbereiding en nabespreking van de ISI-les die de studenten gaven in hun stageklas. De analyse richt zich op de ontwikkeling van vakinhoudelijke kennis.

De vakinhoudelijke ISI-kennis hebben we geprobeerd met de volgende activiteiten te bevorderen. De eerste activiteit was een opdracht om in de media een voorbeeld te zoeken waarin op basis van een steekproef een inferentie wordt gemaakt over een populatie en deze inferentie kritisch te evalueren. Op deze manier wilden we de studenten bewust maken van de relevantie van het onderwerp. In de tweede activiteit demonstreerde ik door middel van een PC-simulatie dat bij grotere steekproefgroottes, de resultaten van steekproeven op elkaar gaan lijken en dat daardoor een inferentie mogelijk is. De derde activiteit betrof een voorbeeldles die de studenten ook zelf op de basisschool gaven (afbeelding 1). In deze les voerden de studenten zelf een onderzoekje uit. De les zag er als volgt uit: Voorin de klas stond een grote stapel kinderboeken. De vraag daarbij was welk woord het meest voorkomend zou zijn in deze stapel kinderboeken. De enormiteit en zichtbaarheid van de stapel boeken zou het trekken van een steekproef moeten uitlokken. In groepjes bedachten studenten welke woorden mogelijk het meest voorkomend zouden kunnen zijn. In een klassengesprek werd vervolgens een gezamenlijke top vijf van meest waarschijnlijke woorden gemaakt. In

► Afbeelding 1. Samenvatting van het lesplan 'Wat is het meest voorkomende woord?'

kleine groepjes en vervolgens plenair werd daarna besproken op welke manier een steekproef uit de stapel boeken gekozen zou kunnen worden. In groepjes werden data verzameld op basis van de overeengekomen methode van steekproeftrekking. Dit resulteerde per groepje in een antwoord op de vraag naar het meest voorkomende woord in de stapel boeken, en antwoorden op de vraag naar de zekerheid van hun conclusie en manieren om de zekerheid te vergroten. De afzonderlijke steekproefdata werden vervolgens samengevoegd en vergeleken met de resultaten van de groepjes om zo een gesprek over steekproefvariabiliteit op gang te brengen. De vierde activiteit bestond uit een opdracht waarmee het idee van de *equiprobability bias* besproken werd. Deze bias komt voort uit de veronderstelling dat elke uitkomst even waarschijnlijk is.

Inleiding

1. Introduceer het onderwerp (meest voorkomende woorden in kinderboeken) door te bespreken of je alle woorden moet kennen om een boek te begrijpen.
2. Laat leerlingen in kleine groepjes een hypothetische top 3 formuleren van woorden die meest gebruikt zou kunnen worden. Bereik in een klassengesprek overeenstemming over een top 5 en stuur richting een aannemelijke top 5..
3. Vertel de leerlingen dat ze een onderzoekje gaan uitvoeren. Toon de onderzoeksvraag: 'Wat is het meest gebruikte woord in de kinderboeken in de schoolbibliotheek?'

Onderzoeksontwerp en dataverzameling

1. Vraag: 'Hoe kunnen we een antwoord op de onderzoeksvraag kunnen vinden?' Probeer het trekken van een steekproef uit te lokken, bijvoorbeeld door te vragen of alle boeken doorgenomen moeten worden.
2. Stuur het gesprek richting een geschikte manier om een steekproef te trekken. Bijvoorbeeld: een (random) steekproef van boeken en pagina's of regels uit de stapel boeken. Bereik hierover overeenstemming.
3. Geef de leerlingen de hand-out en leg uit hoe ze de data moeten verzamelen: 'Lees het afgesproken aantal pagina's/regels en turf het aantal keer dat elk van de woorden uit de top 5 gebruikt wordt. Tel de totalen.'
4. Laat de leerlingen dit uitvoeren in groepjes van 2 of 3. Leerlingen beantwoorden de vragen op de handout.
5. Laat elk groepje hun resultaat vergelijken met dat van een ander groepje en laat ze tot een gezamenlijk antwoord komen. Benadruk dat er maar 1 woord het vaakst voor kan komen. Laat, als dat mogelijk is, groepjes met verschillende conclusies hun resultaten vergelijken.

Conclusie

1. Hang de grafieken op. Vraag naar de antwoorden van de leerlingen. Organiseer een 'cognitief conflict' door verschillende antwoorden tegenover elkaar te zetten. Stuur het gesprek richting het samenvoegen van alle steekproeven tot 1 steekproef.
2. Maak in het voorbeeld Excel-bestand een staafdiagram van de totalen van alle groepjes.
3. Toon de gezamenlijke grafiek. Probeer tot een gezamenlijk antwoord te komen.
4. Bespreek de zekerheid van de conclusie. Bespreek eventueel of ze denken dat als ze andere boeken uit de stapel/bibliotheek hadden gepakt, het antwoord anders zou zijn. Bespreek hoe de zekerheid van de conclusie vergroot kan worden
5. Evalueer de les door leerlingen te laten benoemen wat ze geleerd hebben.

Voor de analyse maakten we gebruik van een voormeting en een identieke nameting, die gebaseerd waren op de vragenlijst die we gebruikt hadden in de eerste studie. Tijdens de interventie hebben we video- en geluidsopnamen gemaakt en het schriftelijk werk verzameld. Op basis van kwalitatieve en kwantitatieve analyses van deze data hebben we een overzicht gemaakt van de ontwikkeling van de ISI-kennis van de studenten en de mogelijke rol van de activiteiten hierin.

Op een aantal punten was de interventie geslaagd. Aan het einde van de interventie was bijna 90% het ermee eens dat het mogelijk is om een inferentie te maken op basis van een steekproef. De studenten maakten ook bewust inferenties, in tegenstelling tot de vorige studies. De opdracht om in de media te zoeken naar een voorbeeld van een inferentie heeft hier waarschijnlijk mede voor gezorgd, want meer dan driekwart van de studenten besteedde in deze opdracht aan inferenties. Uit de nameting bleek dat driekwart ook leek te begrijpen waardoor het mogelijk is een inferentie te maken. Hierbij hielp dat de opdracht over steekproeven in de media als introductie op de PC-simulatie fungeerde, omdat deze opdracht bij studenten de vraag opriep hoe het mogelijk is om op basis van een steekproef een uitspraak te doen over de populatie, als steekproefuitkomsten variëren. De PC-simulatie bleek studenten vervolgens te helpen om dit te begrijpen.

Op een aantal punten was de interventie minder geslaagd. Ten eerste, hoewel een meerderheid onderschreef dat aselechte steekproeftrekking tot een representatieve steekproef zou leiden, koos bijna iedereen voor een selecte methode als ze zelf een steekproef selecteerden. Ten tweede zagen we geen ontwikkeling in de kennis over wanneer een steekproef voldoende groot is. Ten derde bleek dat, hoewel er onder de studenten overeenstemming bestond dat een grotere steekproef tot meer zekerheid over een inferentie leidt, ze handvatten misten om woorden te geven aan de onzekerheid van hun inferenties.

Kennis van ISI van pabostudenten tijdens lesgeven over ISI op de basisschool

Als pabostudenten over ISI lesgeven, moeten ze hun kennis van ISI kunnen vertalen naar de praktijk van het lesgeven. De transfer van kennis die is opgedaan op de pabo naar de praktijk van het lesgeven is echter niet vanzelfsprekend (Ball et al., 2001; Rowland, Turner, Thwaites, & Hurst, 2009). In de laatste studie heb ik onderzocht in hoeverre het drie studenten die deelnamen aan de interventie op de pabo lukt hun ISI-kennis te laten zien als ze de les 'Wat is het meest voorkomende woord?' geven (De Vetten, Schoonenboom, Keijzer, & Van Oers, 2019a).

Methoden

De studie vormde een meervoudige casestudy van drie studenten, Celine, Demi en Alfred. Demi en Alfred hadden op de havo Wiskunde A gehad, waarin inferentiële statistiek op een formele manier wordt aangeboden. Alle drie participanten scoorden bovengemiddeld op de voormeting, waarbij ze op alle componenten van ISI, behalve steekproefgrootte en methoden van steekproeftrekking, voldoende kennis toonden. Celine gaf les aan groep 5, Demi aan groep 7 en Alfred aan groep 5/6. Op basis van video- en geluidsopnamen van de lessen, notities van de observator en reflectie-interviews die direct na afloop van de les gehouden zijn met de studenten hebben we de ISI-kennis die de studenten tijdens de les lieten zien beschreven. Vanuit een overzicht van hun getoonde ISI-kennis hebben we gezocht naar thema's in de ISI-kennis. Vervolgens hebben we gezocht naar verklaringen om de thema's te begrijpen, waarbij de reflectie-interviews een belangrijke bron vormden. Deze verklaringen hebben we gezocht in 1) de kennis die ze toonden tijdens de voormeting van de interventie en de interventie zelf, 2) factoren gerelateerd aan het ontwerp van de gegeven les, 3) vakdidactische overwegingen, bijvoorbeeld overtuigingen van de studenten wat hun leerlingen wel of niet van ISI zouden kunnen begrijpen, en 4) algemeen pedagogische overwegingen, zoals beslissingen omtrent tijdmanagement.

Resultaten

We bespreken drie belangrijke thema's die terugkwamen in de ISI-kennis van Celine, Demi en Alfred terwijl ze lesgeven over ISI. Deze thema's kwamen bij alle drie studenten terug, hoewel meestal op een verschillende manier.

Thema 1: Gebruik van data als bewijs en bewust generaliseren naar de populatie

Niet als tijdens de interventie op de pabo gebruikten de studenten de data als bewijs en maakten ze generalisaties. Enerzijds kwam dit waarschijnlijk doordat dat het lesplan hen goede mogelijkheden bood om te redeneren over steekproeven en generalisaties te maken. Anderzijds voerden de studenten niet blind het lesplan uit. Het leek erop dat ze de ideeën over ISI geïnternaliseerd hadden en bewust generalisaties maakten. Dit bleek bijvoorbeeld toen Demi aan het einde van de les het leerdoel samenvatte, zonder dat het lesplan dit voorschreef:

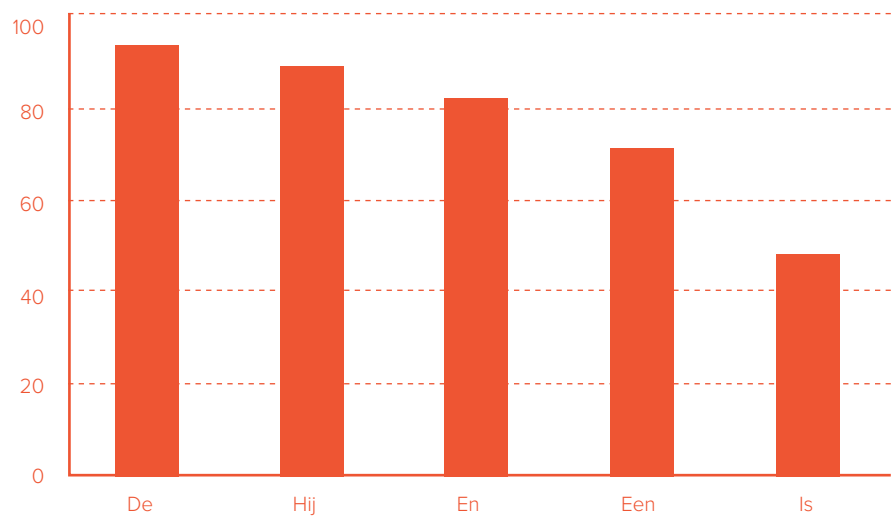
Want wat wij vandaag hebben gedaan is een steekproef. Wij hebben een deel gepakt van de hele boekenkast [...], dat zijn wij gaan onderzoeken. Daar is dit uitgekomen en daar hebben wij een conclusie van getrokken. Dus wij hebben gezegd, we hebben dit gelezen en dit geldt voor de hele boekenkast.

Ook Celine beoogde met de leerlingen inferenties te maken, zoals blijkt uit het reflectie-interview: '...ik had het gevoel, toen ze aan het werk waren met - en ze het idee hadden - dat het echt was: Ja, we doen een steekproef nu.'

Thema 2: Problemen met uitleggen waardoor het maken van inferenties mogelijk is

Het tweede thema gaat over de moeite die de studenten hadden om uit te leggen waardoor het maken van inferenties mogelijk is. Ik beschrijf dit per student, maar uiteindelijk bleek bij allen een onvoldoende beheersing van ISI een belangrijke rol te hebben gespeeld. Ook het ontwerp van de les bleek niet aan te sluiten bij de kern van de uitleg.

► Afbeelding 2. Resultaten van gezamenlijke steekproef in les van Celine



Celine: Inferenties maken als het verschil tussen nummer 1 en 2 klein is

Celine gebruikte in de bespreking van de uitkomsten van de samengevoegde steekproeven aanvankelijk goede argumenten. Ze stemde bijvoorbeeld in met een leerling die opmerkte dat het kleine verschil tussen de nummer 1 en 2 in de steekproef de inferentie onzeker maakte (afbeelding 2):

De 'de' en 'hij' liggen heel dicht bij elkaar, dus dat maakt het wel een beetje lastig want als misschien een groter verschil was, dan hadden we wel kunnen doen, dan hadden we dat misschien wat zekerder kunnen zeggen hè. Dat bedoelde je helemaal goed.

Daarna ging het echter mis. Voortbouwend op bovenstaande observatie vroeg Celine de leerlingen hun hand op te steken als ze vinden dat alle boeken op de stapel gelezen moeten worden. Tot haar verrassing staken alle leerlingen hun hand op. Een leerling legde uit dat alle boeken gelezen moeten worden omdat de boeken die niet gelezen zijn misschien minder vaak 'de' bevat:

[...] er waren best wel veel boeken die we niet hebben gelezen en er zijn allemaal heel veel boeken, ja, eh, er zijn ook veel boeken in Nederland en als daar niet zoveel 'de' in voorkomt [...] Dan kan je niet meteen zeggen: 'ja, 'de' is het meest voorkomende woordje in een boek.' Dat kan je niet zomaar zeggen. Je kan, je kan niet zomaar weten.

Hierop had Celine geen antwoord. Na enige momenten van ongemakkelijke aarzeling eindigde ze abrupt de les.

Uit het reflectie-interview komen drie redenen naar voren waarom Celine er niet in slaagde om te reageren op het idee van de klas om alle boeken te lezen. De eerste reden was organistorisch van

aard: De tijd ontbrak haar. De tweede reden was dat ze compleet verrast was dat alle leerlingen aangaven dat alle boeken gelezen moesten worden. De derde reden was dat ze het idee had dat het herhalen van de uitleg weinig zin had. Onder de derde reden ligt mogelijk een andere reden: In de les werden wel verschillende kleinere steekproeven gebruikt, maar was er maar een grote steekproef. Het lesplan bood Celine geen tools om uit te leggen dat bij voldoende grote steekproeven de meeste steekproeven hetzelfde resultaat opleveren, zelfs als het verschil tussen nummer 1 en 2 klein is. Daarvoor zouden er meerdere grote steekproeven met elkaar vergeleken moeten worden.

Demi: Het negeren van de onzekerheid van inferenties

Demi gebruikte enkele correcte argumenten, zoals dat de steekproef voldoende groot was en dat bij samenvoegen van de kleine steekproeven de invloed van verschillende boeken tegen elkaar opgeheven wordt. Zij liet echter een sterke neiging zien om de onzekerheid van de inferentie te bagatelliseren. Zo keurde zij het idee af om nog een boek te bekijken om de zekerheid te vergroten. Met name aan het einde van de les kwam de neiging om zich te stellig uit te drukken naar voren. Op dat moment claimde dat ze wist dat 'de' het meest voorkomende woord in de hele populatie was, zonder duidelijk te maken hoe ze dat wist:

Demi: Oké, als ik jullie nou vertel dat 'de' het meest voorkomende woord is. Dus ik zeg, 'de' is het meest voorkomende woord en ik vertel dat als wij meer boeken zouden gaan lezen dat 'de' nog steeds op 1 staat. [...] Maar als je dat nou weet en denkt van goh, als ik die hoeveelheid boeken groter maak, blijft nummer 1 nog steeds hetzelfde?

Leerling: Nee.

[...]

Leerling: Ik snap het niet.

Demi: Ik vertel jullie dat als wij de hele boekenkast zouden lezen [...], dat nog steeds 'de' op 1 komt.

Leerling: Nee.

Leerling: Jawel.

Demi: Het is zo. [...]

Leerling: Ja, maar ik snap het niet. Ja, ik snap hem niet.

Demi: Oké, de laatste keer, ik ga het nog een keer proberen. Wij hebben een boekenkast, dit is de boekenkast. Wij hebben zo'n klein stukje van die boekenkast getest, dus daarin hebben wij gekeken welke woorden er het meest voorkomen, toch? Zover, zijn we, oké. Hier, dit is er uitgekomen. 'De' staat op één en 'een' staat twee. Als wij de hele boekenkast zouden lezen, gaan we natuurlijk niet doen, dan komt daar hetzelfde uit.

Leerling: Huh?

Demi: Dan komt er hetzelfde uit, dan staat 'de' nog steeds op 1.

Ze leek hier te suggereren dat het maken van een inferentie mogelijk is omdat een andere steekproef een soortgelijk *zal*, in plaats van *kan*, opleveren. Tijdens het reflectie-interview gaf Demi aan dat haar claim dat ze wist dat 'de' het meest voorkomende woord een verspreking was. Ze had willen zeggen dat als 'de' in de steekproef het meest voorkomend is, dat ze konden aannemen dat ook in de populatie 'de' het meest voorkomende woord zou zijn. Toch laat dit citaat zien dat hoe ze zekerheidstaal (...dan krijgen we hetzelfde resultaat, 'Het is zo', 'Ik vertel jullie...') gebruikte om haar leerlingen te overtuigen.

Tijdens het reflectie-interview gaf ze aan het gevoel te hebben ISI niet onder de knie te hebben. Ze had haar les gedegen voorbereid om beter grip op de materie te krijgen: 'Dus ik heb gisteren [het hele lesplan] door zitten bladeren, zodat ik zekere antwoorden zou hebben, zodat ik zelf ook wist wat het was.' Blijkbaar had ze tevergeefs naar een bevredigend antwoord gezocht. Net als Celine kan Demi de argumenten gemist hebben om de vraag te beantwoorden waardoor het maken van inferenties mogelijk is. Dit had haar er mogelijk toe aangezet om onzekerheid te vermijden. Deze verklaring is in lijn met momenten in de voormeting en de interventie waarin ze zich ook te stellig uitdrukte en waaruit ook een neiging naar voren kwam om te streven naar volkomen zekerheid.

Alfred: Incomplete en incorrecte uitleg en doelloos doorvragen

Alfred liep tegen twee problemen in zijn uitleg waardoor het mogelijk is inferenties te maken. In de eerste plaats was zijn onderbouwing incompleet en deels incorrect. Hij gebruikte bijvoorbeeld het

incorrecte argument dat de steekproefgrootte afhankelijk is van de populatiegrootte. Dit argument is niet alleen op zichzelf genomen incorrect, maar het is ook contraproductief om te onderbouwen dat de gebruikte steekproef gebruikt kan worden voor een inferentie. Correct was zijn argument dat bij een groter verschil in de steekproef tussen de nummer 1 en 2 het waarschijnlijker is dat ook in de populatie de nummer 1 het meest voorkomende woord is. Een argument dat in de bespreking echter ontbrak was dat men kan verwachten dat als de steekproeven voldoende groot zijn de meeste steekproeven hetzelfde meest voorkomende woord zullen hebben. In de tweede plaats had de bespreking van de zekerheid van de inferenties geen richting. Alfred vroeg doelloos door en stemde in met tegenstrijdige uitspraken van de leerlingen. Eerst stemde hij bijvoorbeeld in met een leerling die 50% zeker was omdat niet alle boeken gelezen waren. Daarna stemde hij in met een leerling die 99% zeker was omdat andere boeken waarschijnlijk hetzelfde resultaat zouden opleveren. De instemming leek hiermee eerder een pedagogisch doel te dienen. In de hele bespreking over onzekerheid toonde hij geen kennis die evident tot zijn ISI-kennis behoorde.

Tijdens het reflectie-interview voerde Alfred zijn gebrekkige ISI-kennis op als reden waarom het gesprek over de zekerheid, maar ook over de steekproefgrootte doelloos was:

Ik denk dat dat een beetje te maken heeft met dat ik het onderwerp zelf nog zo lastig vind. Dat ik het, dat ik het dan eigenlijk niet goed kan uitleggen, dat ik er niet vanaf genoeg invalshoeken erin kan komen, snel kan inspelen op wat kinderen zeggen.

Deze uitleg werd ook gestaafd door zijn gedrag tijdens de les: Hij hield het lesplan vast en leek erin te zoeken naar hoe hij verder moest gaan. In tegenstelling tot Demi leek Alfred in de voorbereiding van de les zich ook minder ingespannen te hebben om zich de stof eigen te maken. Zijn lesplan was namelijk grotendeels een kopie van het lesplan dat door de docent ter beschikking was gesteld.

Conclusie thema 2

Alle drie studenten hadden moeite uit te leggen waardoor het mogelijk is inferenties te maken. Deze moeite is deels toe te schrijven aan hun gebrekkige ISI-kennis, zoals zij zelf ook aangaven. In het bijzonder ontbrak het hen aan de tools om uit te leggen hoe steekproefvariatie werkt in de context van de les die ze gaven.

Thema 3: Veel aandacht voor steekproefgrootte, weinig aandacht voor methode van steekproeftrekking

Alle drie studenten besteedden uitgebreid aandacht aan de grootte van de steekproef en zorgden ervoor dat de steekproef groot genoeg was en binnen afzienbare tijd geselecteerd kan worden, bijvoorbeeld door elk kind een pagina of tien regels uit een boek te laten onderzoeken. Alleen Celine deed echter een poging de methode van steekproeftrekking te bespreken. In haar les kwamen leerlingen echter niet met bruikbare suggesties en ze besloot er daarom verder niet op door te gaan en leerlingen zelf een boek van de stapel te laten pakken. Tijdens het reflectie-interview bleek dat ze hierbij goed doorhad dat dit een goede manier was om de steekproef te trekken: 'Nou, neem lekker allemaal maar een random boek, dan laten we het gewoon lekker. Het is genoeg zo.' Demi en Alfred bespraken de methode van steekproeftrekking niet. Demi liet ook de leerlingen zelf een boek pakken, wat ook in een min of meer aselechte steekproef resulteerde. Daarnaast gaf ze aan dat de achterkant van een boek niet representatief is voor de binnenkant ervan. Alfred liet zijn leerlingen het boek uit hun laasje gebruiken, die over het algemeen van een lager niveau waren die dan de boeken op de stapel, zodat deze steekproef mogelijk niet representatief was.

Een eerste mogelijke verklaring waarom de studenten veel aandacht aan de steekproefgrootte besteedden en niet aan de methode van steekproeftrekking was dat voor een goed verloop van de les de steekproef niet te groot moest zijn, zodat de data snel verzameld konden worden. Het bespreken van de methode van steekproeftrekking was voor een goed verloop niet nodig: Hoe de boeken ook gekozen zouden worden, de meeste methoden waren praktisch mogelijk en het zou niet erg opvallen als de steekproef niet representatief was. Een tweede verklaring zou kunnen zijn dat in het lesplan weinig aandacht besteed werd aan hoe de methode van steekproeftrekking bespreken kon worden en wat verschillende typen boeken voor invloed zouden kunnen hebben op het voorkomen van bepaalde woorden in die boeken.

► Afbeelding 3. Overzicht lessenserie over ISI voor de pabo

Een ISI-lessenserie voor de pabo

Nu statistiek, en mogelijk ook ISI, een grotere rol lijkt te gaan krijgen in rekenen-wiskunde op de basisschool, wordt de vraag hoe pabo-studenten ISI kunnen leren urgent. Hieronder schets ik een lessenserie die in grote lijnen de lessenserie volgt die ik gebruik heb in mijn interventie op de pabo (afbeelding 3). De lessenserie beslaat vijf bijeenkomsten, die variëren in duur van 45 tot 90 minuten. Onderdeel van de lessenserie is dat studenten zelf een ISI-les geven in de bovenbouw van de basisschool. Voorbereiding en evaluatie van deze les zijn ook deel van de lessenserie.

Bijeenkomst op de pabo	Duur (min)
Bijeenkomst 0: Instructie huiswerkopdracht 'Steekproeven in de media'	10
Bijeenkomst 1:	45
• Bespreking 'Steekproeven in de media'	
• Instructie huiswerkopdracht 'Inferenties op school'	
Bijeenkomst 2: Voorbeeldles 'Wat is het meest voorkomende woord?'	60
Bijeenkomst 3: PC-simulaties	60-90
Bijeenkomst 4: Voorbereiding stageles	60
Op stage geven studenten de les 'Wat is het meest voorkomende woord?'	
Bijeenkomst 5:	60
• Evaluatie stagelessen	
• Bespreking 'Inferenties op school'	

Studenten krijgen als voorbereiding op de lessenserie de volgende huiswerkopdracht mee: 'Zoek in de media naar een voorbeeld waarin op basis van een steekproef een inferentie wordt gemaakt over een populatie. Schrijf een korte kritische evaluatie: In hoeverre is de conclusie gerechtvaardigd op basis van het uitgevoerde onderzoek?' Doel van deze opdracht is studenten bewust te maken dat inferenties vaak gemaakt worden, van het onderscheid tussen steekproef en populatie en om het probleem dat verschillende steekproeven verschillende resultaten op kunnen leveren op tafel te krijgen. Ook helpt de opdracht hen bewust te maken welke ISI-kennis ze wel en niet hebben. In de volgende activiteit zal hierop voortgebouwd worden.

In de eerste bijeenkomst op de pabo wordt de huiswerkopdracht besproken. In groepjes bespreken studenten vragen over opvallende fouten, steekproefgrootte, methoden van steekproefrekening en de mogelijkheid om inferenties te maken. De bespreking wordt afgesloten met een individuele en gezamenlijke inventarisatie van onderwerpen waarover de studenten kennis missen. Vervolgens krijgen de studenten een vervolgoopdracht die tijdens de duur van de lessenserie zelfstandig moeten maken: 'Zoek in je stageschool een voorbeeld waar een van de schoolmedewerkers een informele statistische inferentie maakt. Beschrijf dit voorbeeld en schrijf een kritische evaluatie van de inferentie op basis van de kennis die je hebt opgedaan tijdens de lessenserie.'

In de tweede bijeenkomst geeft de opleider de les 'Wat is het meest voorkomende woord?' als voorbeeld hoe op de basisschool een ISI-les gegeven kan worden (afbeelding 4). Deze les bevordert daarnaast de ISI-kennis van de studenten zelf. Deze les is licht aangepast ten opzichte van de versie die tijdens de interventie is gebruikt. In de eerste plaats wordt nu expliciet voortgeborduurd op de activiteit 'Steekproeven in de media', waarin de vraag opgeworpen werd waardoor het maken van inferenties mogelijk is. Als studenten op de basisschool de les 'Wat is het meest voorkomende woord?' geven kunnen ze iets soortgelijks doen door aan het begin van de les een artikel uit de media te bespreken waarin een uitspraak op basis van een steekproef gedaan wordt. Op deze manier kan ook in de les die de pabostudenten op de basisschool geven het probleem dat verschillende steekproeven verschillende resultaten kunnen opleveren op tafel komen. Een tweede, meer technische aanpassing, is dat de focus niet meer ligt op welk deel van de steekproeven hetzelfde meest voorkomende woord (de modus) oplevert, maar op de frequenties van de twee woorden die het meest gevonden zijn. De leraar inventariseert hoe vaak elke leerling deze twee woorden gevonden heeft. Het voordeel hiervan is dat frequenties na samenvoeging al bij kleinere steekproefgroottes

► Afbeelding 4. Samenvatting van het aangepaste lesplan 'Wat is het meest voorkomende woord?'. Deze les kunnen pabostudenten ook op de basisschool gebruiken, waarbij ze kunnen starten met het bespreken van een voorbeeld van steekproeven uit de media

naar hetzelfde resultaat convergeren dan modi. Door te kijken naar de frequenties in plaats van de modi kan de leraar dus gemakkelijker laten zien dat een inferentie op basis van een steekproef mogelijk is, omdat andere steekproeven van dezelfde grootte ongeveer hetzelfde resultaat opleveren.

Inleiding

1. Blik terug op de conclusie van de activiteit 'Steekproeven in de media': Hoe kun je op basis van een steekproef een uitspraak doen over de populatie? Zou een andere steekproef niet een ander resultaat opleveren? Vertel dat het doel van de les is te onderzoeken wanneer steekproeven een heel ander resultaat oplevert en wanneer niet.
2. Introduceer het onderwerp (meest voorkomende woorden in kinderboeken) door te bespreken of je alle woorden moet kennen om een boek te begrijpen.
3. Laat studenten in kleine groepjes een hypothetische top 3 formuleren van woorden die meest gebruikt zou kunnen worden. Bereik in een klassengesprek overeenstemming over een top 5 en stuur richting een aannemelijke top 5.
4. Vertel de studenten dat ze een onderzoekje gaan uitvoeren. Toon de onderzoeksvraag: 'Wat is het meest gebruikte woord in de stapel kinderboeken?'

Onderzoeksontwerp en dataverzameling

1. Vraag: 'Hoe kunnen we een antwoord op de onderzoeksvraag vinden?' Probeer het trekken van een steekproef uit te lokken, bijvoorbeeld door te vragen of alle boeken doorgenomen moeten worden.
2. Bespreek hoe en hoeveel boeken, pagina's en woorden geselecteerd moeten worden. Stuur het gesprek richting een geschikte manier om een steekproef te trekken, waarbij gestreefd wordt naar een relatief klein aantal woorden per student. Bijvoorbeeld: elk kind de eerste 50 woorden van een willekeurige pagina van een willekeurig boek.
3. Geef de studenten de hand-out en leg uit hoe ze de data moeten verzamelen: 'Lees het afgesproken aantal woorden en turf het aantal keer dat elk van de woorden uit de top 5. Tel de totalen.' Laat studenten dit individueel uitvoeren.
4. Studenten verzamelen de data en beantwoorden de vragen op de hand-out (Wat is denk je het meest voorkomende woord in de stapel kinderboeken is? Hoe zeker ben je van je antwoord? Hoe zou je nog zekerder van je antwoord kunnen worden?)

Bespreking individuele steekproeven

1. Inventariseer wie welk woord als meest voorkomend heeft gevonden. Benoem dat verschillende steekproeven tot verschillende uitkomsten leiden en zoom in op de twee woorden die het meest als modus zijn gevonden (woord 1 en woord 2).
2. Vraag elke student naar de frequentie van woord 1 en woord 2. Maak hiervan een grafiek bijvoorbeeld door kruisjes in twee kleuren te zetten boven een getallenlijn. Wijs op de grote overlap tussen de frequenties van woord 1 en woord 2.
3. Vraag: 'Weten we het nu zeker genoeg om iets over de hele stapel boeken te kunnen zeggen?' Stuur het gesprek naar: 'De uitkomsten variëren nog best sterk: De ene keer is het dit, de andere keer is dat. Ik durf het niet te zeggen.'
4. Vraag: 'Hoe zouden we het zekerder kunnen weten?' Stuur het gesprek naar het nemen van een grotere steekproef. Vraag: 'Wat denken jullie dat er gebeurt met de resultaten als de steekproeven twee keer zo groot zouden zijn?' Stel voor dit uit te gaan zoeken door studenten in tweetallen hun steekproeven samen te laten voegen.

In tweetallen individuele steekproeven samenvoegen

1. Laat studenten in tweetallen hun steekproeven samenvoegen.
2. Inventariseer de nieuwe totalen van woord 1 en woord 2 en maak een nieuwe grafiek onder de eerste grafiek. Nu is het aantal steekproeven de helft.

Bespreking samengevoegde steekproeven en conclusie

1. Laat zien dat de frequenties van de samengevoegde steekproeven dichter bij elkaar liggen dan de frequenties van de individuele steekproeven en dat er minder overlap is tussen de frequenties van woord 1 en woord 2.
2. Vraag opnieuw wie er nu zeker genoeg is om iets over de hele stapel boeken te kunnen zeggen.
3. Vraag wat ze denken wat er gebeurt als de steekproeven nogmaals worden samengevoegd, dus als de steekproeven nog groter worden. (Antwoord: Hoe groter de steekproeven, hoe meer de verschillende resultaten van woord 1 op elkaar gaan lijken en hoe minder overlap er is tussen de resultaten van woord 1 en woord 2.)
4. Voeg alle steekproeven samen tot een grote steekproef. Vraag: 'We hebben nu een heel grote steekproef van alle onderzochte woorden. Wat verwachten jullie: Wat zou er uit komen als we nog een keer zo veel woorden zouden onderzoeken?' (Antwoord: Als je steekproef maar groot genoeg is, zal een andere steekproef ongeveer hetzelfde opleveren.)
5. Leg een link met het besproken artikel in de les 'Steekproeven in de media' door te bespreken dat omdat daar de steekproef (wel/niet) groot genoeg is, het (wel/pas bij een grotere steekproef) mogelijk is een uitspraak te doen over de hele populatie van [zoals gebruikt in artikel].
6. Vertel een korte conclusie. Bijvoorbeeld: 'In de krant of op internet worden heel vaak uitspraken gedaan op basis van een steekproef. Dan weet je nu waar je op moet letten: Hoe groot is die steekproef? En hebben ze de steekproef wel goed geselecteerd? Als de steekproef groot genoeg is, dan kun je dus een uitspraak doen over alle [vul populatie uit artikel in] of boeken, ook al is die uitspraak altijd een beetje onzeker. Maar handig is het wel: Je kunt met voldoende zekerheid iets zeggen over iedereen zonder dat je iedereen hoeft te ondervragen!'

In de derde bijeenkomst wordt de overstap gemaakt van het handmatig herhaald steekproeftrekken van bijeenkomst 2 naar het simuleren van herhaald steekproeftrekken met behulp van de PC. Hierbij wordt gebruik gemaakt van een programma zoals VUstat, dat gratis beschikbaar is. Studenten experimenteren met simulaties van herhaald steekproeftrekken en ze verkennen het effect van steekproefgrootte op uitkomsten van verschillende steekproeven. De situatie van de voorbeeldles kan als uitgangspunt genomen worden en worden gemodelleerd in het simulatieprogramma. Door middel van een nog te ontwikkelen activiteit kunnen de gevolgen besproken worden van aselechte en selecte methoden van steekproeftrekking op de representativiteit van steekproeven.

In de vierde bijeenkomst wordt het geven van de les 'Wat is het meest voorkomende woord?' voorbesproken. De didactiek van ISI, met aandacht voor onder andere valkuilen, tips en typische redeneringen van leerlingen, komen aan bod. Vervolgens voeren studenten de les uit op hun stageschool aan de hand van het aangereikte lesplan. Tijdens de vijfde bijeenkomst wordt de stageles geëvalueerd. Eventueel kan ervoor gekozen worden, zoals ik in mijn onderzoek heb gedaan, eerst de ene helft van de klas de les te laten geven en een tussentijdse evaluatie te houden. Op deze manier kan de tweede helft van de klas van de ervaringen van de eerste helft leren. In de laatste bijeenkomst kan ook de activiteit 'Inferenties op school' besproken worden. Deze bespreking kan zichtbaar maken wat er over ISI geleerd is. Eventueel kan deze activiteit ook als toetsinstrument ingezet worden (zie hieronder).

Toetsing

Er zijn verschillende mogelijkheden om de ISI-kennis (vakinhoudelijk en vakdidactisch) te toetsen. Te denken valt aan de volgende toetsvormen:

- Om de vakdidactische kennis te toetsen kan een verslag van de gegeven stageles gebruikt worden. Mogelijke toetscriteria zijn:
 - o In hoeverre slaagt de student erin door middel van open vragen aanpakken bij leerlingen uit te lokken?
 - o Hoe legt de student uit waardoor het maken van inferenties mogelijk is?
 - o In hoeverre blijkt uit de reacties van de leerlingen dat de student erin geslaagd is het

► Afbeelding 5. Gedeelte van Taak 2 uit het instrument gebruikt in De Vetten, et al. (in druk)

probleem helder te maken dat verschillende steekproeven verschillende resultaten kunnen opleveren?

- Om de vakinhoudelijke kennis te toetsen kan het instrument dat wij in onze vragenlijststudie hebben gebruikt ingezet worden. Dit instrument bestaat uit open en gesloten vragen. Afbeelding 5 toont een deel van de tweede taak waarmee onder andere kennis van methoden van steekproeftrekking werd gemeten. Het volledige instrument is opgenomen als bijlage 1.
- Om de vakinhoudelijke kennis te toetsen kan de activiteit 'Inferenties op school' gebruikt worden. Er kan bijvoorbeeld getoetst worden in hoeverre de drie ISI-componenten (zie Inleiding) betrokken worden in de evaluatie van de inferentie.

Het Ministerie van Onderwijs wil onderzoek doen naar de vraag wie er in Nederland in het algemeen meer plezier hebben in rekenen: jongens of meisjes. Het Ministerie vraagt jou een vragenlijst voor te leggen aan een groep jongens en aan een groep meisjes uit groep 5 tot en met 8 van de basisschool.

Vraag: Op welke wijze zou je de groep jongens en de groep meisjes selecteren aan wie je de vragenlijst wilt voorleggen, zodat het mogelijk is een uitspraak te doen over alle jongens en meisjes uit groep 5 tot en met 8 van de basisschool?

Antwoord + toelichting:

Je krijgt uitspraken te lezen van andere studenten die de opdracht hebben gemaakt. Geef bij elke uitspraak aan of je hun redenering correct vindt.

Demy: 'Ik zou de vragenlijst voorleggen aan de kinderen in mijn klas. Zij vormen een goede afspiegeling van alle jongens en meisjes in Nederland.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Nick: 'Ik selecteer uit alle basisschoolkinderen (groep 5 t/m 8) willekeurig een groep jongens en een groep meisjes. Om deze selectie te kunnen maken, hoef ik niet te weten wat de achtergrond van de geselecteerde kinderen is en op wat voor soort school ze zitten.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Conclusie

Doel van mijn promotieonderzoek was om te onderzoeken hoe de ISI-kennis van pabostudenten bevorderd kan worden. Over het algemeen is de ontwikkeling in ISI-kennis van pabostudenten tijdens een korte interventie veelbelovend. Aan het einde van de interventie leek ongeveer driekwart van de studenten kernelementen van ISI te begrijpen: Ze onderkenden de waarde van data als bewijs, toonden basaal begrip van hoe steekproeven variëren en hoe het mogelijk is om op basis van een steekproef een uitspraak over de populatie te doen. De observatie van drie van de deelnemende studenten die een ISI-les gaven liet zien dat zij in staat zijn zelf inferenties te maken. De observaties maakten echter ook duidelijk dat de uitdaging vooral ligt in hoe pabostudenten leerlingen kunnen uitleggen waardoor het mogelijk is om op basis van een steekproef een uitspraak te doen over de populatie.

Wat betreft vervolgonderzoek is een eerste onderwerp hoe pabostudenten geleerd kan worden erop te vertrouwen dat aselechte steekproeftrekking in een representatieve steekproef resulteert. Mijn onderzoek liet namelijk zien dat pabostudenten een opvallende neiging hebben tot het verkiezen van selecte methoden om een steekproef te selecteren. Een tweede thema betreft de lengte van de lessenserie. We hebben gekozen voor een korte lessenserie voor de pabo, zodat deze gemakkelijk implementeerbaar is. Dit heeft mogelijk als keerzijde dat pabostudenten minder diepe kennis in ISI ontwikkelen. Hierdoor moeten ze wellicht in het lesgeven over ISI in de basisschool meer gestuurd worden. Tot slot is ontwerponderzoek nodig om de voorgestelde lessenserie verder

te ontwikkelen, te testen en te evalueren. Hierin is samenwerking met pabo's onmisbaar. Hiermee nodig ik u dan ook van harte uit om samen te werken in verdere ontwikkeling van onderwijs over ISI voor pabo en primair onderwijs!

Als ISI inderdaad onderdeel van het nieuwe curriculum van het primair onderwijs wordt, dan zullen ook (aanstaand) leraren zich moeten verdiepen in de vakinhoudelijke en vakdidactische kennis van ISI. Ik hoop dat met mijn onderzoek een basis is gelegd voor verdere ontwikkeling van ISI-onderwijs voor pabo en primair onderwijs, zodat de leraren van de toekomst kunnen leren zelf kritisch tegenover uitspraken op basis van data te staan en dat zij deze kritische houding ook aan hun leerlingen kunnen meegeven.

Literatuur

- Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching (4th edition)* (pp. 433-456). New York, NJ: MacMillan.
- Ben-Zvi, D., Bakker, A., & Makar, K. (2015). Learning to reason from samples. *Educational Studies in Mathematics*, 88(3), 291-303.
- De Vetten, A. (2018). *From sample to population. Pre-service primary school teachers learning informal statistical inference*. (Doctoral dissertation), Retrieved from <http://dare.ubvu.vu.nl/handle/1871/55880>
- De Vetten, A., Schoonenboom, J., Keijzer, R., & Van Oers, B. (2018). The development of informal statistical inference content knowledge of pre-service primary school teachers during a teacher college intervention. *Educational Studies in Mathematics*, 99(2), 217-234. doi:10.1007/s10649-018-9823-6
- De Vetten, A., Schoonenboom, J., Keijzer, R., & Van Oers, B. (2019a). *Knowledge in action. The informal statistical inference content knowledge of three pre-service teachers while teaching ISI in upper-primary school*. Manuscript in preparation.
- De Vetten, A., Schoonenboom, J., Keijzer, R., & Van Oers, B. (2019b). Pre-service teachers and informal statistical inference: Exploring their reasoning during a growing samples activity. In G. Burrill & D. Ben-Zvi (Eds.), *Topics and trends in current statistics education research: International perspectives*. New York, NY: Springer.
- De Vetten, A., Schoonenboom, J., Keijzer, R., & Van Oers, B. (in druk). Pre-service primary school teachers' knowledge of informal statistical inference. *Journal of Mathematics Teacher Education*. doi:10.1007/s10857-018-9403-9
- Gravemeijer, K., Stephan, M., Julie, C., Lin, F.-L., & Ohtani, M. (2017). What mathematics education may prepare students for the society of the future? *International Journal of Science and Mathematics Education*, 15(1), 105-123.
- Groth, R. E., & Meletiou-Mavrotheris, M. (2018). Research on statistics teachers' cognitive and affective characteristics. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 327-355). Cham, Switzerland: Springer.
- Liu, Y., & Grusky, D. B. (2013). The payoff to skill in the third industrial revolution. *American Journal of Sociology*, 118(5), 1330-1374.
- Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, 8(1), 82-105.
- Ontwikkelteam Rekenen & Wiskunde. (2019). *Conceptvoorstellen leergebied rekenen en wiskunde*. Retrieved from <https://curriculum.nu/wp-content/uploads/2019/05/Conceptvoorstellen-Rekenen-en-Wiskunde.pdf>.
- Rowland, T., Turner, F., Thwaites, A., & Hurst, C. (2009). *Developing primary mathematics teaching: Reflecting on practice with the knowledge quartet*. London, UK: Sage.

In today's society, it is important that children learn to critically reflect on generalizations from samples. The foundation for this ability can be laid in primary school, by engaging primary school students in making informal inferences about a population based on sample data (informal statistical inference, ISI). This paper reports on the results of my PhD research on how pre-service primary school teachers can learn ISI. The results show that a majority of pre-service teachers who participated in a short intervention seemed to understand the core ISI elements. However, during teaching ISI in primary classrooms, pre-service teachers faced difficulties when trying to explain why making inferences is possible. Based on these results, I propose a lesson series for ISI for primary education teacher colleges.

BIJLAGE 1

Vragenlijst Redeneren over steekproeven

Opdracht 1: Het rekenplezier van jongens en meisjes

In je klas wordt een onderzoek besproken naar het plezier in rekenen van jongens en meisjes in groep 5 tot en met 8 van de basisschool in Nederland in het algemeen. Dit onderzoek is uitgevoerd door vierdejaars pabostudenten.

De vierdejaars studenten hebben voor hun onderzoek aan 1000 jongens en 1000 meisjes een vragenlijst voorgelegd met daarin de vraag hoeveel plezier ze hebben in rekenen. Het is niet bekend hoe de onder-vraagde jongens en meisjes precies geselecteerd zijn, maar wel is bekend dat de 1000 jongens en 1000 meisjes een zo goed mogelijke afspiegeling vormen van alle jongens en meisjes in de basisschoolleeftijd (groep 5 t/m 8) in Nederland.

De conclusie van de vierdejaars is dat er geen verschil is in rekenplezier tussen jongens en meisjes. Een groot deel van je klasgenoten is verbaasd over deze conclusie: In de klas op hun stageschool is er wel een verschil in rekenplezier: jongens hebben meer plezier in rekenen dan meisjes.

Vraag: Met welke conclusie ben je het eens: met de conclusie van de vierdejaars studenten of met de conclusie van je klasgenoten?

Antwoord + toelichting:

Klasgenoten Laura en Freek zijn het eens met de vierdejaars studenten die het onderzocht hebben. Hieronder staat waarom. Geef bij beide uitspraken aan of je hun argument correct vindt.

Laura: 'Ik ben het eens met de vierdejaars, omdat zij hun conclusie baseren op een grote, goed gekozen groep jongens en meisjes.'

- Correct argument (1)
- Incorrect argument (2)

Freek: 'Het rekenplezier is anders voor elk kind en hangt niet af van of een kind een jongen of meisje is.'

- Correct argument (1)
- Incorrect argument (2)

Klasgenoten Iris en Romy zijn het juist eens met de rest van je klas. Hieronder staat waarom. Geef bij beide uitspraken aan of je hun argument correct vindt.

Iris: 'Ook in mijn klas is het zo dat jongens meer plezier in rekenen hebben dan meisjes.'

- Correct argument (1)
- Incorrect argument (2)

Romy: 'Het is algemeen bekend dat jongens meer plezier hebben in rekenen dan meisjes.'

- Correct argument (1)
- Incorrect argument (2)

Opdracht 2: Selecteren van een groep jongens en een groep meisjes

Het Ministerie van Onderwijs wil onderzoek doen naar de vraag wie er in Nederland in het algemeen meer plezier hebben in rekenen: jongens of meisjes. Het Ministerie vraagt jou een vragenlijst voor te leggen aan een groep jongens en aan een groep meisjes uit groep 5 tot en met 8 van de basisschool.

Vraag: Op welke wijze zou je de groep jongens en de groep meisjes selecteren aan wie je de vragenlijst wilt voorleggen, zodat het mogelijk is een uitspraak te doen over alle jongens en meisjes uit groep 5 tot en met 8 van de basisschool?

Antwoord + toelichting:

Je krijgt uitspraken te lezen van andere studenten die de opdracht hebben gemaakt. Geef bij elke uitspraak aan of je hun redenering correct vindt.

Demy: 'Ik zou de vragenlijst voorleggen aan de kinderen in mijn klas. Zij vormen een goede afspiegeling van alle jongens en meisjes in Nederland.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Nick: 'Ik selecteer uit alle basisschoolkinderen (groep 5 t/m 8) willekeurig een groep jongens en een groep meisjes. Om deze selectie te kunnen maken, hoef ik niet te weten wat de achtergrond van de geselecteerde kinderen is en op wat voor soort school ze zitten.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Michelle: 'Ik selecteer een groep van 200 jongens en een groep van 200 meisjes. Want volgens mij is een groep van 200 jongens en een groep van 200 meisjes groot genoeg om iets te kunnen zeggen over jongens en meisjes in het algemeen.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Rick: 'Om iets te kunnen te zeggen over jongens en meisjes in het algemeen, is het beter een groep van 100.000 jongens en een groep van 100.000 meisjes te selecteren dan een groep van 3.000 jongens en een groep van 3.000 meisjes.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Naomi: 'Om iets te kunnen zeggen over jongens en meisjes in het algemeen, moet de groep jongens ongeveer even groot zijn als de groep meisjes. Een groep van 1200 jongens en een groep van 500 meisjes kan ik bijvoorbeeld niet met elkaar vergelijken.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Emma: 'Om iets te kunnen zeggen over jongens en meisjes in het algemeen, kies ik van elke groep (5 t/m 8) even veel jongens als meisjes, van elk schooltype even veel jongens als meisjes, van elk rekenniveau even veel jongens als meisjes, et cetera.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Lotte: 'Welke groepen we ook selecteren, uiteindelijk kunnen we niets zeggen over jongens en meisjes in het algemeen. Als we andere jongens en meisjes zouden hebben ondervraagd, kan de uitkomst heel anders zijn.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Het is bekend dat er een aantal factoren is dat grote invloed heeft op het rekenplezier van kinderen, zoals de rekenvaardigheid en leeftijd. Gebruik deze informatie om te reageren op de volgende uitspraak:

Eva: 'Ook al hebben we een zorgvuldige selectie gemaakt van jongens en meisjes, we weten nooit zeker of de uitkomst van het onderzoek geldt voor jongens en meisjes in het algemeen, omdat de uitkomst ook veroorzaakt kan worden door verschillen in bijvoorbeeld rekenvaardigheid of leeftijd.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Opgdracht 3: Het rekenplezier van jongens

Aan 20 jongens is de stelling 'Ik heb plezier in rekenen' voorgelegd. Ze konden de stelling beantwoorden op een schaal van 0 tot 100, waarbij 0 betekent dat ze helemaal geen plezier hebben in rekenen en 100 dat ze heel veel plezier hebben in rekenen.

Men heeft er zoveel mogelijk voor gezorgd dat er van diverse groepen, van diverse schooltypen, van diverse intelligentieniveaus, etc. jongens zijn geselecteerd, zodat de 20 jongens een zo goed mogelijke afspiegeling vormen van alle jongens in de basisschoolleeftijd (groep 5 t/m 8) in Nederland.

Op het volgende scherm zie je de gegevens van het rekenplezier van de 20 jongens. Elk bolletje staat voor 1 jongen.

We vragen je een gedachte-experiment te doen: Stel dat we nog een keer 20 jongens op dezelfde manier zouden selecteren en ondervragen, hoe denk je dat de grafiek er dan uit zou kunnen zien? Als je deze informatie goed gelezen hebt, klik dan op volgende (om technische redenen kun je deze informatie niet teruglezen).



Stel dat we nog een keer 20 jongens op dezelfde manier zouden selecteren en ondervragen, hoe denk je dat de grafiek er dan uit zou kunnen zien? Klik 20 bolletjes aan. Je kunt een reeds geplaatst bolletje verwijderen door er nogmaals op te klikken. Op de volgende pagina kun je eventueel toelichting geven. (Let op: tijdens plaatsen bolletjes niet naar beneden scrollen!)

Toelichting (niet verplicht):

Andere studenten hebben ook dit gedachte-experiment gedaan. Je krijgt hun uitleg te lezen waarom ze de grafiek op een bepaalde manier hebben afgemaakt. Geef aan of je hun werkwijze correct vindt.

Max: 'Als we nog 20 jongens op dezelfde manier zouden selecteren en ondervragen, dan verwacht ik dat de grafiek er als volgt uit gaat zien:

Ik denk dat voor alle jongens geldt dat er ongeveer even veel jongens beneden de meest voorkomende waarde (60) scoren als erboven. Daarom verdeel ik de extra 20 bolletjes zo, dat de totale verdeling gelijkmatig en symmetrisch is.'

- Correcte werkwijze (1)
- Incorrecte werkwijze (2)

Amber: 'Als we nog 20 jongens op dezelfde manier selecteren en ondervragen, dan verwacht ik dat de grafiek er als volgt uit gaat zien:

We hebben nu 40 in plaats van 20 bolletjes. Dat zijn er twee keer zo veel. Ik verdubbel alles.'

- Correcte werkwijze (1)
- Incorrecte werkwijze (2)

Niels: 'Als we nog 20 jongens op dezelfde manier zouden selecteren en ondervragen, dan verwacht ik dat de grafiek er als volgt uit gaat zien:

Ik reken het gemiddelde van de eerste 20 uit: dat is ongeveer 65. Ik denk dat het gemiddelde van de 40 jongens in totaal ook ongeveer 65 zal zijn. Ik teken dus de extra 20 zo dat het gemiddelde ongeveer gelijk blijft.'

- Correcte werkwijze (1)
- Incorrecte werkwijze (2)

Fatima: 'Als we nog 20 jongens op dezelfde manier selecteren en ondervragen, dan verwacht ik dat de grafiek er als volgt uit gaat zien:

Ik maak in ieder geval de verdeling iets breder: ik denk dat er bij de 20 extra jongens waarschijnlijk een of enkele jongens zitten die lager scoren dan 50.'

- Correcte werkwijze (1)
- Incorrecte werkwijze (2)

Bram: 'Als we nog 20 jongens op dezelfde manier selecteren en ondervragen, dan verwacht ik dat de grafiek er als volgt uit gaat zien:

Ik denk dat hoe meer bolletjes er zijn, hoe vloeiender de verdeling van de bolletjes is. Daarom verdeel ik de extra bolletjes zo, dat er minder pieken en dalen zijn.'

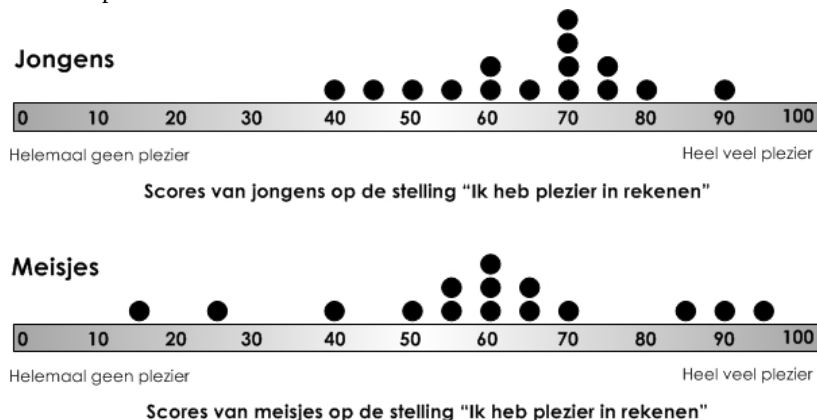
- Correcte werkwijze (1)
- Incorrecte werkwijze (2)

Opdracht 4: Het rekenplezier van jongens en meisjes (2)

Om een antwoord te vinden op de vraag of er in het algemeen een verschil is in rekenplezier tussen jongens en meisjes, is aan 15 jongens en 15 meisjes de stelling 'Ik heb plezier in rekenen' voorgelegd. Ze konden de stelling beantwoorden op een schaal van 0 tot 100, waarbij 0 betekent dat ze helemaal geen plezier hebben in rekenen en 100 dat ze heel veel plezier hebben in rekenen.

Men heeft er zoveel mogelijk voor gezorgd dat er van diverse groepen, van diverse schooltypen, van diverse intelligentieniveaus, etc. jongens en meisjes zijn geselecteerd, zodat de 15 jongens en 15 meisjes een zo goed mogelijke afspiegeling vormen van alle jongens en meisjes in de basisschooleeftijd (groep 5 t/m 8) in Nederland.

In onderstaande grafiek zie je de gegevens van het rekenplezier van de 15 jongens en de 15 meisjes. Elk bolletje staat voor 1 persoon.



Beantwoord op basis van de informatie van de 15 jongens en 15 meisjes de volgende vraag voor jongens en meisjes in de basisschooleeftijd (groep 5 t/m 8) in Nederland in het algemeen: Van wie is het plezier in rekenen in het algemeen groter: van jongens of van meisjes? Of is er geen verschil? Licht je antwoord toe.

Antwoord + toelichting:

Je krijgt uitspraken te lezen van andere studenten die de opdracht hebben gemaakt. Geef bij elke uitspraak aan of je hun redenering correct vindt.

Dani: 'Jongens hebben in het algemeen iets meer plezier in rekenen dan meisjes, omdat in de grafieken de meest voorkomende waarde bij de jongens 70 is en dat is hoger dan de meest voorkomende waarde bij de meisjes (60).'

- Correcte redenering (1)
- Incorrecte redenering (2)

Max: 'Jongens hebben in het algemeen iets meer plezier in rekenen dan meisjes, omdat er maar twee jongens negatief (score lager dan 50) zijn, terwijl er drie meisjes negatief zijn.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Mike: 'De scores van zowel jongens als meisjes liggen voor een groot deel tussen 40 en 90. Kortom, de grafieken van de jongens en meisjes overlappen voor een groot deel. Ik denk daarom dat er geen verschil is in rekenplezier tussen jongens en meisjes in het algemeen.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Simon: 'Het gemiddelde van de 15 jongens is 65 en het gemiddelde van de 15 meisjes is 59,3. Mijn conclusie baseer ik in ieder geval op dit verschil.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Johanna: 'Ik baseer mijn conclusie in ieder geval ook op hoe de waarden van jongens en meisjes verspreid zijn: de waarden van de meisjes zijn erg verspreid.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Britt: 'Er zijn maar 15 jongens en 15 meisjes ondervraagd. We kunnen daarom helemaal niets zeggen over het verschil in rekenplezier voor jongens en meisjes in het algemeen.'

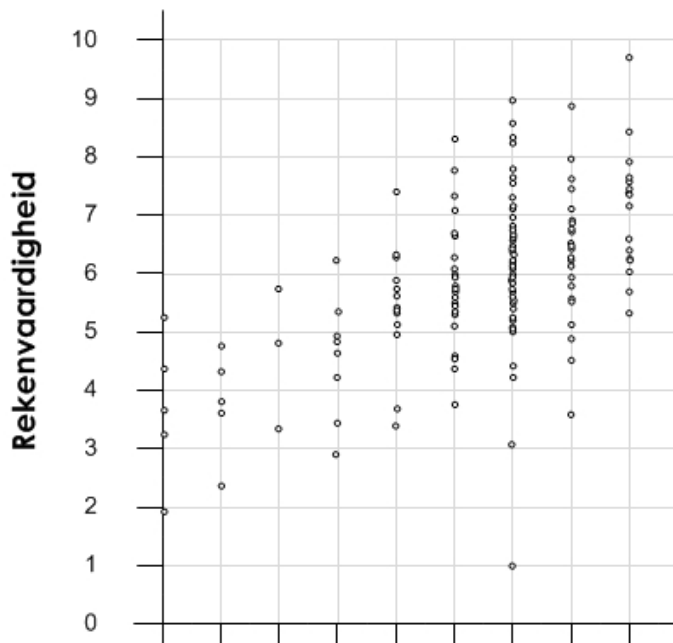
- Correcte redenering (1)
- Incorrecte redenering (2)

Peter: 'Ik weet zeker dat de conclusie die ik op basis van deze 15 jongens en 15 meisjes trek over het verschil in rekenplezier tussen jongens en meisjes in het algemeen klopt.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Opdracht 5: Rekenplezier en rekenvaardigheid

Van 150 meisjes uit groep 5 tot en met 8 van de basisschool is, naast hun rekenplezier, ook hun rekenvaardigheid bekend. In onderstaande grafiek is voor elk van de 150 meisjes het rekenplezier en de rekenvaardigheid weergegeven. Bijvoorbeeld voor een meisje met rekenplezier 30 en rekenvaardigheid 5, is een stipje gezet door bij 30 op de horizontale as naar boven te gaan tot aan 5 op de verticale as. Dit is voor elk meisje gedaan.



We vragen je een gedachtenexperiment te doen: Stel dat we nog een meisje ondervragen. Haar score voor rekenplezier is 60. Wat denk jij dat haar score voor rekenvaardigheid ongeveer zou kunnen zijn? Licht je antwoord toe.

Antwoord + toelichting:

Je krijgt uitspraken te lezen van andere studenten die de opdracht hebben gemaakt. Geef bij elke uitspraak aan of je hun redenering correct vindt.

Daphne: 'Uit de grafiek blijkt dat voor alle meisjes met 60 voor rekenplezier, de scores voor rekenvaardigheid tussen 1 en 9 liggen. Elke score tussen 1 en 9 voor dit extra meisje is dus even goed mogelijk.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Pieter: 'Uit de grafiek blijkt dat voor alle meisjes met 60 voor rekenplezier, de scores voor rekenvaardigheid tussen 1 en 9 liggen. Het midden hiervan is 5. Ik denk de score van het extra meisje in de buurt van 5 ligt.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Cindy: 'Ik baseer mijn antwoord in ieder geval op waar de scores voor rekenvaardigheid van de meerderheid van de meisjes met 60 voor rekenplezier liggen. Dat is ongeveer tussen 5 en 7.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Femke: 'Je weet nooit helemaal zeker wat de score van het extra meisje is. Daarom kijk ik naar welke scores de meeste kans maken. In dit geval is de kans het grootste dat de score van het extra meisje ligt tussen 5 en 7.'

- Correcte redenering (1)
- Incorrecte redenering (2)

Ceder: 'Ik baseer mijn antwoord in ieder geval op het gemiddelde van de scores voor rekenvaardigheid van de meisjes met 60 voor rekenplezier. Dat gemiddelde is ongeveer 6,2.'

- Correcte redenering (1)
- Incorrecte redenering (2)