

Samenvatting thesis

‘Introducing Statistical Inference: Design and Evaluation of a Learning Trajectory’

Marianne van Dijke-Droogers

De overweldigende hoeveelheid data, grafieken en voorspellingen met betrekking tot de COVID-pandemie in het afgelopen jaar, illustreert het essentiële belang van statistiek. De afgelopen decennia is het gebruik van data enorm toegenomen vanwege technologische vernieuwingen die het mogelijk maken om eenvoudig data te verzamelen, op te slaan, te analyseren en representeren. Op basis van data worden ingrijpende beslissingen genomen en uitspraken gedaan, zowel door burgers als professionals. Het is daarom van belang om de statistische geletterdheid van onze leerlingen te ontwikkelen. Dit houdt in dat leerlingen toegerust worden om statistische informatie te interpreteren, kritisch te beoordelen en hier conclusies uit te trekken (Gal, 2002).

Het werken met inferenties vormt het hart van de statistiek. Het is essentieel dat leerlingen het vermogen ontwikkelen om inferenties te interpreteren. Bij inferenties worden op basis van steekproefdata conclusies getrokken over een groter geheel of proces. Deze conclusies gaan vergezeld van onzekerheid omdat niet alles of iedereen is onderzocht. Het interpreteren van deze onzekerheid en het duiden van de waarschijnlijkheid van de conclusie is veelzijdig en complex.

In veel landen, waaronder Nederland, wordt statistische inferentie daarom pas behandeld in de bovenbouw van het voortgezet onderwijs of in het hoger onderwijs¹. Uit onderzoek blijkt dat inferentiële statistiek hier een struikelblok is voor veel leerlingen en studenten. De moeilijkheden van leerlingen worden met name veroorzaakt door een beperkt begrip van kernconcepten die nodig zijn voor inferenties (Castro Soto et al., 2007; Konold & Pollatsek, 2002) zoals steekproef, variatie en verdelingen. Deze conceptuele problemen worden verergerd door een sterk onderwijsfocus op het aanleren van complexe, formele procedures.

Om de moeilijkheden van leerlingen te overbruggen is in de afgelopen decennia gezocht naar *informele* onderwijsbenaderingen om conceptueel begrip te promoten. Het aanbieden van informele inferentiële activiteiten op jongere leeftijd zou het leren van de complexere inferentiële statistiek op latere leeftijd kunnen vereenvoudigen (Zieffler et al., 2008). Het gaat hierbij om het trekken van conclusies vanuit informele statistische kennis, dus niet vanuit formele procedures zoals hypothese toetsen of berekeningen met de normale verdeling. Makar en Rubin (2009) definiëren informele statistische inferentie in de volgende principes: het generaliseren van steekproefdata naar een groter geheel; data als bewijs van deze generalisatie; redeneren over de waarschijnlijkheid van deze generalisatie. Nieuwe digitale middelen bieden mogelijkheden voor het simuleren van steekproeven, waarmee leerlingen op informeel niveau de kernconcepten voor statistische inferentie kunnen onderzoeken.

Het gebruik van technologie is onmisbaar voor het doen en leren van statistiek (Gal, 2002; Thijs, Fisser, & Van der Hoeven, 2014). De inzet van recente digitale leeromgevingen met opties voor

¹ Tevens geldt voor ons land dat de leerlingen met een technisch profiel in de bovenbouw—vanaf vwo 4—helemaal geen inferentiële statistiek krijgen, tenzij ze wiskunde D kiezen.

statistisch modelleren, zoals VUstat en TinkerPlots, biedt een informele aanpak om het begrip van statistische concepten en modellen te verdiepen (Biehler, Frischemeier, & Podworny, 2017). Inzicht in statistische modellen is van fundamenteel belang voor het interpreteren van statistische inferenties (Manor & Ben-Zvi, 2017). Statistische modellen helpen om de waarschijnlijkheid van op steekproefdata gebaseerde conclusies te duiden. Digitale middelen voor het simuleren van steekproefdata uit populatiemodellen maken concepten visueel en toegankelijk. Het modelleren met zulke digitale middelen is veelbelovend voor het statistiekonderwijs nu en in de toekomst.

Kortom: het onderwijzen van statistische inferentie is belangrijk maar ook moeilijk. Het inbedden van informele statistische inferentie in eerdere leerjaren lijkt veelbelovend, met name in combinatie met het gebruik van digitale leermiddelen. Er is echter nog weinig bekend over hoe we ons huidige onderbouwcurriculum kunnen uitbreiden met een goed onderbouwd leertraject. Dit onderzoeksproject beoogt kennis te verwerven over een theoretisch en empirisch gefundeerd leertraject om statistische inferentie te introduceren bij vwo 3-leerlingen (Grade 9).

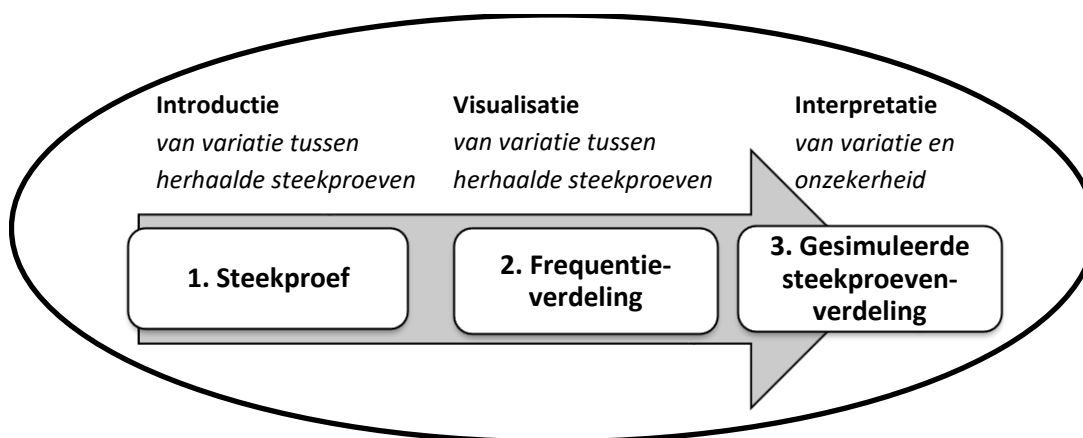
Dit onderzoeksproject volgde een ontwerpgerichte aanpak (Bakker, 2018). Deze aanpak kenmerkt zich door een cyclisch proces waarin onderwijsmateriaal voor leeromgevingen wordt ontworpen, geïmplementeerd en geëvalueerd, voor vervolgcycli van (her)ontwerp en testen (McKenney & Reeves, 2012). In de beginfase richtten we ons vooral op de ontwikkeling van een theoretisch gefundeerd ontwerp, met daarin een specificatie van beoogde leerdoelen en de uitwerking hiervan in een—op dat moment nog hypothetisch—leertraject. Naarmate het onderzoek vorderde, werden meerdere interventies met het leertraject uitgevoerd in de lespraktijk en geëvalueerd. Deze interventies werden in iedere cyclus opgeschaald in zowel de lengte van het leertraject als in het aantal deelnemers. In dit onderzoek zijn drie cycli doorlopen: beginnend met een onderwijsexperiment in één klas, via een interventie in drie klassen, naar een interventie in 13 klassen op verschillende scholen. Daarnaast is tussen cyclus 2 en 3 een verdiepende casestudie uitgevoerd naar het leren van en met technologie. Deze verdiepende casestudie richtte zich op de samenhang tussen het leren van gebruikstechnieken voor een digitale tool en het ontwikkelen van conceptueel statistisch begrip.

Hoofdstuk 2. Herhaalde steekproeven met een black box als opstap naar statistische inferentie

Dit hoofdstuk presenteert de resultaten uit de eerste cyclus—en daarmee de eerste studie—van dit ontwerponderzoek. Succesvolle implementatie van theorie in de onderwijspraktijk impliceert het stapsgewijze ontwerp en de evaluatie in echte klaslokalen van krachtige leertrajecten die ons huidige begrip van effectief leren belichamen (De Corte, 2000). De eerste cyclus richtte zich daarom op het ontwerp, de implementatie en de evaluatie van het eerste deel van het leertraject: leerstap 1 tot en met 3. Allereerst werd op basis van literatuurstudie een hypothetisch leertraject (Simon, 1995) ontwikkeld voor het introduceren van drie kernbegrippen voor statistische inferentie: steekproef, frequentieverdeling en gesimuleerde steekproevenverdeling (ICT). Figuur 1 toont een overzicht van de drie kernbegrippen.

Een hypothetisch leertraject bestaat uit leerdoelen voor de leerlingen, een beschrijving van leeractiviteiten met bijbehorende hulpmiddelen, materialen en taakstructuren, leerlingkenmerken, en onderwijsmethoden die leiden tot het vereiste leerproces en de beoogde leerdoelen (Sandoval, 2014; Simon, 1995). Het door ons ontworpen hypothetische leertraject werd vervolgens geïmplementeerd in één klas met 20 vwo 3-leerlingen. Voor de evaluatie van het traject werden bij elke leerstap

indicatoren opgesteld over observeerbaar leergedrag van leerlingen die de hypothese van iedere stap ondersteunen.



Figuur 1. Overzicht van de drie kernbegrippen voor statistische inferentie, zoals ingebed in stappen 1 tot en met 3 van het leertraject

De hypothese in leerstap 1 was dat leerlingen zich bewust zouden worden van steekproefvariatie door het uitvoeren van experimenten met een fysieke black box gevuld met balletjes. Door het uitvoeren van herhaalde experimenten met een klein en groot kijkvenster, konden ze het effect van herhaalde steekproeven en steekproefomvang op de schatting van de populatie (inhoud black box) exploreren. De resultaten toonden aan dat de met de hypothese verbonden indicatoren werden waargenomen. De eerste leerstap stelde de leerlingen in staat om in korte tijd te redeneren met steekproefdata, inclusief het (informeel) interpreteren van variatie en onzekerheid. Zie Figuur 2 voor een impressie van leerstappen 1 tot en met 3.

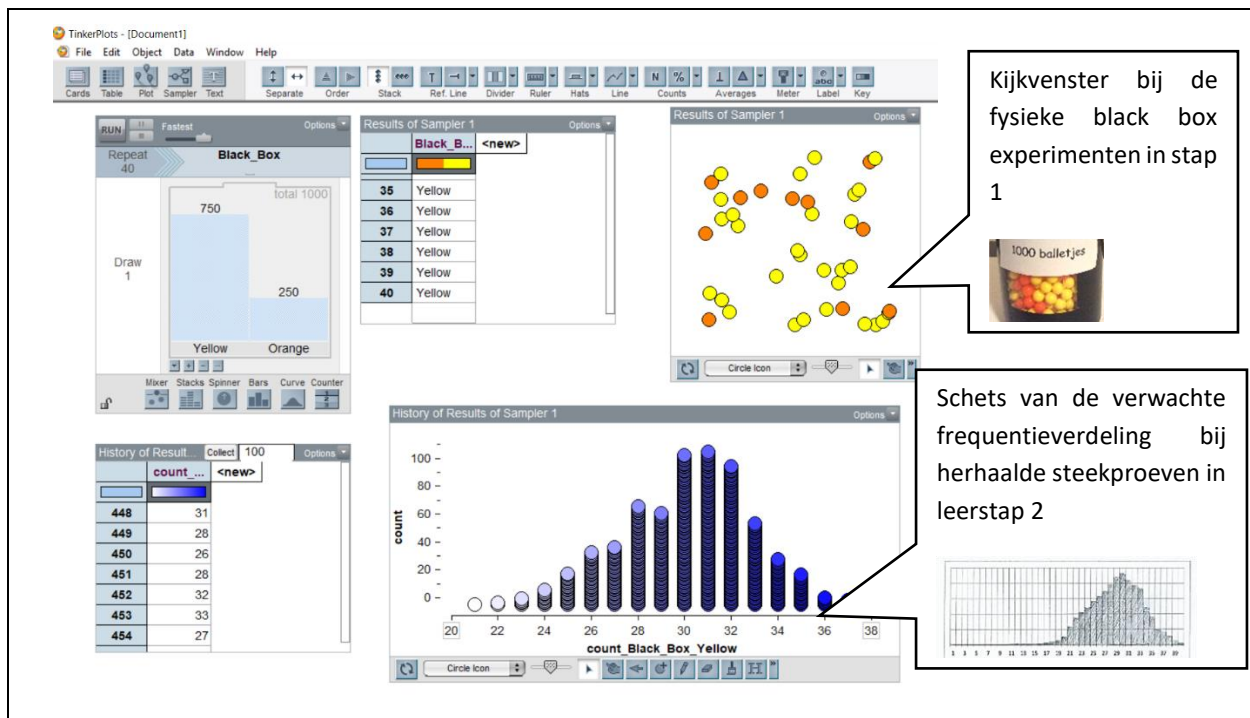
Leerstap 2 van het leertraject was gericht op het concept van de frequentieverdeling bij herhaalde steekproeven. De hypothese was dat leerlingen deze frequentieverdeling allereerst zouden interpreteren als een (visualisatie) *model van* de verkregen resultaten bij een bepaalde black box en dat de ontworpen leeractiviteiten hen zouden stimuleren om de conceptuele overstap te maken naar het gebruik hiervan als *model voor* het interpreteren van variatie en onzekerheid—deze conceptuele overstap van een *model van* een specifieke wiskundige situatie naar een *model voor* een netwerk aan gerelateerde wiskundige situaties is bekend als *emergent modeling* (Gravemeijer, 1999). In deze leerstap stonden twee activiteiten centraal. Ten eerste het schetsen van de verwachte frequentieverdeling bij 100.000 herhaalde steekproeven met de fysieke black box uit de eerste leerstap, en ten tweede het gebruik van een dergelijke frequentieverdeling om de waarschijnlijkheid van specifieke steekproefresultaten te bepalen. Uit de resultaten van deze leerstap bleek dat de meeste indicatoren werden waargenomen. Leerlingen waren in staat om een correcte schets te maken en de kans op specifieke steekproefresultaten te bepalen in de context van de black box—bijvoorbeeld het bepalen van de kans op een steekproefresultaat van meer dan 35 gele balletjes bij een steekproef van 40 uit een populatie met een proportie van 75%.



Figuur 2. Impressies van leerstappen 1 tot en met 3. Links twee foto's van leerstap 1, rechtsboven leerstap 2 en rechtsonder leerstap 3

In leerstap 3 lag de focus op het conceptualiseren van de gesimuleerde steekproevenverdeling. Hiervoor gebruikten leerlingen statistisch modelleren in een digitale omgeving om herhaalde steekproeven met de bijbehorende steekproevenverdeling te simuleren. De hypothese was dat leerlingen zouden begrijpen dat deze gesimuleerde verdeling kan worden gebruikt als een model voor het interpreteren van variatie en onzekerheid. In deze leerstap neemt de computer als het ware hun handwerk uit de eerste twee leerstappen over. De resultaten van leerstap 3 toonden aan dat ook hier de indicatoren die de hypothese ondersteunen, werden waargenomen.

Op basis van de bevindingen in deze studie vermoedden we dat de sterke samenhang en opbouw tussen de drie leerstappen het voor leerlingen mogelijk maakte om deze probleemloos te doorlopen. Vanuit hun concrete ervaringen met steekproefvariatie in leerstap 1, gevolgd door het visualiseren van de opschaling van dit experiment in leerstap 2, konden leerlingen gemakkelijk de overgang maken naar het modelleren en interpreteren van de gesimuleerde steekproevenverdeling in leerstap 3. Zie figuur 3 voor een illustratie van deze samenhang tussen leerstap 1 tot en met 3. Deze eerste drie stappen van het leertraject gaven leerlingen het benodigde inzicht in hoe een steekproevenverdeling ontstaat en hoe deze kan worden gebruikt als model voor het interpreteren van variatie en onzekerheid. Deze bevindingen suggereerden een veelbelovende manier om leerlingen te laten kennismaken met (informele) statistische inferentie.



Figuur 3. De samenhang tussen het werken in de digitale omgeving van TinkerPlots in leerstap 3, en de fysieke black box-activiteiten in leerstappen 1 en 2

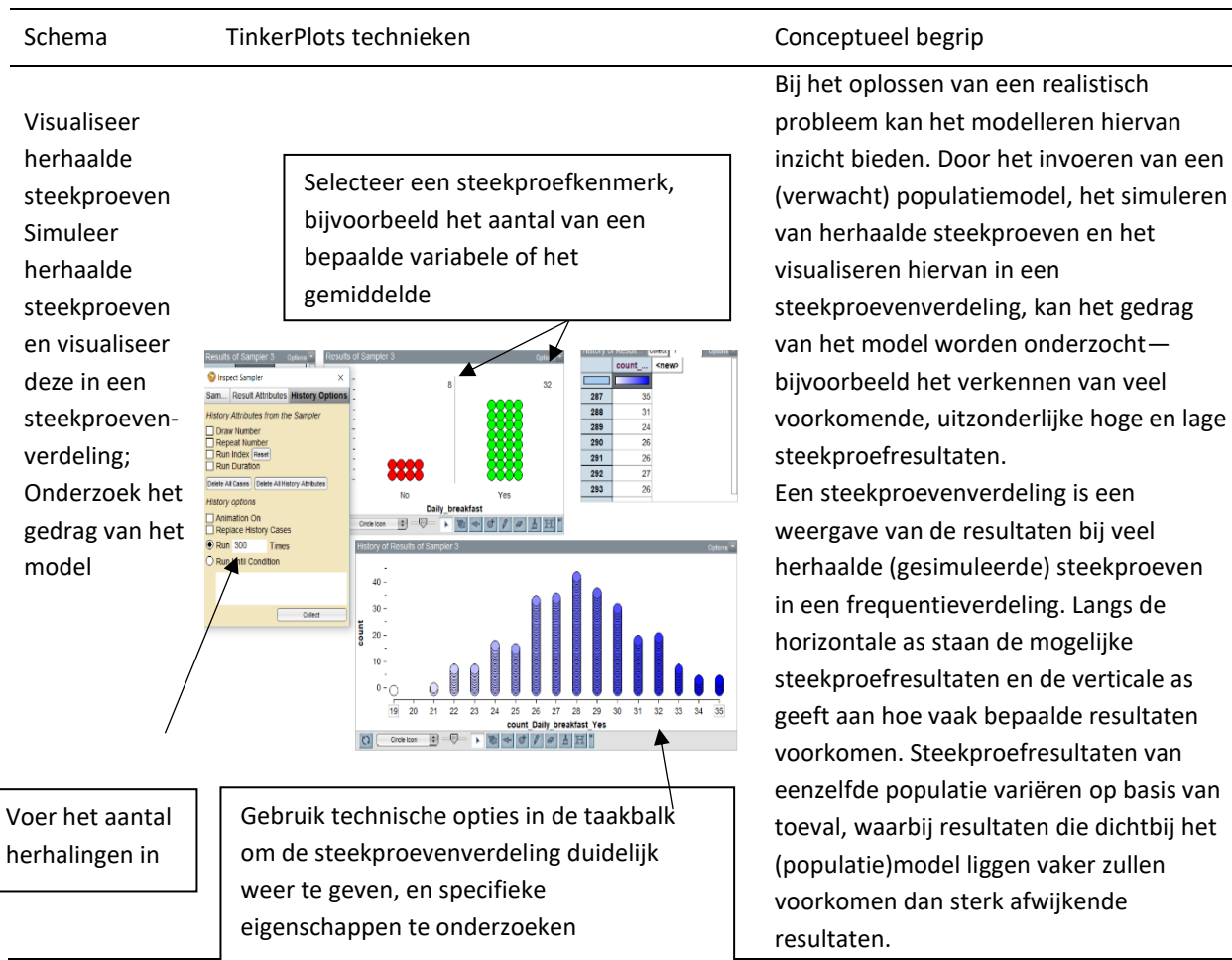
Hoofdstuk 3. Statistische modellerprocessen bekeken door de lens van instrumentele genese

Om meer inzicht te krijgen in het leren van en met technologie in leerstap 3 en verder werd een verdiepende casestudie uitgevoerd. Inzicht in leren met digitale middelen is voorwaardelijk om deze effectief te kunnen inzetten voor het bereiken van beoogde leerdoelen. Digitale leermiddelen voor statistiek, zoals TinkerPlots, bieden mogelijkheden voor statistisch modelleren via een informele aanpak. Deze digitale middelen faciliteren leerlingen om populatiemodellen te bouwen en deze modellen te gebruiken om steekproefdata te simuleren. Dit statistisch modelleren bevordert het inzicht in concepten en modellen die fundamenteel zijn voor statistische inferentie (Biehler et al., 2017; Manor & Ben-Zvi, 2017).

Vanuit wiskundeonderwijs is bekend dat het aanleren van gebruikstechnieken voor een digitale tool en het ontwikkelen van conceptueel begrip met elkaar verweven zijn. Tot nu toe heeft deze verwevenheid van gebruik technisch en conceptueel begrip, bekend als instrumentele genese (Artigue, 2002), weinig aandacht gekregen in onderzoek naar statistiekonderwijs met digitale middelen. Deze verdiepende casestudie richtte zich daarom op de toepasbaarheid van het theoretisch perspectief van instrumentele genese binnen statistiekonderwijs, en meer specifiek bij het statistisch modelleren in de digitale omgeving van TinkerPlots.

Een geschikte fase om de instrumentele genese van leerlingen te onderzoeken is na de introductie van de tool en de concepten, bij het toepassen van de verworven kennis in nieuwe situaties. Deze fase vindt plaats in leerstap 4 van het leertraject. De data voor dit onderzoek bestonden uit video- en audio-opnames van twee laboratoriumsessies met in totaal 28 leerlingen uit vwo 3 bij het uitvoeren van leeractiviteiten in stap 4 van het traject. In het bijzonder analyseerden we hoe de ontwikkeling van (gebruiks)technieken en conceptueel begrip verweven waren in de instrumentatieschema's die leerlingen ontwikkelden. We identificeerden zes instrumentatieschema's,

A tot en met F, voor statistisch modelleren met TinkerPlots. Figuur 4 illustreert als voorbeeld instrumentatieschema C. De linkerzijde van het figuur bevat een beknopte beschrijving van het schema, in het midden is een schermafbeelding vanuit TinkerPlots weergegeven met een duiding van de gebruikte technieken, en de rechterzijde beschrijft het conceptueel begrip dat in dit schema aan de orde is.



Figuur 4. Voorbeeld van instrumentatieschema C voor statistische modelleerprocessen met TinkerPlots

We observeerden een sterke verwevenheid tussen het aanleren van technieken en het ontwikkelen van conceptueel begrip. Technieken voor het gebruik van TinkerPlots hielpen de leerlingen om context-onafhankelijke technische patronen te ontdekken, die de belangrijke conceptuele overstap van een *model van* naar een *model voor* (Gravemeijer, 1999) bevorderden. Meer concreet betekende dit dat leerlingen ontdekten dat gebruikstechnieken in specifieke contexten meer algemeen, dus contextonafhankelijk, toegepast konden worden. Dit ging gepaard met meer abstracte statistische terminologie—bijvoorbeeld het invoeren van de steekproefomvang in plaats van het aantal bevraagde leerlingen. Omgekeerd leidde het conceptuele begrip van de leerlingen tot de verkenning van meer geavanceerde digitale technieken. Deze bevindingen toonden aan dat investeren in het aanleren van digitale technieken tegelijkertijd een positief effect heeft op het ontwikkelen van statistisch begrip.

Hoofdstuk 4. Introductie in statistische inferentie: Ontwerp van een theoretische en empirisch onderbouwd leertraject

In dit hoofdstuk worden de resultaten van de derde studie—gebaseerd op onderzoekcyclus 3—gepresenteerd. Op basis van de eerste twee onderzoekcycli, de verdiepende casestudie en aanvullend literatuuronderzoek werd het (hypothetische) leertraject (her)ontworpen voor de derde cyclus. Deze cyclus omvatte het gehele traject van acht leerstappen, opgesplitst in twee vergelijkbare delen van vier: (1) experimenteren met een fysieke black box, (2) visualiseren van verdelingen, (3) onderzoeken van steekproevenverdelingen met behulp van simulatiesoftware, (4) interpreteren van steekproevenverdelingen voor inferenties in realistische contexten. De stappen 1 tot en met 4 zijn alleen gericht op categoriale data en in de stappen 5 tot en met 8 wordt gewerkt met numerieke data. Een overzicht van het gehele leertraject is weergegeven in tabel 1.

Het leertraject werd geïmplementeerd in een interventie onder 267 leerlingen in 13 klassen op verschillende scholen. De tijdsomvang van het leertraject bestond uit zes lessen per deel, met een totaal van 12 lessen. We analyseerden de posttestresultaten van de leerlingen na de interventie om te onderzoeken *of* het traject inderdaad de beoogde leerstapgerelateerde doelen voor statistische inferentie stimuleerde. Om de posttestresultaten te kunnen interpreteren werden deze vergeleken met die van een vergelijkingsgroep ($n = 217$) die het reguliere vwo 3-curriculum gevolgd had. De reguliere aanpak bestond uit 10–12 lessen gericht op beschrijvende statistiek. Tevens analyseerden we de werkbladen van de leerlingen tijdens elke leerstap om te onderzoeken *hoe* het stapsgewijze traject het leerproces bevorderde.

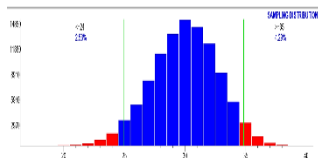
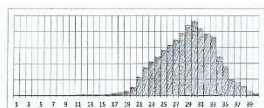
De posttestresultaten toonden aan dat leerlingen die les kregen vanuit het leertraject significant hoger scoorden op alle specifiek leerstapgerelateerde doelen uit het leertraject dan leerlingen van een vergelijkingsgroep ($n = 217$) die het reguliere curriculum volgden. Deze leerdoelen omvatten in leerstap 1 en 5 het gebruik van steekproeven, in leerstap 2 en 6 het visualiseren van verdelingen, in leerstap 3 en 7 het effect van herhaalde steekproeven en steekproefomvang, en in leerstap 4 en 8 het interpreteren van inferenties in realistische contexten. De analyse van de werkbladen van de leerlingen toonde aan dat het traject van acht stappen bevorderlijk was voor het leerproces van de meeste leerlingen.

De bevindingen in deze studie toonden aan *dat* en *hoe* het ontworpen leertraject werkt. Een aanpak gebaseerd op herhaalde steekproeven met een black box gecombineerd met statistisch modelleren in de digitale omgeving van TinkerPlots, bleek vruchtbaar voor het introduceren van statistische inferentie. Beide ideeën hebben tevens potentieel voor inbedding in meer complexe vervolgactiviteiten, zoals het toetsen van hypothesen en het vergelijken van groepen. Deze bevindingen suggereren dat huidige statistiekcurricula met een beschrijvende focus kunnen worden uitgebreid met inferentiële activiteiten, zodat beter kan worden geanticipeerd op vervolgstappen van leerlingen in het statistiekonderwijs.

Tabel 1

Overzicht van stap 1 – 8 van het leertraject

Leerstap	Beschrijving	Voorbeeld van activiteiten	Leerdoel	Opbouw in leerstappen
<i>Categorale data</i>				
1. Experimenteren met een fysieke black box	Fysiek black box met balletjes experiment (met klein en groot kijkvenster)	Schat het aantal gele balletjes in de black box, gevuld met 1000 balletjes, door het (herhaaldelijk) schudden en onderzoeken van het aantal zichtbare balletjes	Leerlingen formuleren inferenties en maken kennis met concepten als steekproef, steekproefomvang, steekproefvariatie, frequentie en centrum- en spreidingsmaten, in de context van herhaalde steekproeven met een fysieke black box.	Leerlingen ervaren dat steekproefresultaten variëren en dat een grotere steekproefomvang en meer herhalingen leiden tot een betere schatting van de populatie.
2. Visualiseren van verdelingen	Grafiek als model (visualisatie) voor de frequentieverdeling bij herhaalde steekproeven met een black box	Schets de frequentieverdeling die je verwacht als het black box-experiment met groot venster 100.000 keer wordt herhaald.	Leerlingen kunnen een visualisatie tekenen van de verwachte steekproevenverdeling bij herhaalde steekproeven. Leerlingen interpreteren de steekproevenverdeling om inferenties te formuleren over de waarschijnlijkheid van specifieke steekproefresultaten.	Vervolgvrage: Wat gebeurt er als we de steekproefomvang en het aantal herhaalde steekproeven verder vergroten? Het uitvoeren van meer en grotere steekproeven is tijdrovend, daarom kan een gedachtenexperiment helpen.
3. Modelleren van een black box (ICT)	Simulaties van herhaalde steekproeven met een gemodelleerde black box in een steekproevenverdeling,	Gebruik TinkerPlots om de meest voorkomende steekproefresultaten te bepalen voor een black box gevuld met 750 gele en 250	Leerlingen gebruiken statistisch modelleren in de digitale omgeving van TinkerPlots om (on)waarschijnlijke steekproefresultaten te bepalen, in de	De steekproevenverdeling van herhaalde steekproeven kan gebruikt worden om de waarschijnlijkheid van specifieke steekproefresultaten te bepalen. Vervolgvrage: Hoe kunnen we snel en eenvoudig de steekproevenverdeling van herhaalde steekproeven verkrijgen? Het gebruik van technologie kan helpen.



voor het bepalen en interpreteren van de waarschijnlijkheid van specifieke resultaten

oranje balletje, bij een steekproefomvang van 40

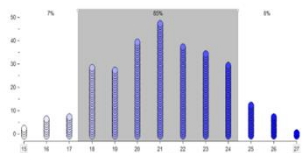
context van een black box [Statistisch modelleren omvat het invoeren van een model, simuleren van (herhaalde) steekproeven, visualiseren van de steekproevenverdeling en het interpreteren van de resultaten]

Statistisch modelleren—inclusief het interpreteren van de steekproevenverdeling bij herhaalde steekproeven—kan gebruikt worden om de waarschijnlijkheid van specifieke steekproefresultaten te bepalen.

Vervolg vraag: Kan statistisch modelleren meer algemeen gebruikt worden, in andere situaties en contexten?

4. Modelleren van realistische contexten (ICT)

Modelleren van een specifieke situatie in TinkerPlots, voor het simuleren en interpreteren van de steekproevenverdeling en het beter begrijpen van de situatie (inclusief de bijbehorende onzekerheid)



Gebruik TinkerPlots om te bepalen welke steekproefresultaten je kunt verwachten als een steekproef met 30 leerlingen wordt uitgevoerd op een school met 300 leerlingen, bij onderzoek naar het aantal leerlingen dat dagelijks ontbijt (gegeven dat gemiddeld 70% van de VO-leerlingen dagelijks ontbijt)

Leerlingen gebruiken statistisch modelleren in de digitale omgeving van TinkerPlots voor het formuleren van inferenties, in de context van realistische probleemstellingen

Vanuit stap 1 tot en met 4 ontstaat de vraag hoe statistisch modelleren gebruikt kan worden bij andere—niet categoriale—data.

Numerieke data

5. Experimenteren met een fysieke black box

Fysiek black box met briefjes experiment. (De black box is gevuld met 4000 briefjes. Elk briefje bevat gegevens over het geslacht en de lichaamslengte van een leerling uit klas 2, bijvoorbeeld: jongen—155 cm)



Neem een steekproef van 40 briefjes en vat de steekproefdata samen (bereken centrum- en spreidingsmaten, gebruik visualisaties). Maak een schatting van de lichaamslengte van de 4000 leerlingen uit klas 2

Leerlingen formuleren inferenties in de context van een fysieke black box met briefjes (geslacht en lichaamslengte van 4000 leerlingen uit klas 2) met inachtneming van steekproefomvang, steekproefvariatie en centrummaten

Leerlingen bespreken hoe numerieke data van herhaalde steekproeven gebruikt kunnen worden om inferenties over de populatie te formuleren.

Vervolg vraag: Hoe kan de onderliggende populatieverdeling—de lichaamslengte

6. Visualiseren van verdelingen



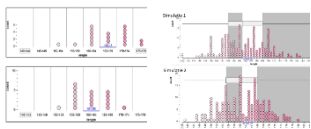
Verwachtingen over de populatie (lengte van 4000 leerlingen) samenvatten en visualiseren op basis van de gevonden steekproefdata in leerstap 5

Schets de frequentieverdeling die je verwacht voor de totale populatie, op basis van de gevonden steekproefresultaten in leerstap 5

Leerlingen schetsen een visualisatie van de populatieverdeling die ze verwachten op basis van de gevonden steekproefresultaten. Leerlingen maken inferenties over de populatieverdeling en het populatiegemiddelde.

van de 4000 leerlingen op de briefjes in de black box—weergegeven worden op basis van de gevonden steekproefresultaten?

7. Modelleren van een black box (ICT)



Experimenteren met simulaties van (herhaalde) steekproeven, bij variatie in omvang en aantal herhalingen, in de context van de black box met briefjes uit leerstap 5 en 6

Gebruik TinkerPlots om de meest voorkomende steekproefresultaten te bepalen voor de (gegeven) gemodelleerde black box met briefjes die hoort bij leerstap 5

Leerlingen gebruiken statistisch modelleren in de digitale omgeving van TinkerPlots om (on)waarschijnlijke steekproefresultaten te bepalen, in de context van de black box met briefjes. [Statistisch modelleren omvat hier het simuleren van (herhaalde) steekproeven vanuit een *gegeven* model, visualiseren van de steekproevenverdeling van het steekproefgemiddelde, en het interpreteren van de resultaten]

Leerlingen formuleren inferenties over het populatiegemiddelde en de populatieverdeling op basis van de gevonden steekproefdata.

Vervolg vraag: Wat zijn de effecten van grotere en meer steekproeven op de schatting van het populatiegemiddelde en de populatieverdeling?

Het gebruik van technologie kan helpen bij het onderzoeken van deze effecten

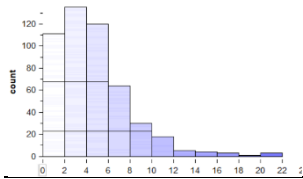
Vanuit leerstap 7 ontstaat de vraag hoe het statistisch modelleren met numerieke data kan worden toegepast in andere contexten en situaties

Modelleren van een realistische situatie in

Gebruik TinkerPlots voor het simuleren van herhaalde

Leerlingen gebruiken statistisch modelleren in de digitale omgeving

8. Modelleren van realistische contexten (ICT)



TinkerPlots, voor het simuleren en interpreteren van de steekproevenverdeling, en het beter begrijpen van de situatie (inclusief de bijbehorende onzekerheid).

steekproeven (omvang 200) vanuit een verborgen populatiemodel van 4000 personen om de sporttijd van deze personen te onderzoeken en bepalen

van TinkerPlots om inferenties te formuleren, in realistische contexten en probleemstellingen

Hoofdstuk 5. De effecten van het leertraject voor statistische inferentie op de statistische geletterdheid van leerlingen

Als laatste studie in dit onderzoeksproject, werden de effecten van het leertraject op de statistische geletterdheid van leerlingen onderzocht. Statistische geletterdheid kan onderverdeeld worden in drie domeinen: (1) Statistische inferentie, (2) Grafieken en variatie, (3) Kans en gemiddelde. Het ontworpen leertraject was vooral gericht op statistische inferentie, het eerste domein. We vermoedden echter dat het leertraject ook positieve invloed zou hebben op de andere twee domeinen. Het reguliere vwo 3-curriculum was enkel gericht op domein twee en drie. Voor de evaluatie werd gebruik gemaakt van een pre-post onderzoeksopzet met de interventiegroep ($n = 267$) uit de vorige studie—de derde onderzoek cyclus. Om de significante leerwinst van deze groep te interpreteren, hebben we de resultaten van deze leerlingen vergeleken met een nationale baseline en internationale prestaties. Voor de nationale baseline gebruikten we de pre- en posttestresultaten van de vergelijkingsgroep ($n = 217$) uit de vorige studie die het reguliere leerplan van vwo 3 volgde, en de internationale vergelijking werd gedaan aan de hand van een Australische studie met vergelijkbare testopzet.

Tabel 2

Gemiddelde leerlingscore op de domeinen van statistische geletterdheid bij de pre- en posttest voor de interventie- en vergelijkingsgroep, inclusief vooruitgang van pre naar post

		Interventiegroep ($n = 267$) M (SD)	Vergelijkingsgroep ($n = 217$) M (SD)	Interventiegroep – vergelijkingsgroep M(I) – M(C)
Pretest	SG	2.60 (0.61)	2.97 (0.68)	– 0.37***
	SI	2.45 (0.65)	2.72 (0.71)	– 0.27***
	GV	2.07 (0.63)	2.29 (0.58)	– 0.22***
	KG	3.29 (1.38)	3.92 (1.31)	– 0.63***
Posttest	SG	3.28 (0.69)	2.95 (0.78)	+ 0.33***
	SI	3.34 (0.84)	2.67 (0.84)	+ 0.67***
	GV	2.59 (0.81)	2.38 (0.88)	+ 0.21*
	KG	3.92 (0.88)	3.80 (1.06)	+ 0.12
Pre naar post	SG	+ 0.68 (0.86)***	– 0.02 (0.73)	0.70***
	SI	+ 0.89 (0.92)***	– 0.04 (0.71)	0.93***
	GV	+ 0.52 (0.98)***	+ 0.09 (0.94)	0.43***
	KG	+ 0.63 (1.53)*	– 0.11 (1.45)	0.74***

* $p < .05$, ** $p < .005$, en *** $p < .0005$

Noot. SG = statistische geletterdheid; SI, GV, KG zijn de drie domeinen binnen SG; SI = statistische inferentie; GV = grafieken en variatie; KG = kans en gemiddelde.

De nationale vergelijking van testresultaten toonde aan dat de interventiegroep significant hoger scoorde op statistische geletterdheid, en in het bijzonder op het domein van statistische inferentie. Tevens vonden we aanzienlijk positieve effecten voor de andere twee domeinen. Hoewel het leertraject niet gericht was op de andere domeinen, bleek het leertraject—bestaande uit een onderzoekgerichte aanpak met digitale middelen en meer complexe leeractiviteiten voor statistische inferentie—ook hier een positief effect te hebben. Tabel 2 toont een overzicht van de resultaten voor de interventie- en vergelijkingsgroep. Hierbij moet vermeld worden dat de leerlingen van de vergelijkingsgroep de reguliere statistieklessen voorafgaand aan de pretest gevolgd hadden, waardoor

hun pretestresultaten hoger zijn dan die van de interventiegroep. Tussen de pre- en posttest volgde deze vergelijkingsgroep geen statistieklessen, wat zichtbaar is in de gelijkwaardige resultaten op beide tests voor deze groep.

De vergelijking met de internationale studie toonde aan dat de posttestresultaten van de interventiegroep met 14–15-jarigen op statistische geletterdheid het meest overeenkwamen met die van Australische leerlingen in Grade 7–8 met een leeftijd van ongeveer 13 jaar. De resultaten van de vergelijkingsgroep met 14–15-jarigen waren het meest vergelijkbaar met die van Australische leerlingen in Grade 6–7 met een leeftijd van ongeveer 12 jaar. Uit deze internationale vergelijking kunnen we opnieuw concluderen dat de interventiegroep aanzienlijk hoger scoorde, met ongeveer één leerjaar verschil, dan de vergelijkingsgroep. Opvallend is dat de resultaten van beide groepen met leerlingen in de leeftijd van 14–15 jaar overeenkwamen met die van aanzienlijk jongere Australische leerlingen. Vermoedelijk komt dit doordat het statistiekaanbod in Australië uitgebreider is dan in Nederland.

De bevindingen vanuit zowel de nationale als internationale vergelijking toonden aan dat het leertraject een significant positief effect had op de statistische geletterdheid van de leerlingen, en in het bijzonder op het domein van statistische inferenties. Tevens signaleerden we positieve effecten voor de andere domeinen. Op basis hiervan kunnen we constateren dat huidige statistiecurricula met een sterk beschrijvende focus verrijkt kunnen worden met een inferentiële focus—in ieder geval voor de onderbouw van het vwo. Het voordeel hiervan is dat leerlingen meer leren over statistische inferenties en niet minder over de andere domeinen van statistische geletterdheid, om zo beter te anticiperen op vervolgstappen van de leerling binnen statistiekonderwijs.

Hoofdstuk 6. Algemene discussie

Dit onderzoeksproject heeft kennis opgeleverd over essentiële vernieuwingen in statistiekonderwijs. Theoretische inzichten werden ontwikkeld in nauwe samenhang met een praktisch onderwijsontwerp. Deze inzichten waren zowel inhoudelijk als methodologisch van aard.

Inhoudelijke bijdrage

Op inhoudelijk gebied draagt dit onderzoek bij aan inzicht in de samenhang tussen de ontwikkeling van statistische inferentie en statistische geletterdheid. Statistische inferentie wordt beschouwd als een complex domein van statistische geletterdheid, wat vaak pas op latere leeftijd wordt aangeboden. De resultaten in dit onderzoek toonden aan dat het ontworpen leertraject met (informele) inferentiële activiteiten een significant positief effect had op het domein van statistische inferentie, en eveneens op de andere twee domeinen van statistische geletterdheid—de domeinen grafieken en variatie, en gemiddelde en kans, beiden met een beschrijvende focus. Dit positieve effect van (informele) inferentiële activiteiten op de andere domeinen van statistische geletterdheid pleit voor het eerder introduceren hiervan. Leerlingen ontwikkelen dan al op vroege leeftijd statistische concepten die noodzakelijk zijn voor statistische inferentie en voor statistische geletterdheid. Het integreren van (informele) statistische inferentie bij de huidige aanpak voor statistische geletterdheid kan zo leiden tot een duurzame verandering in het leren van leerlingen. Het grote voordeel hiervan is dat leerlingen meer leren over inferenties, en hierdoor beter worden voorbereid op hun vervolgstappen in statistiekonderwijs.

Een tweede inhoudelijke bijdrage van het onderzoek betreft het netwerken van theorieën. Het statistiekonderwijs wordt steeds meer gezien als onderscheidend van het wiskundeonderwijs, met

eigen perspectieven op onderwijzen en leren (Groth, 2015). Het integreren van onderwijsperspectieven vanuit verschillende disciplines is wenselijk om tot nieuwe kennis en inzichten te komen. Dit onderzoek draagt bij door theoretische perspectieven uit onderzoek naar wiskundeonderwijs te integreren in onderzoek naar statistiekonderwijs. Het theoretisch perspectief van Realistisch Wiskundeonderwijs (Freudenthal, 1983) werd gebruikt bij het ontwerp van dit leertraject voor statistiek. Op basis van de ontwerpheuristieken vanuit deze theorie werd het black box-paradigma uitgewerkt in concrete leeractiviteiten. Het black box-paradigma bleek effectief als leidende activiteit binnen de leerstappen van het traject. Het theoretisch perspectief van Instrumentele Genese werd gebruikt voor onderzoek naar het leren van en met technologie. Het toepassen van dit perspectief leidde tot inzicht in hoe leerlingen uit vwo 3 concepten ontwikkelen bij het statistisch modelleren in TinkerPlots. Vanuit deze bevindingen lijkt het theoretisch perspectief van instrumentele genese breder inzetbaar binnen onderzoek naar statistiekonderwijs, zoals bij de inzet van andere digitale middelen en in andere onderwijsleerjaren en niveaus.

Methodologische bijdrage

Op methodologisch gebied draag dit onderzoek bij door te laten zien hoe de complexiteit die gepaard gaat bij het experimenteren met innovatief onderwijsmateriaal, overwonnen kan worden door gebruik te maken van ontwerpgericht onderzoek (Bakker, 2018). Een ontwerpgerichte aanpak met een cyclische opschaling in zowel het aantal deelnemers als in de lengte van het leertraject bleek effectief voor het ontwerp en de evaluatie van het innovatieve leertraject. De start met een kleinschalige interventie in de eigen klas van de docent-onderzoeker maakte het mogelijk om de leerdoelen voor het traject te expliciteren en de haalbaarheid ervan te beproeven. De evaluatie was hier vooral gericht op de eerste drie stappen van het leertraject. In deze stappen werd het fundament gelegd van het leertraject en de resultaten uit deze cyclus werden dan ook als uitgangspunt gebruikt voor het ontwerp van de vervolgstappen. In cyclus 2 werd opgeschaald naar drie klassen met 60 leerlingen. De evaluatie was hier voornamelijk gericht op leerstap 4. Aangezien leerstap 5 tot en met 8 een vergelijkbare aanpak en benadering hadden als leerstap 1 tot en met 4, konden we door deze stapsgewijze opschaling een constructief ontwerp realiseren.

In cyclus 3 werd een kwantitatieve benadering gebruikt om de effecten van het leertraject te onderzoeken. Een kwantitatieve aanpak wordt zelden gecombineerd met ontwerpgericht onderzoek. Het kwantificeren, en daarmee samenhangend het opschalen naar een grote groep deelnemers, is een intensief proces. Bij het kwantificeren van de effecten van een leertraject is het van belang dat alle materialen eenduidig, compleet en haalbaar zijn, zodat het traject op de beoogde wijze door docenten kan worden uitgevoerd. Voor de evaluatie van het traject werd een pre-posttestaanpak met een interventie- en vergelijkingsgroep gebruikt. Bij de analyse van de testresultaten werd gekeken naar de prestaties van de leerlingen voor statistische geletterdheid, en tevens naar hun score op leerstapgerichte items. Daarnaast werden in cyclus 3 de werkbladen van leerlingen uit de interventiegroep geanalyseerd. Deze aanpak maakte het mogelijk om empirisch aan te tonen *dat* het leertraject werkt, en tevens *hoe* het leertraject werkt.

Beperkingen van het onderzoek

Zoals elke studie heeft ook dit onderzoek uiteraard beperkingen. In dit onderzoek hebben we aangetoond dat het leertraject voor het introduceren van statistische geletterdheid een positief effect heeft op het leren van statistische inferentie bij vwo 3-leerlingen. Het is echter mogelijk dat andere aanpakken ook werken, waardoor niet zeker is of dit leertraject ook de meest effectieve manier is. Het

onderzoek toont echter wel aan dat het ontworpen leertraject werkt. Bij het evalueren van de effecten van het leertraject is het moeilijk om de generaliseerbaarheid en causaliteit te waarborgen. Door te werken met een grote groep leerlingen met verschillende docenten op diverse scholen- bieden de resultaten een sterke indicatie dat het doorlopen van het leertraject (bij uitvoering zoals beoogd) een positief effect heeft op het leren van leerlingen.

Aanbevelingen voor vervolgonderzoek en de onderwijspraktijk

Vanuit deze studie doen we een aantal aanbevelingen voor vervolgonderzoek en de lespraktijk. Het aanvullen van bestaande statistiecurricula met (informele) statistische inferentie lijkt haalbaar en wenselijk. Het veranderen van bestaande curricula is echter complex. Meer onderzoek is nodig voor succesvolle implementatie. Het paradigma van de black box lijkt tevens toepasbaar voor andere onderwijsniveaus en ook voor meer complexe vervolgactiviteiten zoals hypothese toetsen. Het ontwikkelen van efficiënte leertrajecten hiervoor vereist nader onderzoek. De inzet van technologie is onmisbaar voor statistiek (onderwijs), en voor statistische inferentie in het bijzonder. Wiskundedocenten zijn vaak onervaren in het gebruik van digitale leermiddelen in de les. Tevens worden statistieklessen in de onderbouw van het voortgezet onderwijs vaak verzorgd door tweedegraads wiskundedocenten die onervaren zijn in het doceren van inferentiële statistiek. Onderzoek naar hoe docenten toegerust kunnen worden voor het doceren van inferentiële statistiek met behulp van technologie is wenselijk. Daarnaast kampen veel scholen nog met praktische beperkingen bij de inzet van computers en het installeren van software. Onderzoek naar mogelijkheden om deze praktische obstakels te beperken kan het gebruik van technologie in (statistiek)onderwijs bevorderen. Tot slot veroorzaakte de COVID-pandemie en bijbehorende schoolsluiting een overweldigende toename van technologie in de onderwijspraktijk. Deze actuele ontwikkeling vraagt om onderzoek naar duurzame onderwijsvernieuwingen waarin het gebruik van technologie geïntegreerd kan worden in het reguliere onderwijssysteem.

Persoonlijke reflectie als docent-onderzoeker

Dit onderzoeksproject heeft een rijke bijdrage geleverd aan mijn professionele ontwikkeling als docent en als onderzoeker. Dit onderzoek heeft mijn docentschap op zowel micro-, meso- als macroniveau (Akkerman & Bruining, 2016) versterkt. Op microniveau in mijn eigen lespraktijk als docent heeft dit traject inzicht gegeven in leerprocessen van leerlingen en hoe deze bij het lesgeven gepromoot kunnen worden. Op mesoniveau als docent in de school heeft dit traject geleid tot een meer analytische blik op het schoolsysteem, en op vernieuwende (inter)nationale onderwijsaanpakken en materialen. Op macroniveau van de (regionale en landelijke) onderwijswereld zijn de onderzoeksresultaten via verschillende docentworkshops en artikelen in vaktijdschriften voor wiskundedocenten gedeeld. Diverse docenten zijn vervolgens zelf aan de slag gegaan met het ontworpen leertraject in allerlei onderwijsniveaus—zoals in de vwo-bovenbouw en in het hbo. Deze ervaringen vormden een waardevol vervolg en aanvulling op dit onderzoeksproject. Als onderzoeker heb ik mijn competenties en passie voor onderzoek doen kunnen ontwikkelen. Het functioneren in een wetenschappelijke omgeving heeft mijn kijk op onderwijsonderzoek verdiept en verbreed. Tevens heeft de intensieve samenwerking met internationale collega's mijn visie op onderzoek in allerlei opzichten verruimd. Samenvattend heeft dit traject mijn brede professionele functioneren versterkt—zowel in de klas, binnen de school als binnen de (inter)nationale onderwijs- en onderzoekswereld.