

# ***Verbetering toetspraktijk: onderzoek naar en ontwikkeling van een diagnostische toets aftrekken in groep 5***

*Floor Scheltens (Cito), Jorine Vermeulen (Cito/Universiteit Twente) en Marjolein Nieuwenhuizen (Cito/Universiteit Utrecht)*

Panamaconferentie rekenen-wiskunde 2015

[floor.scheltens@cito.nl](mailto:floor.scheltens@cito.nl), [jorine.vermeulen@cito.nl](mailto:jorine.vermeulen@cito.nl), [marjolein.nieuwenhuizen@cito.nl](mailto:marjolein.nieuwenhuizen@cito.nl)

# Outline

1. Achtergrond project.
2. Projectdoel.
3. Inleiding diagnostisch assessment.
4. 4 Inhoudelijke vragen.
  - a) Vooronderzoek.
  - b) Experiment.
  - c) Vervolgonderzoek.

# Project consortium

- NWO- research & Development project.
- Classroom Assessment – Universiteit Utrecht/Marja van den Heuvel-Panhuizen.
- Samenwerking Universiteit Twente, Cito en Universiteit Utrecht.
- Ontwikkeling van instrumenten (en technieken).

# Wat moet de opbrengst zijn?

*“Eenvoudig te gebruiken diagnostische instrumenten, die een niet te grote tijdsinvestering van de leerkracht vergen. Daarbij moeten ze de leerkracht helpen bij het analyseren van leerlingwerk en kennis geven over de oplossingsstrategieën van de leerlingen.”*

# Vragen die we hier beantwoorden

1. Welke misconcepties hebben leerlingen in groep 5?
2. Welke toetsvragen geven inzicht in de misconcepties van groep-5-leerlingen?
3. Wat vinden leerkrachten van de rapportage van het diagnostische instrument?
4. Verdwijnen deze misconcepties door het onderwijs?

# Onderzoek tussen de jaren '70 en '90

(Brown, & Burton, 1978; Brown, & VanLehn, 1980; VanLehn, 1990)

- Antwoordpatronen en fouten niet stabiel over tijd.
- Geen rekening houden met taakkenmerken.
- Assessments hadden geen formatieve functie.
- Gebaseerd op Amerikaans reken-wiskundeonderwijs.

# Onderzoek > jaren '90

(e.g., Narciss & Huth, 2006)

- Digitale assessmentmethoden.
- Bug-related feedback.
- Meerdere items van hetzelfde type.
- Maar: nadruk op procedure en niet op het begrip van de onderliggende concepten.

# Diagnostiek rekenen-wiskunde Groep 5: Vooronderzoek

- “Klachten”: veel voorkomende fouten optellen en aftrekken.
- Probleem: buggy-strategieën.
- Verklaring: misconcepties.
- Behandeling: aangepast leertraject.



# 1. Welke misconcepties hebben leerlingen in groep 5?

1. Welke fouten maken leerlingen op LVS M5?
2. Is het mogelijk uit het antwoordpatroon van leerlingen strategiegebruik te herkennen op het gebied van aftrekken in de LVS-toets E4, M5 en E5?
3. Is er een verband tussen vaardigheid en een bepaald type fouten?
4. Hoe ontwikkelen fouten zich over de meetmomenten heen?

# Methode vooronderzoek

## Analyses van dataretour van de digitale afnames LVS E4, M5, E5

- schooljaren 2008/2009 tot en met 2012/2013
- E4 13.096 leerlingen
- M5 16.375 leerlingen
- E5 12.760 leerlingen

*Welke fouten maken leerlingen op LVS M5?*

Analyse foutenfrequentie op de opgaven toets M5

	A	B	C	D	E	F	G	H	I	J	K	L
1	item_nr	n	antgoed	ant1	ant2	ant3	ant4	pct_goed	pct_ant2	pct_ant3	pct_ant4	pct_ant5
2	d52172	16375	300	200	290	30	295	81,8	5,2	2,5	0,8	0,8
3	d52137	16375	298	302	297	300	101	75,2	3,7	3,5	2,8	1,8
4	d52074	16375	4	6	3	2	8	75,7	9,6	3,2	1,9	1,8
5	d52185	16375	60	50	40	53	70	86,6	4,3	2,0	0,6	0,6
6	d52171	16375	78	68	58	88	75	88,7	2,1	1,6	1,3	0,8
7	d52038	16375	28	24	3	32	21	80,9	2,4	1,7	1,6	1,5
8	d52062	16375	54	50	2	26	52	63,0	4,6	3,8	3,4	2,6
9	d52126	16375	17	72	18	16	19	51,9	7,1	3,9	3,4	3,1
10	d52188	16375	320	420	400	480	220	54,4	20,6	2,6	1,6	1,5
11	d52072	16375	48	40	42	46	56	71,7	4,3	1,8	1,6	1,6
12	d52056	16375	8	7	6	9	?	68,7	5,2	3,3	2,0	1,6
13	d52386	16375	28	32	22	38	16	62,8	11,5	3,4	2,8	2,6
14	d52394	16375	16	26	24	20	64	67,1	10,5	4,7	4,0	1,8

# Vooronderzoek stap 2

voorbeeld 1

Chi<sup>2</sup>-toetsen: sal  
meest voorkome

Remco spaart voor een nieuwe computer.  
Hij heeft al 680 euro gespaard.  
Hoeveel euro komt hij tekort?

$$76 - 48 =$$



	<b>Foute antw.</b>	<b>32</b>	<b>22</b>	<b>Totaal</b>
420	518	147	665	
400	106	9	115	
<b>Totaal</b>	<b>624</b>	<b>156</b>	<b>780</b>	

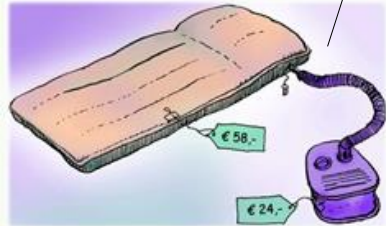


# Vooronderzoek stap 2

voorbeeld 2

Hoeveel euro kosten het luchtbed en de pomp samen? (€58,- + €24,-)

$$76 - 48 =$$



Hoeveel euro kosten het luchtbed en de pomp samen?

Foute antw.	32	22	Totaal
72	90	22	112
92	29	6	35
Totaal	119	156	147

# Vooronderzoek stap 2

- Veel afhankelijkheden (96 van de 756 = 12,7%)
- Vooral binnen de optel- en aftrekopgaven (39 van de 210 = 18,6%)
- Nauwelijks tussen vermenigvuldig- en deelopgaven (2 van de 72 = 2,7%)

## Voorlopige conclusie:

- Leerlingen maken in de toets M5 systematisch fouten in optel- en aftrekopgaven

Is het mogelijk uit het antwoordpatroon van leerlingen strategiegebruik te herkennen op het gebied van aftrekken in de LVS-toets E4, M5 en E5?

- Inhoudelijke analyses van de fouten.
- Analyse van de itemkenmerken.

Conclusie:

- Leerlingen die de toets M5 maken, maken systematisch fouten in aftrekken over het tien- (en honderdtal)tal -> “splitsfouten”

# Waarom keuze naam “splitsfouten”?

- Leerlingen lijken fouten te maken bij het inwisselen van eenheden (bij over het tiental rekenen) en tientallen (bij over het honderdtal rekenen).
- Deze fouten treden vaak op bij het gebruiken van een splitsaanpak.



# Vooronderzoek stap 4

Analyse fouten bij optellen en aftrekken LVS E4 en LVS E5

Conclusie:

- In E4 worden naast “splitsfouten” ook “telfouten” gemaakt.
- In E5 zitten weinig opgaven waarbij leerlingen over het tien- (honderd)tal moeten rekenen waardoor deze fouten niet zichtbaar gemaakt kunnen worden

# Vooronderzoek stap 5

- 8708 leerlingen hebben zowel E4 als M5 gemaakt.
- 13,3% maakt nooit “splitsfouten”.
- 16% maakte geen “splitsfouten” in E4, maar wel in E5.
- 22% maakt “splitsfouten” in E4, maar niet in E5.

Is er een verband tussen vaardigheid en een bepaald type fouten?

Vijf groepen verdeeld I tot en met V.

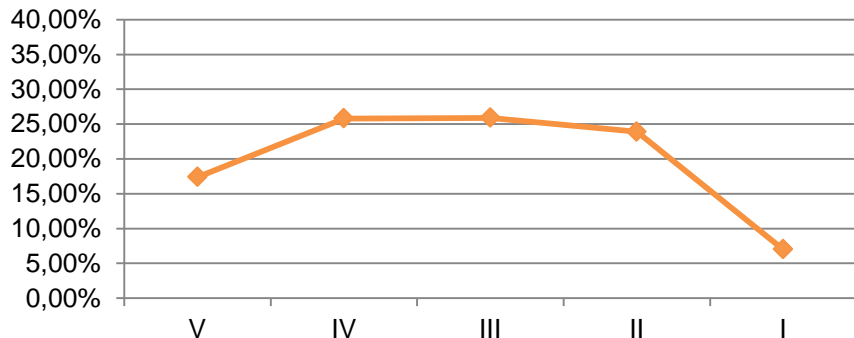
# Percentage splitsfouten per rekenvaardigheidsniveau

## *Lenen van nul*

LVS E4 50 – 38 =

N = 3822

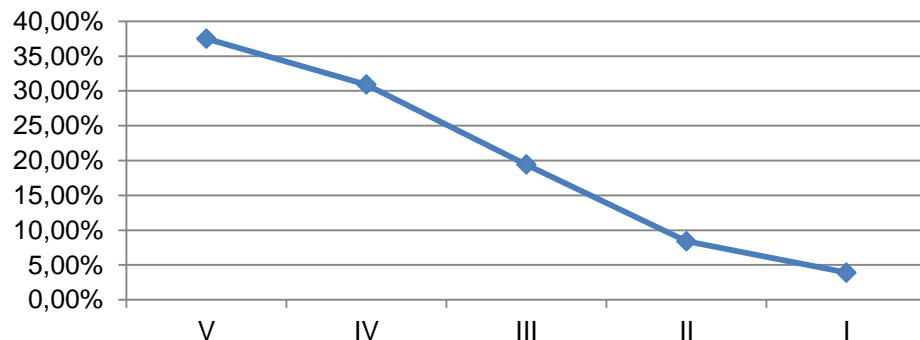
Percentage splitsfouten per normeringsgroep  
d44032



LVS M5 40 – 24 =

N = 2716

Percentage splitsfouten per normeringsgroep  
d52394



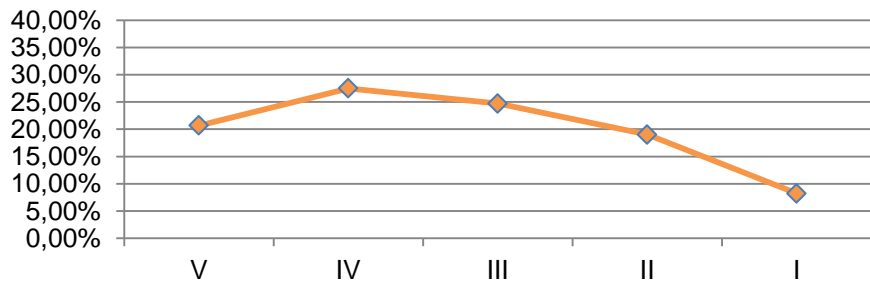
# Percentage splitsfouten per rekenvaardigheidsniveau

## *Tekort aan eenheden*

E4 48 – 39 =

N = 1070

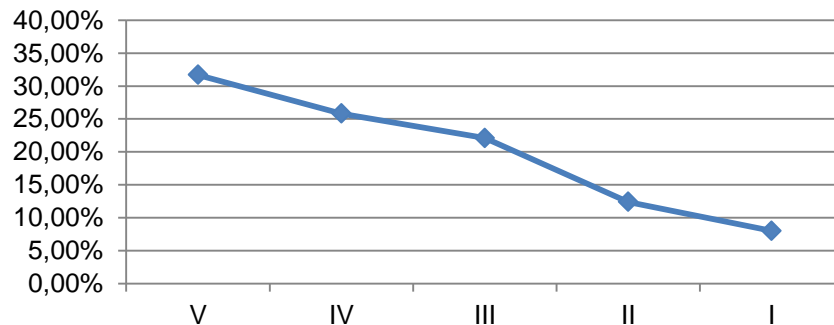
Percentage splitsfouten per normeringsgroep  
d44617



M5 76 – 48 =

N = 2884

Percentage splitsfouten per  
normeringsgroep d52386



# Gevolgen verdere onderzoek

- Alle leerlingen maken splitsfouten, maar met verschillende mate.
- M5 signaal.
- Meting nodig in E5.

# Experiment M5-E5 Participanten: leraren en scholen

26 scholen: 12 experimenteel en 14 controle.

36 leraren: 13 experimenteel en 23 controle.

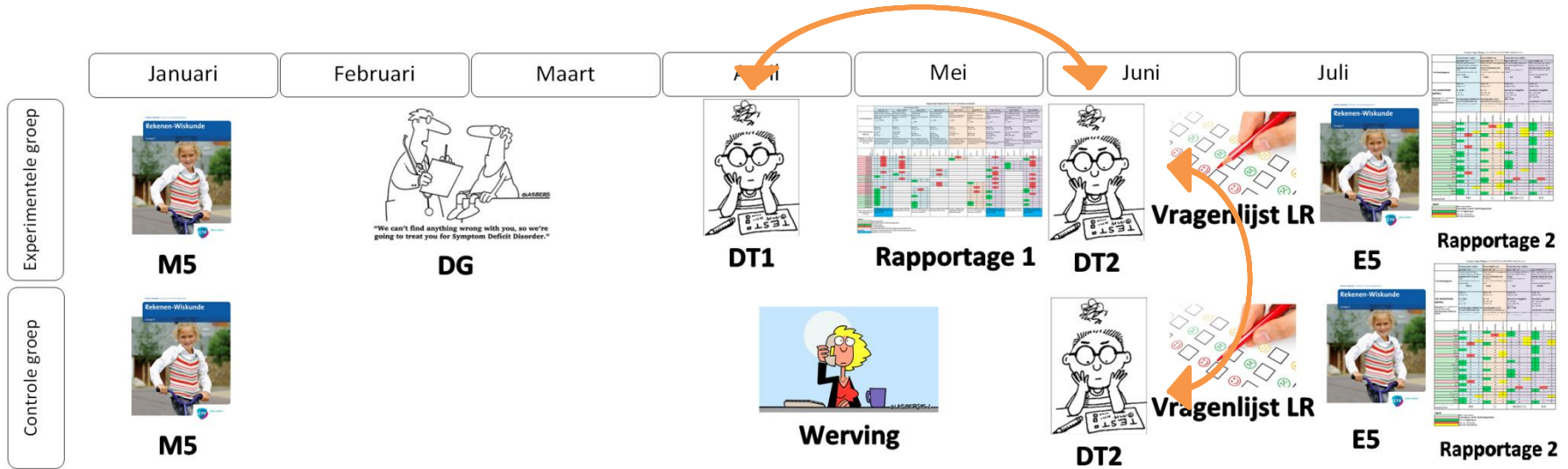
- Lesmethodes
- Speerpunt Rekenen-Wiskunde  
16,7% (n = 6)
- Leeftijd  
25 – 63 jaar Gem. 41,69 (SD = 12,11)
- Leservaring  
1 – 42 jaar Gem. 17,49 jaar (SD = 12,55)
- Differentiatie  
38,9% instructie en oefenstof en 55,6% alleen oefenstof

# Experiment M5-E5 Participanten: leerlingen

- N = 783 groep 5 leerlingen
- 8 jaar en 11 maanden  
(n = 778, sd = 6 maanden)
- Experimentele groep (N = 283)
  - 50,2% jongen
- Controle groep (N = 500)
  - 40,8% jongen



# Experiment M5-E5



# Constructie Diagnostische Toets 1 Stap 1

Klonen aftrekitems M5 LVS toets:

- $76 - 48 =$
- $1000 - 680 =$
- $700 - 32 =$



Remco spaart voor een nieuwe computer. Hij heeft al 680 euro gespaard. Hoeveel euro komt hij tekort?

# Constructie Diagnostische Toets 1 Stap 2

Constructie nieuwe types a.d.h.v. de volgende regels:

1. Geen 8/9 of 80/90.
2. Afstand groter dan 10.
3. Een splitsfout mag niet tot het juiste antwoord leiden: vb.  $61 - 26 =$  ;  $1 - 6 =$  min 5 en  $6 - 1$  is 5. Het maken van een “smaller-from-larger” fout resulteert in het juiste antwoord.

# Resultaat stap 1 en 2: 9 item types

## Lenen (zonder nullen)

Type	Overschrijding	Voorbeeld
1* **	10	$83 - 26 =$
2	10	$453 - 127 =$
3	100	$347 - 62 =$
4	100	$634 - 251 =$

## Lenen van nul

Type	Overschrijding	Voorbeeld
5	10	$70 - 43 =$
6 **	100	$406 - 22 =$
7* **	Beide	$400 - 27 =$

## 1000 – veelvoud van 10

Type	Overschrijding	Voorbeeld
8	10	$1000 - 70 =$
9* **	Beide	$1000 - 340 =$

\* LVS-M5 toets bevat vergelijkbare opgaven

\*\* LVS-E5 toets bevat vergelijkbare opgaven

- Probleem: toetslengte en aantal types
- Oplossing: adaptieve samenstelling papieren toets.

# Constructie Diagnostische Toets 1 Stap 3

Adaptieve samenstelling:

- Koppeling M5; gekloonde items (type 1,7 en 9)
- Antwoorden 4 items M5
- Goed/fout bepaalt opgavetype
- 11 versies met 5 opgavetypen
- 6 items van elk type (tot. 30 items)
- 50/50 context en kaal
- 1/3 meerkeuze

# Stellen van de diagnose

- Per opgavetype
- 2 of meer splitsfouten
- Geautomatiseerde scoring
- Leerkracht rapportage

# Experiment M5-E5 Resultaten: Rapportage

## Voorbeeldopgave

type 1: 83 -

Annemiek heeft 83 kippen.  
Ze doet 27 knikkers in  
bakje en de rest in  
bakje.  
Hoeveel knikkers kon  
groene bakje?

Welke leerlingen hebben deze  
opgaven gemaakt? (zie ook  
handleiding)

## Veel voorkomende splitsfout

Opmerking: het is mogelijk  
dat de leerling via een andere  
oplossingswijze op dezelfde  
fout uitkomt.

$$60 + 4 = 64$$

De leerling begrijpt  
voor een tekort in de  
een tiental moet lenen  
als oplossing de een

+	alle opgaven goed
5	2 of meer splitsfouten
3	leerling die de leerling dit opgave
toetsversie	Leerlingen met dezelfde toetsversie hebben de
	opgavetype zit ook in de M5-toets.

naam	toetsversie	goed	splitsfout	andere fout
Emre A.	2	0	2	4
Daniela	5	3	2	1
Mandy	5	6	0	0
Elanur	6	4	2	0
Gennaro	6	3	0	3
Oumnia	6	5	0	1
Arda	7	5	0	1
Mohammed-Amine	9	5	1	0
Alicia	9	6	0	0
Simon	10	4	1	1
Emre S.	10	5	0	1
Jente	11	6	0	0
Emre K.	11	6	0	0
Naaf	11	6	0	0
Olaf	11	5	1	0
Yasmin S.	11	6	0	0

type 3: 136-52

136 liedjes op zijn  
staan.  
2 af.  
immers houdt Karsten

type 4: 816 - 755

Gertine heeft een boerderij met  
816 bruine en witte kippen.  
755 kippen zijn bruin.  
Hoeveel kippen zijn wit?

de Leerlingen die geen fout  
maakten in één van de vier  
aftrekopgaven in de M5 toets (zie  
handleiding).

$$= 124$$

begrijpt niet dat het  
tallen opgelost moet  
de 100 in te  
draait als oplossing  
om.

$$800 - 700 = 100$$

$$100 + 40 + 1 = 141$$

De leerling begrijpt niet dat het  
tekort in tientallen opgelost moet  
worden door de 100 in te  
wisselen en draait als oplossing  
de tientallen om.

### 3. Wat vinden leerkrachten van de rapportage van het diagnostische instrument?

#### Sterke punten:

- Kleurgebruik.
- Details.
- Voorbeeldopgaven.

#### Verbeterpunten:

- Onduidelijkheid over lege cellen.
- Grootte van de rapportage (A3).
- Leerlingen met veel andere fouten toch groen.



# Enkele leraarreacties

*“In één oogopslag de hele groep in beeld. Kleurgebruik is erg overzichtelijk. Kort en bondig.”*

*“Fijn zo'n voorbeeld erbij!” (sterkpunt voorbeeldopgaven)*

*“Dit [informatie over M5 opgaven] vond ik te veel aan informatie en bracht me niet zo veel.”*

*“Complimenten voor de uitwerking van de informatie, toets en rapportage! Zeer zorgvuldig en uitgebreid.”*

*“Soms lastig het voor het juiste kind af te lezen omdat het zo'n lang blad is schuif je al snel een regel naar beneden.”*

*“Ik vond de rapportage veel informatie geven over de rekenmoeilijkheden. Ik ga het doorgeven aan de volgende leerkracht van groep 5 zodat zij daar in haar lessen rekening mee kan houden. Ik ga de moeilijkheden nog oefenen met de kinderen maar daar heb ik nog geen tijd voor gehad.”*

# Constructie tweede diagnostische toets

- Lengte DT1 te lang → van 30 naar 24 opgaven.
- Niet adaptief.
- Link M5, DT1 en E5 opgaven.
- DT2:
  - Type 1:  $83 - 26 =$
  - Type 6:  $406 - 22 =$
  - Type 7:  $400 - 27 =$
  - Type 9:  $1000 - 340 =$

# Herontwerp rapportage: nieuwe legenda

Vooruitgang t.o.v. DT 1

**Rapportage 1**

Legenda	
	Geen misconceptie.
	Misconceptie bij één of meerdere opgavetypen.
6	Alle opgaven goed.
5	2 of meer splitsfouten.
	Lege cel betekent dat de leerling dit opgavetype niet gemaakt heeft.
toetsversie	Leerlingen met dezelfde toetsversie hebben dezelfde opgavetypen gemaakt.
	Opgavetype zit ook in de M5-toets.

**Rapportage 2**

Legenda	
	Geen misconceptie.
	Misconceptie bij één of meerdere opgavetypen.
	Alle 6 de opgaven goed.
	2 of meer andere fouten.
*	De leerling heeft niet deelgenomen aan toets 1.
**	De leerling heeft niet deelgenomen aan toets 2.

**Rapportage 2**

Kolom splitsfout	
	Misconceptie in toets 1.
	geen misconceptie in toets 1.
	niet afgenomen in toets 1.
	Misconceptie in toets 1, misconceptie in toets 2.

N.V.T. controle groep, geen DT1

# Toepassing nieuwe legenda

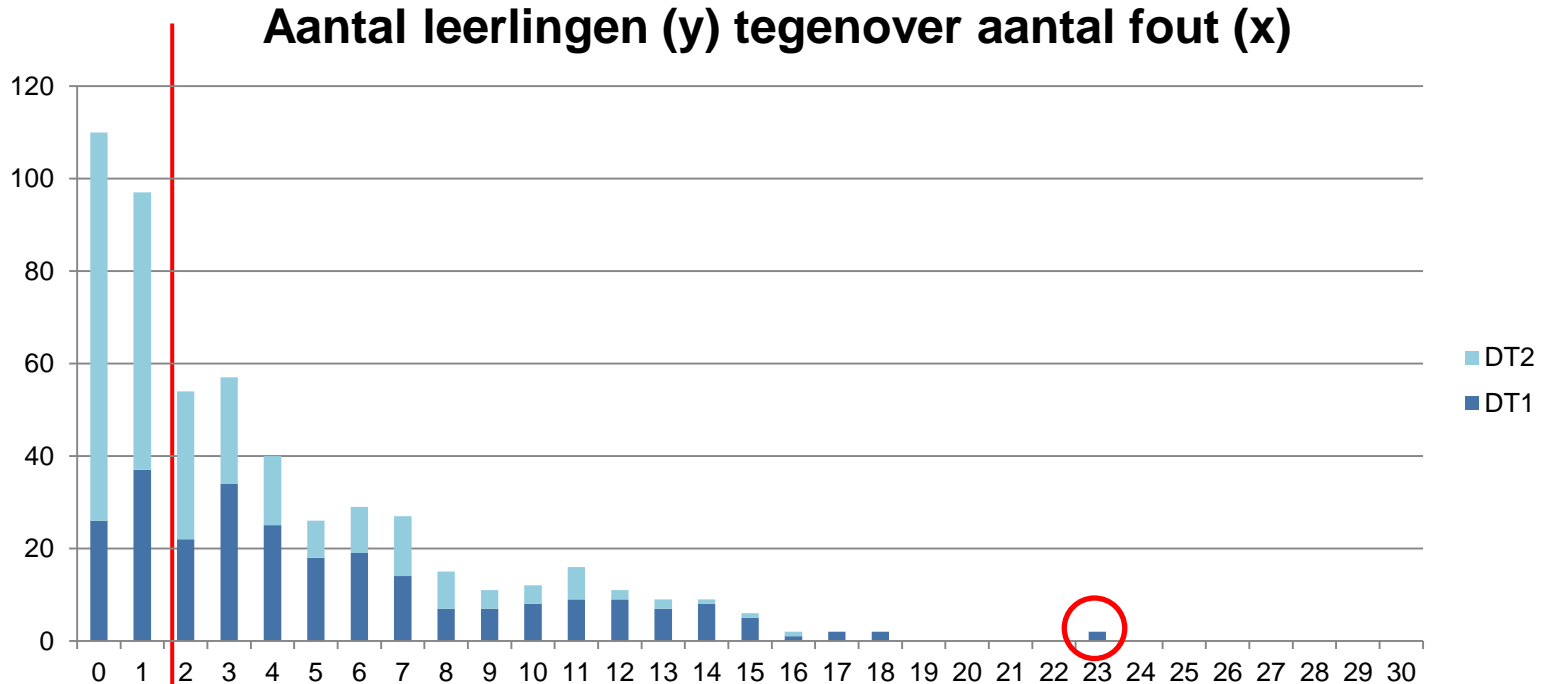
naam	toetsversie	goed	splitsfout	andere fout
Sara G	1	4	1	1
Jenny	1	6	0	0
Sara v	1	2	0	4
Alma	3	2	2	6
Renske	2	2	2	2
Gwen	3	5	1	0
Felijn	3	6	0	0
Hamza	5	3	2	1
Lila	5	4	0	2
Nuria	5	6	0	0
Bjorn	6	4	2	0
Eva	6	3	1	2
Ariëlla	6	6	0	0
Sherenity	8	4	2	0
Carlijn	9	6	0	0
Livia	9	6	0	0
Manouk	9	6	0	0
Ava	10	3	2	1
Troy	10	6	0	0
Noëlle	10	6	0	0
Caspar	11	4	2	0
Domenique	11	2	3	1
Quinten	11	3	2	1
Jasmine	11	5	1	0
Matthijs	11	6	0	0
Bowien	11	6	0	0
Mica	11	6	0	0
Beshoi	11	6	0	0
Sander	11	4	1	1

Naam	goed	splitsfout	andere fout
Aishia	6	0	0
Bram te Linde	6	0	0
Bram te R.	4	2	0
Eelkje	6	0	0
Fedon	6	0	0
Janneke	5	0	1
Jasper	4	1	1
Jochem F.	1	2	0
Jochem V.	3	1	2
Jori	6	0	0
Juliette	5	1	0
Kim	5	1	0
Maarten	5	1	0
Madelon	6	0	0
Malijn	6	0	0
Maud	3	2	1
Max	6	0	0
Nami	5	1	0
Nienke	3	3	0
Pam	5	1	0
Stijn	6	0	0
Thijm	6	0	0
Tygo	6	0	0
Wouter	5	0	1



## 4. Verdwijnen deze misconcepties door het onderwijs?

1. Vergelijking DT1 en DT2 experimentele scholen.
2. Vergelijking experimentele en controle scholen DT2.

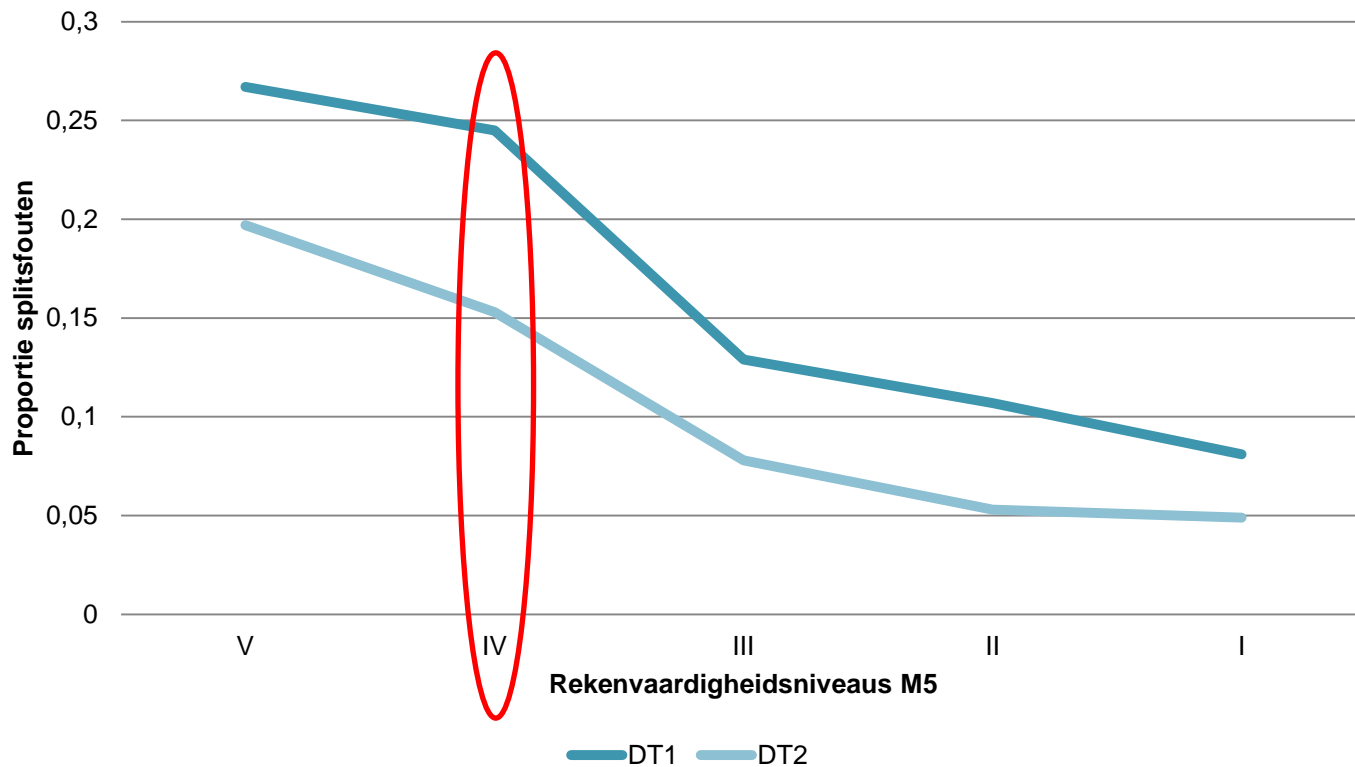


DT1 n = 199 leerlingen (76,0%)

DT2 n = 131 leerlingen (47,6%)

# Afname splitsfouten tussen DT1 en DT2

4.1



1. Vergelijking DT1 en DT2 experimentele scholen.
2. Vergelijking experimentele en controle scholen DT2.



- Aantal leerlingen met een misconceptie DT2:
  - Experimentele scholen: 131 (47,6%)
  - Controle scholen: 240 (49,8%)
- Rekenvaardigheid
  - M5:  $C > E$  (matching onsuccesvol)
  - E5:  $C > E$
  - Groei: E: Gem = 6,67 (SD = 15,99)
  - Groei: E: Gem = 5,01 (SD = 14,87)
    - Niet significant.

# Conclusies

- Splitsfouten nemen af over tijd.
- Het foutenpatroon varieert met vaardigheidsgroep.
- Rapportage droeg niet bij aan afname fouten.
  - Mogelijk te korte interventie tijd.
  - Leraren zeer enthousiast, grote bereidheid voor deelname aan vervolgonderzoek.

# Vervolgonderzoek

- Analyses per opgavetype.
- E4 – B5 toets; meer ruimte voor interventie.
- Alternatieve rapportage mogelijkheden.

# Vragen?

Floor: [floor.scheltens@cito.nl](mailto:floor.scheltens@cito.nl)

Jorine: [jorine.vermeulen@cito.nl](mailto:jorine.vermeulen@cito.nl)

Marjolein: [marjolein.nieuwenhuizen@cito.nl](mailto:marjolein.nieuwenhuizen@cito.nl)



Vind ons op LinkedIn!



Neem een visitekaartje mee!

# Referenties

- Brown, J., & Burton, R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science*, 2(2), 155–192. doi:10.1016/S0364-0213(78)80004-4
- Brown, J., & VanLehn, K. (1980). Repair theory: A generative theory of bugs in procedural skills. *Cognitive Science*, 4(4), 379–426. doi:10.1016/S0364-0213(80)80010-3
- Narciss, S., & Huth, K. (2006). Fostering achievement and motivation with bug-related tutoring feedback in a computer-based training for written subtraction. *Learning and Instruction*, 16, 310–322. doi:10.1016/j.learninstruc.2006.07.003
- Leighton J. P., & Gierl, M. J. (Eds) (2007). *Cognitive diagnostic assessment for education: Theory and applications*. Cambridge, UK: Cambridge University Press.
- VanLehn, K. (1990). *Mind bugs. The origins of procedural misconceptions*. Cambridge, MA: Massachusetts Institute of Technology.
- Thurlings, M., Vermeulen, M., Bastiaens, T., & Stijnen, S. (2013). Understanding feedback: A learning theory perspective. *Educational Research Review*, 9, 1–15. doi:10.1016/j.edurev.2012.11.004