

Clusteropgaven in de rekentoetsen

Tussenrapport mei 2015

Eindrapport verwacht najaar 2015



Inhoudsopgave

INLEIDING	3
VOORLOPIGE OPBRENGSTEN DEELONDERZOEK 1	4
MOEILJKHEID	4
NIET BEANTWOORDE VRAGEN	4
RESPONSTIJD	5
SAMENHANG MET DE REST VAN DE TOETS	5
CONCLUSIE	6

Inleiding

De commissie Bosker (2014) heeft geadviseerd om nader te onderzoeken of het mogelijk is om meerdere vragen bij één context te stellen. Dit advies komt voort uit de gedachte dat clusteropgaven mogelijk een verbetering van de toets vormen. De veronderstelling is hierbij dat het voor leerlingen prettig is wanneer zij zich niet zo vaak hoeven in te leven in een nieuwe context. Op basis hiervan is de volgende onderzoeksvraag geformuleerd:

Maakt het clusteren van vragen het voor leerlingen beter mogelijk om te laten zien wat ze kunnen doordat de mentale belasting verminderd wordt door het aantal contexten te verminderen?

Om dit nader te onderzoeken zijn twee deelonderzoeken ontworpen:

1. een analyse van de prestaties van leerlingen op clustervragen in een reguliere afname;
2. een analyse van prestaties van leerlingen op vragen die in een cluster of als losse vragen worden aangeboden.

Voor deelonderzoek 1 zijn in maart en april 2015 geclusterde opgaven opgenomen in de rekentoetsen in vo en rekenexamens in mbo. De resultaten van de maartafname 2015 van de rekentoets vo en van de mbo- rekenexamens in P3 in april 2015 zijn geanalyseerd. In afwachting van het definitieve rapport over het onderzoek naar clusteropgaven zijn er op basis van de afnamegegevens uit deelonderzoek 1 analyses uitgevoerd met betrekking tot het functioneren van de clusteropgaven ten opzichte van reguliere contextopgaven. Vanuit de onderzoeksvraag zoals die hierboven is weergegeven, bestuderen we vier uitkomstvariabelen:

- moeilijkheid van de opgave
- het percentage vragen dat niet is beantwoord
- benodigde responstijd voor de opgave
- de samenhang van de score op de opgave met de score op de rest van de toets.

Bij ieder van deze uitkomstvariabelen bekijken we a) wat het gemiddelde is voor de reguliere contextopgaven, b) wat het gemiddelde is voor het cluster, en c) wat het gemiddelde is per clusterpositie, dus of dit de eerste, het tweede of het derde vraag binnen een cluster betreft.

Deelonderzoek 2 wordt in mei en juni 2015 uitgevoerd. Dit onderzoek zal ook worden gebruikt om een aantal vermoedens die zijn gerezen bij de analyse van deelonderzoek 1, nader te bestuderen. De resultaten daarvan worden dit najaar in het eindrapport besproken.

Voorlopige opbrengsten deelonderzoek 1

In dit rapport beperken we ons tot voorlopige resultaten van het eerste deelonderzoek. Daarbij gaan we kort in op elk van de vier uitkomstvariabelen en tonen we de resultaten in tabellen. Deze resultaten geven input aan de analyses van deelonderzoek 2 en hebben om die reden een voorlopig karakter. Er zijn na afloop van de afnames in deelonderzoek 1 algemene vragenlijsten afgenomen waarin ook een vraag over clustervragen werd gesteld. Hieruit kwam naar voren dat kandidaten en docenten geen voorkeur of afkeur van clustervragen hebben.

Moeilijkheid

Tabel 1 laat zien dat de geclusterde vragen qua moeilijkheid niet veel afwijken van de reguliere contextvragen.

De hypothese dat een kandidaat bij de tweede en derde vraag meer bekend raakt met de context en dat hij daarom op de tweede en derde vraag beter scoort, wordt voorlopig verworpen. Het tegendeel blijkt het geval: de laatste vraag wordt relatief slechter gemaakt dan de eerste twee (hoewel het verschil bij COE 2F verwaarloosbaar klein is). Waarschijnlijk zijn de tweede en derde vragen in zichzelf al moeilijker. Dit wordt in deelonderzoek 2 nader onderzocht.

Tabel 1. Gemiddelde p-waarden uitgesplitst naar rekentoets/rekenexamen en cluster

	RVO 2F	RVO 3F	COE 2F	COE 3F
Contextvragen	61,7	60,1	47,2	44,3
Clusters	60,4	64,5	55,2	35,8
Cluster - 1	71,8	77,1	52,8	51,6
Cluster - 2	68,1	61,7	60,1	29,3
Cluster - 3	41,3	54,9	52,7	26,6

Niet beantwoorde vragen

Het percentage kandidaten dat een vraag onbeantwoord laat, loopt op binnen een cluster. Dit zou veroorzaakt kunnen worden door een oplopende moeilijkheid van de vragen maar ook door het 'afhaken' van leerlingen. Met afhaken wordt bedoeld dat leerlingen geen zin meer hebben in het beantwoorden van nog meer vragen bij dezelfde context. Ook dit vermoeden zal in deelonderzoek 2 nader onderzocht worden.

Tabel 2. Percentage kandidaten dat geen antwoord heeft gegeven uitgesplitst naar rekentoets/rekenexamen en cluster

	RVO 2F	RVO 3F	COE 2F	COE 3F
Contextvragen	1,1	1,7	3,1	4,7
Clusters	1,6	1,8	2,6	6,4
Cluster - 1	0,4	1,5	2,2	4,1
Cluster - 2	1,4	1,7	2,1	6,3
Cluster - 3	3,0	2,3	3,7	8,8

Responstijd

Bij drie van de vier onderzochte rekentoetsen hebben de geclusterde vragen gemiddeld per vraag een langere responstijd dan de reguliere contextvragen. De clustering van de vragen levert de kandidaten vooralsnog geen tijdwinst op. De verklaring hiervoor zou kunnen liggen in het feit dat de contexten van de clustervragen groter en complexer zijn dan reguliere contexten. Bij clustercontexten moeten meerdere vragen gesteld kunnen worden en dat kan alleen wanneer de context meerdere aanknopingspunten biedt. Dit vermoeden zal nader onderzocht worden in deelonderzoek 2.

Behalve bij de RVO 2F-rekentoets zien we een patroon dat voor de opeenvolgende vragen in een cluster steeds minder tijd nodig is. Behalve in de RVO 3F-rekentoets is de besparing van de hoeveelheid tijd op de tweede en derde vraag onvoldoende om de extra benodigde tijd voor de eerste vraag binnen het cluster in te halen.

Tabel 3. Responstijden in seconden uitgesplitst naar rekentoets/rekenexamen en cluster

	RVO 2F	RVO 3F	COE 2F	COE 3F
Contextvragen	73,7	165,7	88,2	154,0
Clusters	75,1	122,7	79,3	162,4
Cluster - 1	65,4	144,2	90,5	175,3
Cluster - 2	67,2	124,8	73,6	165,7
Cluster - 3	92,8	99,0	73,5	146,1

Samenhang met de rest van de toets

Met betrekking tot de samenhang van de geclusterde vragen met de rest van de toets is er geen effect gevonden van de clustering van vragen. Hiermee wordt bedoeld dat de score op de clustervragen even sterk correleert met de totale toetsscore als de score op een willekeurig ander drietal vragen. De samenhang is ongeveer even sterk voor geclusterde vragen als voor reguliere vragen. Dit betekent dat clustervragen geen betere voorspellers zijn voor de eindscore op de rekentoets/rekenexamen dan andere contextvragen. Bovendien is er geen patroon te bespeuren dat samenhangt met de positie van de vraag.

Tabel 4. Vraag-toetscorrelaties uitgesplitst naar rekentoets/rekenexamen en positie in het cluster (correlatiecoëfficiënt $r \times 100$)

	RVO 2F	RVO 3F	COE 2F	COE 3F
Contextvragen	33,1	32,0	43,7	39,9
Clusters	35,3	34,8	39,7	38,9
Cluster - 1	28,6	36,1	41,8	45,6
Cluster - 2	33,9	35,4	30,7	40,8
Cluster - 3	43,4	32,7	46,7	44,1

Conclusie

De eerste bevindingen tonen aan dat clustervragen niet beter of sneller gemaakt worden. Resultaten uit het tweede deel van het onderzoek zullen meer inzicht moeten geven in hoeverre leerlingen baat hebben bij clustervragen. Op basis van de uitkomsten kan worden bepaald of clustervragen wel of niet in de rekentoets zouden moeten worden gehandhaafd.