

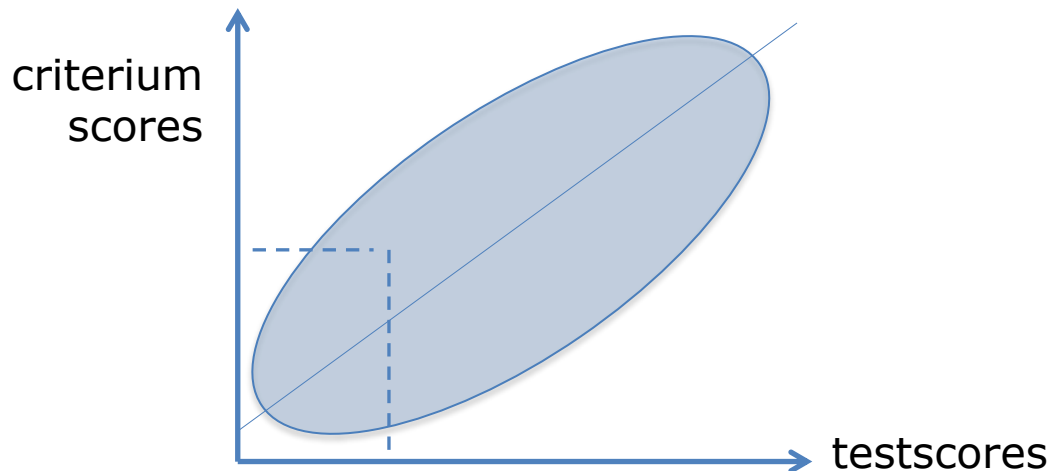
Testtheoretische analyse van ijkingstoetsen

'the elephant in the room'



Testtheorie – correlatie tussen score behaald op ijkingsstoets en behaalde studieresultaten in 1^e jaar

- Ijkingstoets -> leerlingen behalen een **testscore** (bv. 13/20).
- **Criteriumscore**: een score die de studieresultaten in 1^e jaar weergeeft (bv. *percentage behaalde studiepunten na 1^e en 2^e zittijd examens*).
- De **correlatie** is de samenhang tussen de testcores en criteriumscores. Bij positieve samenhang is de correlatie een decimaal getal tussen 0 en 1 (0: *geen samenhang, ellips -> cirkel*; 1: *perfecte samenhang, ellips -> rechte*).
- In de internationale literatuur zijn geen dergelijke proeven bekend met een correlatie hoger dan 0,50 - 0,55



Correlatie voor tot nu toe ontwikkelde ijkingsstoetsen

- VLIR-studie over de diverse ijkingsstoetsen afgenomen in de periode 2013-2017 (*ijkingsstoetsen ingenieurswet. niet meegerekend*):

correlatie met percentage behaalde studiepunten in 1^e jaar (na 2 zittijden): tussen **0,20** en **0,35** (gemiddeld: 0,29)

- Studie door Vanderoost J. et al. (2015) over de ijkingsstoets ingenieurswetenschappen 2013-2014; 350 studenten

correlatie: **± 0,50**

- **De ijkingsstoets ingenieurswetenschappen vertoont een correlatie gelijk aan de best gerapporteerde correlaties in internationale literatuur**

De andere toetsen die tot nu toe uitgetoetst werden blijven daar behoorlijk onder

Andere correlaties ter vergelijking

	correlatie	studie
intelligentie	0,23	Poropat (<i>meta-analyse</i>) Richardson: gelijkaardig
doorzetten-inspannen	0,27- 0,35	Sitzmann; Credé; Richardson (<i>meta-analyses</i>)
aspiratie	0,44	Sitzmann (<i>meta-analyse</i>)
Luci-test (taaltest)	0,37	studie KU Leuven
studiekeuzeadvies klassenraad SO	0,47	studie UHasselt (<i>Limburgse scholen en UHasselt/KU Leuven</i>)
puntengemiddelde SO	0,47	Westrick (<i>meta-analyse</i>), Trapmann; Robbins: gelijkaardig
ijkingstoetsen: - handelsingenieur - biologie-chemie- geologie... - bio-ingenieurswet.	0,20 0,30 0,33	VLIR-studie (voorgaande slide)

Testtheorie: criteriumgrens en selectiegrens

Criteriumgrens en selectiegrens bepalen

Criteriumgrens: wordt bepaald door welke doelstelling we willen bereiken

- **$\geq 80\%$ van studiepunten (SP) verwerven**
(studenten die bijna zeker zullen slagen, desnoods na een individueel traject, positief advies geven)
- **$> 30\%$ van studiepunten (SP) verwerven**
(studenten die bijna zeker niet zullen slagen negatief advies geven)

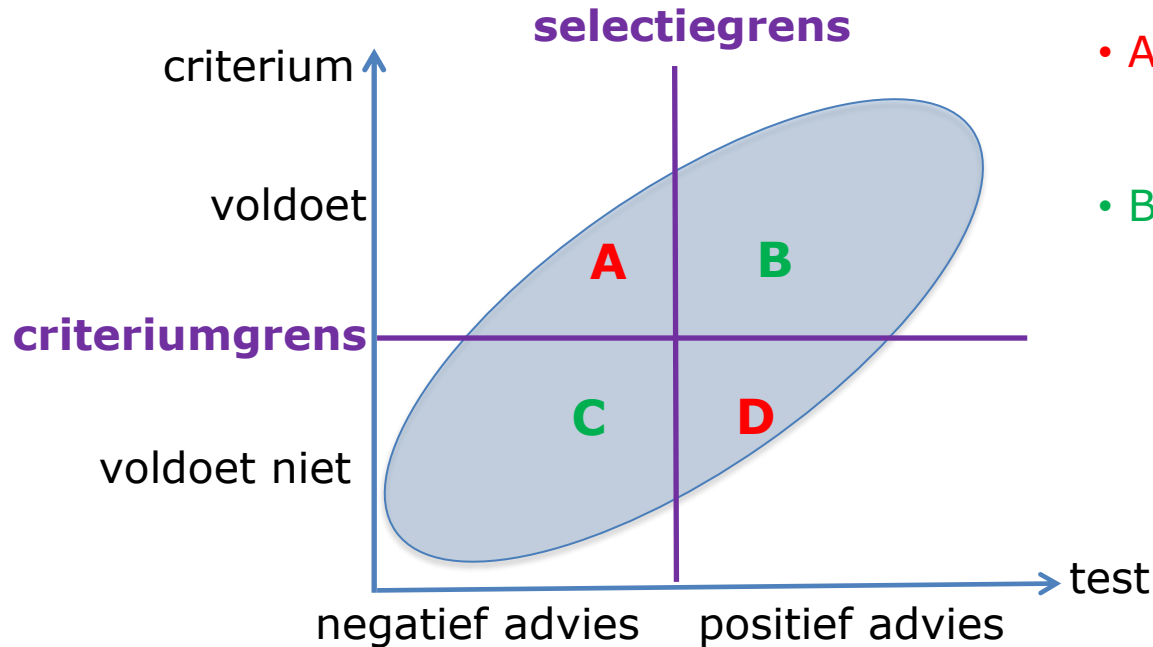
Selectiegrens: percentage studenten dat we een negatief advies willen geven op basis van de scores op de toets

bijv. stel dat de ijkingsstoets op 20 punten gequoteerd wordt;

selectiegrens: $< 8/20$ behaald, dan negatief advies

Testtheorie – je zal altijd foutieve adviezen geven, maar hoeveel ?

Adviezen op basis van proef



- A en D foutieve adviezen
A= vals negatief,
D= vals positief
- B en C correcte adviezen

Als correlatie, criteriumgrens en selectiegrens bekend zijn, kan het percentage studenten in de vakken A, B, C en D berekend worden met testtheorie.

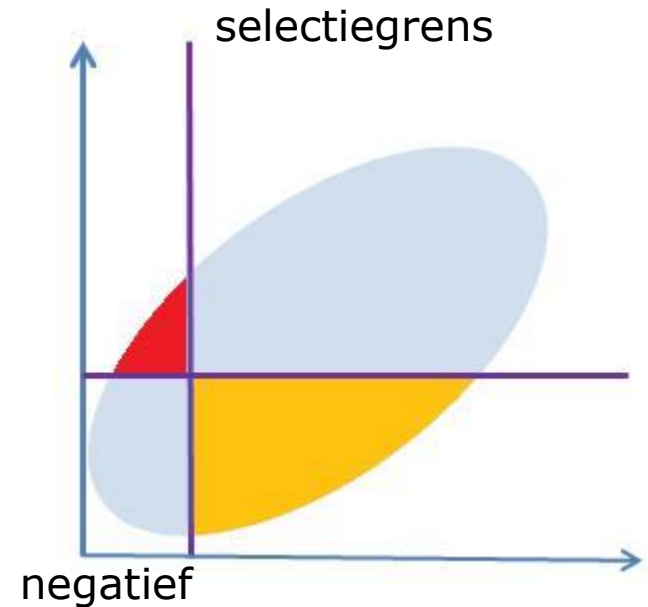
Testtheorie: berekening foutieve adviezen als we alleen de beste studenten een positief advies willen geven

criteriumgrens: $\geq 80\%$ SP verwerven

(studenten die bijna zeker zullen slagen een positief advies geven)

selectiegrens: bijv. 8/20 waarbij 20% van de deelnemers een negatief advies zou krijgen op basis van de ijkingsstoets

berekening foutieve adviezen:



	negatief advies na ijkingsstoets en toch $\geq 80\%$ verwerven	positief advies na ijkingsstoets en toch < 80% verwerven
correlatie 0,30	34%	46%
correlatie 0,50	22%	43%

Testtheorie – berekening foutieve adviezen als we alleen zeer zwakke studenten een negatief advies willen geven

selectiegrens zelfde (20% krijgt negatief advies)

criteriumgrens: $> 30\%$ SP verwerven

(studenten die bijna zeker niet zullen slagen een negatief advies geven)

berekening foutieve adviezen:

	negatief advies na ijkingstoets en toch $> 30\%$ verwerven	positief advies na ijkingstoets en toch $\leq 30\%$ verwerven
correlatie 0,30	61%	21,5%
correlatie 0,50	49%	18,5%



Worden de voorspelde foutieve adviezen ook echt waargenomen bij de ijkingsstoets ?

Ja, bij elke analyse, bij elke ijkingsstoets (*als je de analyse maakt...*)!

Voorbeeld: ijkingsstoets wiskunde-fysica-informatica

Betreft leerlingen die deelnamen aan deze toets in 2013-2017 en inschreven in een van deze opleidingen; aantal: 237

- Correlatie tussen ijkingsstoets en studierendement 1^e jaar: **0,29**
- Criteriumgrens: geslaagd in 1^e jaar (na 2 zittijden)
Selectiegrens: **5/20** (*19% behaalt <5/20 en krijgt negatief advies*)

	negatief advies (<5/20) en toch geslaagd in 1 ^e jaar	positief advies (≥5/20) en toch niet geslaagd in 1 ^e jaar
reële data ijkingsstoets	29%	46%
testtheorie voorspelt	34%	46%

Ijkingstoetsen: de (onvermijdelijke) bias voor gender en SES is ook hier aanwezig

Deelnemers aan ijkingstoetsen in 2019 (VLIR-studie)

- **Man/vrouw**

	deelgenomen	geslaagd op toets
vrouw	1500	37,8%
man	2204	57,6%

- **Sociaal-economische status**

indicator-leerling: moeder geen diploma hoger secundair onderwijs of beursgerechtigde leerling

	deelgenomen	geslaagd op toets
indicator leerling	337	39,5%
niet-indicator	3249	50,6%

Nood aan validatiekader voor ijkingstoetsen

Ijkingstoetsen hebben inherente en onvermijdelijke handicaps (beperkte correlatie, foutieve adviezen, biaseffecten voor gender en SES,...).

Dit speelt ons niet alleen parten als men studiesucces wil voorspellen of go/no-go adviezen wil geven, maar evenzeer als men een remediëringsadvies wil geven. Als de correlatie te laag is met slagen in 1^e jaar zal ook het geadviseerde remediëringstraject niet het juiste zijn, met alle gevolgen vandien.

Er is dringend nood aan een **validatiekader** dat aangeeft wat we aanvaardbaar vinden:

- criteriumgrens bepalen (doelstelling bepalen)
- minimum % studenten dat slaagt op de toets
- ondergrens voor aanvaardbare minimum correlatie
- bovengrens voor aanvaardbaar % foutieve adviezen
- bovengrens voor aanvaardbaar bias-effect
- ...

Referenties

Nauwelaerts, E., Doumen, S., & De Schepper, L. (2014). Effectiviteit van toelatingsproeven. *Tijdschrift voor Onderwijsrecht en Onderwijsbeleid*, 1-2, 79-85.

Nauwelaerts, E., & Doumen, S. (2015). Selectie- en oriënteringsproeven doorstaan de toetsing niet. Op zoek naar een andere aanpak. *TH&MA: Tijdschrift voor Hoger onderwijs & Management*, 2015 (2), 18-25.

Nauwelaerts, E., & Doumen, S. (2016). De belangrijkste studentfactoren voor studiesucces in het hoger onderwijs: een grootschalig literatuuronderzoek. *Tijdschrift voor Onderwijsrecht en Onderwijsbeleid*, 2015-2016 (5), 373-385.

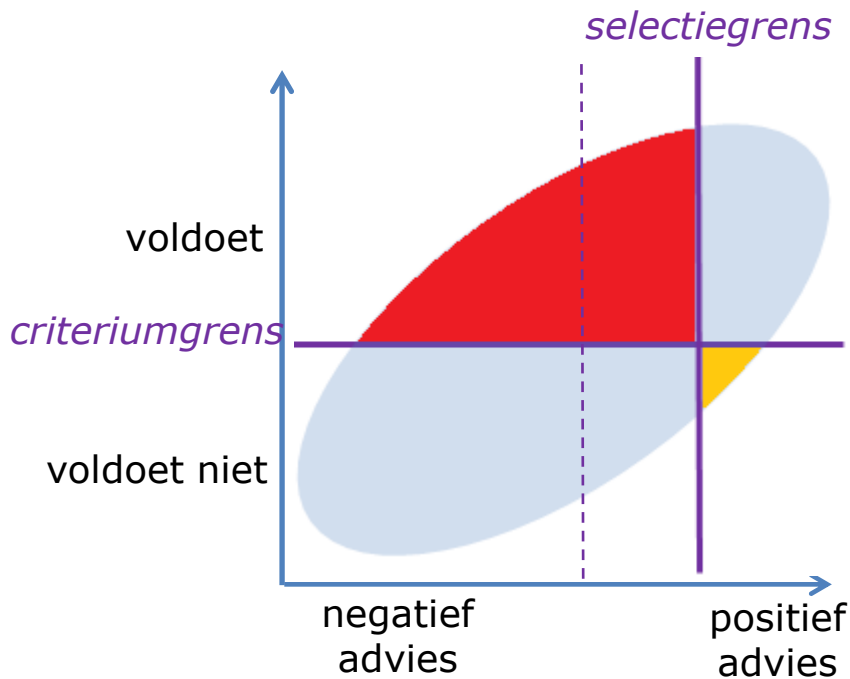
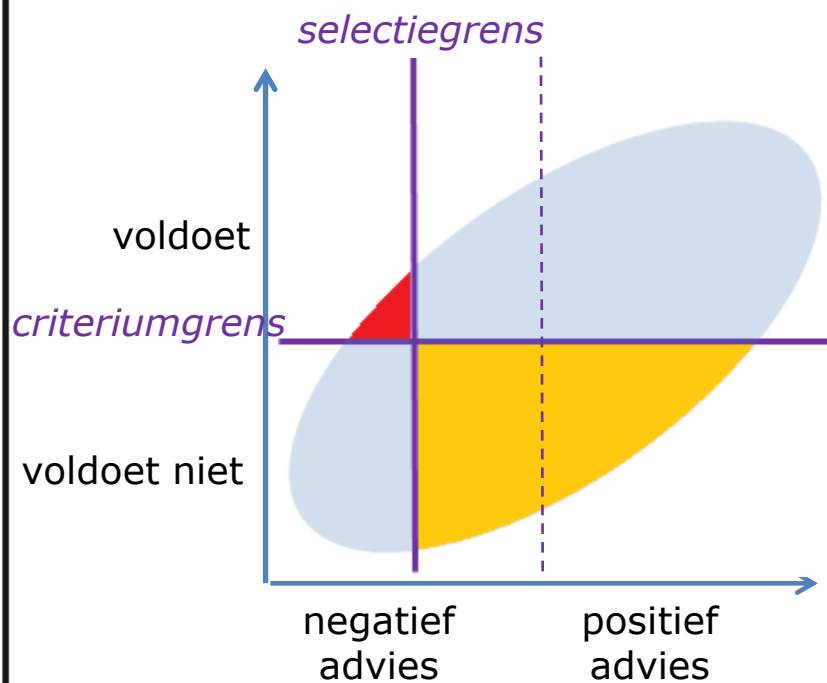
Nauwelaerts, E., Doumen, S., Verhaert, G., Molenberghs, G., & De Schepper, L. (2018). Een nieuw oriënteringsinstrument – met hoge correlatie met reëel studiesucces – voor de begeleiding van de studiekeuze naar de universiteit toe. *Tijdschrift voor Onderwijsrecht en Onderwijsbeleid (T.O.R.B.)*, 2018-2019 (1-2), 90-96.



UHASSELT

KNOWLEDGE IN ACTION

Verschillende selectiegrenzen



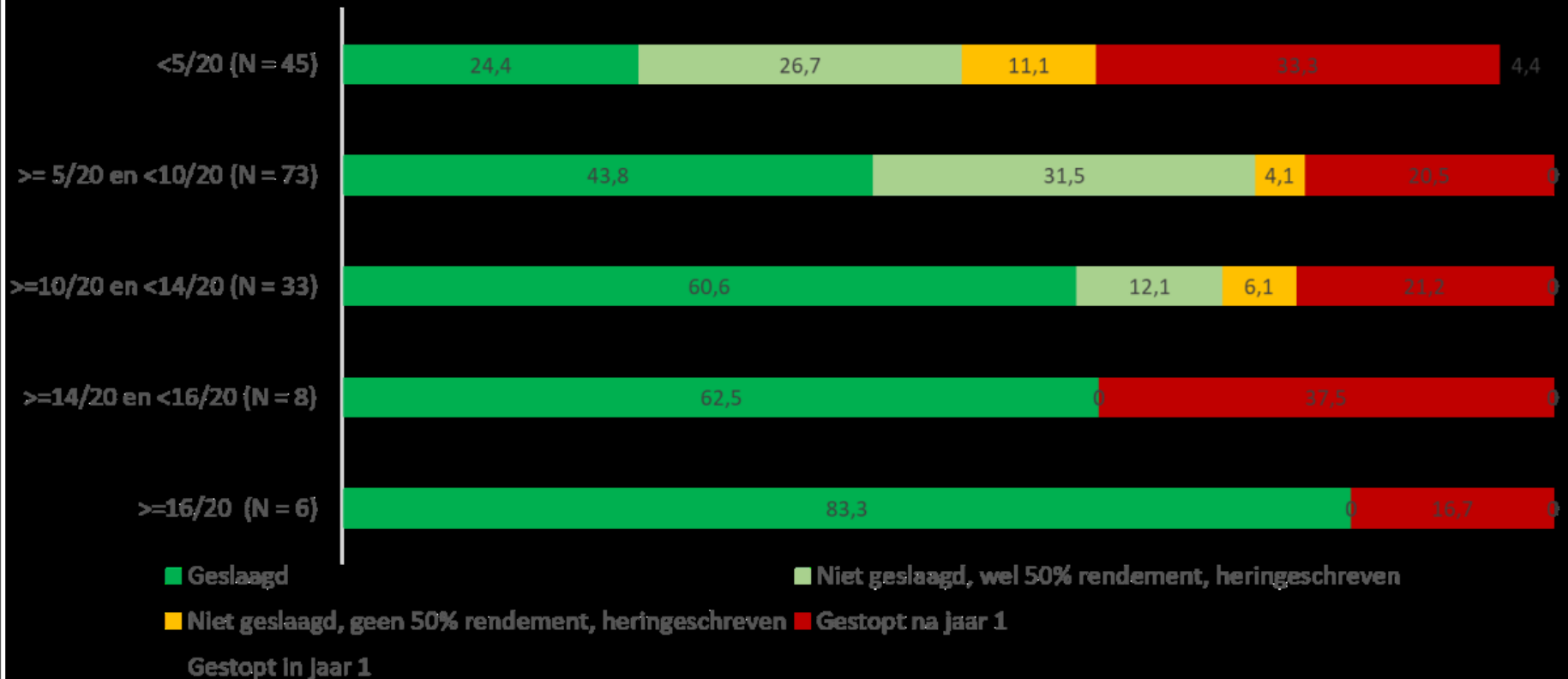
bv. **toelatingsproef arts/tandarts**



Ijkingstoets biologie-chemie-geologie-

Periode 2013-2017

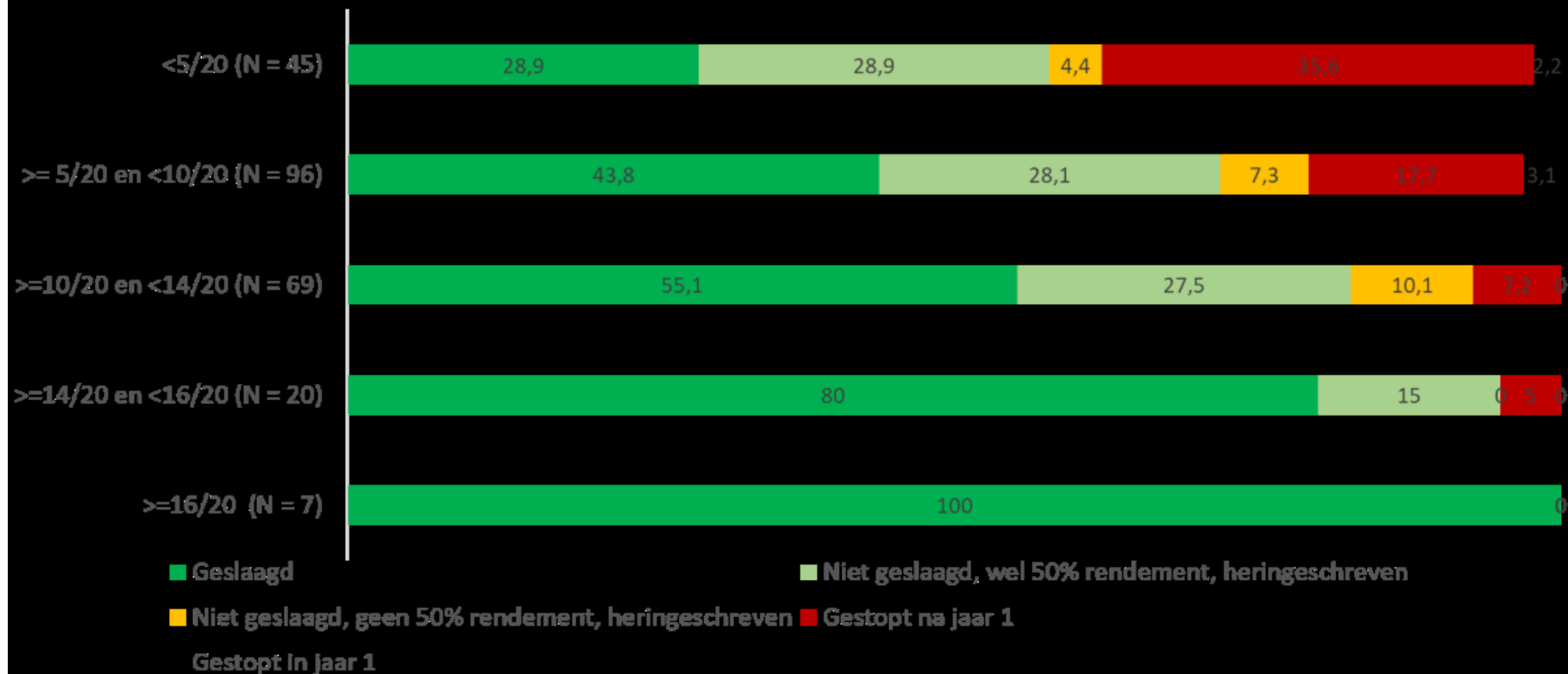
Resultaat eerste jaar opleiding (CBBGG 2013-2017, N = 165)



Ijkingstoets wiskunde-fysica-informatica

Periode 2013-2017

Resultaat eerste jaar opleiding (Wiskunde, Informatica, Fysica 2013-2017, N = 237)



UHASSELT

KNOWLEDGE IN ACTION