

Wanneer is het een voldoende?

Normeren in RemindoToets

Uit een goede toets komen toetsscores. Maar wat betekent zo'n score nu precies? Geeft 55 van de 100 punten behalen een voldoende, of 60 punten, of misschien 80 punten? Wanneer is het een voldoende en wanneer is het een onvoldoende? Deze vragen worden beantwoord door de toets te normeren. Het resultaat van normering is vaak een cijfer en een bepaling bij welk punt precies de kandidaat een voldoende behaalt. Die keuze kan grote gevolgen hebben. Daarom is het belangrijk dat je zorgvuldig overweegt hoe je de toets wilt normeren.

Er zijn twee manieren van normeren: relatief en absoluut. Daarnaast biedt RemindoToets extra nog een compromismethode: Cohen-Schotanus.

1. Relatief normeren

Relatief normeren doe je na afloop van een toetsafname. Het relatief normeren is ooit bedacht toen bleek dat kandidaten soms de dupe werden van slechte toetsen. Bij het relatief normeren wordt gekeken naar hoe de toets daadwerkelijk is gemaakt, naar de punten die door de kandidaten zijn behaald. Alle toetsscores worden met elkaar vergeleken. Je stelt dan bijvoorbeeld vast dat de 10% kandidaten die het hoogst hebben gescoord op het examen een 10 krijgen, de 10% daaronder een 9, enz. Omdat bij het relatief normeren de score van een kandidaat afhankelijk is van de score van de groep, is het van belang dat de scores representatief zijn voor wat te verwachten is van de doelgroep. Dat betekent dat je relatief normeren alleen kunt toepassen wanneer er sprake is van een grote groep kandidaten. Scholen komen dan vaak in de problemen bij herkansingsexamens, waar minder kandidaten voor zijn die over het algemeen minder goed scoren op een examen. Daarnaast is relatief normeren erg fraudegevoelig: kandidaten zouden vooraf kunnen afspreken om delen van de lesstof niet te leren, zodat de normering minder streng verloopt.

2. Absoluut normeren

Bij het absoluut normeren stel je voorafgaand aan de toetsafname de norm vast. Je meet de toetsscore met een vooraf vastgesteld criterium. Meestal bestaat dit criterium uit een percentage van het aantal punten dat behaald moet worden voor een voldoende. Kandidaten weten dan vooraf precies wat nodig is om de toets te halen. Het nadeel van absoluut normeren zijn de onregelmatige slagingspercentages. Het is een flinke uitdaging om bijvoorbeeld ieder jaar een examen te maken van gelijke zwaarte. Daarom kan het ene jaar een slagingspercentage van 70% hebben, terwijl het jaar erop maar 40% slaagt voor de toets.

In RemindoToets moet je een normering invoeren, voorafgaand aan de toetsafname. Na de toetsafname kun je de normering vervolgens aanpassen als je dat wilt, dus kun je in theorie kiezen voor beide methoden.

Veel toetsontwikkelaars kiezen bij het normeren voor de absolute normering. Ze willen de toetsuitslagen niet laten afhangen van het niveau van de groep examenkandidaten of van de

kwaliteit van het onderwijs dat eraan vooraf ging. Bij een beroep past vaak een bepaalde norm en het diploma of certificaat wordt pas gegeven als de kandidaat aan deze norm voldoet.

Toch is de praktijk vaak genuanceerder. Hoe vaak wordt een beroep nu zo concreet dat het ene vak binnen een opleiding nu precies de doorslag geeft tussen iemand die het beroep goed kan uitoefenen en iemand die dat niet kan? En hoe zeker ben je van de kwaliteit van de toets? Zijn de vragen van de originele toets even moeilijk als die van de herkansingstoets bijvoorbeeld?

Voor dit soort vraagstukken zijn er compromismethoden ontworpen. Eén van de meest gebruikte voorbeelden hiervan is de methode Cohen-Schotanus. RemindoToets geeft de mogelijkheid om deze normeringsmethode na de toetsafname toe te passen.

Dr. Janke Cohen-Schotanus

De grondlegger van de Cohen-Schotanus-methode is Janke Cohen-Schotanus. Ze studeerde psychologie aan de Rijksuniversiteit Groningen en promoveerde in 1994 tot doctor in de Medische Wetenschappen op het proefschrift 'Effecten van curriculumveranderingen, studiewaardering, studeergedrag, kennis en studiedoorstroom en een veranderd medisch curriculum'. Dr. J. Cohen-Schotanus werkt op dit moment als hoofd van het Centrum voor Innovatie en Onderzoek van Medisch Onderwijs in het UMCG. Zij richt zich in haar onderzoek met name op nieuw ontwikkelde onderwijsactiviteiten in de initiële en de vervolgopleidingen. Aandachtspunten hierin zijn de invloed van toetsing op studeergedrag, doorstroom en rendement, en de effectiviteit van het leren op de klinische werkplek. "Een cesuurbepaling waarbij geen rekening wordt gehouden met de moeilijkheid van de toets, deugt per definitie niet. We laten veel studenten onterecht zakken", stelde Cohen-Schotanus in een interview met het vakblad Examens (augustus 2017).

3. Cohen-Schotanus

De methode Cohen-Schotanus is een compromismethode. In de praktijk bepaal je voorafgaand aan de toetsafname de norm (absoluut normeren). Na de toetsafname kun je de Cohen-Schotanus-methode gebruiken om de norm te nuanceren. De methode kijkt naar de maximale score die is behaald op de toets. Dus in geval van de 100 punten, is er wellicht niemand die echt 100 punten heeft behaald, maar wel 98 punten. Cohen-Schotanus rekent dan uit wat de gemiddelde score is van de beste 5% van de kandidaten tijdens deze toetsafname. Dat wordt vervolgens beschouwd als de maximale score. Waarbij de cesuurgrenzen iets verlegd worden.

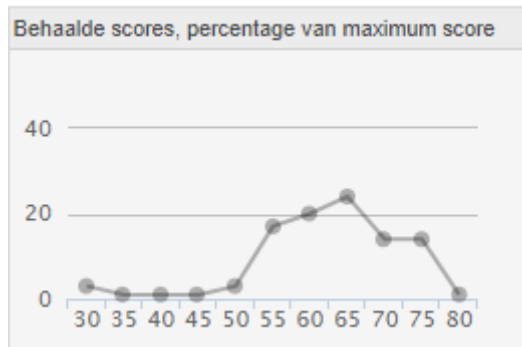
Een voorbeeld:

Elisa maakt een toets Anatomie voor 75 studenten. Voor het examen van 100 meerkeuzevragen kan maximaal 100 punten gehaald worden. Iedere vraag levert 1 punt op.

Relatief normeren

Vooraf wordt niet bepaald wanneer de kandidaten een voldoende hebben. Dit kan dus niet worden gecommuniceerd. Dat getuigt van weinig transparantie.

De toets Anatomie geeft deze uitslag:



Figuur 1 frequentieverdeling toets Anatomie

Je ziet dat de studenten hebben gescoord tussen de 30 en 80 punten. Elisa bepaalt dat 10% van de best scorende studenten een 10 verdienen van daaruit rekent ze omlaag.

10% van de studenten scoren tussen de 78 en 80 punten.

Ze zet de cesuur als volgt neer:

Bij 78 punten of hoger krijgen de studenten een 10. Om het cijfer te berekenen gebruikt ze deze formule, waarbij studenten maximaal een tien kunnen halen:

$$cijfer = \frac{puntenaantal}{7,8}$$

- De grootste groep studenten heeft 65 punten gehaald en de studenten krijgen hiermee een 8,3.
- De laagst scorende student krijgt met 30 punten een 3,8.

Bij deze methode gaan we ervan uit dat je kunt verwachten dat in de groep kandidaten een 10 zou vallen. In dit geval, met een groep van 75 kandidaten, is dat discutabel. Met een groep van bijvoorbeeld 1000 kandidaten, zou je dat met een grotere zekerheid kunnen zeggen en zou deze methode beter passen.

Relatief normeren kan op veel verschillende manieren gebeuren. Om de betrouwbaarheid van de toets te waarborgen, kun je dit alleen toepassen wanneer er sprake is van een grote groep kandidaten en wanneer er in het toetsbeleid duidelijk is aangegeven welke normeringsmethode wordt toegepast, zodat dit voor iedere keer dat dezelfde toets wordt afgenomen, hetzelfde is. In het voorbeeld is een toetsafname van 75 kandidaten gebruikt. Je kunt je bij dit aantal afvragen of het wel

realistisch is dat je mag verwachten dat er altijd een 10 valt. In de praktijk wordt vaak een minimum van 50 kandidaten aangehouden om relatief te normeren.

Begint de cesuur bij 0 of 1?

Welke wijze van normeren ook wordt gebruikt, je moet je altijd afvragen waar je de normering wilt laten beginnen. Kunnen kandidaten ook een 0 halen voor een toets of is de minimale score een 1?

Het lijkt een kleine vraag, maar het heeft effect op de cesuurgrens. Dus bepaalt het in sommige gevallen of die ene kandidaat het examen net wel of net niet heeft gehaald.

Wanneer een 1?

Het is heel gebruikelijk om kandidaten bij het deelnemen aan een toets de eerste punt(en) cadeau te geven, dus de normering te beginnen bij 1 (voor 'het juist invullen van hun naam'). Omdat het veel gebeurt, lijkt het meer de standaard te zijn. En omdat je graag kandidaten eerlijk en betrouwbaar wilt toetsen, zul je in de praktijk vaak de cesuur beginnen bij 1.

Wanneer een 0?

Wanneer de cesuurgrens doorslaggevend is voor de beroepsbeoefening. Bijvoorbeeld wanneer de kandidaat 100% van de toetsscore moet behalen, dan is het niet verstandig om de eerste punten cadeau te geven. Dan start je de opbouw van de cesuur bij 0.

Absoluut normeren

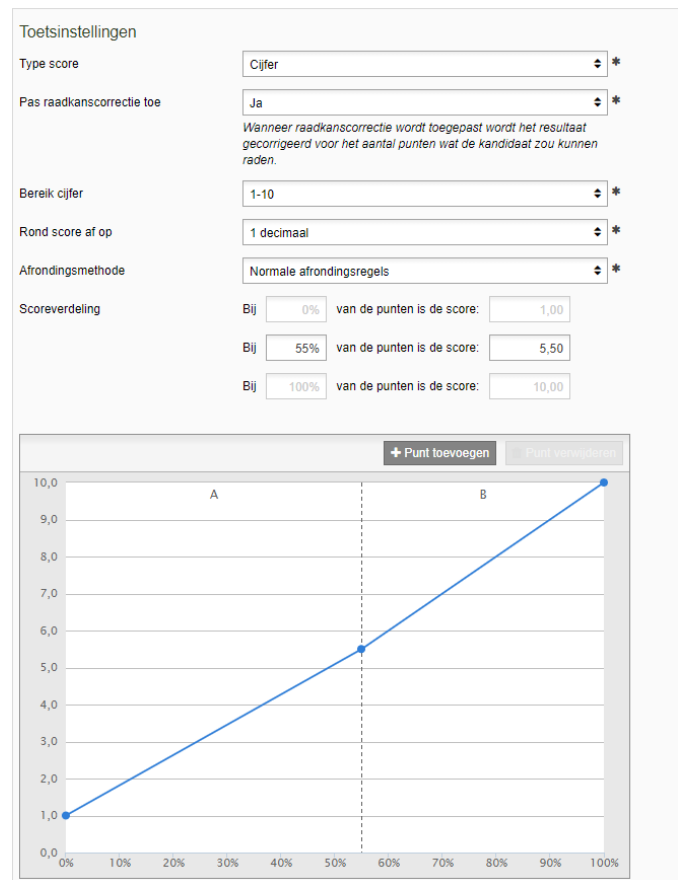
Om voor de toets Anatomie absolute normering toe te passen, is het allereerst belangrijk dat Elisa bepaalt wat de grens moet zijn tussen het zakken en slagen. Veel mensen denken dat 55 van de 100 punten automatisch een voldoende moet opleveren, maar dat hoeft niet zo te zijn. Hoe belangrijker de toets is in het hele curriculum van de opleiding en hoe minder vaak dit wordt getoetst, hoe urgenter de toets is en hoe hoger de cesuurgrens. Gaat het om een toets voor verpleegkundigen en is dit de enige toets waarmee kennis over de anatomie wordt geëxamineerd, dan is het een gewichtige en kritieke toets over een belangrijk onderwerp en mag de cesuurgrens relatief hoog liggen. Komt anatomie als toetsonderdeel verder in de opleiding nog een keer aan bod. Of weegt het in het betreffende leerjaar nog niet zo zwaar, dan hoeft de cesuurgrens niet zo hoog te liggen en kun je de conclusie trekken dat 55 punten een voldoende oplevert.

Of de cesuurgrens rond het midden ligt, is afhankelijk van het antwoord op de volgende vraag:

Hoe belangrijk is het dat de kandidaat de examenstof op dit moment (volledig) beheerst?

Om tot een goede absolute normering te komen, heeft Elisa meerdere mogelijkheden, waarvan we hier twee uitwerken:

1. Ze stelt vast dat 55% van de punten een 5,5 moet opleveren en voert een raadkansberekening uit. De raadkansberekening voert ze uit, omdat er sprake is van een toets met alleen gesloten vragen. Iemand zonder kennis van het onderwerp zou toch punten kunnen behalen, simpelweg door antwoordopties willekeurig aan te klikken. Daar wil Elisa zoveel mogelijk rekening mee houden. In RemindoToets kan zij geautomatiseerd de raadkans mee laten wegen. Ze stelt de cesuur dan in zoals in figuur 2 wordt getoond.

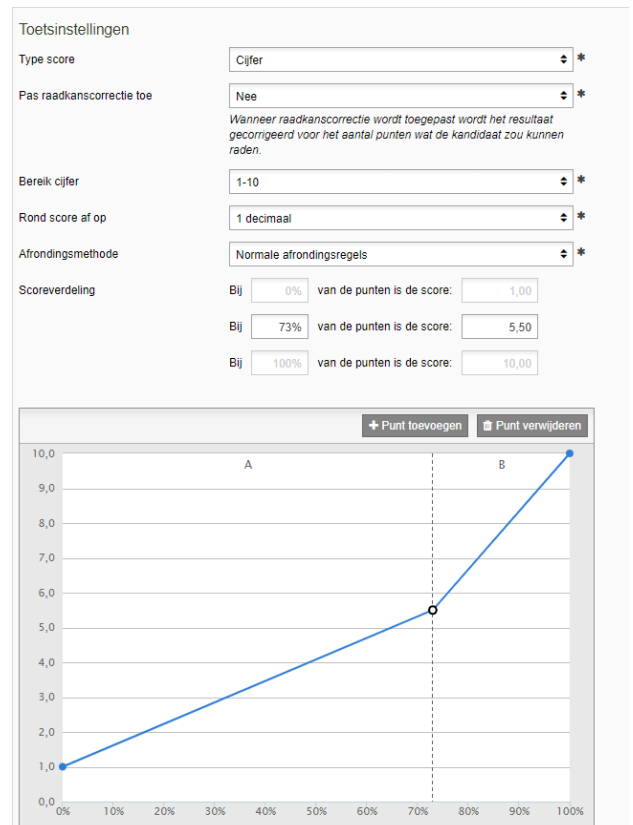


Figuur 2 absoluut normeren

Het nadeel van deze methode is dat de toetsvragen gelijkwaardig worden beschouwd. Hoe goed je de toets ook maakt, vragen zijn nooit helemaal gelijkwaardig. De ene vraag is nu eenmaal moeilijker dan de andere vraag. Om dit mee te kunnen wegen zijn er nog andere methoden waarmee Elisa tot een goede absolute cesuur kan komen.

2. **Angoff-methode:** De Angoff-methode is gericht op de vraag: wat is de kans dat de grenskandidaat deze vraag juist beantwoordt? Voor de Angoff-methode gaat Elisa met een team collega's om de tafel en bekijken ze de 100 vragen. Ze neemt daarbij een typische 5,5-student in gedachten. Eén uit haar klas of één die altijd in het verleden net een 5,5 scoorde. Met deze persoon in gedachten, bekijkt ze de toetsvragen en stelt ze de vraag: hoe groot is de kans dat de 5,5 student deze vraag juist beantwoordt? Voor de ene vraag zal dat 80% zijn voor de andere vraag 30% en zo doet ze dit voor alle 100 vragen. Het percentage ligt nooit lager dan 25%, omdat dit de raadkans al is. Je hebt tenslotte minimaal 25% kans om het juiste antwoord te raden (tenzij het een strikvraag is). Dit doet ze samen met een team collega's, zodat er een objectief beeld ontstaat van de vragen. Het gemiddelde van deze percentages bepaalt waar de 5,5 wordt geplaatst. Omdat ze bij de vraag al rekening houdt

met de raadkans, zet ze de raadkanscorrectie uit. Wanneer Elisa deze methode toepast, komt er een gemiddelde uit van bijvoorbeeld 72%. Ze stelt de cesuur dan in, zoals getoond in figuur 3.



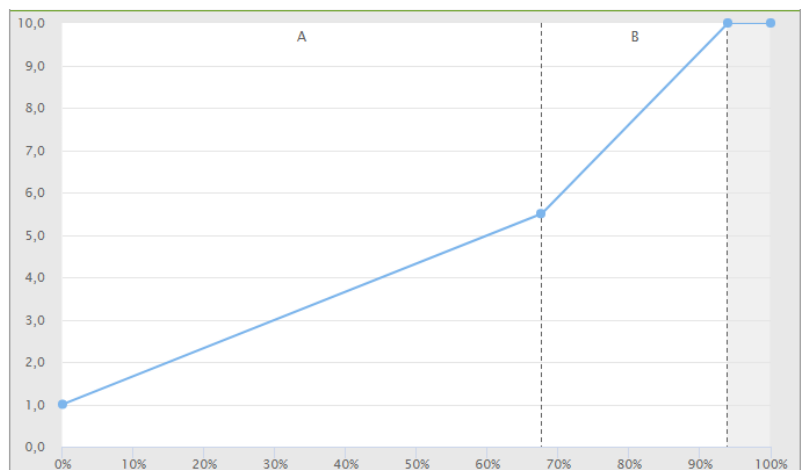
Figuur 3 Angoff methode

Cohen-Schotanus-methode

De Cohen-Schotanus-methode geeft een compromis tussen de absolute en relatieve wijze van normeren. Met deze methode gaat Elisa vooraf uit van een absolute cesuurmethode, zoals één van de bovenstaande methodes. Vervolgens wordt de toets afgenomen en kijkt Elisa na de toetsafname naar de toetsresultaten. De Cohen-Schotanus gaat uit van de groep best scorende kandidaten en stelt de maximale score van de toets bij naar wat er maximaal is gescoord.

In dit voorbeeld heeft Elisa de Angoff methode gebruikt als absolute cesuurmethode en de cesuur staat ingesteld zoals in figuur 3, de cesuurgrens ligt op 72%.

In de planning van de toets, gaat ze naar de betreffende toetsafname en wijzigt de matrijsinstellingen. Daarbij geeft ze aan dat ze de Cohen-Schotanus-methode wilt toepassen. RemindoToets geeft automatisch het 95^e percentiel aan van de scores die



Figuur 4 Cohen Schotanus methode

zijn behaald in de toets. Dat betekent dat 5% hoger heeft gescoord dan deze waarde en 94% lager heeft gescoord dan deze waarde. Wanneer Elisa dit doet, ziet ze dat het 95^e percentiel staat op 94%. Dat betekent dat de 5% best-scorende kandidaten gemiddeld 94 van de 100 punten hebben gescoord. Volgens de Cohen Schotanus-methode zou dit de maximale score moeten zijn. De normering schuift daarmee op en ziet eruit zoals figuur 4. Daarbij geeft RemindoToets direct inzicht in wat dit betekent voor de uitslagen van de kandidaten.

De toets normeren

Stap 1: Bepaal als organisatie het beleid.

Allereerst is het belangrijk om vast te stellen welke manier van normeren je als organisatie wilt hanteren. Het is belangrijk om deze afspraken samen te maken, zodat iedere toetsontwikkelaar zich kan conformeren aan de keuzes. Alle normeringsmethoden hebben voor- en nadelen. Het is belangrijk dat iedereen binnen de organisatie hetzelfde doet als het om cesuur bepalen gaat. Dan wordt steeds dezelfde meetlat gebruikt in de toetsen.

Stap 2: Voer de normering in.

Hoe je de normering invoert, is afhankelijk van de normeringsmethode die is gekozen.

1. Relatief normeren

Wanneer is gekozen voor het relatief normeren, voer je vooraf de normering op basis van uw eigen inschatting. Waar denk je dat de cesuurgrens zou moeten liggen? Je zet bij de toetsmatrijsinstellingen de zichtbaarheid op 'nee', zodat de kandidaten na afloop van de toets geen uitslag in beeld krijgen. Deze uitslag weet je namelijk nog niet.

2. Absoluut normeren

- Op eigen inzicht

Wanneer je absoluut normeert op eigen inzicht, bekijk je welke rol deze toets heeft in het gehele curriculum. Het is slim wanneer je deze keuzes bespreekt met collega's, zodat je eventuele blinde vlekken kunt voorkomen. Je voert de normering in en zet de raadkanscorrectie aan.

- Angoff methode

Wanneer is gekozen voor de Angoff-methode, verzamel je eerst een aantal deskundigen waarmee je gezamenlijk kijkt naar de vragen, volgens deze methode. Je voert op basis van de uitslag van deze sessie de normering in en zet de raadkanscorrectie uit.

Stap 3: Neem de toets af en bekijk de toetsanalyse.

De toetsanalyse zegt veel over de kwaliteit van de toets en de normering. Deze stap sla je nooit over. Op welke wijze je de toets ook hebt genormeerd, je moet altijd iets doen met de toetsanalysegegevens.

Wanneer je hebt gekozen voor relatief normeren kun je nu de normering bepalen. Dit doe je in de afname-omgeving door de toetsmatrijinstellingen aan te passen.

Stap 4: Pas eventueel de normering aan of pas de Cohen-Schotanus-methode toe.

Wanneer je de toetsgegevens hebt geanalyseerd kan dit betekenen dat je achteraf nog iets wijzigt in de normering. Dat doe je bijvoorbeeld als blijkt dat het antwoordmodel niet klopt. Of als de vraag onduidelijk is gesteld, waardoor niemand hem juist heeft beantwoord. Je kunt dan achteraf een vraag niet laten meetellen in de normering. Of je wijzigt het antwoordmodel en rekent meer antwoorden goed. Als een vraag niet goed blijkt te zijn (te moeilijk gesteld, verkeerd antwoordmodel, onduidelijk gesteld of hij hoort er niet tussen) dan pas je de vraag aan of je laat hem niet meetellen. Valt de uitslag je tegen (ten opzichte van voorgaande jaren) en is er geen maximale score behaald, dan kun je ervoor kiezen om de Cohen-Schotanus-methode gebruiken toe te passen.

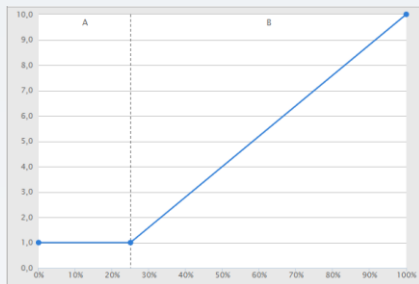
Stap 5: Stel de toetsresultaten vast.

Op het moment dat je tevreden bent over de normering, stel je de toetsresultaten vast en kun je deze vermelden.

Raadkanscorrectie

Bij het afnemen van gesloten vragen, is er altijd sprake van een raadkans. De enige uitzondering hierop is de invulvraag, bij alle andere gesloten vragen kun je door gokken/raden punten behalen. Bij meerkeuzevragen is dat voor veel mensen heel logisch. Er is bij 4 antwoordopties 25% kans dat de kandidaat de juiste kiest, simpelweg doordat de kans 1 op 4 is. Maar ook bij andere gesloten vragen (zoals de matrixvraag, de hotspotvraag, de meervoudig-juistvraag en de matchingvraag) is er sprake van een raadkans. Als toetsontwikkelaar moet je bij gesloten vragen altijd rekening houden met de raadkans. Kandidaten die de leerstof niet beheersen, kunnen per slot van rekening het juiste antwoord raden. Je wilt met de toets onderscheid maken tussen kandidaten die de kennis wel hebben en kandidaten die de kennis niet hebben.

In RemindoToets kun je geautomatiseerd en handmatig de raadkanscorrectie toepassen. Bij 100 meerkeuzevragen (4 antwoordopties) heeft de kandidaat 25% kans om het juiste antwoord te raden. De scorelijn komt er dan, als je hem handmatig instelt, als volgt uit te zien:



Let op: wanneer je de raadkanscorrectie op 'ja' zet, hoef je de grafiek niet meer te wijzigen en laat je een lineaire lijn staan. Je ziet de raadkanscorrectie dus niet, maar RemindoToets verrekenet deze dan op de achtergrond.

Wij wensen je veel succes bij het normeren van de toetsen.