



TIP 2: COMPUTERGESTUURD ADAPTIEF TOETSEN (CAT)

In deze ToetsTip behandelen we een onderwerp, waarmee u totnogtoe misschien nog niet veel te maken hebt gehad. Toch vinden we het zinvol het onderwerp ‘adaptief toetsen met behulp van de computer’ aan de orde te stellen. De ontwikkelingen met betrekking tot computergestuurd toetsen gaan snel en het lijkt ons goed u daarvan op de hoogte te stellen. Deze ToetsTip is wat theoretischer van aard dan eerdere ToetsTips. Ons doel is niet om tips te geven om zelf adaptieve toetsen te maken, maar om achtergrondinformatie te geven over deze wijze van toetsen. U leest wat adaptief toetsen eigenlijk is, aan welke eisen een computergestuurde adaptieve toets moet voldoen en wat de voordelen en de nadelen zijn. We besluiten met een aantal tips over demo’s, links, literatuur en presentaties van/over adaptieve taaltoetsen.

1. Omgaan met niveaoverschillen in een heterogene groep

Als studenten zich aanmelden bij een taal cursus Nederlands aan uw instelling, dan probeert u of één van uw collega’s te bepalen of de aankomend student een absolute beginner is of al (enigszins) gevorderd. Bij de student die gevorderd is, neemt u doorgaans een intake toets af. U wilt immers weten hoe goed deze student het Nederlands beheerst. Op basis van de resultaten van die intake toets beslist uw instelling in welke cursusgroep de student het beste geplaatst kan worden.

Omdat het aantal groepen dat u kunt vormen, in de praktijk beperkt is, zult u toch te maken krijgen met een heterogene groep; er zullen zowel zwakkere als betere studenten in uw groep zitten. Dat vereist veel flexibiliteit en motivatie, niet alleen van uw kant, maar ook van de zijde van de student. U probeert bij de aanbieding en verwerking van de leerstof zoveel mogelijk aan te sluiten bij niveaoverschillen binnen uw groep, wat niet altijd makkelijk of mogelijk is. Uw studenten zullen soms ondervinden dat de lesstof net een beetje te moeilijk of te makkelijk is, maar weten vaak dat dit niet anders kan en proberen toch gemotiveerd te blijven.

Ook als u toetsen afneemt of beoordeelt, speelt een rol dat u te maken hebt met een heterogene groep. Het is voor u niet altijd mogelijk elke student een toets aan te bieden die goed past bij zijn of haar niveau van dat moment. Met een adaptieve toets kunt u aan dit probleem tegemoetkomen.

2. Wat is een computergestuurde adaptieve toets (CAT)?

Grofweg gezien zijn er vier methoden om toetsen af te nemen:

1. pen-en-papier-toetsen
2. conventionele beeldschermtoetsen
3. computergestuurde niet-adaptieve toetsen
4. computergestuurde adaptieve toetsen.

Bij conventionele beeldschermtoetsen (computerondersteund toetsen) print men de toetsen uit en de kandidaat krijgt vervolgens de pen-en-papier-toets voorgelegd. Soms kan men voor de samenstelling van zo’n toets gebruikmaken van een niet-geautomatiseerde itembank.

Bij computergestuurde toetsen wordt de toets op het beeldscherm gepresenteerd. De kandidaat beantwoordt de vragen op elektronische wijze en de computer berekent het resultaat. Vaak wordt hierbij gebruikgemaakt van geautomatiseerde itembanken waaruit de toets ‘getrokken’ wordt. Een bekend computerprogramma waarmee dergelijke toetsen gemaakt kunnen worden, is Question Mark Perception (QMP)

Een speciale vorm van computergestuurd toetsen is computergestuurd adaptief toetsen (CAT). Bij deze toetsvorm past de toets zich tijdens de toetsafname aan het niveau aan van de student die op dat moment de toets maakt. Om die aanpassing mogelijk te maken wordt bij adaptief toetsen altijd gebruikgemaakt van een computergestuurde itembank. In die itembank zit een flinke hoeveelheid items (vragen/opdrachten). Die items hebben een label, waarin opgenomen is tot welke categorie een item hoort, welke moeilijkheidsgraad binnen die categorie dat item heeft en met welke aanvullende code het item geïdentificeerd kan worden (bijv. nummer).

Het volgende vereenvoudigde voorbeeld van een mogelijke CAT-toets Begrijpend Lezen is slechts bedoeld om u het algemene principe van het adaptief toetsen duidelijk te maken. Het voorbeeld is niet gebaseerd op een bestaande CAT-toets. In de werkelijkheid zal een dergelijke toets er waarschijnlijk heel anders uitzien, bijvoorbeeld wat betreft:

- het aantal items waarmee de categorieën in de itembank gevuld worden,
- het aantal items dat nodig is om het beginniveau nader vast te stellen,
- het aantal items dat meetelt bij de berekening van het niveau van het volgende item dat de student krijgt aangeboden (zelfde, lagere of hogere niveau),
- het aantal items dat meetelt voor het eindresultaat.

Bij het ontwerpen van een CAT-programma zullen daarover beslissingen moeten worden genomen. Het gaat te ver om in deze ToetsTip te behandelen hoe dat in zijn werk gaat en met welke specifieke (technische) eisen men daarbij rekening moet houden.

Voorbeeld CAT Begrijpend Lezen

Een docente wil meten op welk niveau haar studenten functioneren wat betreft het lezen van:

1. reclamefolders (categorie RECL)
2. instructies (categorie INST)
3. nieuwsberichten (categorie NIEU).

In de itembank zitten van elke categorie 30 items per niveau.

voorbeeld eenvoudige itembank

niveaus itemcodes categorie	A1	A2	B1	B2
reclamefolders	RECL-A1-1 RECL-A1-2 RECL-A1-3 RECL-A1-4 RECL-A1-30	RECL-A2-1 RECL-A2-2 RECL-A2-3 RECL-A2-4 RECL-A2-30	RECL-B1-1 RECL-B1-2 RECL-B1-3 RECL-B1-4 RECL-B1-30	RECL-B2-1 RECL-B2-2 RECL-B2-3 RECL-B2-4 RECL-B2-30
instructies	INST-A1-1 INST-A1-2 INST-A1-3 INST-A1-4 INST-A1-30	INST-A2-1 INST-A2-2 INST-A2-3 INST-A2-4 INST-A2-30	INST-B1-1 INST-B1-2 INST-B1-3 INST-B1-4 INST-B1-30	INST-B2-1 INST-B2-2 INST-B2-3 INST-B2-4 INST-B2-30
nieuwsberichten	NIEU-A1-1 NIEU-A1-2 NIEU-A1-3 NIEU-A1-4 NIEU-A1-30	NIEU-A2-1 NIEU-A2-2 NIEU-A2-3 NIEU-A2-4 NIEU-A2-30	NIEU-B1-1 NIEU-B1-2 NIEU-B1-3 NIEU-B1-4 NIEU-B1-30	NIEU-B2-1 NIEU-B2-2 NIEU-B2-3 NIEU-B2-4 NIEU-B2-30

Hoe werkt nu de adaptieve toetsing in dit voorbeeld?

Er wordt door de docente een schatting gemaakt van het niveau van de student wat betreft het lezen van reclamefolders, instructies en nieuwsberichten. Als dat niveau B1 is, dan begint de student met het maken van een aantal items op niveau B1, bijvoorbeeld met zes items (uit elke categorie twee). Deze items worden door de computer willekeurig uit de betreffende categorieën van de itembank getrokken. Ze dienen alleen om het beginniveau nader te bepalen en tellen niet mee bij het eindoordeel. Na het beantwoorden van deze zes vragen, berekent de computer de uitslag en past vervolgens het oorspronkelijke beginniveau van de toets aan. Studenten die bijvoorbeeld nul, één of twee items goed hebben beantwoord, beginnen op niveau A2, studenten die drie, vier of vijf vragen goed hebben beantwoord, beginnen op niveau B1 en studenten die zes items goed hebben gemaakt, krijgen als eerste een vraag op niveau B2 aangeboden.

De computer berekent na elke beantwoorde vraag steeds opnieuw het niveau, bijvoorbeeld aan de hand van de antwoorden op de laatste twee of drie vragen. Als dat niveau hoger of lager is dan het niveau van de laatst gestelde vraag, dan wordt de volgende vraag op basis van dat 'nieuwe' niveau uit de itembank getrokken.

voorbeeld CAT-toets + uitslag drie studenten

X = goed - = fout

Toets Cleides	antw.
RECL-A2-1	X
RECL-A2-25	X
RECL-B1-30	-
RECL-A2-4	X
RECL-A2-30	X
INST-A2-13	X
INST-A2-22	X
INST-B1-9	X
INST-B1-14	X
INST-B2-3	-
NIEU-A2-7	X
NIEU-A2-2	X
NIEU-B1-3	-
NIEU-A2-16	X
NIEU-A2-19	X

Toets Janec	antw.
RECL-B1-14	X
RECL-B1-22	X
RECL-B2-3	X
RECL-B2-24	-
RECL-B1-10	X
INST-B1-1	X
INST-B1-12	X
INST-B2-13	-
INST-B1-4	X
INST-B1-25	X
NIEU-B1-16	-
NIEU-B1-23	-
NIEU-A2-12	X
NIEU-A2-14	X
NIEU-B1-30	-

Toets Katja	antw.
RECL-B1-1	X
RECL-B1-12	X
RECL-B2-3	-
RECL-B2-4	X
RECL-B1-3	X
INST-B1-15	X
INST-B1-11	X
INST-B2-23	X
INST-B2-6	-
INST-B1-5	X
NIEU-B1-10	X
NIEU-B1-7	X
NIEU-B2-30	X
NIEU-B2-4	-
NIEU-B1-3	X

Totaal uitslag: niveau hoog A2

Totaal uitslag: niveau laag B1

Totaal uitslag: niveau hoog B1

Mogelijke conclusies uit deze uitslag:

- Student Cleides functioneert eigenlijk op niveau A2 wat betreft Begrijpend Lezen; alleen wat betreft het begrijpen van instructies haalt ze niveau B1. Deze student zal extra aandacht moeten krijgen van de docent wat betreft het lezen van reclames en nieuwsberichten.
- De docent zal bij student Janec extra aandacht moeten geven aan het lezen van nieuwsberichten.

3. Eisen aan adaptief toetsen

3.1 Technische specificaties

Als u zelf een computergestuurde adaptieve toets wilt ontwikkelen en daarvoor software gaat aanschaffen, moet u er rekening mee houden dat uw computers aan bepaalde technische specificaties moeten voldoen. Die technische eisen hangen samen met de wijze waarop de items worden aangeboden (bijv. geschreven tekst, plaatjes, audio, video, animaties) en van het adaptatie-algoritme dat gebruikt wordt bij het adaptief toetsen. Het adaptatie-algoritme is het stukje software dat ervoor zorgt dat de items voor de toets één voor één uit de itembank getrokken worden. Dat gebeurt op basis van de toetsmatrijs en op basis van een algoritmisch berekeningsmodel, waarbij de resultaten tot dan toe in de berekening worden meegenomen. Ook moet er software beschikbaar zijn voor het presenteren van items op het scherm, het evalueren van de antwoorden van de student, het schatten van het vaardigheidsniveau van de student, het verzorgen van rapportages en het onderhouden van de itembank.

3.2 Toetsmatrijs: representativiteit

Een goede toets moet representatief zijn voor het onderwijs dat gegeven wordt. Bij adaptief toetsen is het van essentieel belang te werken met toetsmatrijzen. Een toetsmatrijs is een soort blauwdruk van een toets en bestaat minimaal uit:

- onderwijsdoel
- te toetsen deeldoelen
- aantal items per deeldoel (eventueel aangevuld met aantal te behalen punten per item),
- totaal aantal items.

De keuze voor het aantal items en/of het aantal punten dat de student per item kan scoren, weerspiegelt de relatieve relevantie van het deeldoel dat met het item getoetst wordt. In het voorbeeld hieronder van een toetsmatrijs voor een traditionele pen-en-papier-toets kan de leerling voor elk item 1 punt behalen. Het aantal items per deeldoel varieert echter. Blijkbaar vinden de ontwerpers van deze toets dat het kunnen opzoeken in het woordenboek relevanter is dan het kunnen kiezen van de geschikte informatiebron..

Voorbeeld toetsmatrijs traditionele toets Studievaardigheden groep 8 basisonderwijs

Onderwijsdoel: Naslagwerken kunnen gebruiken			
code deeldoel	omschrijving deeldoel	aantal items toets	te behalen punten
6.1.1	kunnen alfabetiseren	2	2x1
6.1.2	kunnen kiezen van geschikte informatiebron	2	2x1
6.1.3	Kunnen kiezen van adequate zoekwoorden	3	3x1
6.1.4	kunnen opzoeken in woordenboek	6	6x1
6.1.5	kunnen opzoeken in encyclopedie	3	3x1
6.1.6	kunnen opzoeken op internet	4	4x1
Totaal aantal items per deelttoets		20	20

Als u items wilt maken voor een itembank die gebruikt wordt voor adaptief toetsen, dan zult u eerst voor alle onderwerpen die u wilt toetsen een toetsmatrijs moeten samenstellen.

Vervolgens maakt u voor elke categorie van de toetsmatrijs items. U geeft elk item een label, waarin minimaal zijn opgenomen: code deeldoel, niveau/moeilijkheidsgraad en volgnummer item binnen deze categorie. Het vierde item voor categorie 6.1.4 op niveau eind groep 8 uit het voorbeeld hierboven krijgt bijvoorbeeld de volgende code: 6.1.4.E8.4.

Niet alleen de items, maar ook de toetsmatrijzen worden opgenomen in de itembank. Op basis van een toetsmatrijs en aan de hand van het adaptatie-algoritme worden de items voor de bijbehorende toets uit de bank getrokken.

4. Voordelen van adaptief toetsen

4.1 Meting van niveau en diagnostiek

Met adaptief toetsen kan preciezer worden gemeten wat de student beheerst en op welk niveau. De docent weet wat de eventuele lacunes zijn en kan zijn of haar programma beter laten aansluiten bij het niveau van de student.

4.2 Moeilijkheidsgraad en motivatie studenten

Studenten worden meer gemotiveerd, omdat ze niet gefrustreerd raken door toetsen die te makkelijk of te moeilijk zijn. Adaptief toetsen vermindert op die manier angst voor toetsen. Als in het voorbeeld hiervoor de student Cleides een toets met alleen maar vragen op niveau B1 had moeten maken, had ze waarschijnlijk weinig vragen correct beantwoord. Dit zou een negatieve invloed kunnen hebben op haar motivatie.

4.3 Flexibele instroom

De CAT-toets kan gebruikt worden voor de intake van nieuwe studenten die al enigszins gevorderd zijn. Deze aankomende studenten kunnen vervolgens instromen in een passende cursusgroep. Omdat elke student op ieder moment een adaptieve toets kan maken, is instroom op verschillende momenten tijdens het cursusjaar mogelijk.

4.4 Betrouwbaarheid en toetslengte

Om betrouwbaar te kunnen meten wat de student beheerst, moeten conventionele toetsen (computerondersteunde beeldschermtoetsen of pen-en-papier-toetsen) een flinke hoeveelheid items bevatten. Op die manier kan men scores vermijden die teveel beïnvloed worden door 'toevalligheden', zoals wel/geen kennis van of affiniteit met dat specifieke thema. Uit onderzoek is gebleken dat men bij CAT-toetsen gemiddeld de helft minder items nodig heeft om betrouwbaar te kunnen meten. Dat houdt verband met het feit dat bij CAT-toetsen de vragen veel preciezer zijn afgestemd op het beheersingsniveau van de student. Bij sommige CAT-toetsen is het programma zelfs zo ingesteld dat de toetslengte afhangt van de beantwoording door de student. Bij de ene student zullen minder items nodig zijn om een betrouwbare uitslag te krijgen dan bij de andere student. De toets is dan afgelopen als het programma vaststelt dat een betrouwbare uitslag is bereikt.

4.5 Beoordeling en tijdstip uitslag

En voordeel van computergestuurde toetsen is dat de beoordeling objectief gebeurt; de computer beoordeelt. Bijkomend voordeel is dat dit de docent tijd bespaart. Omdat de computer meteen beoordeelt, hoeft de student niet op de uitslag te wachten. Onmiddellijke feedback werkt stimulerend. Als de CAT-toets als intake-toets wordt gebruikt, kan de student snel instromen.

4.6 Geheimhouding

Het heeft voor een student geen zin om andere studenten te vertellen welke vragen hij of zij heeft moeten beantwoorden. Elke student maakt immers een andere toets, want voor elke student worden steeds andere vragen geselecteerd op basis van zijn antwoord op eerdere vragen. Geheimhouding speelt dus geen rol.

5. Nadelen van adaptief toetsen

5.1 Investering CAT-toetsen

Het ontwerpen van een CAT-programma vergt veel tijd en deskundigheid, en is zeer kostbaar. Allereerst moeten er computers beschikbaar zijn die aan bepaalde technische specificaties voldoen. Daarnaast moet de benodigde software aangeschaft worden en moet het programma ontworpen of gekocht worden waar de itembank op zal 'draaien', het adaptatie-algoritme. De itembank moet vervolgens gevuld worden met een groot aantal items. Daarvoor wordt vaak een beroep gedaan op docenten, die op dat moment veel meer tijd moeten investeren dan ze nodig zouden hebben voor het maken van een traditionele toets. Er is dus een flinke investering aan geld en menskracht nodig, die pas later terugverdiend kan worden.

5.2 Toetsdoelen en toetsvormen

Met CAT kan men niet alles toetsen. Bij CAT-toetsen kan alleen gebruikgemaakt worden van gesloten of half open toetsvormen. Zo kunnen receptieve taalvaardigheden als lezen en luisteren, en taalkennis als woordenschat, spelling en grammatica, wel getoetst worden, maar productieve taalvaardigheden als spreken, schrijven en mondelinge interactie, nauwelijks. Het invullen van een formulier zou bijvoorbeeld wel getoetst kunnen worden, maar alleen als de antwoorden vastliggen.

5.3 Vrijheid student

Bij traditionele toetsen kunnen studenten vaak zelf uitmaken in welke volgorde ze de vragen beantwoorden, bij CAT-toetsen bepaalt het programma welke vraag de student op dat moment voorgeschoteld krijgt.

6. Tips voor interessante demo's, literatuur en links

Tip 1 demo *Itemdito*

Met **Itemdito**, een adaptieve NT2-toets, kan leesvaardigheid adaptief worden getoetst. De kandidaat moet gesloten vragen over functionele inputteksten beantwoorden. De items zijn geschaald volgens ERK-niveau. Via de website kunt u een demo downloaden. Daarin zijn onder meer gegevens over een leerlingvolgsysteem te zien en twee leesitems bij een reclamefolder. Zie: <http://www.itemdito.nl>

Tip 2 demo *Taal voor kleuters*

Een demo van de adaptieve woordenschattoets 'Taal voor kleuters' die in een leerlingvolgsysteem wordt gebruikt, vindt u op:
http://www.citogroep.nl/po/lovs/lvs/producten/digitaal/eind_fr_adaptief.htm

Tip 3 voorbeelditems *Taaltoets Pabo 2007/2008*

Op http://www.cito.com/pend/ho_pabo/eind_fr.htm vindt u informatie over de adaptieve 'Taaltoets Pabo 2007/2008'. De onderdelen van deze toets zijn spelling, formuleren, interpunctie en basisgrammatica. Er staan enkele voorbeeld-items op de site.

Tip 4 uitgebreide informatie over CAT + links overige demo's

Uitgebreidere informatie over CAT-toetsen is te vinden in publicaties en presentaties van Gerard Straetmans van het Cito. In het artikel '*Computers en toetsing: voor elkaar gemaakt!*' op <http://toetswijzer.kennisnet.nl/html/cbt/cbt.htm> geeft hij heel veel achtergrondinformatie. Daarnaast vindt u hierin voorbeelden van adaptieve toetsen met bijbehorende links, ook voor andere vakken dan het Nederlands.

Zie ook presentatie: Straetmans, G., *Computergestuurde adaptieve intake-toets voor rekenen en wiskunde*, Landelijke Dag Studievaardigheden '97 Open Universiteit, op: www.open.ou.nl/lds97/adaptintaketoets.htm

OPDRACHTEN

Situatie: Bij uw instelling gaat men een itembank voor adaptief toetsen inrichten. U werkt daaraan mee.

Opdracht 1

- Maak met behulp van het model hieronder een toetsmatrijs voor het toetsen van een aantal leerdoelen *Luistervaardigheid* uit uw eigen lespraktijk die geschikt zijn om adaptief getoetst te worden.

Opdracht 2

- Voor welke ERK-niveaus zouden er dan items in deze itembank moeten zitten voor deze toetsing? Waarom?

U kunt uw ervaringen met deze opdracht naar ons doorsturen voor feedback. Wanneer u vragen hebt over dit onderwerp, kunt u deze natuurlijk ook altijd aan ons stellen. Onze contactgegevens staan onder aan deze ToetsTip.

toetsmatrijs

Onderwijsdoel:			
code deeldoel	omschrijving deeldoel	aantal items	aantal te behalen punten
Totaal aantal items			

TOETSTIP 3 SEPTEMBER 2008

In de volgende ToetsTip gaan we dieper in op het gebruik van authentiek materiaal in toetsen.

Vragen, opmerkingen en suggesties

Opmerkingen over ToetsTip 2:

Suggesties voor volgende ToetsTips:

Inhoudelijke opmerkingen over evaluatie in het algemeen:

Praktische bemerkingen over de ToetsTip:

Voor vragen & reacties één adres:

het CNaVT-secretariaat

Certificaat Nederlands als Vreemde Taal
Blijde-Inkomststraat 7
B-3000 Leuven
België
0032 (0) 16 53 55 16

cnavt@arts.kuleuven.be