

# Activiteit 5

---

## Twintig keer raden — Informatie theorie

### Samenvatting

Hoeveel informatie zit er in een boek van 1000 pagina's? Zit er meer informatie in een telefoonboek van 1000 bladzijden, of in een stapel van 1000 lege vellen papier of in de boeken van Harry Potter? Als we dit kunnen meten dan kunnen we een inschatting maken van hoeveel ruimte nodig is om de informatie op te slaan. Kan je bijvoorbeeld de volgende zin lezen?

*n dz zn mssn d klnkrs.*

Waarschijnlijk wel, want er zit niet veel 'informatie' in de klinkers. Deze activiteit laat een manier zien om de hoeveelheid informatie te meten.

### Kerdoelen

- Rekenen: groep 5 en hoger. Getallen ontdekken, groter dan, kleiner dan en verspreiding.
- Algebra groep 5 en hoger. Patronen en reeksen.
- Taal

### Vaardigheden

- Getallen vergelijken en het werken met reeksen.
- Deductie
- Vragen stellen

### Leeftijd

- 10 jaar en ouder.

### Materialen

- Voor de eerste activiteit is geen materiaal nodig.
- Er is een uitbreidingsactiviteit waar de leerling het volgende voor nodig heeft: *Werkblad: Beslisbomen*

# Twintig keer raden

---

## Discussie

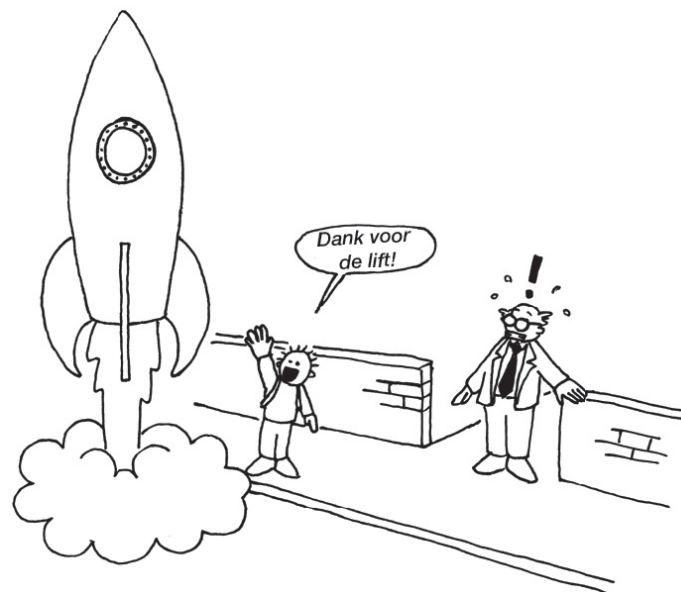
1. Bespreek met de leerlingen wat zij denken dat informatie is.
2. Hoe kunnen we meten hoeveel informatie er in een boek staat? Is het aantal pagina's of het aantal woorden belangrijk? Kan het ene boek meer informatie bevatten dan het andere? Wat als het een heel saai of stom boek is of juist een heel interessant of heel leuk boek? Zou een boek met bijvoorbeeld 400 pagina's vol met "bla, bla, bla, bla" meer of minder informatie bevatten dan bijvoorbeeld een telefoonboek?

Leg uit dat computerdeskundigen de informatie meten door te kijken hoe verrassend een bericht (of boek!) is. Als iemand je iets vertelt dat je al weet - bijvoorbeeld wanneer een vriend die altijd naar school loopt tegen je zegt dat hij "naar school is gelopen" - dan geeft hij je weinig informatie, omdat dat niet verrassend nieuws is. Als hij je zou vertellen dat hij vandaag met de heli­copter naar school is gekomen dan is dat heel verrassend en bevat dat bericht dus veel informatie.

Hoe kan de verrassingswaarde van een bericht worden gemeten?

Een manier om de verrassingswaarde te meten is, is na te gaan hoe moeilijk het is om de informatie te raden. Als je vriend je vraagt: "Raad eens hoe ik naar school ben gekomen vandaag," en hij heeft gelopen, dan raad je dat waarschijnlijk in een keer. Het zal waarschijnlijk een paar keer raden kosten voor je bij een heli­copter zou zijn, laat staan een ruimteschip.

De hoeveelheid informatie die een bericht bevat wordt dus gemeten door te kijken hoe makkelijk het is om het te raden. Het volgende spelletje geeft ons een idee over hoe het werkt.



## Twintig vragen activiteit

---

Dit is een aanpassing van het 20-vragenspel (zie [www.matsoft.nl/20vragen](http://www.matsoft.nl/20vragen)). Leerlingen mogen vragen stellen aan een ander en deze mag alleen met ja en nee antwoorden tot het antwoord is geraden. Iedere vraag mag gesteld worden, maar het antwoord kan dus alleen ja of nee zijn.

Bijvoorbeeld, de vraag: “Ik denk aan”:

- een getal tussen 1 en 100
- een getal tussen 1 en 1000
- een getal tussen 1 en 1.000.000.
- elk heel getal
- een reeks van 6 getallen in een patroon (aanpassen aan het niveau van de groep). Raad in de volgorde van de eerste tot de laatste (bijvoorbeeld 2,4,6,8,10).

Tel het aantal vragen dat gesteld wordt om het antwoord te achterhalen. Hiermee kan je de hoeveelheid “informatie” meten.

### Nabespreking

Welke strategie heb je gebruikt? Welke zijn het beste?

Leg uit dat het maar zeven keer raden kost om een getal tussen één en honderd te raden als je de verzameling mogelijke uitkomsten iedere keer halveert. Bijvoorbeeld:

- |                          |      |
|--------------------------|------|
| 1. Is het minder dan 50? | Ja.  |
| 2. Is het minder dan 25? | Nee. |
| 3. Is het minder dan 37? | Nee. |
| 4. Is het minder dan 43? | Ja.  |
| 5. Is het minder dan 40? | Nee. |
| 6. Is het minder dan 41? | Nee. |
| 7. Dan moet het 42 zijn! | JA!  |

Bijzonder is dat als de verzameling mogelijke uitkomsten tot duizend, in plaats van honderd, is dit niet betekent dat je tien keer zo vaak moet raden, maar dat er maar drie vragen meer nodig zijn. Iedere keer als het domein van de getallen verdubbelt heb je maar één vraag meer nodig om het antwoord te vinden.

### **Uitbreiding: Hoe veel informatie zit er in een bericht?**

Computerdeskundigen raden niet (alleen) met getallen - ze kunnen ook raden welke letter waarschijnlijk de volgende zal zijn in een woord of een zin.

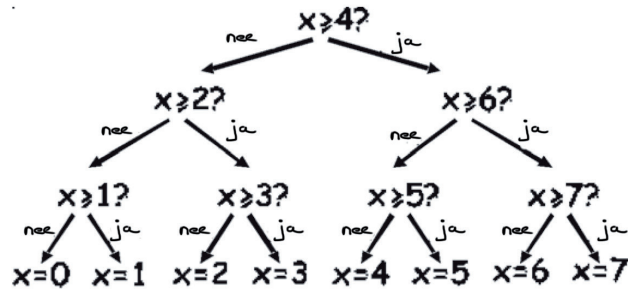
Probeer het raadsel te spelen met een korte zin van 4 tot 6 woorden. De letters moeten geraden worden in de juiste volgorde, van de eerste naar de laatste. Laat iemand alle letters noteren die geraden zijn en houd bij hoeveel keer raden het kost om iedere letter te vinden. Alle vragen die beantwoord kunnen worden met “ja” of “nee” kunnen gebruikt worden. Dus de vragen zouden kunnen zijn: “Is het een T?”, “Is het een klinker?”, “Komt het voor de M in het alfabet?” Een spatie telt ook als letter en moet dus ook geraden worden. Raad om de beurt en kijk welke delen van het bericht het makkelijkst te raden zijn.

## Werkblad: Beslisboom

---

Wanneer de strategie om zo snel mogelijk bij het antwoord te komen bekend is, is het zelfs niet nodig om de vragen te stellen. De persoon die het getal in zijn hoofd heeft, weet welke vragen er komen en hoeft alleen het antwoord te geven zonder de vraag te horen. Een beslisboom maakt dit duidelijk.

Hier is een 'beslisboom' om een getal te raden tussen 0 en 7:



Welk pad van ja/nee beslissingen volg je om het getal 5 te 'raden'?

Hoe veel ja/nee beslissingen heb je nodig om tot elk willekeurig getal te komen?

Nu gebeurt er iets merkwaardigs en fascinerends. Onder de getallen (0,1,2,3...) in de laatste rij van de boom schrijf je het getal nog een keer maar nu in binaire code (zie Activiteit 1, ook bij binair staat het kleinste getal rechts).

Kijk nu nog eens goed naar de boom. Als nee = 0 en ja = 1, wat zie je dan?

In het getalraadspel hiervoor probeerde we de vragen zo te kiezen dat de volgorde van antwoorden precies hetzelfde was als in deze beslisboom.

Maak nu je eigen beslisboom om een getal tussen 0 en 15 te raden.

### Extra voor experts:

Wat voor soort boom zou je gebruiken om iemands leeftijd te raden? En wat voor boom om de volgende letter te raden in een zin?

## Waar gaat dit eigenlijk over?

---

Een gevierde Amerikaanse wiskundige (daarnaast jongleur en eenwieler) genaamd Claude Shannon deed veel experimenten met dit spel. Hij mat de hoeveelheid informatie in bits, iedere ja/nee vraag is op te vatten als 1/0 bit. Hij vond dat de hoeveelheid informatie die in een bericht staat, afhangt van wat je al weet. Soms kunnen we zulke vragen stellen dat een heleboel andere vragen overbodig worden. In dit geval is de hoeveelheid informatie in een bericht laag. Bijvoorbeeld de hoeveelheid informatie bij het tossen van een munt is normaal gesproken maar één bit: kop of munt. Maar als de munt niet zuiver is en het negen van de tien keer kop wordt, is de hoeveelheid informatie niet één bit maar, geloof het of niet, minder dan één bit!



Hoe kom je erachter wat de uitkomst is van een toss met minder dan een simpele ja/nee vraag? Heel simpel - stel vragen als “zijn de volgende twee tossen beide kop?” Voor een reeks van tossen met de oneerlijke munt zal het antwoord in 80% van de gevallen “ja” zijn. In de 20% van de keren waarin het antwoord “nee” is hoef je maar twee extra vragen te vragen. Maar gemiddeld vraag je minder dan één vraag per toss!

Shannon noemde de informatiedichtheid van een bericht “entropie”. Entropie hangt niet alleen af van het aantal mogelijke uitkomsten - in het geval van een munt opgooien, 2 - maar ook om de kans dat het gebeurt. Bij ongewone gebeurtenissen, of verrassende informatie hebben we meer vragen nodig om achter het bericht te komen omdat ze ons informatie vertellen die we niet al wisten - denk aan het naar school gaan met een helicopter.

De entropie van een bericht is erg belangrijk voor computerdeskundigen. De ruimte om een bericht op te slaan kun je niet kleiner maken dan zijn entropie en de beste compressieprogramma's zijn een equivalent van het raadspel. Omdat een computerprogramma de ‘vragen’ stelt, kan de lijst met gestelde vragen altijd gereproduceerd worden, dus als de antwoorden (bits) maar opgeslagen worden, kunnen we de informatie altijd weer herconstrueren! De beste compressieprogramma's kunnen tekstbestanden tot een kwart van hun originele grootte terugbrengen en dat scheelt een hoop opslagruimte!

Deze raadmethode kan ook gebruikt worden om een computerprogramma te maken dat raadt wat de gebruiker wil gaan typen. Bijvoorbeeld de bekende zoeksuggesties van zoekprogramma's of autocorrect op telefoons.