

Hoe interpreteren we statistische informatie over onzekerheid?

Martin Knotters¹, Cees van Woerkum²

¹ESG, Centrum Bodem

²SSG, Communicatiestrategie

Wageningen University and Research Centre



The monster of uncertainty



naar: Van der Sluijs, J., 2005. Uncertainty as a monster in the science-policy interface: four coping strategies. *Water Science & Technology* **52**(6): 87-92.

Hoe tem je het monster van onzekerheid?

1. Monster exorcism: onzekerheid uitbannen, of tenminste flink reduceren
2. Monster adaptation: onzekerheid kwantificeren, benutten bij beslissingen
3. Monster embracement: wetenschap/spiritualiteit
4. Monster assimilation: transparantie over onzekerheden bij controversiële onderwerpen (*unknown unknowns*, voorzorgsbeginsel, *post normal science*)
5. Monster denial ...?

Hoe tem je het monster van onzekerheid?

1. **Monster exorcism:** onzekerheid uitbannen, of tenminste flink reduceren
2. **Monster adaptation:** onzekerheid kwantificeren, benutten bij beslissingen
3. **Monster embracement:** wetenschap/spiritualiteit
4. **Monster assimilation:** transparantie over onzekerheden bij controversiële onderwerpen (*unknown unknowns*, voorzorgsbeginsel, *post normal science*)
5. **Monster denial ...?**

Hoe tem je het monster van onzekerheid?

1. **Monster exorcism**: onzekerheid uitbannen, of tenminste flink reduceren
2. **Monster adaptation**: onzekerheid kwantificeren, benutten bij beslissingen
3. **Monster embracement**: wetenschap/spiritualiteit
4. **Monster assimilation**: transparantie over onzekerheden bij controversiële onderwerpen (*unknown unknowns*, voorzorgsbeginsel, *post normal science*)
5. **Monster denial** ...?

Hoe tem je het monster van onzekerheid?

1. **Monster exorcism**: onzekerheid uitbannen, of tenminste flink reduceren
2. **Monster adaptation**: onzekerheid kwantificeren, benutten bij beslissingen
3. **Monster embracement**: wetenschap/spiritualiteit
4. **Monster assimilation**: transparantie over onzekerheden bij controversiële onderwerpen (*unknown unknowns*, voorzorgsbeginsel, *post normal science*)
5. **Monster denial** ...?

Hoe tem je het monster van onzekerheid?

1. **Monster exorcism**: onzekerheid uitbannen, of tenminste flink reduceren
2. **Monster adaptation**: onzekerheid kwantificeren, benutten bij beslissingen
3. **Monster embracement**: wetenschap/spiritualiteit
4. **Monster assimilation**: transparantie over onzekerheden bij controversiële onderwerpen (*unknown unknowns*, voorzorgsbeginsel, *post normal science*)
5. **Monster denial ...?**

Hoe tem je het monster van onzekerheid?

1. **Monster exorcism**: onzekerheid uitbannen, of tenminste flink reduceren
2. **Monster adaptation**: onzekerheid kwantificeren, benutten bij beslissingen
3. **Monster embracement**: wetenschap/spiritualiteit
4. **Monster assimilation**: transparantie over onzekerheden bij controversiële onderwerpen (*unknown unknowns*, voorzorgsbeginsel, *post normal science*)
5. **Monster denial ...?**

Hoe tem je het monster van onzekerheid?

- 1.
2. **Monster adaptation:** onzekerheid kwantificeren, benutten bij beslissingen
- 3.
- 4.
- 5.

Doel:

“de discussie te voeren over de vraag hoe je het beste kunt informeren over statistisch gekwantificeerde onzekerheid.”



Waarom statistische informatie over onzekerheid?

1. Als de houding van mensen tegenover risico's belangrijk is (gezondheid, inkomen; *risk aversion*, *ostrich effect*).
2. Als informatie uit verschillende bronnen, met verschillende nauwkeurigheid, moet worden gecombineerd.
3. Als er moet een beslissing moet worden gemaakt over het (verder) reduceren van onzekerheid (*value of information*, *data worth analysis*).
4. Als een beslissing op basis van 'de beste schatting' niet optimaal is (bijv. onderzoek naar effecten van *global warming* op gewasproductie).

(naar: Morgan, M.G. en M. Henrion, 1990. *Uncertainty. A guide to dealing with uncertainty in quantitative risk and policy analysis*. New York, Cambridge University Press.)

Waarom statistische informatie over onzekerheid?

1. Als de houding van mensen tegenover risico's belangrijk is (gezondheid, inkomen; *risk aversion*, *ostrich effect*).
2. Als informatie uit verschillende bronnen, met verschillende nauwkeurigheid, moet worden gecombineerd.
3. Als er moet een beslissing moet worden gemaakt over het (verder) reduceren van onzekerheid (*value of information*, *data worth analysis*).
4. Als een beslissing op basis van 'de beste schatting' niet optimaal is (bijv. onderzoek naar effecten van *global warming* op gewasproductie).

(naar: Morgan, M.G. en M. Henrion, 1990. *Uncertainty. A guide to dealing with uncertainty in quantitative risk and policy analysis*. New York, Cambridge University Press.)

Waarom statistische informatie over onzekerheid?

1. Als de houding van mensen tegenover risico's belangrijk is (gezondheid, inkomen; *risk aversion*, *ostrich effect*).
2. Als informatie uit verschillende bronnen, met verschillende nauwkeurigheid, moet worden gecombineerd.
3. Als er moet een beslissing moet worden gemaakt over het (verder) reduceren van onzekerheid (*value of information*, *data worth analysis*).
4. Als een beslissing op basis van 'de beste schatting' niet optimaal is (bijv. onderzoek naar effecten van *global warming* op gewasproductie).

(naar: Morgan, M.G. en M. Henrion, 1990. *Uncertainty. A guide to dealing with uncertainty in quantitative risk and policy analysis*. New York, Cambridge University Press.)

Waarom statistische informatie over onzekerheid?

1. Als de houding van mensen tegenover risico's belangrijk is (gezondheid, inkomen; *risk aversion*, *ostrich effect*).
2. Als informatie uit verschillende bronnen, met verschillende nauwkeurigheid, moet worden gecombineerd.
3. Als er moet een beslissing moet worden gemaakt over het (verder) reduceren van onzekerheid (*value of information*, *data worth analysis*).
4. Als een beslissing op basis van 'de beste schatting' niet optimaal is (bijv. onderzoek naar effecten van *global warming* op gewasproductie).

(naar: Morgan, M.G. en M. Henrion, 1990. *Uncertainty. A guide to dealing with uncertainty in quantitative risk and policy analysis*. New York, Cambridge University Press.)

Waarom statistische informatie over onzekerheid?

1. Als de houding van mensen tegenover risico's belangrijk is (gezondheid, inkomen; *risk aversion*, *ostrich effect*).
2. Als informatie uit verschillende bronnen, met verschillende nauwkeurigheid, moet worden gecombineerd.
3. Als er moet een beslissing moet worden gemaakt over het (verder) reduceren van onzekerheid (*value of information*, *data worth analysis*).
4. Als een beslissing op basis van 'de beste schatting' niet optimaal is (bijv. onderzoek naar effecten van *global warming* op gewasproductie).

(naar: Morgan, M.G. en M. Henrion, 1990. *Uncertainty. A guide to dealing with uncertainty in quantitative risk and policy analysis*. New York, Cambridge University Press.)

Waarom statistische informatie over onzekerheid?

5. Bij controversiële beslisprocessen.
6. Bij toetsing aan (wettelijke) normen, om het risico van foute conclusies te beheersen (bijv. KRW).
7. Bij het beoordelen van het effect van ingrepen, maatregelen (*targets, accountability*).

Waarom statistische informatie over onzekerheid?

5. Bij controversiële beslisprocessen.
6. Bij toetsing aan (wettelijke) normen, om het risico van foute conclusies te beheersen (bijv. KRW).
7. Bij het beoordelen van het effect van ingrepen, maatregelen (*targets, accountability*).

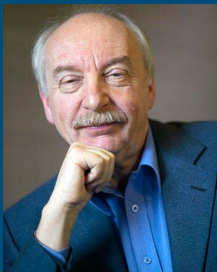
Waarom statistische informatie over onzekerheid?

5. Bij controversiële beslisprocessen.
6. Bij toetsing aan (wettelijke) normen, om het risico van foute conclusies te beheersen (bijv. KRW).
7. Bij het beoordelen van het effect van ingrepen, maatregelen (*targets, accountability*).

Waarom statistische informatie over onzekerheid?

5. Bij controversiële beslisprocessen.
6. Bij toetsing aan (wettelijke) normen, om het risico van foute conclusies te beheersen (bijv. KRW).
7. Bij het beoordelen van het effect van ingrepen, maatregelen (*targets, accountability*).

Hoe presenteert je statistische informatie over onzekerheid?



Gerd Gigerenzer

“Dat dokters en patiënten statistische informatie niet begrijpen ligt niet aan hen, maar aan de slechte presentatie.”

“Slecht gepresenteerde statistische informatie kan verkeerde communicatie van risico's veroorzaken, met ernstige gevolgen.”

Morgen 30% kans op regen

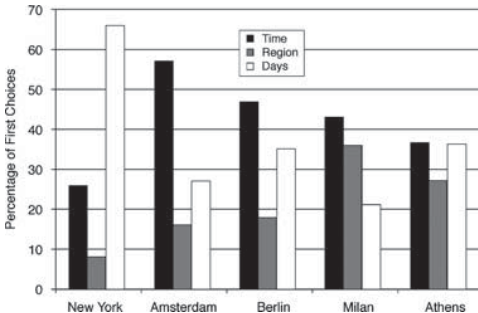


Fig. 1. First choice. People in New York ($n = 103$), Amsterdam ($n = 117$), Berlin ($n = 219$), Milan ($n = 203$), and Athens ($n = 108$) were asked what the statement “There is a 30% chance of rain tomorrow” refers to. The three alternatives were “It will rain tomorrow for 30% of the time,” “in 30% of the region,” and “on 30% of the days like tomorrow.”

Gigerenzer, G., R. Hertwig, E. van den Broek, B. Fasolo en K.V. Katsikopoulos, 2005. “A 30% chance of rain tomorrow”: how does the public understand probabilistic weather forecasts? *Risk Analysis* 25(3): 623-629.

Interpretatie van statistische informatie over onzekerheid

Analytisch	'Experiëntieel'
Langzaam	Snel
<i>Logical: reason oriented</i>	<i>Affective: pleasure-pain oriented</i>
Logische verbanden	Associaties
Bewuste beoordeling van gebeurtenissen	Gevoelsmatige beoordeling o.b.v. ervaringen
Abstracte symbolen, woorden en getallen	Beelden, metaforen, verhalen
Logica en bewijs	'Ervaren is geloven'

Slovic, P., M. L. Finucane, E. Peters and D. G. MacGregor, 2004. Risk as analysis and risk as feelings: some thoughts about affect, reason, risk, and rationality. *Risk Analysis* 24: 311-322.

Street Calculus



Interpretatie van statistische informatie

Enkele voorbeelden:

1. Statistische significantietoetsen
2. 'Bandbreedtes', betrouwbaarheidsintervallen

Interpretatie van statistische informatie

Enkele voorbeelden:

1. Statistische significantietoetsen
2. 'Bandbreedtes', betrouwbaarheidsintervallen

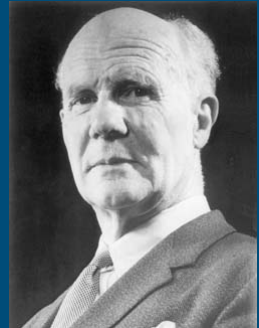
Statistische significantietoetsen (NHST, SST)



Sir Ronald A. Fisher



Jerzy Neyman



Egon S. Pearson

De controverse over significantietoetsen

Fisher	Neyman&Pearson
Conclusies trekken	Beslissingen nemen
P -waarden	Kansen op fouten: α , β
–	Onderscheidend vermogen (<i>power</i>)
Informereren	Beslissen



Enkele misinterpretaties bij significantietoetsen

1. $P > 0.05$: “Er is geen effect, verschil, verband etc.”
2. “ P is de kans dat H_0 waar is.”

Echter: P is de kans op de data (of extremere), gegeven de nulhypothese: $P(D|H)$.

$$P(D|H) \neq P(H|D)$$

Enkele misinterpretaties bij significantietoetsen

1. $P > 0.05$: “Er is geen effect, verschil, verband etc.”
2. “ P is de kans dat H_0 waar is.”

Echter: P is de kans op de data (of extremere), gegeven de nulhypothese: $P(D|H)$.

$$P(D|H) \neq P(H|D)$$

Enkele misinterpretaties bij significantietoetsen

1. $P > 0.05$: “Er is geen effect, verschil, verband etc.”
2. “ P is de kans dat H_0 waar is.”

Echter: P is de kans op de data (of extremere), gegeven de nulhypothese: $P(D|H)$.

$$P(D|H) \neq P(H|D)$$

Enkele misinterpretaties bij significantietoetsen

1. $P > 0.05$: “Er is geen effect, verschil, verband etc.”
2. “ P is de kans dat H_0 waar is.”

Echter: P is de kans op de data (of extremere), gegeven de nulhypothese: $P(D|H)$.

$$P(D|H) \neq P(H|D)$$

Enkele misinterpretaties bij significantietoetsen

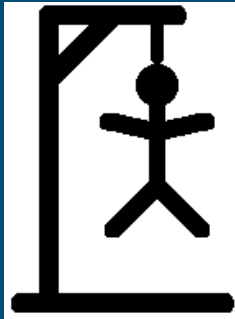
1. $P > 0.05$: “Er is geen effect, verschil, verband etc.”
2. “ P is de kans dat H_0 waar is.”

Echter: P is de kans op de data (of extremere), gegeven de nulhypothese: $P(D|H)$.

$$P(D|H) \neq P(H|D)$$

De 'omdraaifout'

$$P(D|H) \neq P(H|D)$$



De kans dat iemand dood is, gegeven dat hij is opgehangen,
is niet gelijk aan
de kans dat iemand is opgehangen, gegeven dat hij dood is.

3. “statistische significantie = praktische relevantie”

Een klein effect kan *statistisch significant* zijn, bijvoorbeeld omdat je veel waarnemingen hebt verzameld, maar daarmee is het effect nog niet *praktisch relevant*.

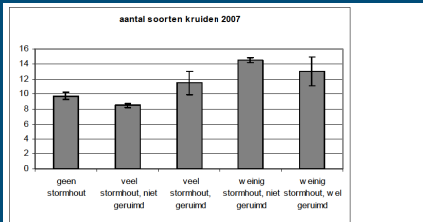
(Mis)interpretatie resultaten significantietoetsen

Is een medicijn effectief?

Testresultaat	omvang effect	conclusie	% statistici
significant	groot	effectief	96
		weet niet	4
niet-significant	klein	niet effectief	84
		weet niet	16
significant	klein	effectief	12
		niet effectief	80
		weet niet	8
niet-significant	groot	effectief	12
		niet effectief	36
		weet niet	56

Lecoutre, M.-P., J. Poitevineau en B. Lecoutre, 2003. Even statisticians are not immune to misinterpretation of Null Hypothesis Significance Tests. *International Journal of Psychology* **38**: 37-45.

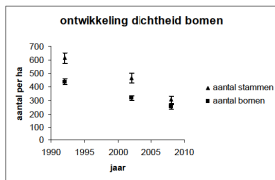
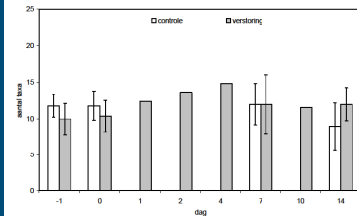
Grafische presentatie van onzekerheid



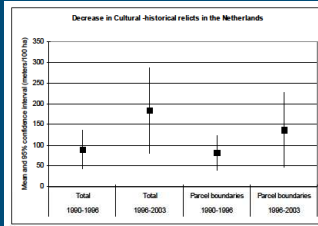
Figuur 5. Het gemiddeld aantal soorten kruiden ($\pm$ standaardfout) van de opstanden in het Zandenbos in 2007.

Figuur 7.

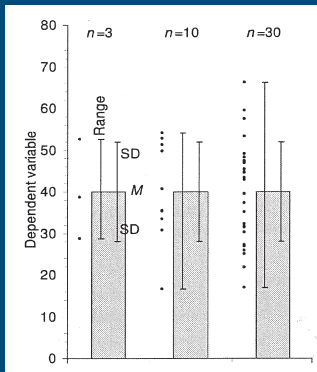
Aantal taxa in het controle- en verstoringstraject (met standaarddeviatie).



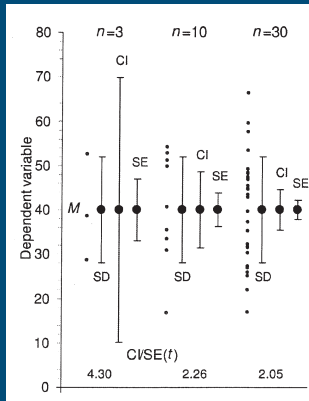
Figuur 2.1. Ontwikkeling van de dichtheid aan bomen en stammen in de zeventien steekproefcirkels in de Keizersdijk die driemaal werden geïnventariseerd.



Bandbreedtes, foutenmarges, betrouwbaarheidsintervallen

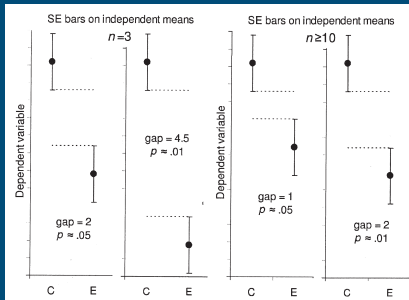


descriptive

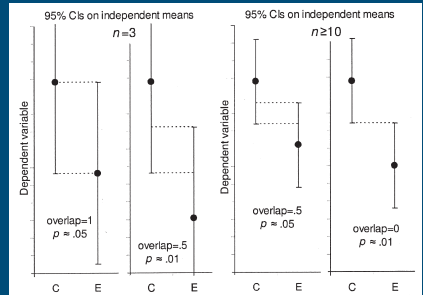


inferential

Bandbreedtes, foutenmarges, betrouwbaarheidsintervallen



standaardfouten



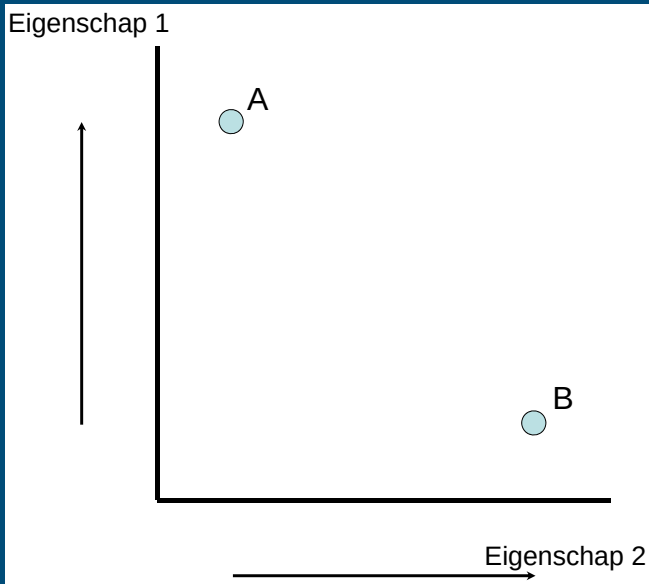
95%-betrouwbaarheidsintervallen

Cumming, G. en S. Finch, 2005. Inference by eye. Confidence intervals and how to read pictures of data. *American Psychologist* **60**(2): 170-180.

Rationaliteit van beslissingen

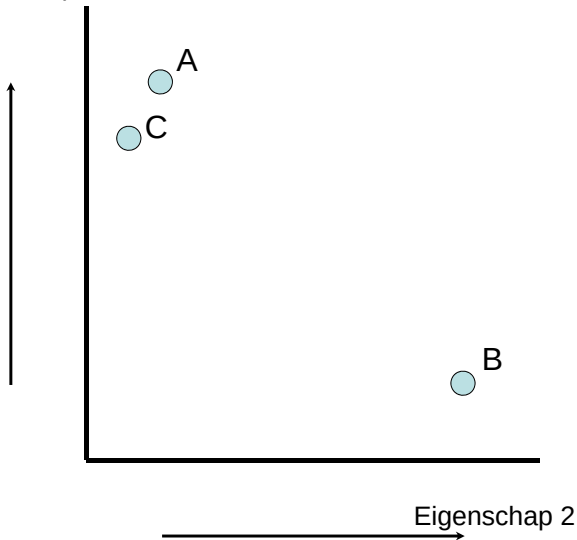
- ▶ Daniel Kahneman, Amos Tversky
- ▶ Sociale psychologie, behavioral economics, besliskunde
- ▶ Beslissingen zijn vaak niet rationeel
- ▶ Onzekerheid: extra 'complicatie'

Relativiteit en irrationaliteit



Relativiteit en irrationaliteit, *decoy-effect*

Eigenschap 1



Relativiteit en irrationaliteit

Tabel S.1 De mate waarin wordt voldaan aan de door de ministeries van VROM, LNV en V&W gestelde eisen.

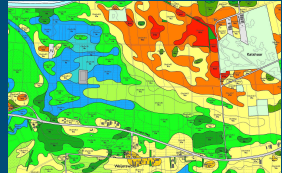
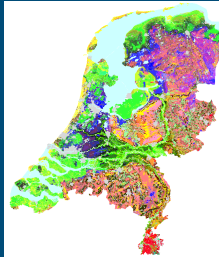
Voldoen aan:	<i>Scenario 1</i>	<i>Scenario 2</i>	<i>Scenario 3</i>
Wettelijke EU verplichtingen (obliged to know)	Beperkt	Uitgebreid	Beperkt
Eisen vanuit nationaal beleid (need to know)	Uitgebreid	Beperkt	Beperkt
Mogelijke toekomstige eisen (might need to know in the future)	Actief	Passief	Passief

Scenario	1	2	3
Kostenbesparing (x1000 euro)	270	750	1600

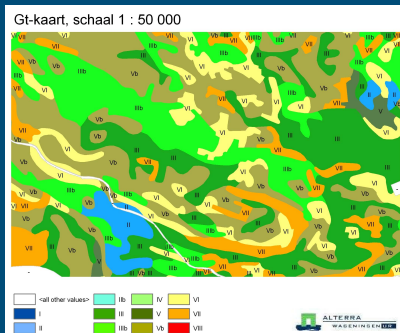
Relativiteit en irrationaliteit

optie	auto		
	1	2	3
airconditioning	X		X
schuifdak		X	X
sportvelgen		X	X
bijbetalen (euro)	270	750	1600

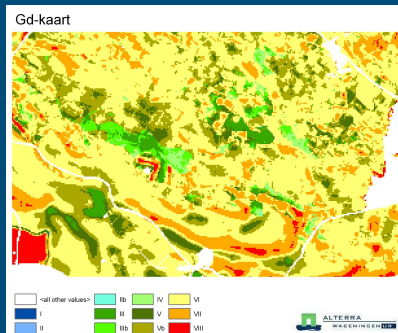
Eerste indruk (imprint)



Eerste indruk, verwachting en irrationaliteit

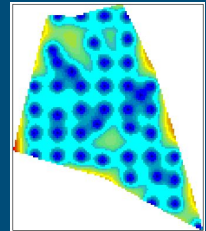
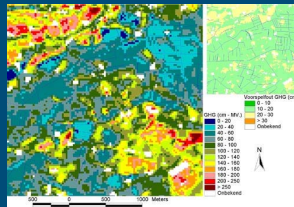
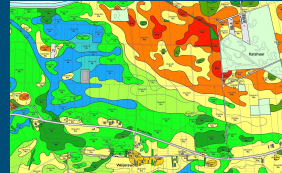
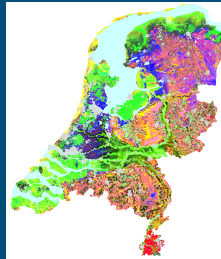


zuiverheid = 20 %



zuiverheid = 62 %

Gestratificeerde kanssteekproef, $n = 92$



Enkele overwegingen

1. Ligt de nadruk op informeren of op beslissen?
2. Benutting van statistische informatie over onzekerheid hangt af van de presentatie.
3. Meer aandacht voor experiëntiële verwerking van informatie kan presentatie en benutting van statistische informatie over onzekerheid bevorderen.
4. Meer aandacht voor het beslisproces kan presentatie en benutting van statistische informatie over onzekerheid bevorderen.

Enkele overwegingen

1. Ligt de nadruk op informeren of op beslissen?
2. Benutting van statistische informatie over onzekerheid hangt af van de presentatie.
3. Meer aandacht voor experiëntiële verwerking van informatie kan presentatie en benutting van statistische informatie over onzekerheid bevorderen.
4. Meer aandacht voor het beslisproces kan presentatie en benutting van statistische informatie over onzekerheid bevorderen.

Enkele overwegingen

1. Ligt de nadruk op informeren of op beslissen?
2. Benutting van statistische informatie over onzekerheid hangt af van de presentatie.
3. Meer aandacht voor experiëntiële verwerking van informatie kan presentatie en benutting van statistische informatie over onzekerheid bevorderen.
4. Meer aandacht voor het beslisproces kan presentatie en benutting van statistische informatie over onzekerheid bevorderen.

Enkele overwegingen

1. Ligt de nadruk op informeren of op beslissen?
2. Benutting van statistische informatie over onzekerheid hangt af van de presentatie.
3. Meer aandacht voor experiëntiële verwerking van informatie kan presentatie en benutting van statistische informatie over onzekerheid bevorderen.
4. Meer aandacht voor het beslisproces kan presentatie en benutting van statistische informatie over onzekerheid bevorderen.

Enkele overwegingen

1. Ligt de nadruk op informeren of op beslissen?
2. Benutting van statistische informatie over onzekerheid hangt af van de presentatie.
3. Meer aandacht voor experiëntiële verwerking van informatie kan presentatie en benutting van statistische informatie over onzekerheid bevorderen.
4. Meer aandacht voor het beslisproces kan presentatie en benutting van statistische informatie over onzekerheid bevorderen.