



Aalto-yliopisto
Perustieteiden
korkeakoulu

Asiantuntijatiedon määrittäminen todennäköisyyspohjaiseen ristivaikutusarviointiin (valmiin työn esittely)

Andrea Lyly

27.08.2020

Ohjaaja: *DI Juho Roponen*

Valvoja: *Prof. Ahti Salo*

Työn saa tallentaa ja julkistaa Aalto-yliopiston avoimilla verkkosivuilla. Muilta osin kaikki oikeudet pidätetään.

Tausta

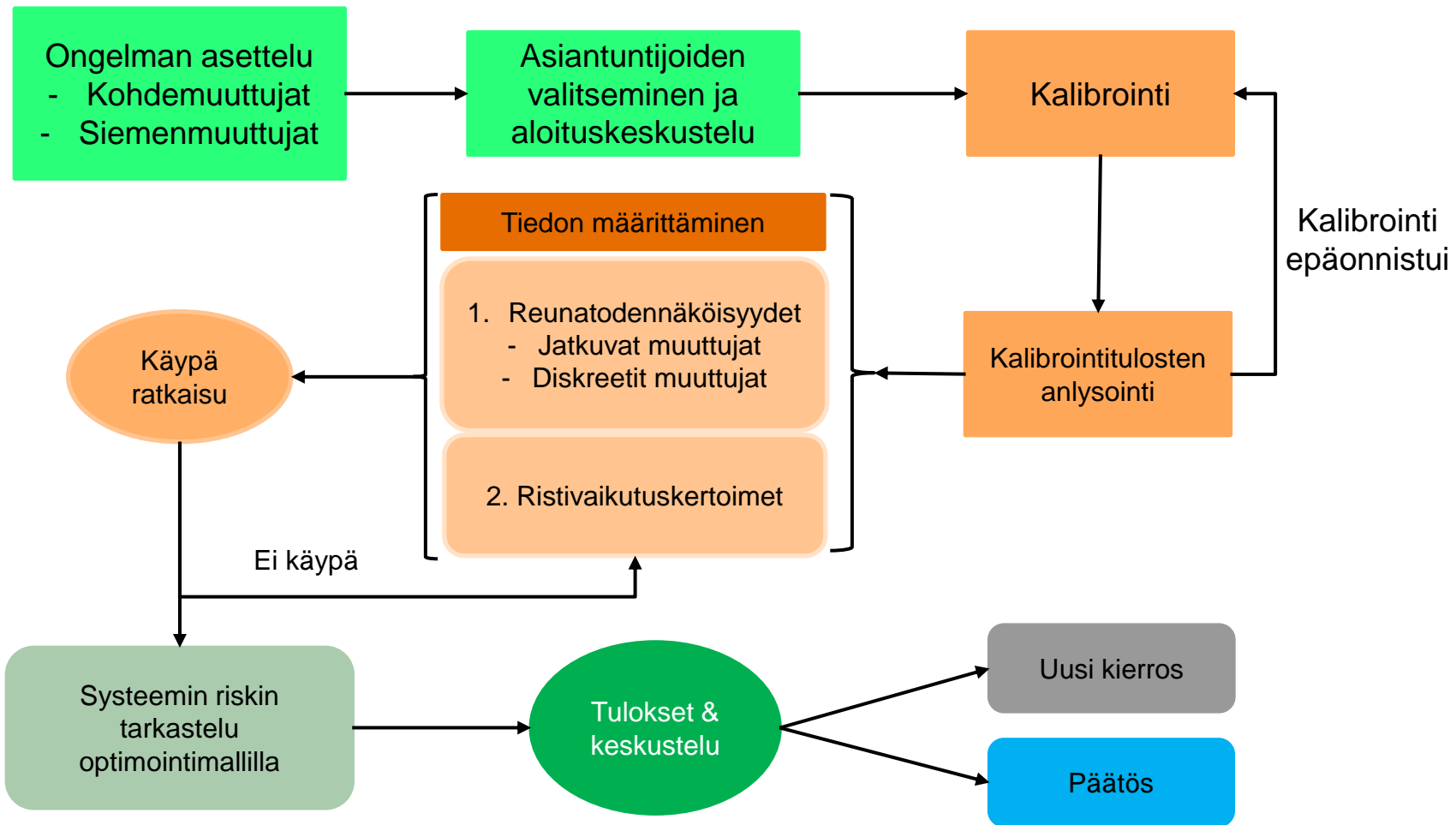
- Riskejä voidaan arvioida skenaario- ja ristivaikutusmenetelmillä
- Skenaariot määritetään epävarmuustekijöiden toteuminen yhdistelminä
- Skenaarioiden määrä $N = \prod_{i=1}^k K_i$, missä k on epävarmuustekijöiden määrä ja K_i on epävarmuustekijän i toteumien määrä
- Suuri määrä skenaarioita hankaloittaa skenaariokohtaisten todennäköisyyksien arvioimista

Ristivaikutusanalyysi

- Tarkastellaan toteumaparien vaikutuksia toisiinsa
- Arvioitavien kohteiden määrä kasvaa hitaammin
- Toteumien reunatodennäköisyydet ja ristivaikutukset määrittävät yhteisjakauman (huom. ei yksiselitteisesti)
- Odotusarvoinen riski ratkaistaan optimointimallilla

		Epävarmuustekijä 1		Epävarmuustekijä 2		Epävarmuustekijä 3	
		Toteumat		1	2	1	2
Epävarmuustekijä 1	1	Black	Black	Blue	Blue	Blue	Blue
	2	Black	Black	Blue	Blue	Blue	Blue
Epävarmuustekijä 2	1	Grey	Grey	Black	Black	Blue	Blue
	2	Grey	Grey	Black	Black	Blue	Blue
Epävarmuustekijä 3	1	Grey	Grey	Grey	Grey	Black	Black
	2	Grey	Grey	Grey	Grey	Black	Black

Prosessi asiantuntijatiedon määrittämiseen



Prosessin tausta ja ongelman asettelu

- Asiantuntija-arvioita käytetään kun tilastollista dataa ei ole saatavilla
- Prosessissa käytetään Cooken arviointimallia ja Delfoi menetelmää
- Määritetään epävarmuustekijät ja niiden toteumat
- **Kohdemuuttujat** muodotuvat toteumien reunatodennäköisyyksistä ja totetumaparien ristivaikutuksista
- Eri kohdemuuttujatyypeille generoidaan 8-20 **siemenmuuttujaa**, joita käytetään kalibroinnissa
- Esimerkki kysymyksestä siemenmuuttujaan: “Arvioi uusien Covid-19 tapausten lukumäärä Suomessa viikolla 35 (24.8.2020 – 30.8.2020).”

Työpaja: Ongelman asettelu

- Prosessia testattiin työpajassa, jonka tavoitteena oli arvioida Covid-19 pandemian vaikutuksia Aalto-yliopiston kolmannen vuoden kandidaattiopiskelijan valmistumiseen lukuvuodella 2020-2021
- Kohdemuuttujina oli neljän epävarmuustekijän toteumat

Epävarmuustekijä	Toteuma 1	Toteuma 2	Toteuma 3
Covid-19 tilanne (Diskreetti)	Vakaa	Kohtalainen	Kriittinen
Kurssijärjestelyt (Jatkuva)	0-10% etänä	11-80% etänä	81-100% etänä
Opiskelijan suoriutuminen (Jatkuva)	Korkea: > 5 op muutos	Normaali: ± 5 op muutos	Matala: < -5 op muutos
Taloudellinen tilanne (Jatkuva)	Parantunut: > 50 € käyttövaraa/kk	Normaali: ± 50 € käyttövaraa/kk	Huonontunut: < -50 € käyttövaraa/kk

Työpaja: Asiantuntijat

- Mukana viisi asiantuntijaksi valittua kolmannen vuoden opiskelijaa Aalto-yliopistosta eri koulutusohjelmista

Asiantuntija	Tausta
1	Tietotekniikka
2	Kone- ja rakennustekniikka
3	Teknillinen fysiikka ja matematiikka
4	Tuotantotalous
5	Kemian tekniikka

- Asiantuntijat vastasivat etäyhteyksillä prosessia varten valmistettuun kyselylomakkeeseen opastetusti
 - Kierroksia suoritettiin kaksi, jossa toisessa annettiin tarkentavia arvioita

Kalibrointi, informatiivisuus ja painottaminen

- Empiirisiä jakaumia r siemenmuuttujista verrataan odotusarvoisiin jakaumiin p Kullback-Leibler divergenssillä:

$$I(r, p) = \sum_{i=1}^t r_i \ln\left(\frac{r_i}{p_i}\right)$$

- Kalibroinnissa testataan kuinka hyvin asiantuntija osasi arvioida siemenmuuttujia
- Informatiivisuutta testataan niin, että siinä verrataan asiantuntijan arvioita siemenmuuttujista ja kohdemuuttujista suhteessa tasajakaumaan
- Painottaminen lasketaan kalibroinnin ja informatiivisuuden tulosta, ja lopuksi painot normalisoidaan

Työpaja: Kalibrointi, informatiivisuus ja painottaminen

- Todennäköisyydet

Asiantuntija	1	2	3	4	5
Kalibrointi	0.185	0.0426	0.5405	0.0877	0.5405
Informatiivisuus	0.4775	0.9929	0.1443	0.6516	0.4296
Painotus	0.1775	0.0849	0.1566	0.1147	0.4663

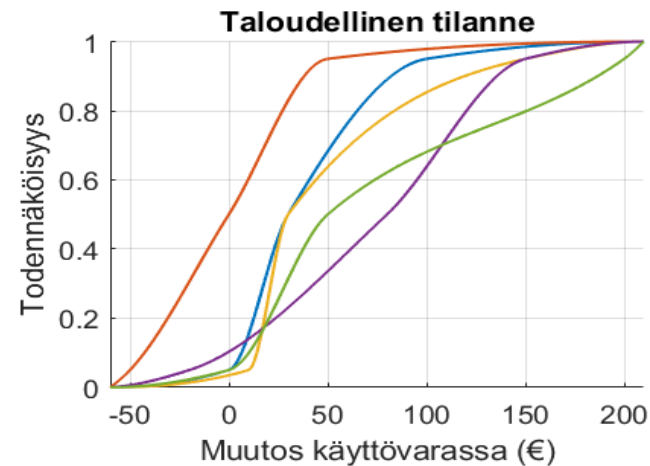
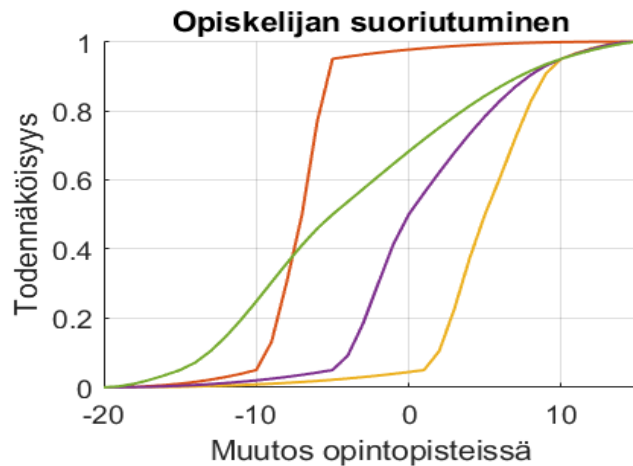
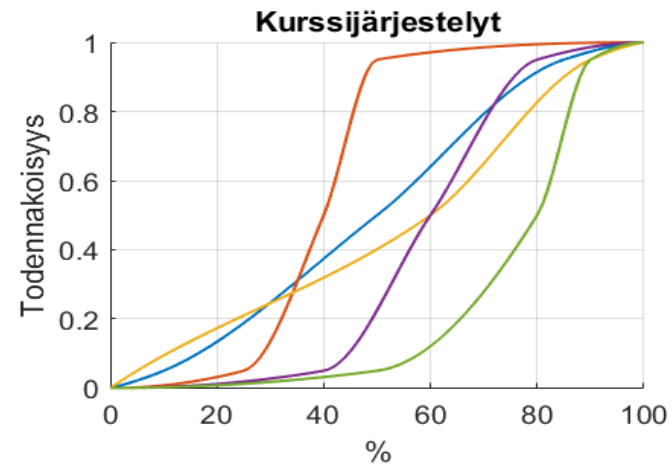
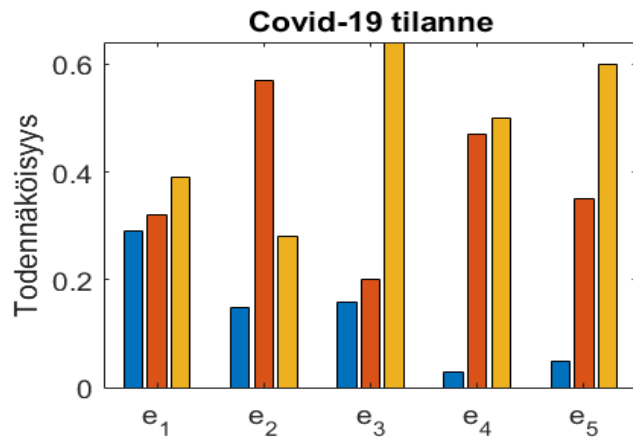
- Ristivaikutukset

Asiantuntija	1	2	3	4	5
Kalibrointi	0.168	0.031	0.503	0.503	0.398
Informatiivisuus	1.6661	1.614	1.5618	1.7183	1.4575
Painotus	0.1093	0.0195	0.3069	0.3376	0.2266

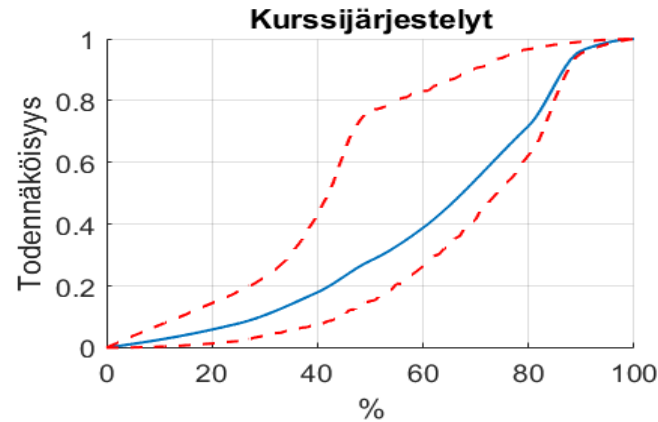
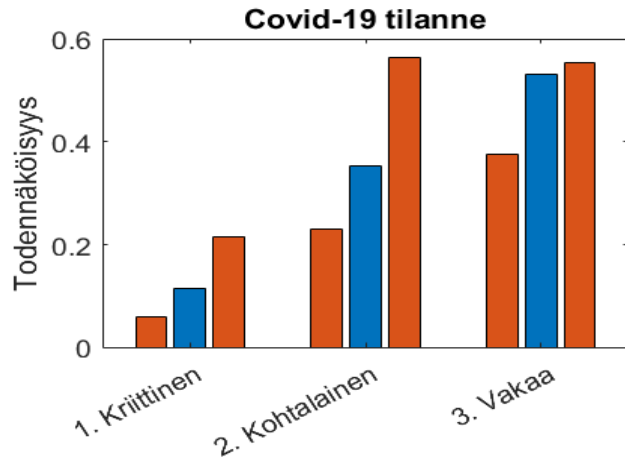
Tiedon määrittäminen (1/2): Todennäköisyydet

- Jatkuvat muuttujat
 - Määritetään asiantuntijoiden arvioista kohdemuuttujan 5% alarajasta, 95% ylärajasta ja 50% mediaanista
 - Asiantuntijoiden arviot interpoloidaan lineaarisella tai kvadraattisella menetelmällä
- Diskreetit muuttujat
 - Määritetään arpapeli menetelmällä, missä ideana on, että asiantuntija on neutraali kahden arpapelin välillä
- Luottamusvälit arvioille lasketaan painotetuilla keskiarvoilla Bootstrap algoritmia käyttäen

Työpaja: Arviot kohdemuuttujiin

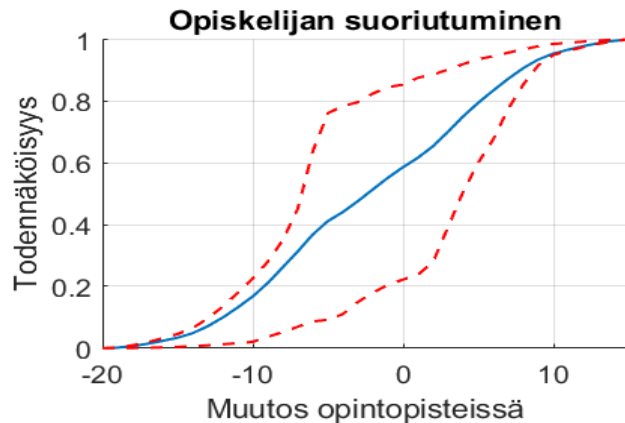


Työpaja: Todennäköisyydet ja 95% luottamusvälit

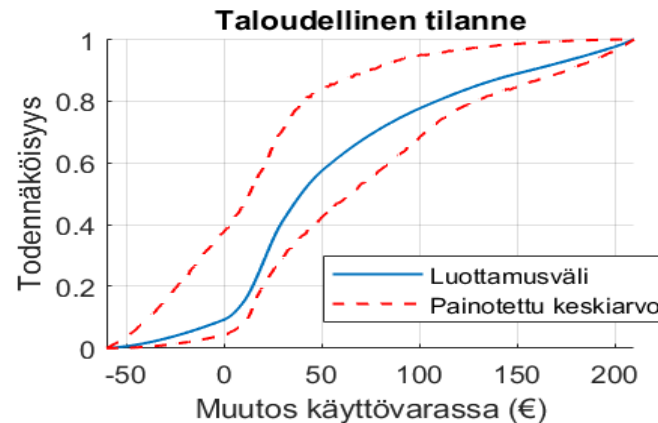


	Covid-19
Kriittinen	0.0591 – 0.2162
Kohtalainen	0.2310 – 0.5639
Vakaa	0.3770 – 0.5528

	Kurssijärjestelyt
0-10% etä	0.0026 – 0.0660
11-80% etä	0.6212 – 0.9006
81-100% etä	0.0334 – 0.3762



	Suoriutuminen
< -5 op	0.0920 – 0.7615
± 5 op	0.1743 – 0.5089
> 5 op	0.0642 – 0.3990



	Taloudellinen
< -50 €/kk	0.0011 – 0.0369
± 50 €/kk	0.4216 – 0.8043
> 50 €/kk	0.1588 – 0.5773

Tiedon määrittäminen (2/2): Ristivaikutuskertoimet

- Asiantuntijat arvioivat totetumaparien ristivaikutuksia kvalitatiivisella johdonmukaisuusskaalalla $c \in [-3,3]$

Taso	Kuvaus
3	Merkittävä vaikutus. B todennäköisyys kasvaa merkittävästi kun A tapahtuu.
2	Vaikuttava. B todennäköisyys kasvaa kohtalaisesti kun A tapahtuu.
1	Lievä vaikutus. B todennäköisyys kasvaa lievästi kun A tapahtuu.
0	Riippumattomia. Toteumilla ei ole suoraa vaikutusta.
-1	Lievä laskeva vaikutus. B todennäköisyys laskee lievästi kun A tapahtuu.
-2	Laskeva vaikutus. B todennäköisyys laskee kohtalaisesti kun A tapahtuu.
-3	Merkittävä laskeva vaikutus. B todennäköisyys laskee merkittävästi kun A tapahtuu.

Tiedon määrittäminen (2/2): Ristivaikutuskertoimet

- Johdonmukaisuuden taso muutetaan numeeriseksi väliksi Q_c käyttämällä $\delta \in R$ parametria

Q_{-3}	Q_{-2}	Q_{-1}	Q_0	Q_1	Q_2	Q_3
$[\delta^{-max}, \delta^{-2}]$	$[\delta^{-2}, \delta^{-1}]$	$[\delta^{-1}, 0.995]$	$\delta^0 \pm 0.005$	$[1.005, \delta^1]$	$[\delta^1, \delta^2]$	$[\delta^2, \delta^{max}]$

- Luottamusvälit ristivaikutuksille lasketaan Bootstrap algoritmilla painotetusta keskiarvosta
- Jos kaikilla asiantuntijoilla sama arvio → Väli otetaan taulukosta
- Määritettyjä asiantuntija-arvioita testataan ratkaisemalla optimointiongelma
 - Jos ratkaisu on käypä → siirrytään tulosten tarkasteluun
 - Jos ratkaisu ei ole käypä → levennetään ristivaikutusarvioita korjausalgoritmilla
 - Jos korjaus ei toimi tai väleistä tulee liian leveitä → Uusi tiedon määrittäminen kierros ristivaikutuksille

Työpaja: Ristivaikutusarviot 99% luottamusvälillä (1/2)



= nostava vaikutus



= laskeva vaikutus



= mahdollisesti riippumattomia

Kurssijärjestelyt

		81-100% etänä	11-80 % etänä	0-10% etänä
Covid-19	Vakaa	0.3386 - 0.8679	1.0050 – 2.6135	1.1064 – 1.5548
	Kohtalainen	1.1186 – 3.2002	0.6805 – 1.3940	0.2316 – 0.8610
	Kriittinen	1.1704 – 4.0651	0.2959 – 1.3825	0.2425 – 0.8422

Opiskelijan suoriutuminen

		< -5 op	±5 op	> 5 op
Covid-19	Vakaa	0.7141 – 0.9950	1.0050 – 2.6072	1.0050 – 1.4312
	Kohtalainen	0.9434 – 1.4482	0.8821 – 0.9950	0.6805 – 1.1525
	Kriittinen	1.2000 – 2.6590	0.7088 – 1.1177	0.2959 – 1.1525

Työpaja: Ristivaikutusarviot 99% luottamusvälillä (2/2)

		Taloudellinen tilanne		
		< -50 €/kk	±50 €/kk	> 50 €/kk
Covid-19	Vakaa	0.7456 – 1.1343	1.0050 – 2.5079	0.9100 – 1.4312
	Kohtalainen	0.7692 – 0.9950	0.8821 – 1.1200	1.0165 – 1.6865
	Kriittinen	0.4603 – 0.9790	0.7205 – 0.9848	1.0257 – 3.8347

		Opiskelijan suoriutuminen		
		< -5 op	±5 op	> 5 op
Kurssi-järjestelyt	0-10% etänä	0.7194 – 0.9950	1.0050 – 2.6120	0.7205 – 0.9950
	11-80% etänä	1.0100 – 2.7232	0.8821 – 1.1305	1.0100 – 2.7289
	81-100 % etänä	1.1064 – 2.7320	0.6744 – 0.9174	1.1064 – 2.7320

- Muut parit riippumattomia eli ristivaikutus on $C = [0.9950 - 1.0050]$

Lopputulos riskianalyysiin

- Riskianalyysissä ratkaistiin ala- ja yläraja todennäköisyydelle, että opiskelijan valmistuminen myöhästyy lukuvuotena 2020-2021 skenaariokohtaisten haittojen avulla
- Alaraja riskille $Risk^{low} = 0.3401$
- Yläraja riskille $Risk^{upp} = 0.3807$
- Tuloksista voidaan päätellä, että kolmannen vuoden kandidaattiopiskelijoista noin kolmasosa valmistuu myöhässä suhteessa tavoiteaikaan (3v) Covid-19 pandemian vaikutuksista johtuen

Tietolähteet

- L. C. Dias, A. Morton ja J. Quigley. Elicitation: The Science and Art of Structuring Judgement. Springer, New York, 2018.
- Seeve, T. 2018. A Structured Method for Identifying and Visualizing Scenarios. Diplomityö. Aalto-yliopisto. Teknillinen fysiikka ja matematiikka. Espoo.
- V. A. Banuls ja M. Turoff. Scenario Construction via Delphi and Cross-Impact Analysis. Technological Forecasting and Social Change, 78(9):1579-1602, 2011.
- N. Dalkey and O. Helmer. An Experimental Application of the Delphi Method to the Use of Experts. Managements Science, 9(3):458-467, 1963.