



HAL
open science

Quantifier des données qualitatives pour identifier des causalités complexes : Présentation de la méthode Fuzzy-set Qualitative Comparative Analysis (FsQCA)

Sébastien Brion

► To cite this version:

Sébastien Brion. Quantifier des données qualitatives pour identifier des causalités complexes : Présentation de la méthode Fuzzy-set Qualitative Comparative Analysis (FsQCA). Doctorat. France. 2018. cel-01756461

HAL Id: cel-01756461

<https://shs.hal.science/cel-01756461>

Submitted on 2 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Quantifier des données qualitatives pour identifier
des causalités complexes :
Présentation de la méthode Fuzzy-set Qualitative
Comparative Analysis (FsQCA)

26 mars 2018 - Séminaire Méthodologique Cret-Log

Sébastien Brion

@brionsebastien

Agenda

1. Présentation de la méthode
2. Les concepts de base
3. Mise en œuvre
4. Avantages et limites

@brionsebastien

1. Présentation de la méthode

@brionsebastien

1. La méthode

Limites des approches qualitatives

- Difficile de proposer une démarche rigoureuse permettant l'analyse de causalités
- Variété des méthodes pour mener des études comparatives

Limites des approches quantitatives

- Peu adaptées aux petits échantillons
- Limitées pour traiter des combinaisons complexes de variables
- Une seule série de causes possible pour expliquer un phénomène (algèbre linéaire)

L'analyse qualitative comparée (FsQCA) tente de répondre à ces limites

« It is surprising, for all that has been said about the value of comparison, that a rigorous comparative methodology has not emerged. The reason of this lack may be the great difficulties that a rigorous comparative methodology would impose » M. Porter, 1970

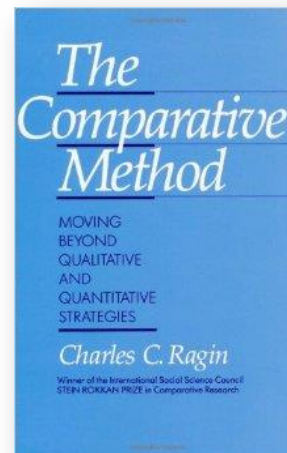
@brionsebastien

1. La méthode

- Définition

La méthode de l'analyse qualitative comparée apporte une **sophistication méthodologique à la comparaison des « cas »** et permet de traiter explicitement des variables et d'effectuer des comparaisons.

La QCA s'appuie sur l'**algèbre booléenne**. FsQCA (Fuzzy set QCA) est un raffinement de QCA en offrant la possibilité de traiter des **variables quasi-continues** plutôt que binaires.



Méthode adaptée à l'étude de situations complexes, car permet de mettre en évidence **plusieurs combinaisons (configurations) de conditions (variables) possibles** pour expliquer un même phénomène.

@brionsebastien

1. La méthode

- 5 variantes de QCA

- 1. Crisp-set QCA (csQCA) (Ragin 1987) – Basés sur des variables binaires
- 2. Fuzzy-set QCA (fsQCA) (Ragin 2000) – Données continues et ordinales
- 3. Multi-value QCA (mvQCA) (Cronqvist and Berg-Schlusser 2008; Thiem 2014) – Données multinomiales
- 4. Temporal QCA (tQCA) (Caren and Panofsky 2005) - basés sur des séquences temporelles (cf. Garcia and Arino 2016 for panel data; Baumgartner and Epple 2014; Fischer & Maggetti 2016)
- 5. Two-step QCA (tsQCA) (Schneider and Wagemann 2006)
 - Permet de distinguer différents contextes, niveaux d'analyses etc...

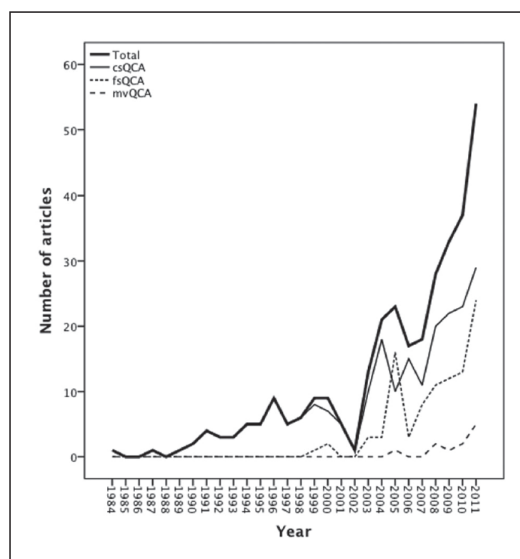


Figure 1. Number of articles by QCA technique, 1984–2011.⁵

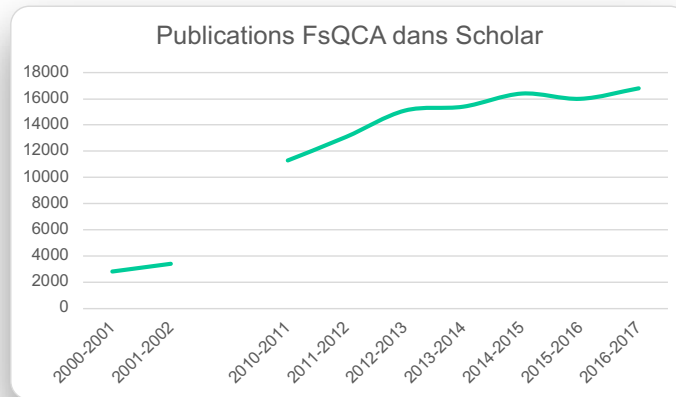
Schneider and Wagemann 2012

Rihoux et al., 2013

@brionsebastien

1. La méthode

Envergure de la méthode :



- Types de revues qui plébiscitent cette démarche : SMJ, CMR, MISQ **JBR**, ...
- Charles Ragin (h : 61 ; 40.000 citations !)

@brionsebastien

1. La méthode

• Contexte d'utilisation :

- Comparaison de cas (cas = individu, projet, entreprise, pays...)
- Etudes exploratoires comme explicatives (validité externe ?)
- Concerne toutes les sciences sociales ou presque
- Fonctionne avec de petits échantillons (10 à 50 cas ou unités d'analyse)

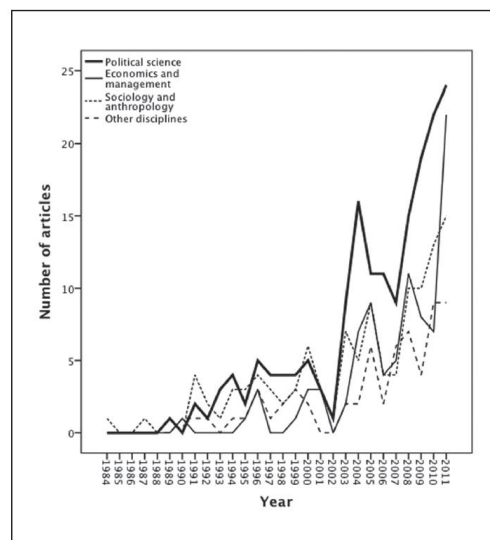


Figure 2. Number of articles by discipline, 1984-2011.⁶

Rihoux et al., 2013

@brionsebastien

1. La méthode

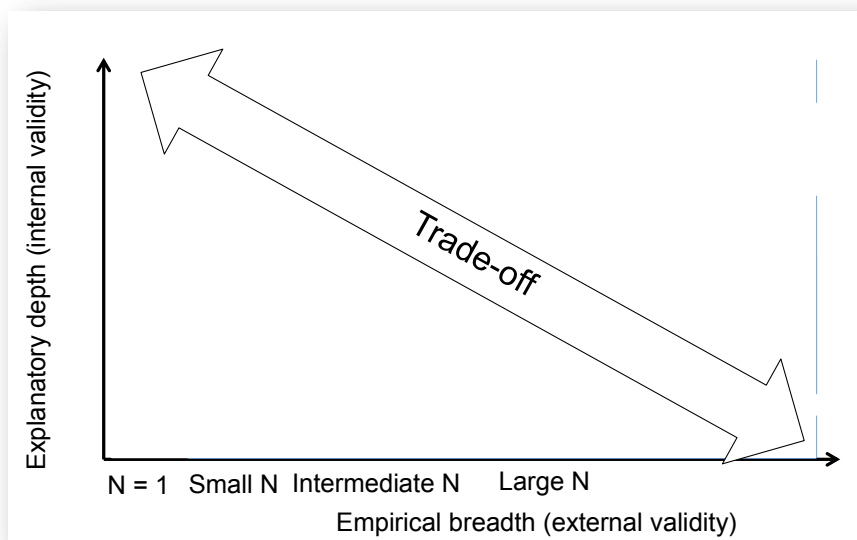
- A quelle type de problématique permet-elle de répondre ?

- Descriptive : Présenter les variables et leurs relations ou créer des typologies empiriques, ou idéaux types (Kvist, 2007)
- **Explicative complexe : développement (exploration ?) d'hypothèses et de théories**
- **Enrichissement des théories existantes (Fiss, 2011)**

@brionsebastien

1. La méthode

Positionnement méthodologique

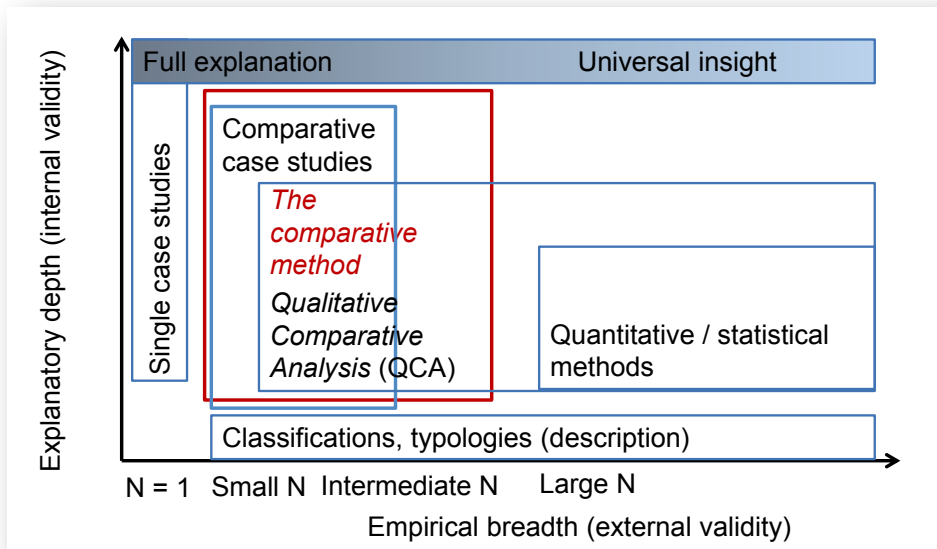


Thomann, 2017

@brionsebastien

1. La méthode

Positionnement méthodologique



Thomann, 2017 @brionsebastien

1. La méthode

Positionnement méthodologique

<i>Quantitative</i>	<i>QCA</i>	<i>Qualitative</i>
Linear causality	Complex causality	Holistic causality
Variable-based	Case-based and comparative	Case-based
Large-N	Medium-N	Small-N
Pattern recognition	Between generality and complexity	High level of detail
Objectifying	Both - systematic and transparent comparison	Interpretive

Verweij, 2015 @brionsebastien

1. La méthode

• L'écosystème des outils d'analyse :

Gratuits et de plus en plus perfectionnés

L'offre logiciel :

<http://www.compasss.org/software.htm>

Les plus utilisés sont multi-OS

Function	Tosmana ^b	Kirq ^c	fs/QCA ^d	fuzzy ^e	QCA3 ^f	QCA ^g
<i>variant</i>						
csQCA	●	●	●	●	●	●
mvQCA	●	○	○	○	●	●
fsQCA	○	●	●	●	●	●
(tQCA)	○	○	●	○	●	●
<i>solution type</i>						
complex	●	●	●	●	●	●
intermediate	○	◐	●	○	◐	●
parsimonious	●	●	●	●	●	●
<i>procedure</i>						
necessity tests	○	●	◐	◐	◐	●
parameters of fit	○	●	●	◐	◐	●
calibration	◐	○	◐	◐	◐	●
factorization	○	○	○	○	○	●
identify (C)SAs	●	○	○	○	●	●
statistical tests	○	○	○	●	◐	○

^a ○ / ◐ / ● = no/partial/full functionality

^b version 1.3.2.0; ^c version 2.1.9; ^d version 2.5; ^e version st0140_2; ^f version 0.0-5; ^g version 1.0-5

Thiem & Dusa 2013: 88.

@brionsebastien

2. Les concepts de base

@brionsebastien

Les concepts à connaître pour aborder la méthode

- Sources de données : des variables aux **conditions**...
- La méthode repose sur les notions de **conditions nécessaires** et de conditions **suffisantes**. Ces notions sont plus larges et plus souples que les effets nets mis en évidence par les tests statistiques
- Empiriquement, les conditions nécessaires et suffisantes sont imparfaites, des **indicateurs sont nécessaire pour saisir l'étendue de ces imperfections**
- La recherche de **causes explicatives** ne s'entend pas au sens commun, elles s'exercent dans le paradigme de la « **complexité causale** »
- Si certaines régularités complexes peuvent être observées (causes), elles dépendent toujours des **particularités des cas**.

@brionsebastien

Sources de données

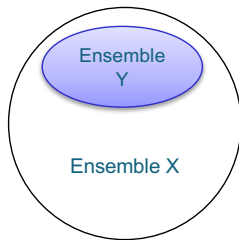
- L'équivalent de la notion de variable est appelée « **condition** » dans QCA.
 - En **algèbre linéaire**, une **variable est en compétition avec les autres variables** pour expliquer un résultat
 - Dans la **théorie des ensembles (algèbre booléen)**, une **condition est le plus souvent combinée** avec d'autres conditions pour expliquer un phénomène
- Ces conditions peuvent être construites à partir des sources de données multiples :
 - Données **quantitatives** (continues, échelle, indice, ratio...)
 - Données **qualitatives** (nominales, textes...)
- Les sources peuvent être **micro, méso** ou **macro**

@brionsebastien

Conditions nécessaire et suffisante dans la théorie des ensembles

Condition nécessaire

Une **condition X est dite nécessaire** lorsque **le résultat Y requiert systématiquement sa présence** dans toutes les combinaisons de conditions qui conduisent au résultat, sa seule présence ne garantit pas le résultat



Y est un sous-ensemble de X

$$Y \leq X$$

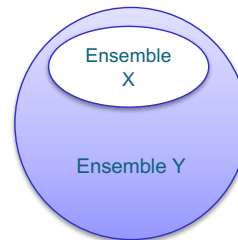
Y implique X

Y apparaît si et seulement si X est présent

Toutefois, si X est nécessaire à Y, cela n'implique pas qu'il soit suffisant

Condition suffisante

Une **condition est dite suffisante** si **les cas qui partagent ces conditions donnent le même résultat**.



X est un sous-ensemble de Y

$$X \leq Y$$

X implique Y

La condition X peut conduire au résultat Y

Schneider and Wagemann 2012

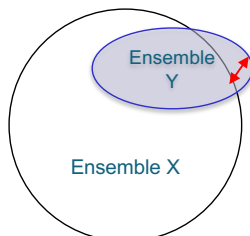
@brionsebastien

Indicateurs clés : *Consistency and Coverage*

Au plan empirique, les **ensembles de conditions** nécessaires et suffisantes sont **imparfaits**

=> Besoin d'un indicateur (**Consistency** – entre 0 et 1) permettant d'apprécier **dans quelle mesure les conditions sont en lien avec le résultat** (paramètre de Fit)

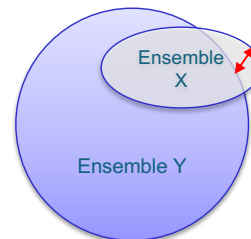
Necessary consistency



X est présent mais pas systématiquement lorsque Y apparaît

Dans ce cas, Y n'est pas un sous-ensemble parfait de X

Sufficiency consistency



X est un sous-ensemble partiel de Y :
l'indicateur de **consistency** mesure ce pourcentage d'occurrence (idéalement \geq à 0.75 %)

Schneider and Wagemann 2012

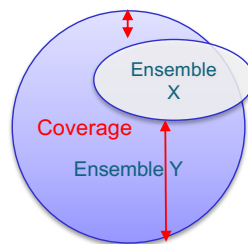
Thomann, 2017

@brionsebastien

Indicateurs clés : *Consistency and Coverage*

Au plan empirique, les **ensembles de conditions** nécessaires et suffisantes sont **imparfaits**

=> Besoin d'un autre indicateur (**Coverage** – entre 0 et 1) permettant d'apprécier le **niveau de couverture empirique du résultat**



Empiriquement, plus les combinaisons de conditions sont nombreuses, moins le taux de couverture empirique est élevé

Part de Y couverte par X :
Le **coverage** indique la proportion de cas concernés par la condition X expliquant Y

Schneider and Wagemann 2012

Thomann, 2017
@brionsebastien

La complexité causale comme propriété des résultats de FsQCA

Éléments constitutifs :

- **Causalité conjonctive**
 - **Une condition** se manifeste généralement à partir de **combinaisons avec d'autres conditions** (notion de configuration, noté* = AND)
- **Equifinalité** : « *Tous les chemins mènent à Rome* »
Un résultat peut être expliqué par plusieurs **configurations mutuellement non exclusives** (+ = OR)
- **Asymétrie causale** :
 - **La présence de X pour Y n'implique pas l'absence de X** (lorsque Y est absent (la réciproque se vérifie rarement par les mêmes conditions)).
 - **Multifinalité**
 - **Une même condition peut conduire à des résultats différents**, tout dépend du contexte dans lequel elle apparaît
 - De la même manière, le rôle causale de l'occurrence d'une condition, ne nous dit rien sur le rôle causal de sa non-occurrence.

Schneider and Wagemann 2012

@brionsebastien

Quelques conditions empiriques...

- Le chercheur doit disposer d'une connaissance préalable des cas.
- Il faut tenter de réduire le nombre de conditions (de variables) :
 - condition $C \leq 8$ (256 combinaisons)
 - idéalement C entre 4 (16 combinaisons) et 6 (64 combinaisons) conditions
- Taille de l'échantillon :
 - $N \geq 10$
 - *Rule of thumb* : $N \geq C \times 3$; plus restrictif : $N \geq 2^C$

Schneider and Wagemann 2012

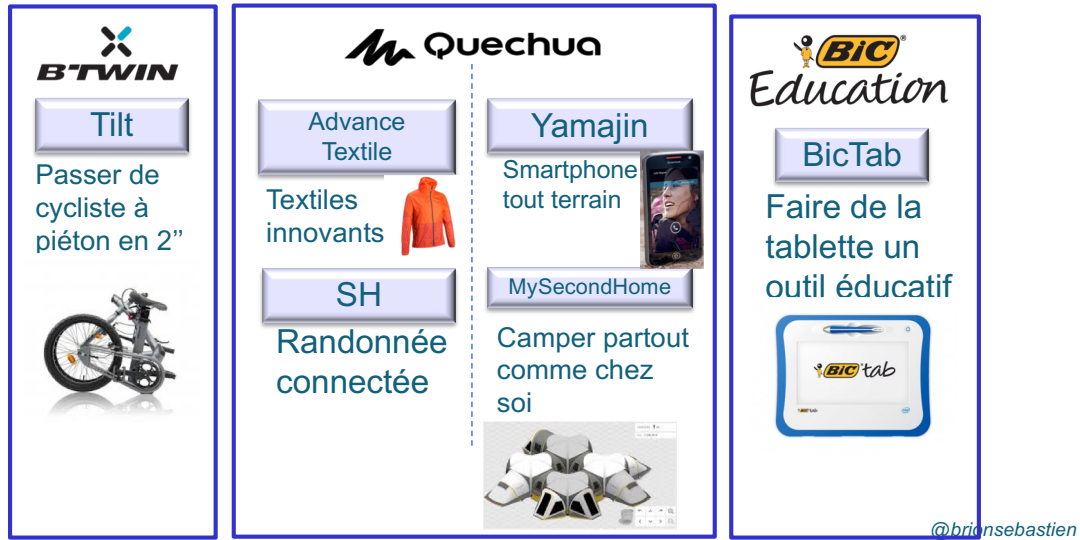
@brionsebastien

3. Mise en œuvre

@brionsebastien

Etude des tensions paradoxales des Start-up Internes innovantes (SUI)

Projet **ANR Kinetic** 2012 : Quelles tensions paradoxales favorisent la mise en œuvre de SUI pour innover ? – Brion & Chollet (2014)



Quelques éléments de la démarche méthodologique....

Sources de données qualitatives :

- 8 SUI + 3 projets d'exploration dédiés, dans 2 entreprises (3 DAS)
- 44 entretiens semi-directifs (70h, tous types d'acteurs, + de 2500 verbatims codés), N=39 retenus pour l'étude (verbatims manquants)
- + Observation non participante
- + Présentation des résultats intermédiaires à l'équipe dirigeante (en 2 étapes) > co-validation des tensions les plus significatives tout au long des projets

Plan d'analyse :

- **Phase 1** : Nvivo + interactions avec les acteurs terrains = bonne connaissance du contenu de chaque cas et des tensions (1^{ère} publication, MI 2017)
- **Phase 2** : FsQCA > existe-t-il des tensions paradoxales récurrentes entre les cas (les ICC étant non significatifs)

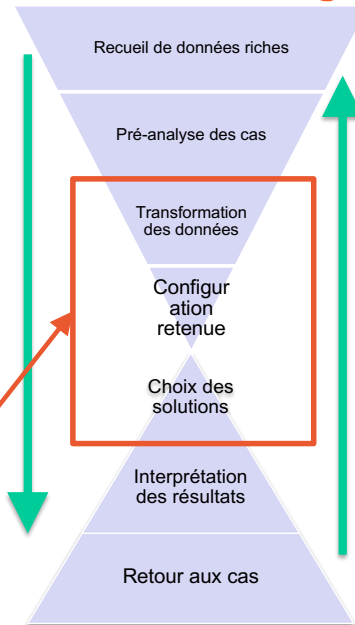
@brionsebastien

Quelques éléments de la démarche méthodologique....

La démarche FsQCA permet de s'appuyer sur un processus itératif

- Une opération de transformation des données est nécessaire
- La richesse des cas sert à interpréter les résultats de FsQCA

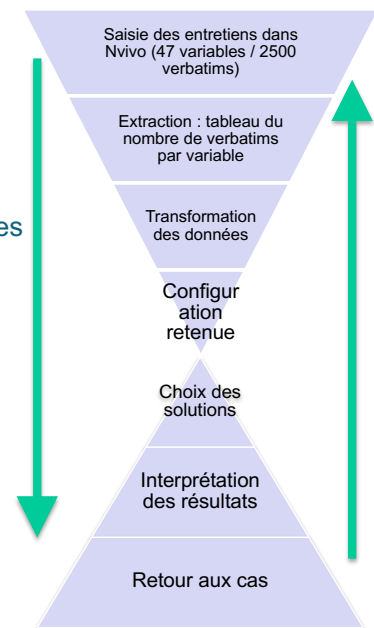
La démarche FsQCA intervient sur ces 3 phases



Adapté de Verweij, 2015 @brionsebastien

Démarche méthodologique pour l'étude

- Préparation des données
- Transformation des données
- Analyse



@brionsebastien

Préparation des données

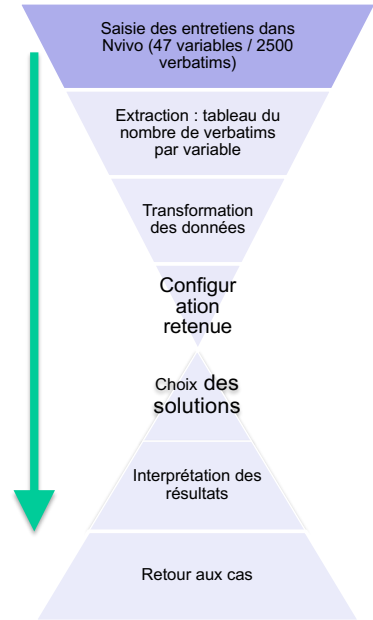
Guide théorique pour le codage des verbatims dans Nvivo

	Propositions	Integrative mechanisms	↔	Generative mechanisms
Organizing	1	Control and support	↔	Autonomy and separation
Belonging	2	Team cohesiveness	↔	Team diversity
Learning	3	Internal knowledge	↔	External knowledge
Performing	4	Short term efficiency	↔	Long term renewal

A partir des verbatims codés sur ces 8 conditions, on cherche à identifier les tensions que sont en mesure de gérer les acteurs, ainsi que celles qu'ils redoutent (efficacité perçue faible)

Quelles combinaisons de tensions paradoxes augmentent l'efficacité perçue des acteurs des SUI ?

Brion & Chollet, 2015



@brionsebastien

Préparation des données

Résultat du codage dans Nvivo des dimensions paradoxales

	A: B2-	B: B3+	C: B3-	D: B3+	E: B4-	F: B4+
Bpiez1	0	0	0	0	0	0
Bpiez2	0	1	2	0	0	0
Bpiez3	1	0	0	0	0	0
Biab1	5	8	8	0	0	0
Biab2	0	3	0	2	0	1
Biab3	0	0	0	1	0	0
Bwatp1	0	3	0	0	0	0
Bwatp2	0	3	0	0	0	0
Bwatp3	1	0	2	0	1	0
Viof1	4	1	0	2	0	2
Vioz2	2	2	0	0	0	0
Veasyl1	1	2	2	2	0	0
Veasyl2	0	1	1	1	2	1
Vvtt1	0	1	0	0	2	2
Vvtt2	0	1	2	0	0	0
Vvtt3	2	0	2	0	0	0
Vvtt4	0	1	0	0	0	0
Vvtt5	0	1	0	0	0	2
Vvtt6	0	0	0	0	0	0
Vvtt7	0	0	0	0	0	2
Vvtt8	0	0	0	0	0	0
Vvtt9	0	0	0	0	0	0
Vvtt10	0	0	0	0	0	0
Vvtt11	0	0	0	0	0	0
Vvtt12	0	0	0	0	0	0
Vvtt13	0	0	0	0	0	0
Vvtt14	0	0	0	0	0	0
Vvtt15	0	0	0	0	0	0
Vvtt16	0	0	0	0	0	0
Vvtt17	0	0	0	0	0	0
Vvtt18	0	0	0	0	0	0
Vvtt19	0	0	0	0	0	0
Vvtt20	0	0	0	0	0	0
Vvtt21	0	0	0	0	0	0
Vvtt22	0	0	0	0	0	0
Vvtt23	0	0	0	0	0	0
Vvtt24	0	0	0	0	0	0
Vvtt25	0	0	0	0	0	0
Vvtt26	0	0	0	0	0	0
Vvtt27	0	0	0	0	0	0
Vvtt28	0	0	0	0	0	0
Vvtt29	0	0	0	0	0	0
Vvtt30	0	0	0	0	0	0
Vvtt31	0	0	0	0	0	0
Vvtt32	0	0	0	0	0	0
Vvtt33	0	0	0	0	0	0
Vvtt34	0	0	0	0	0	0
Vvtt35	0	0	0	0	0	0
Vvtt36	0	0	0	0	0	0
Vvtt37	0	0	0	0	0	0
Vvtt38	0	0	0	0	0	0
Vvtt39	0	0	0	0	0	0
Vvtt40	0	0	0	0	0	0
Vvtt41	0	0	0	0	0	0
Vvtt42	0	0	0	0	0	0
Vvtt43	0	0	0	0	0	0
Vvtt44	0	0	0	0	0	0
Vvtt45	0	0	0	0	0	0
Vvtt46	0	0	0	0	0	0
Vvtt47	0	0	0	0	0	0
Vvtt48	0	0	0	0	0	0
Vvtt49	0	0	0	0	0	0
Vvtt50	0	0	0	0	0	0
Vvtt51	0	0	0	0	0	0
Vvtt52	0	0	0	0	0	0
Vvtt53	0	0	0	0	0	0
Vvtt54	0	0	0	0	0	0
Vvtt55	0	0	0	0	0	0
Vvtt56	0	0	0	0	0	0
Vvtt57	0	0	0	0	0	0
Vvtt58	0	0	0	0	0	0
Vvtt59	0	0	0	0	0	0
Vvtt60	0	0	0	0	0	0
Vvtt61	0	0	0	0	0	0
Vvtt62	0	0	0	0	0	0
Vvtt63	0	0	0	0	0	0
Vvtt64	0	0	0	0	0	0
Vvtt65	0	0	0	0	0	0
Vvtt66	0	0	0	0	0	0
Vvtt67	0	0	0	0	0	0
Vvtt68	0	0	0	0	0	0
Vvtt69	0	0	0	0	0	0
Vvtt70	0	0	0	0	0	0
Vvtt71	0	0	0	0	0	0
Vvtt72	0	0	0	0	0	0
Vvtt73	0	0	0	0	0	0
Vvtt74	0	0	0	0	0	0
Vvtt75	0	0	0	0	0	0
Vvtt76	0	0	0	0	0	0
Vvtt77	0	0	0	0	0	0
Vvtt78	0	0	0	0	0	0
Vvtt79	0	0	0	0	0	0
Vvtt80	0	0	0	0	0	0
Vvtt81	0	0	0	0	0	0
Vvtt82	0	0	0	0	0	0
Vvtt83	0	0	0	0	0	0
Vvtt84	0	0	0	0	0	0
Vvtt85	0	0	0	0	0	0
Vvtt86	0	0	0	0	0	0
Vvtt87	0	0	0	0	0	0
Vvtt88	0	0	0	0	0	0
Vvtt89	0	0	0	0	0	0
Vvtt90	0	0	0	0	0	0
Vvtt91	0	0	0	0	0	0
Vvtt92	0	0	0	0	0	0
Vvtt93	0	0	0	0	0	0
Vvtt94	0	0	0	0	0	0
Vvtt95	0	0	0	0	0	0
Vvtt96	0	0	0	0	0	0
Vvtt97	0	0	0	0	0	0
Vvtt98	0	0	0	0	0	0
Vvtt99	0	0	0	0	0	0
Vvtt100	0	0	0	0	0	0

Saisie des entretiens dans Nvivo (47 variables / 2500 verbatims)

Extraction : tableau du nombre de verbatims par variable

Transformation des données

Configuration retenue

Choix des solutions

Interprétation des résultats

Retour aux cas

@brionsebastien

Transformation des données

Notion de calibrage des données : l'exemple du Freedom House Index

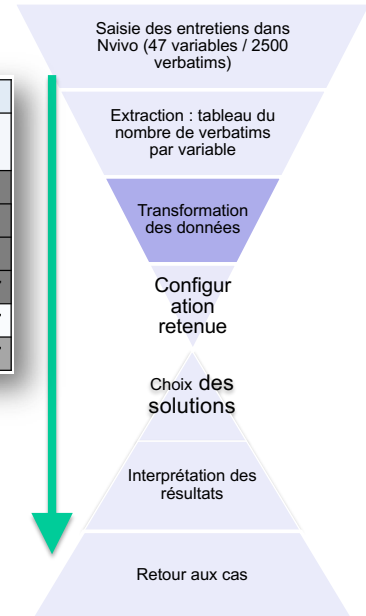
Ex. Membership in the set of «democratic countries»			
Measurement: Freedom house index FHI	<i>Free</i> 1.0 - 2.5	<i>Partly Free</i> 3.0 - 5.0	<i>Not Free</i> 5.5 - 7.0
Crisp set I	Democratic	Undemocratic	
Crisp set II	Democratic		Undemocratic
Fuzzy set I	Fully in: 2.5 1 - 2.5	Crossover 4.0 3 - 3.5 4.5 - 5	
			Fully out: 5.5 5.5-7
Fuzzy set II	Fully in: 1 1 - 5	Crossover 5.25	
			Fully out: 7 5.5 - 7

Pour chaque condition, plusieurs solutions de codage sont possibles

Ce codage détermine à quel ensemble appartient une condition

Le choix du **codage doit être justifié** empiriquement (connaissance des cas) et/ou théoriquement

Schneider and Wagemann 2012



@brionsebastien

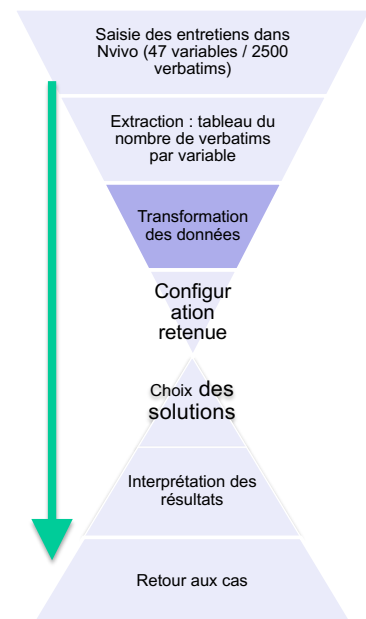
Transformation des données

Plusieurs niveaux de finesse existent dans le codage : dichotomique à continu

CRISP VERSUS FUZZY SETS				
Crisp set	Three-value fuzzy set	Four-value fuzzy set	Six-value fuzzy set	"Continuous" fuzzy set
1 = fully in	1 = fully in	1 = fully in	1 = fully in	1 = fully in
			.9 = mostly but not fully in	Degree of membership is more "in" than "out": $.5 < x_i < 1$
	.5 = neither fully in nor fully out	.75 = more in than out	.7 = more or less in	.5 = cross-over: neither in nor out
		.25 = more out than in	.3 = more or less out	Degree of membership is more "out" than "in": $0 < x_i < .5$
			.1 = mostly but not fully out	
0 = fully out	0 = fully out	0 = fully out	0 = fully out	0 = fully out

Ragin, 2008

Calibrage utilisé dans l'étude



@brionsebastien

Transformation des données

Calibrage des données à partir du comptage des verbatims de Nvivo

cases	perfFuzz	explorFuzz	perfFuzz	b2+Fuzz	b3+Fuzz	I1-Fuzz	I1+Fuzz	I2+Fuzz	I3+Fuzz	o1-Fuzz	o1+Fuzz	o4-Fuzz	o10+Fuzz	o31+Fuzz	p3-Fuzz	p3+Fuzz	p4-Fuzz	p4+Fuzz	
Piezo1	0.01	0.05	0.02	0.05	0.05	0.05	0.27	0.5	0.05	0.5	0.05	0.5	0.05	0.5	0.05	0.5	0.05	0.5	
Piezo2	0.01	0.18	0.05	0.5	0.05	0.05	0.95	0.85	0.05	0.93	0.05	0.05	0.05	0.5	0	0.5	0.05	0.05	
Piezo3	0.01	0.05	0.02	0.5	0.05	0.05	0.27	0.5	0.9	0.5	0.05	0.95	0.05	0.05	0.05	0.05	0.05	0.5	
Bictab4	0.5	0.99	0.95	0.98	0.97	0.5	0.99	0.27	0.5	0.05	0.92	0.5	0.98	0.99	0	0.95	0.05	0.97	
Bictab5	0.82	0.99	0.98	0.78	0.62	0.99	0.98	0.27	0.82	0.05	0.86	0.05	0.95	0.5	0.05	0.98	0.5	0.92	
Bictab6	0.95	0.99	0.99	0.7	0.62	0.05	0.82	0.61	0.9	0.05	0.95	0.05	0.98	0.05	0	0.82	0.05	0.65	
Waterp7	0.01	0.5	0.1	0.7	0.05	0.99	0.98	0.97	0.05	0.89	0.05	0.5	0.05	0.82	0.05	0.5	0.05	0.05	
Waterp8	0.01	0.18	0.05	0.7	0.5	0.05	0.05	0.5	0.05	0.85	0.05	0.5	0.05	0.05	0.5	0.05	0.5	0.05	
Waterp9	0.01	0.5	0.1	0.05	0.05	0.5	0.5	0.5	0.05	0.5	0.05	0.95	0.05	0.05	0.5	0.05	0.5	0.05	
Borign10	0.95	0.05	0.5	0.5	0.62	0.95	0.98	0.93	0.9	0.61	0.5	0.95	0.5	0.95	0	0.05	0.99	0.77	
Borign11	0.82	0.05	0.32	0.61	0.05	0.82	0.95	0.98	0.98	0.7	0.05	0.05	0.5	0.5	0	0.05	0.82	0.5	
Easy12	0.5	0.5	0.5	0.61	0.62	0.5	0.9	0.95	0.9	0.78	0.5	1	0.5	0.99	0.98	0.68	0.05	0.05	
Easy13	0.82	0.82	0.88	0.05	0.5	0.5	0.27	0.95	0.5	0.98	0.65	0.05	0.5	0.05	0	0.5	0.68	0.5	
Rock14	0.95	0.05	0.5	0.5	0.5	0.05	0.5	0.51	0.5	0.05	0.65	1	0.88	0.05	0	0.05	0.05	0.5	
Rock15	0.95	0.18	0.73	0.5	0.5	0.82	0.68	0.85	0.9	0.61	0.05	0.05	0.5	0.05	0	0.05	0.5	0.77	
Rock16	0.82	0.18	0.5	0.05	0.5	0.95	0.12	0.7	0.5	0.99	0.05	0.5	0.5	0.05	0.99	0.05	0.05	0.05	
TIR17	0.5	0.95	0.88	0.5	0.73	0.82	0.68	0.12	0.05	0.5	0.65	1	0.95	0.05	0.5	0.68	0.5	0.05	
TIR18	0.5	0.95	0.88	0.5	0.5	0.05	0.82	0.27	0.05	0.61	0.86	0.95	0.73	0.05	0.5	0.68	0.05	0.05	
TIR19	0.82	0.95	0.95	0.61	0.5	0.5	0.27	0.12	0.5	0.5	0.95	0.05	0.73	0.5	0.05	0.5	0.5	0.05	
Velomob20	0.01	0.99	0.5	0.05	0.05	0.05	0.05	0.5	0.85	0.05	0.61	0.77	0.05	0.98	0.5	0	0.82	0.05	0.05
AT21	0.99	0.18	0.88	0.7	0.73	0.82	0.9	0.5	0.5	0.05	0.95	0.5	0.05	0.95	0.97	0.5	0.93	0.65	
AT22	0.99	0.18	0.88	0.7	0.62	0.95	0.5	0.05	0.82	0.05	0.65	0.05	0.73	0.95	0.5	0.05	0.5	0.05	
AT23	0.95	0.05	0.5	0.61	0.5	0.05	0.12	0.05	0.5	0.05	0.77	0.05	0.98	0.5	0.93	0.05	0.75	0.05	
AT25	0.99	0.18	0.88	0.61	0.62	0.5	0.05	0.12	0.05	0.05	0.97	0.05	0.05	0.05	0.5	0.05	0.05	0.05	
AT27	0.95	0.05	0.5	0.05	0.73	0.82	0.05	0.12	0.68	0.05	0.05	0.05	0.05	0.05	0	0.05	0.05	0.65	
MSH28	0.18	0.82	0.5	0.61	0.62	0.05	0.05	0.05	0.5	0.78	0.5	0.95	0.98	0.05	0.85	0.05	0.75	0.05	
MSH29	0.18	0.99	0.88	0.85	0.5	0.05	0.27	0.61	0.5	0.61	0.05	0.5	0.95	0.82	0.95	0.05	0.68	0.65	
MSH30	0.18	0.95	0.73	0.61	0.05	0.5	0.95	0.5	0.68	0.78	0.5	0.5	0.05	0.05	0.05	0.9	0.87	0.05	
MSH31	0.05	0.95	0.5	0.7	0.62	0.5	0.12	0.05	0.5	0.5	0.05	0.05	0.05	0.05	0.05	0.97	0.68	0.75	0.05
MSH33	0.18	0.82	0.5	0.85	0.73	0.05	0.05	0.05	0.05	0.05	0.5	0.95	0.05	0.05	0.61	0.05	0.05	0.65	
MSH36	0.5	0.82	0.73	0.05	0.05	0.05	0.12	0.82	0.05	0.05	0.12	0.82	0.05	0.5	0.61	0.5	0.05	0.77	
SH37	0.05	0.18	0.1	0.5	0.82	0.05	0.98	0.7	0.5	0.7	0.05	0.05	0.5	0.82	0.95	0.5	0.82	0.77	
SH38	0.18	0.5	0.32	0.05	0.73	0.5	0.95	0.89	0.5	0.05	0.05	0.95	0.88	0.82	0.95	0.05	0.98	0.77	
SH39	0.01	0.05	0.02	0.05	0.05	0.05	0.05	0.27	0.5	0.05	0.05	0.95	0.05	0.05	0.5	0.05	0.05	0.05	
SH40	0.05	0.18	0.1	0.05	0.05	0.5	0.12	0.27	0.05	0.05	0.5	0.5	0.05	0.82	0.61	0.05	0.75	0.05	
Yam41	0.82	0.95	0.95	0.7	0.5	0.5	0.05	0.5	0.68	0.05	0.65	0.5	0.5	0.05	0.05	0.68	0.5	0.5	
Yam42	0.95	0.99	0.99	0.61	0.05	0.05	0.12	0.27	0.05	0.05	0.65	0.5	0.05	0.82	0	0.5	0.05	0.05	
Yam43	0.82	0.95	0.95	0.5	0.62	0.05	0.5	0.27	0.05	0.05	0.86	0.05	0.73	0.5	0.5	0.82	0.05	0.65	
Yam44	0.99	0.82	0.98	0.05	0.73	0.05	0.02	0.12	0.05	0.05	0.77	0.05	0.5	0.82	0	0.82	0.05	0.5	

Saisie des entretiens dans Nvivo (47 variables / 2500 verbatims)

Extraction : tableau du nombre de verbatims par variable

Transformation des données

Configuration retenue

Choix des solutions

Interprétation des résultats

Retour aux cas

@brionsebastien

- **Calibrage** : conditions qui respectent la distribution des fréquences des verbatims.
- **Difficulté** : déterminer à partir de quel nbr de verbatims une variable bascule de part et d'autre du point d'indifférence (0,5) : ici, choix de la médiane des verbatims (cf. note méthodo – Brion (2014))

Configuration retenue : calcul du résultat

Etape 1 : identifier les cas qui partagent les mêmes configurations > Notion de table de vérité (Truth table)

Exemple simple : 8 cas, 3 conditions = $2^3 > 8$ combinaisons possibles

Cette table montre pour chaque cas si les conditions et le résultat sont au-dessus ou en-dessous du seuil d'indifférence (0,5)

C	I	M	Case	Outcome Y
1	0	0	WIER	0
1	1	1	ZUID	1
1	1	1	NOORD	1
1	1	0	LENT	1
0	1	0	WAAL	1
0	0	0	DIEF	1
1	1	1	IJSS	1
1	0	1	PERK	1
1	1	1	SIJT	1
1	1	0	SCHEL	0
1	0	1	DELFT	1
1	1	1	WEST	1
0	1	0	GOUW	0
0	1	1	BROEK	1

Truth table

C	I	M	Outcome Y	Cases
1	1	1	1	ZUID, NOORD, IJSS, SIJT, WEST
1	1	0	C	LENT, SCHEL
1	0	1	1	PERK, DELFT
1	0	0	0	WIER
0	1	1	1	BROEK
0	1	0	C	WAAL, GOUW
0	0	1	---	---
0	0	0	1	DIEF

Adapté de Verweij et al., 2013

Saisie des entretiens dans Nvivo (47 variables / 2500 verbatims)

Extraction : tableau du nombre de verbatims par variable

Transformation des données

Configuration retenue

Choix des solutions

Interprétation des résultats

Retour aux cas

@brionsebastien

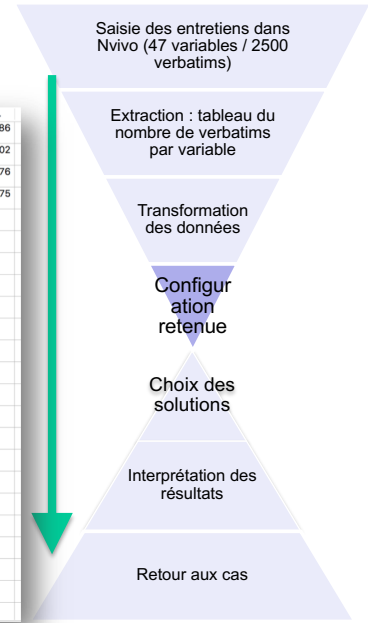
Configuration retenue : calcul du résultat

Etape 1 : Truth table de l'étude des tensions SUI

8 conditions = $2^8 > 256$ combinaisons possibles !

256 lignes !

b2+Fuzz	b3+Fuzz	I1-Fuzz	I1+Fuzz	o4-Fuzz	o10+Fuzz	p3-Fuzz	p4-Fuzz	number	perfgFuzz	raw consist.
1	1	0	1	0	1	0	0	1 (25%)		0.975886
1	1	0	0	1	1	1	1	1 (50%)		0.919802
0	1	1	0	0	0	0	0	1 (75%)		0.91376
1	1	0	0	1	0	1	0	1 (100%)		0.725775
0	0	0	0	0	0	0	0	0 (100%)		
1	0	0	0	0	0	0	0	0 (100%)		
0	1	0	0	0	0	0	0	0 (100%)		
1	1	0	0	0	0	0	0	0 (100%)		
0	0	1	0	0	0	0	0	0 (100%)		
1	0	1	0	0	0	0	0	0 (100%)		
1	1	1	0	0	0	0	0	0 (100%)		
0	0	0	1	0	0	0	0	0 (100%)		
1	0	0	1	0	0	0	0	0 (100%)		
1	1	1	1	0	0	0	0	0 (100%)		
0	0	0	1	0	0	0	0	0 (100%)		
1	0	0	1	1	0	0	0	0 (100%)		
0	1	1	1	0	0	0	0	0 (100%)		
1	1	1	1	1	0	0	0	0 (100%)		
0	0	0	0	1	0	0	0	0 (100%)		
1	0	0	0	1	0	0	0	0 (100%)		
0	1	0	0	0	1	0	0	0 (100%)		
0	1	0	0	0	1	0	0	0 (100%)		
1	1	1	0	0	1	0	0	0 (100%)		
1	1	1	1	1	0	0	0	0 (100%)		
0	0	0	0	0	1	0	0	0 (100%)		
1	0	0	0	1	0	0	0	0 (100%)		
0	1	0	0	0	1	0	0	0 (100%)		
1	1	0	0	1	0	0	0	0 (100%)		



@brionsebastien

Configuration retenue : calcul du résultat

Etape 2 : recherche de solutions > Principe de Minimisation

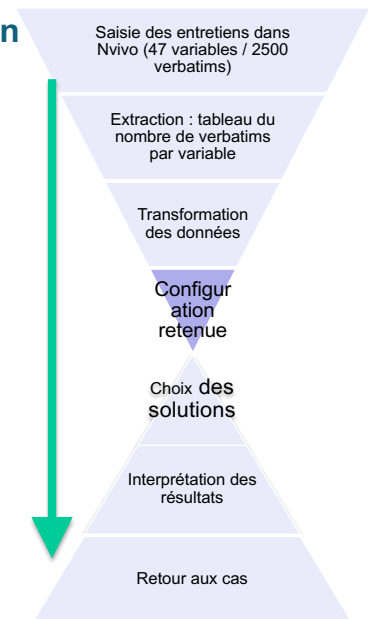
La minimisation consiste à simplifier les combinaisons au regard des principes de l'algèbre booléen (commutativité, distributivité, associativité...)

$$\overline{C} \cdot I \cdot M + C \cdot \overline{I} \cdot M + \overline{C} \cdot I \cdot \overline{M} + \overline{C} \cdot \overline{I} \cdot \overline{M} \rightarrow Y$$

$$\overline{C} \cdot M + I \cdot M + \overline{C} \cdot \overline{I} \cdot \overline{M} \rightarrow Y$$

C	I	M	Outcome Y	Cases
1	1	1	1	ZUID, NOORD, IJSS, SIJT, WEST
1	1	0	C	LENT, SCHEL
1	0	1	1	PERK, DELFT
1	0	0	0	WIER
0	1	1	1	BROEK
0	1	0	C	WAAL, GOUW
0	0	1	---	---
0	0	0	1	DIEF

Verweij et al., 2013



@brionsebastien

Configuration retenue : calcul du résultat

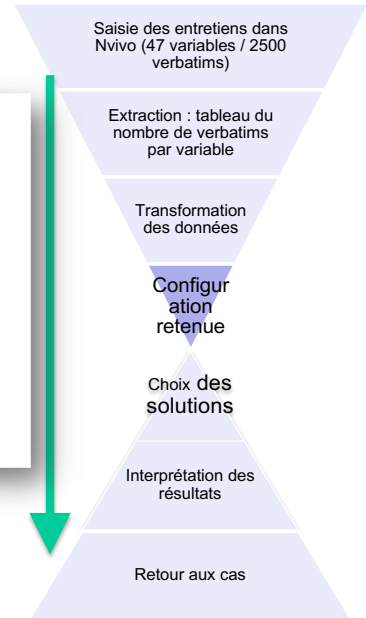
Etape 2 : Minimisation de l'étude > résultats bruts

```

Algorithm: Quine-McCluskey
--- PARSIMONIOUS SOLUTION ---
frequency cutoff: 1
consistency cutoff: 0.891037
*****
*TRUTH TABLE ANALYSIS*
*****
File: /Users/sbrico/Dropbox/recherche/Papiers en cours/QCA/Fuzz44.csv
Model: perfgFuzz = f(b2+Fuzz, b3+Fuzz, l1-Fuzz, l1+Fuzz, o4-Fuzz, o10+Fuzz, p3-Fuzz, p4+Fuzz)
Algorithm: Quine-McCluskey
--- INTERMEDIATE SOLUTION ---
frequency cutoff: 1
consistency cutoff: 0.891037
Assumptions:

```

	raw coverage	unique coverage	consistency
b2+Fuzz*b3+Fuzz*l1+Fuzz*-o4-Fuzz*o10+Fuzz*-p3-Fuzz*p4+Fuzz	0.181034	0.109385	0.98032
-b2+Fuzz*b3+Fuzz*l1-Fuzz*-l1+Fuzz*-o4-Fuzz*-o10+Fuzz*-p3-Fuzz*p4+Fuzz	0.122089	0.0528993	0.92179
b2+Fuzz*b3+Fuzz*-l1-Fuzz*-l1+Fuzz*o4-Fuzz*o10+Fuzz*p3-Fuzz*-p4+Fuzz	0.0986668	0.0476351	0.919132
b2+Fuzz*b3+Fuzz*-l1-Fuzz*-l1+Fuzz*o4-Fuzz*-o10+Fuzz*p3-Fuzz*p4+Fuzz	0.0709876	0.022415	0.891037
solution coverage: 0.314308			
solution consistency: 0.938905			



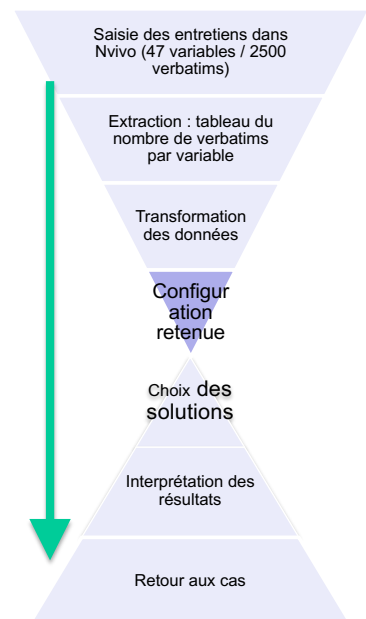
@brionsebastien

Configuration retenue : calcul du résultat

Etape 2 : Minimisation de l'étude > résultats mis en forme

Configurations for achieving perceived performance				
Conditions	Solution 1	Solution 2	Solution 3	Solution 4
Belonging				
B2- : Team cohesiveness	●	⊗	●	●
B3+ : Team diversity	●	●	●	●
Learning				
L1- : Use internal K	○	●	⊗	⊗
L1+ : Use external K	●	⊗	⊗	⊗
Structuring				
O10+ : structural separation	●	⊗	●	⊗
O4- : high frequency control	⊗	⊗	●	●
Performing				
P3- : Use short term perf	⊗	⊗	●	●
P4+ : Multidimensionnal perf	●	●	⊗	●
Raw Cov.	0.181	0.122	0.098	0.071
Uni. Cov	0.109	0.053	0.048	0.022
Consistency	0.980	0.921	0.919	0.891
Cases	3	2	1	1
Solution coverage	0.314			
Solution consistency	0.939			

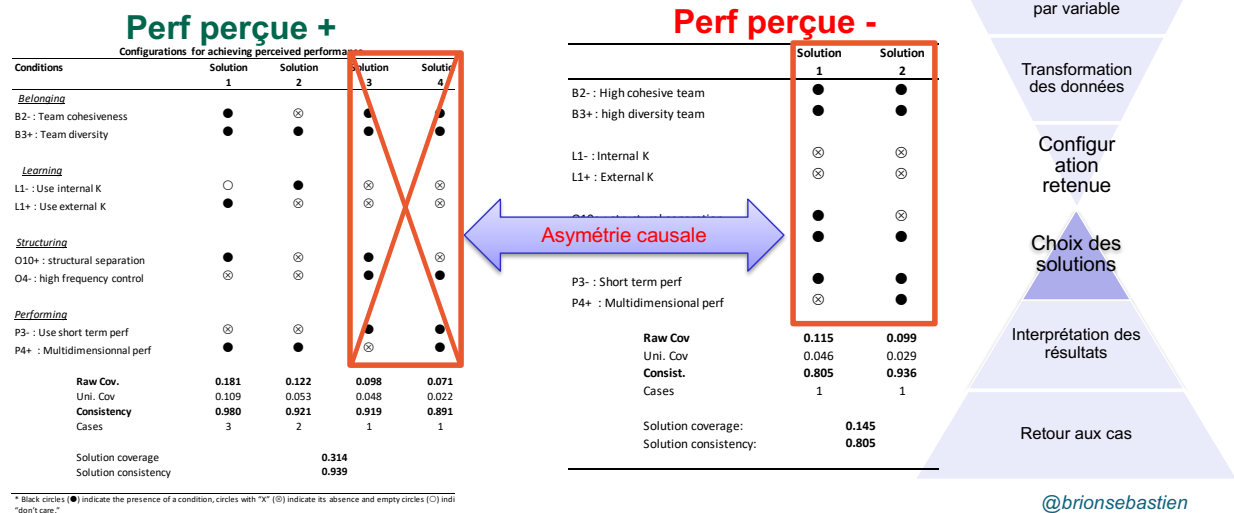
* Black circles (●) indicate the presence of a condition, circles with "X" (⊗) indicate its absence and empty circles (○) indicate "don't care."



@brionsebastien

Choix des solutions

Etape 3 : vérification que des configurations correspondantes à une faible performance perçue (*counter factual*) ne sont pas identiques ou proches de celles perçues comme élevées.

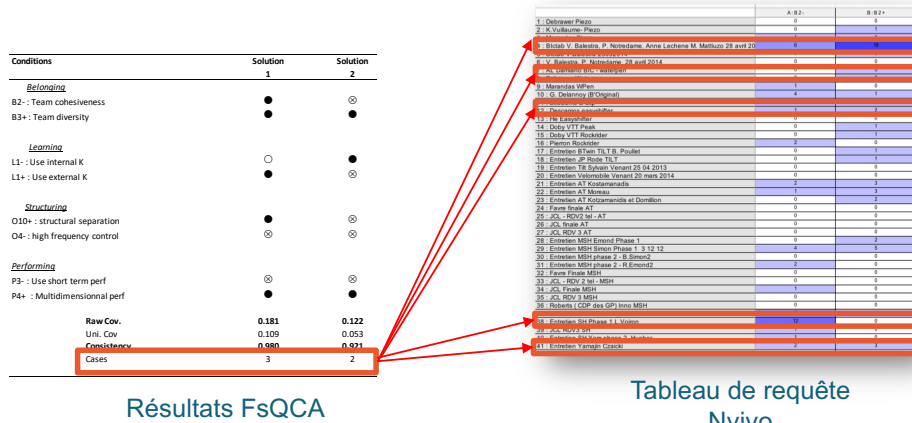


@brionsebastien

Interprétation des résultats / retour aux cas

Mieux comprendre la signification des solutions à l'aide de Nvivo...

FsQCA permet d'identifier les cas concernés par le résultat



@brionsebastien

Interprétation des résultats / retour aux cas

Mieux comprendre la signification des solutions à l'aide de Nvivo...

	A:R2	B:R2+
1. Debrauer Papp	0	0
2. K. Siskione - Papp	0	0
3. H. K. Siskione - Papp	0	0
4. H. K. Siskione - Papp	0	0
5. H. K. Siskione - Papp	0	0
6. H. K. Siskione - Papp	0	0
7. H. K. Siskione - Papp	0	0
8. H. K. Siskione - Papp	0	0
9. H. K. Siskione - Papp	0	0
10. H. K. Siskione - Papp	0	0
11. H. K. Siskione - Papp	0	0
12. H. K. Siskione - Papp	0	0
13. H. K. Siskione - Papp	0	0
14. H. K. Siskione - Papp	0	0
15. H. K. Siskione - Papp	0	0
16. H. K. Siskione - Papp	0	0
17. H. K. Siskione - Papp	0	0
18. H. K. Siskione - Papp	0	0
19. H. K. Siskione - Papp	0	0
20. H. K. Siskione - Papp	0	0
21. H. K. Siskione - Papp	0	0
22. H. K. Siskione - Papp	0	0
23. H. K. Siskione - Papp	0	0
24. H. K. Siskione - Papp	0	0
25. H. K. Siskione - Papp	0	0
26. H. K. Siskione - Papp	0	0
27. H. K. Siskione - Papp	0	0
28. H. K. Siskione - Papp	0	0
29. H. K. Siskione - Papp	0	0
30. H. K. Siskione - Papp	0	0
31. H. K. Siskione - Papp	0	0
32. H. K. Siskione - Papp	0	0
33. H. K. Siskione - Papp	0	0
34. H. K. Siskione - Papp	0	0
35. H. K. Siskione - Papp	0	0
36. H. K. Siskione - Papp	0	0
37. H. K. Siskione - Papp	0	0
38. H. K. Siskione - Papp	0	0
39. H. K. Siskione - Papp	0	0
40. H. K. Siskione - Papp	0	0
41. H. K. Siskione - Papp	0	0

Le tableau de requête Nvivo permet d'afficher la liste des verbatims correspondants aux conditions des configurations

Sélectionner les plus emblématiques permet de donner du sens aux conditions retenues

Conditions	Cases	Data code	Verbatims
Paradoxical dimensions			
Organizing			
Structural separation	Case 4 & 14	O10+	"I think if we had our offices at 10 quarters, we would be polluted by "the existing". Here it is scheduled and it is a little "out of time", there is higher chances that we create disruptions."
Frequent milestone (control)	Case 14 & 32	O4-	"We have 25 meetings, very detailed, which take the form of "micro-councils". So of course, we discuss the strategy of our projects very often. If anything goes off the road, we are quickly notified."
Belonging			
High identifiers of the members of the team	Case 4 & 14 & 23	B2-	"It was a virtuous circle that has gradually emerged, a team spirit has emerged thanks to a strong feeling between us. I analyze it as a common passion. We all have a shared belief in the product."
Team profile and knowledge diversity	Case 4 & 14 & 23	B3+	"In the project we tried to have some diversity in the profiles because we know it will generate lots of ideas."
Learning			
Use of internal knowledge base	Case 22 & 23	L1-	"In terms of production quality purchase, I tapped into internal networks that helped us, people who keep helping us also to work on quality. Finally, I also tap into support functions such as legal or communication."
Use of external new knowledge	Case 14 & 23	L1+	"We looked for partners, especially in the middle of the kite surfing. Companies who had the technology that seemed suited to our product."
Performing			
The results of the projects have to come in the short term	Case 4 & 14	P3-	"We have to manage to balance between the short and long term, because ultimately in the project, the short term will have to finance the long term."
Multi-criteria dimensions were used to evaluate performance (not only financial)	Case 14 & 23 & 32	P4+	"Innovation is so far from what we are used to doing in the core business. With top management we try to find metrics to monitor that are more suited to the life cycles of these new products, which happen to be very different and quite longer."

Saisie des entretiens dans Nvivo (47 variables / 2500 verbatims)

Extraction : tableau du nombre de verbatims par variable

Transformation des données

Configuration retenue

Itérations

Choix des solutions

Interprétation des résultats

Retour aux cas

@brionsebastien

Les avantages

Validité de la méthode :

- Méthode systématique (scientifique ?) capable d'explorer les relations causales qualitatives complexes avec des indicateurs de robustesse (consistency & coverage)
- Certaines combinaisons de variables explicatives proposées par FsQCA ne sont pas visibles dans les méthodes basées sur l'algèbre linéaire :
 - Les analyses basées sur une loi normale écartent les points extrêmes (outliers) qui nous intéressent ici
 - La supermodularité basée sur la théorie de la complémentarité (Milgrom & Robert, 1990 ; 1995) pourrait être assimilée à la FsQCA, ne permet pas d'explorer les combinaisons de variables, car nécessite de spécifier à l'avance ces combinaisons dans un modèle.

@brionsebastien

4. Avantages et limites

@brionsebastien

Les limites

- **Limites**

- Analyse uniquement sur une **seule variable expliquée** à la fois
- Analyse sensible aux particularités des cas (ajouter ou supprimer un cas peut changer fortement les résultats (Goldthorpe, 1997))
- Analyse sensible au choix du *cross-over point* (point d'indifférence)
- Nombre de variables explicatives limitées : configurations possibles = 2^c , **Max 7/8 variables** (*omitted variable bias*, Seawright, 2005)
- **Même avec une bonne connaissance préalable des cas, l'interprétation** des configurations est parfois difficile
- **Risque d'erreur de mesure de FsQCA** – Critiques provenant de spécialistes des proba, pas des théoriciens des ensembles...
- Si l'analyse est alimentée par des verbatims, **temps de traitement long** = codage des textes + calibrage des variables + interprétation des configurations + mise en forme des résultats bruts, retour sur les cas (cf. Fiss, 2007)

@brionsebastien

Conclusion

- Un outil qui ouvre de nouvelles perspectives d'analyse (« faire parler » de vieilles données qualitatives ?)
- Des logiciels libres en GUI, multi-OS (FsQCA, Module QCA dans R,...)
- Un potentiel élevé pour des publications de haut niveau
- La mise en œuvre nécessite un apprentissage long et l'usage de plusieurs logiciels (parfois capricieux)
- Coût d'entrée et temps de traitement élevés pour des résultats parfois décevants

@brionsebastien

Bibliographie

- Brion, S. & B. Chollet, "Organizational tensions in dedicated innovation projects: a fuzzy set configurational approach", 6th AIMS Innovation Workshop, Strasbourg, September 17-18, 2015.
- Duşa, Adrian, and Alrik Thiem. "Enhancing the Minimization of Boolean and Multivalued Output Functions With e QMC." *The Journal of Mathematical Sociology*, 39.2 (2015): 92-108.
- Fiss, P. C. 2007. A set-theoretic approach to organizational configurations. *Academy of Management Review*, 32: 1180–1198.
- Fiss, P. C. (2011). Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of Management Journal*, 54(2), 393-420.
- Goldthorpe J., « Current Issues in Comparative Macrosociology: a Debate on Methodological Issues », in *Comparative Social Research*, vol. 16, 1997, p. 1-26.
- Kvist, J. (2007). "Fuzzy set ideal type analysis." *Journal of Business Research*, 60(5): 474-481.

@brionsebastien

Bibliographie

- Ragin, C. C. 1987. *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.
- Ragin, C. C. 2000. *Fuzzy set social science*. Chicago: University of Chicago Press.
- Ragin, C. C. 2006. Set relations in social research: Evaluating their consistency and courage. *Political Analysis*, 14: 291–310.
- Ragin, C. C. 2008. *Redesigning social inquiry: Fuzzy sets and beyond*. Chicago: University of Chicago Press.
- Ragin, C. C., & Fiss, P.C. 2008. Net effects analysis versus configurational analysis: An empirical demonstration. In C. C. Ragin (Ed.), *Redesigning social inquiry: Fuzzy sets and beyond*: 190–212. Chicago: University of Chicago Press.
- Rihoux, B., Alamos, P., Bol, D., Marx, A. and I. Rezsóhazy (2013). "From niche to mainstream method? A comprehensive mapping of QCA applications in journal articles from 1984 to 2011", *Political Research Quarterly* 66(1): 175-184.
- Seawright J., « Qualitative Comparative Analysis (QCA) vis-à-vis Regression », *Studies in Comparative International Development*, vol. 40, n° 1, 2005,
- Schneider, C.Q., and C. Wagemann (2012). *Set-Theoretic Methods for the Social Sciences. A Guide to Qualitative Comparative Analysis*. New York: Cambridge University Press.
- Schneider, C. Q., and C. Wagemann (2012). „Notions and operations in set theory“. *Set-Theoretic Methods for the Social Sciences. A Guide to Qualitative Comparative Analysis*. New York: Cambridge University Press, 42-56.
- Thiem, A., & Dusa, A. (2013). "QCA: A package for qualitative comparative analysis", *The R Journal*, 5(1): 1-11.
- Verweij, S. (2015). *Once the shovel hits the ground: Evaluating the management of complex implementation processes of public-private partnership infrastructure projects with qualitative comparative analysis*. Rotterdam: Erasmus University Rotterdam.
- Verweij, S. & Gerrits, L.M. (2013). Understanding and researching complexity with qualitative comparative analysis: Evaluating transportation infrastructure projects. *Evaluation*, 19(1), 40-55.

@brionsebastien