

# EGC 2019

du 21 au 25 janvier 2019  
à Metz

**19ème CONFERENCE FRANCOPHONE**

**SUR L'EXTRACTION  
ET LA GESTION DES CONNAISSANCES**



## Sur l'intérêt de l'analyse de la dynamique des processus de classement : un retour d'expériences avec la Q méthode

Claire Gauzente, Pascale Kuntz, Yves Roy, Aurélien Milliat



Lemna



# Motivation

## S5A : Apprentissage

(...)

C : 6 Wissam Siblini, Frank Meyer, Pascale Kuntz  
*CRAFTML, une forêt aléatoire efficace pour l'apprentissage multi-label extrême (présenté à ICML 2018)*

Dataset	$d_x$	$d_y$	$n$
Mediamill	120	101	30993
Bibtex	1836	159	4880
Delicious	500	983	12920
RCV1-2K	47236	2456	623847
EURLex-4K	5000	3993	15539
AmazonCat-13K	203882	13330	1186239
AmazonCat-14K	597540	14588	4398050
Wiki10-31K	101938	30938	14146
Delicious-200K	782585	205443	196606
WikiLSHTC-325K	1617899	325056	1778351
Wikipedia-500K	2381304	501070	1813391
Amazon-670K	135909	670091	490449
Ads-1M	164592	1082898	3917928
Amazon-3M	337067	2812281	1717899
Ads-9M	2082698	8838461	70455530

XML Repository

## S5B : Données dynamiques

L : 17 Claire Gauzente, Pascale Kuntz, Aurélien Milliat, Yves Roy  
*Sur l'intérêt de l'analyse de la dynamique des processus de classement. Un retour d'expérience avec la Q méthode (...)*



Quelques individus

Ramlon S.E. (2011). Q methodology and its position in the mixed-methods continuum. *Operant Subjectivity* 34(3).

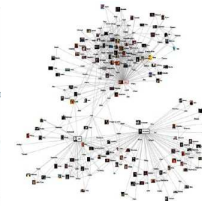
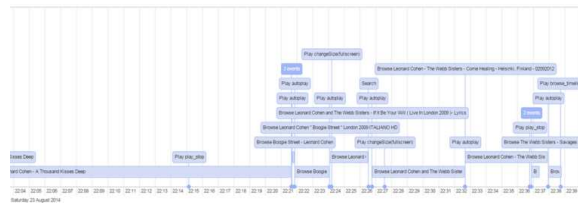


Combiner différentes échelles

# Vers une science humaine « intégrative »

« Internet Studies » : Il ne s'agit pas « d'oublier [...] l'épaisseur sociale des utilisateurs »  
(Denouël et Granjon, 2011)

- Traces numériques
- Enquêtes
- Entretiens ethnographiques
- (...)



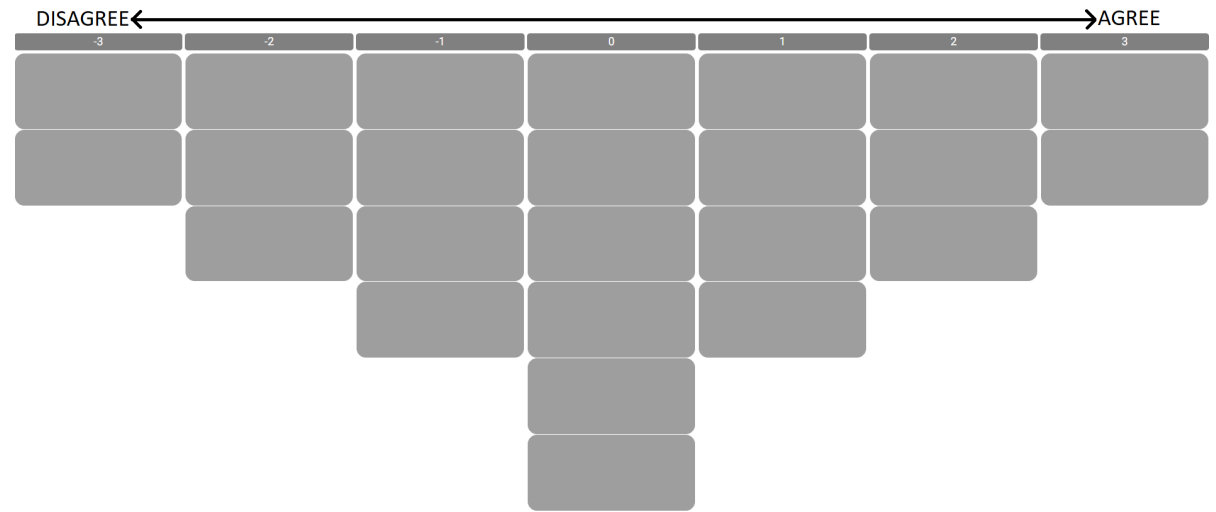
Date	Titre de l'article	Journal	Page
2004	...	...	...
2005	...	...	...
2006	...	...	...
2007	...	...	...
2008	...	...	...
2009	...	...	...
2010	...	...	...
2011	...	...	...
2012	...	...	...
2013	...	...	...
2014	...	...	...
2015	...	...	...
2016	...	...	...
2017	...	...	...
2018	...	...	...
2019	...	...	...
2020	...	...	...
2021	...	...	...
2022	...	...	...
2023	...	...	...
2024	...	...	...



Combiner différentes échelles

# Q method briefly described 1

- Typically, in a Q methodological study people are presented with a sample of statements about some topic.
- Respondents are asked to rank-order the statements from their individual point of view, according to some preference, judgement or feeling about them, mostly using a quasi-normal distribution.
- By Q sorting people give their subjective meaning to the statements.



# Q method briefly described 2

## Experimental framework

24 statements -from a study on augmented reality- (S. Gautthier thesis)

Utiliser la réalité augmentée (RA) avec une webcam, c'est trop compliqué. Avec un téléphone, ça va.

Je ne comprends pas comment ça marche, c'est trop compliqué. Je n'ai pas envie d'essayer.

La RA, c'est pas surprenant. J'ai déjà vu des choses comme ça avant.

L'effet d'étonnement ne va pas durer longtemps.

Ça ne rime à rien, c'est totalement inutile.

C'est mieux d'aller dans les magasins plutôt que de vivre derrière son écran et d'essayer des choses avec la RA.

Je n'utiliserais la RA qu'exceptionnellement, si je n'avais pas de temps pour aller en magasin.

C'est pas intéressant pour voir des objets réels, mais pour visualiser comment certaines situations pourraient évoluer (notre apparence physique, un lieu, une maladie...)

C'est bien seulement pour attirer l'attention. Etc ...

Statements are randomly displayed on the screen

### 2 stages

1. 3 classes: *disagree*, *neutral*, *agree*
2. q-sort with the Gaussian distribution: from -3 « *strongly disagree* » to +3 « *strongly agree* »

# Q method briefly described 3

PCA + rotations

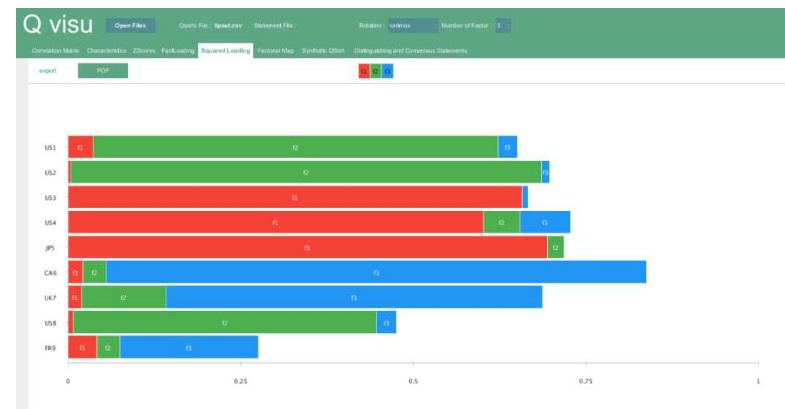
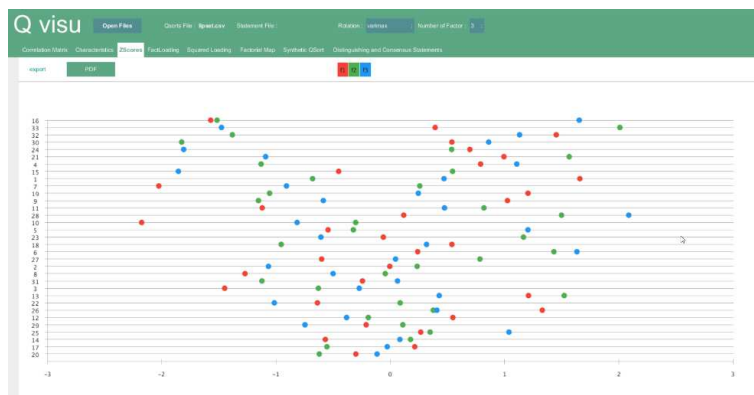
Q visu

Open Files   Qsorts File: ipset.csv   Statement File:   Rotation: varimax   Number of Factor: 3

Correlation Matrix   Characteristics   ZScores   FacLoadings   Squared Loading   Factorial Map   Synthetic QSort   Distinguishing and Consensus Statements

export   PDF

	Average reliability coefficient	Number of loading Q-sorts	Eigenvalues	Percentage of explained variance	Composite reliability	Standard error of factor scores	Sum of explained variance
f1	0,80	3,00	2,09	23,17	0,92	0,28	23,17
f2	0,80	3,00	1,97	21,93	0,92	0,28	45,09
f3	0,80	3,00	1,68	18,68	0,92	0,28	63,78



# Tracking the sorting dynamics – Related work

- **Opinion mining** : hesitation mining (Pathak & Towari, 2016 ; Lu & al, 2007)
- **Multicriteria decision making** : hesitation modeling (Roy, 1988 ; Boyssou, 1989)
- **Survey analysis** (CAMS : Cognitive Aspect of Survey Methodology): response times (Tourangeau, 2003)
- **Behavioral psychology** : understanding mind in motion (Freeman & al, 2011 ; Hehman & al, 2014)
- **Sensory data analysis** : identifying different sorting processes (Cadoret et al. 2011 ; Lê et al., 2016)

T.M. Lê, T.M, F. Husson, S. Lê (2016). *Digital-tracking: interpreting the evolution over time of sensory Dimensions of an individual product space issued from mapping and sorted mapping*, Food Quality and Preferences.

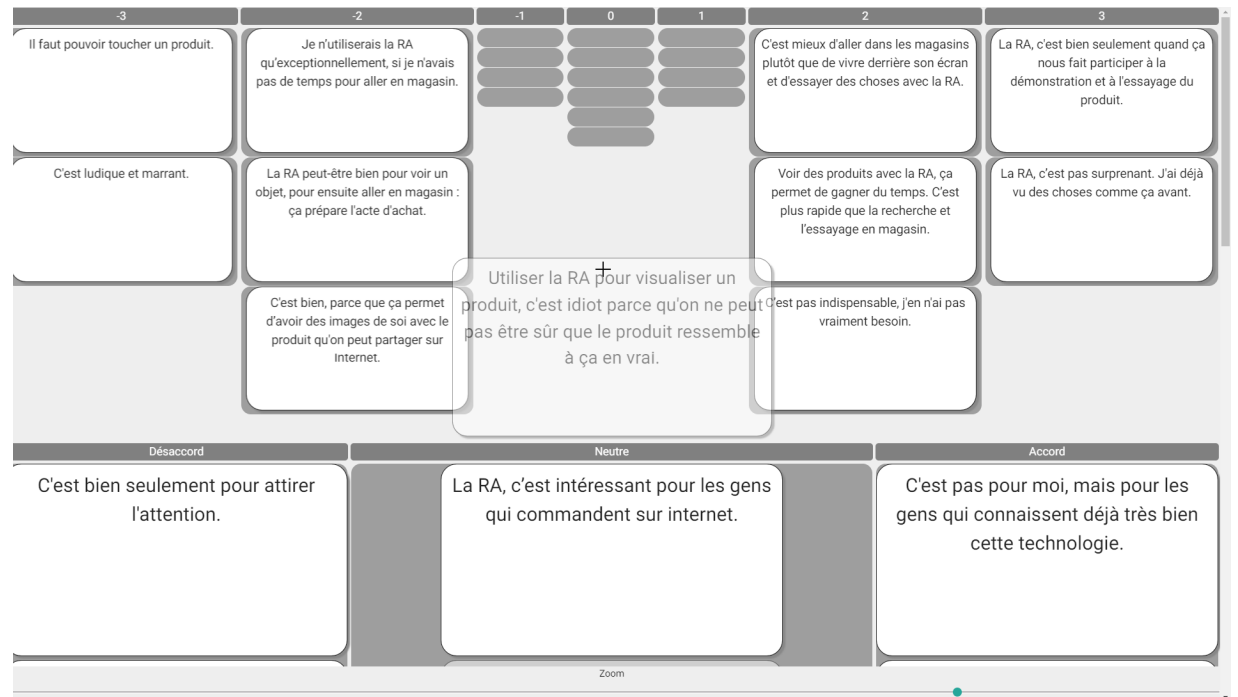
# Tracking the sorting dynamics of the Q sort

## Events :

- mouse click
  - mouse move
  - drag
  - drop
  - scrolling
  - zooming
  - key press
- etc ...

x-y coordinates of each move

**Time** of each event in milliseconds

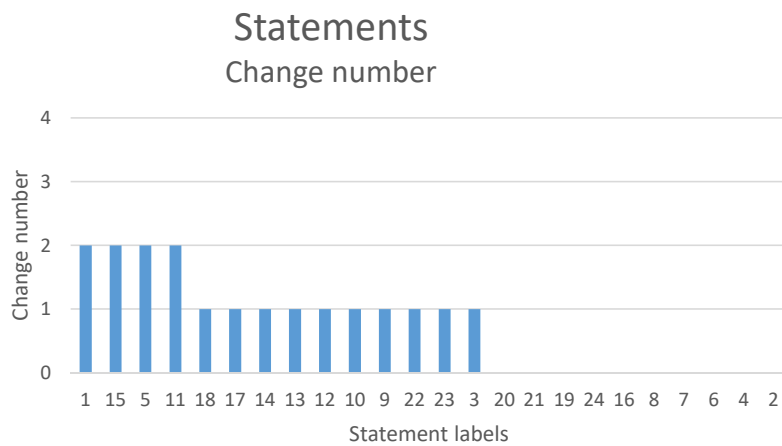


**Q-connect**

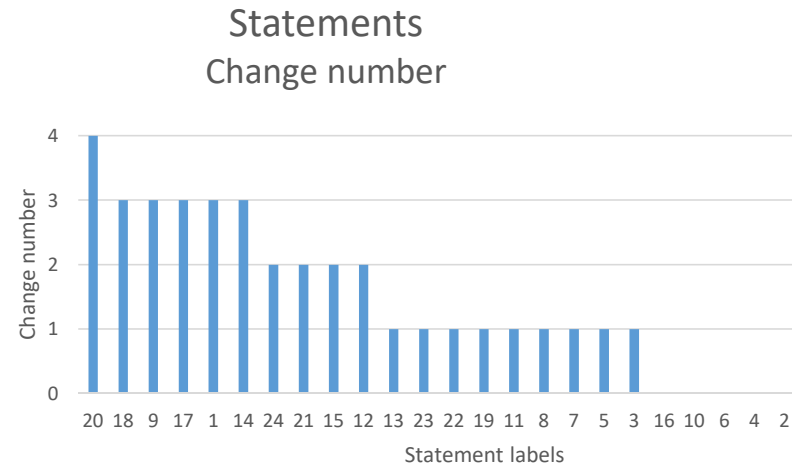


# Exploration of decision evolution 1

**Question:** are there some placement changes of the statements in each sorting stage ?



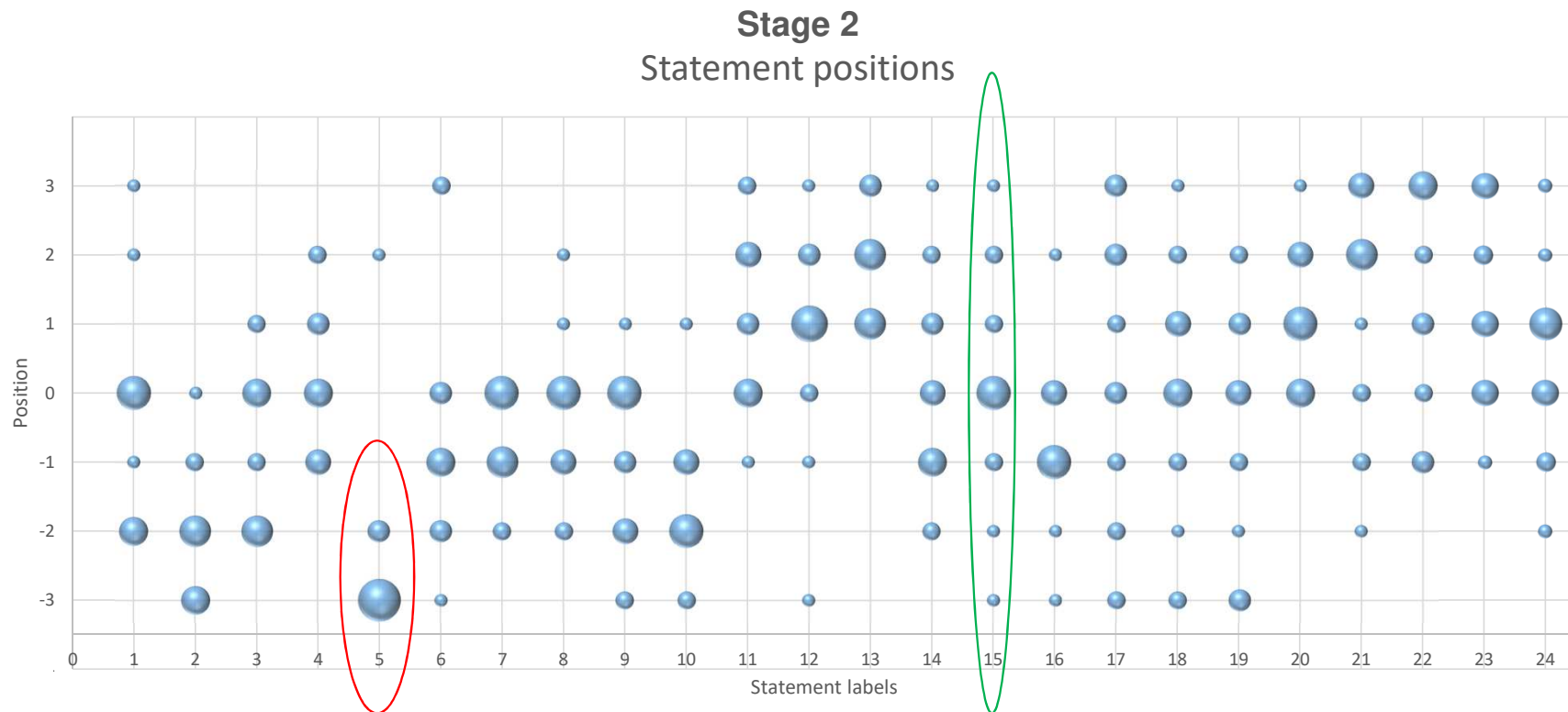
**stage 1**



**stage 2**

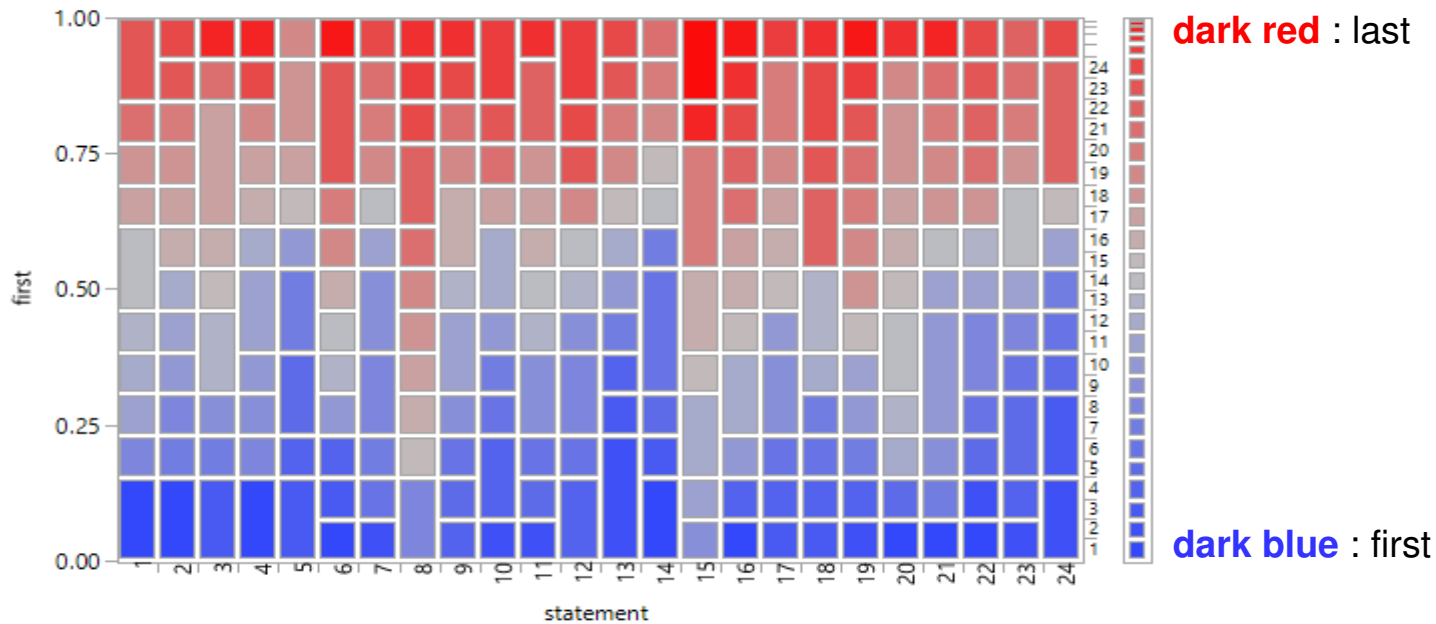
# Exploration of decision evolution 2

**Question:** are some statement placements more « stable » than others ?

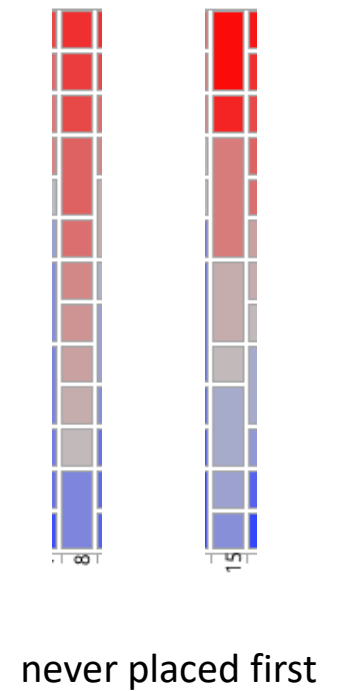


# Exploration of decision evolution 3

**Question:** are some statements placed before others ?



stage 1



never placed first

# Exploration of decision evolution 4

**Question:** what is the evolution of the « neglected » statements ?

**8** : once placed never moved

C'est pas intéressant pour voir des objets réels, mais pour visualiser comment certaines situations pourraient évoluer (notre apparence physique, un lieu, une maladie...)

**15** : has moved several times

Quand on prévisualise un produit chez soi par la RA, on n'a ni le conseil des vendeurs, ni le plaisir d'aller en magasin.

*Hypothesis* : « ambiguous » or double-barelled statements ?

# Exploration of decision evolution 5

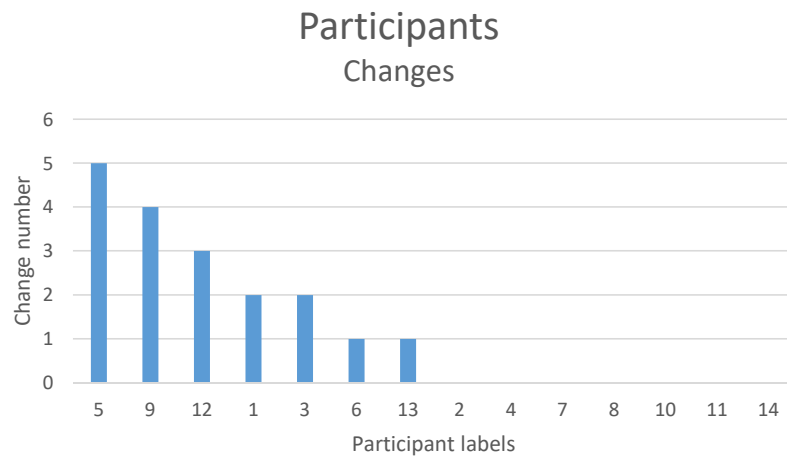
**Question:** what happens between stage 1 and stage 2 ?

		stage 2							Total	
		firstmove								
From		-3	-2	-1	0	1	2	3		
stage 1	disagree	Frequency	28	37	27	13	3	2	1	111
	neutral	Frequency	0	1	23	58	23	1	0	106
	agree	Frequency	0	1	0	10	23	35	26	95

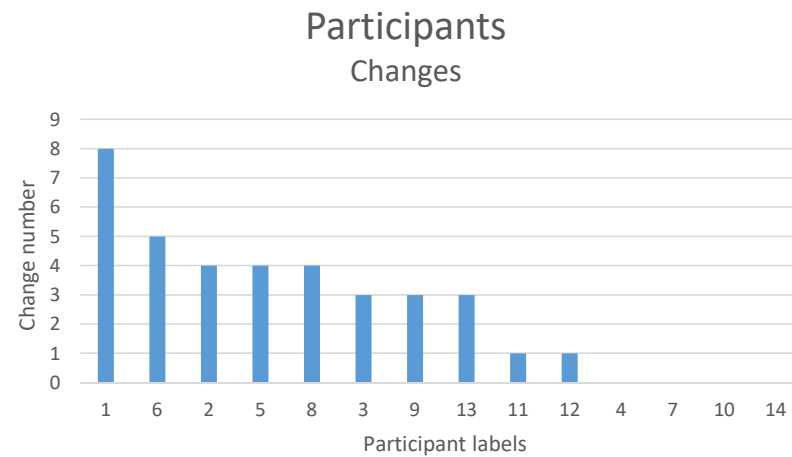
*Hypothesis* : the attitude is mainly built during stage 1

# Exploration of participant attitudes 1

**Question:** are there some differences between the participants concerning the decision changes during the sorting process ?



**stage 1**



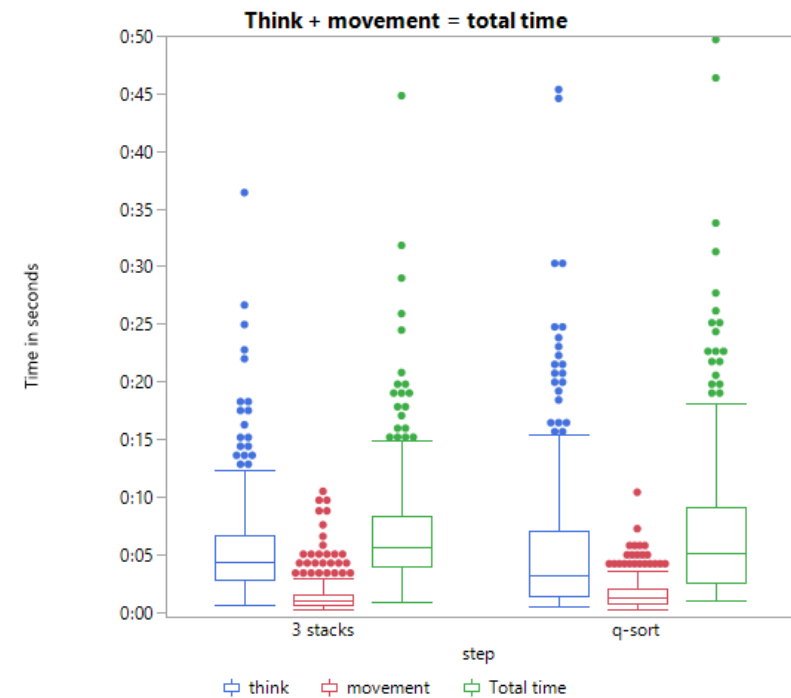
**stage 2**

# Exploration of participant attitudes 2

**Question:** are there some differences between the sorting times ?

For each **displacement**:

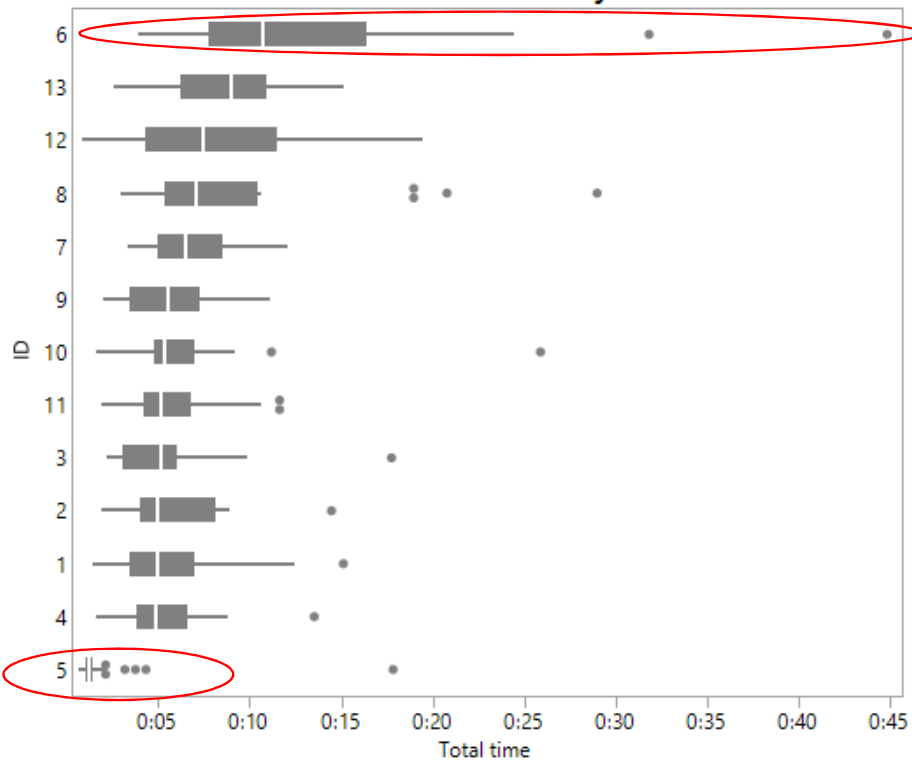
- «reflection time »:  
time measured between the last movement and the displacement start
- «displacement time »:  
time measured between the displacement start and its end
- **total time**: reflection time + displacement time



# Exploration of participant attitudes 3

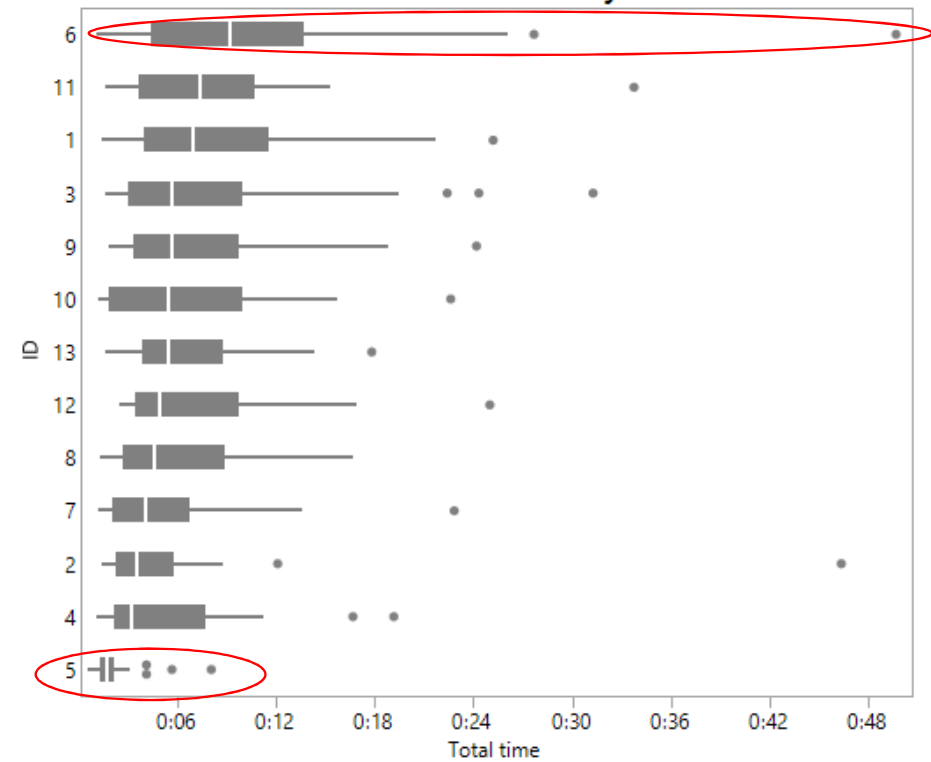
**Question:** are there some differences between the sorting times ?

Total time of movements by ID



stage 1

Total time of movements by ID

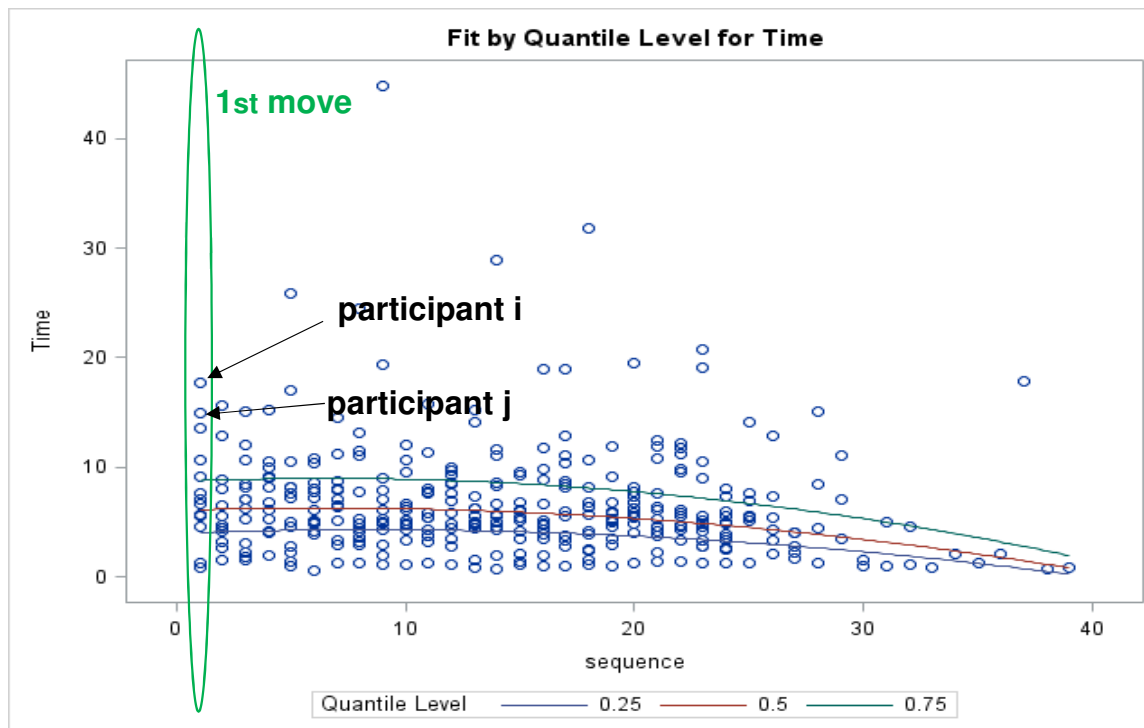


stage 2



# Exploration of participant attitudes 3

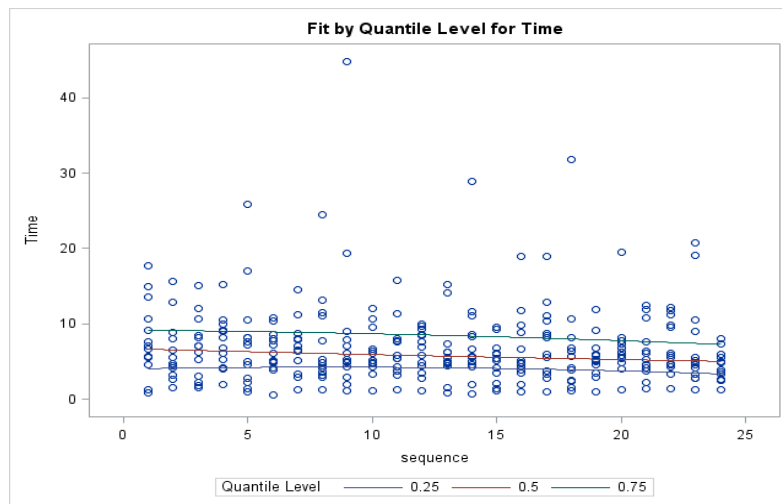
**Question:** how does the sorting behavior evolve ?



**model** : quantile regression with a polynomial of degree 2

**stage 1**

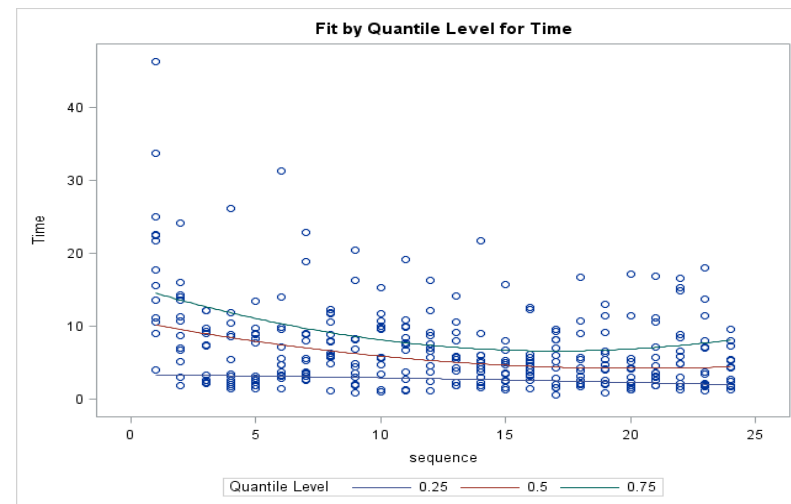
# Exploration of participant attitudes 4



**stage 1**

median stability following  
by a slight decreasing

*Hypothesis:* same attitude for each statement  
with a slight learning of the process



**stage 2**

significant median decreasing

*Hypothesis:* discovering of the sorting constraints at  
the beginning and then acceleration of the decisions  
(attitude mainly built in stage 1)

# Impact on the Q results

**Proposition:** adding weights in PCA to modulate the « outlier » impact (both statements & individuals)

## Impact with statement weighting

Eigen value#	eigen value1	% without weight Sum	Cum % without weight	eigen value2	% with 65% weight	Cum % with 65% weight	eigen value3	% with 35% weight	Cum % with 35% weight
1	18.08	47.05%	47.05%	17.99	47.61%	47.61%	17.91	48.13%	48.13%
2	5.09	13.25%	60.30%	5.05	13.36%	60.97%	5.02	13.48%	61.62%
3	3.49	9.07%	69.36%	3.43	9.08%	70.05%	3.40	9.14%	70.76%
4	2.85	7.42%	76.79%	2.73	7.24%	77.29%	2.66	7.15%	77.91%

## Statistical tests

- Weight effect (random or individual atypicality)  
p-value = 0.047 (for individual 5 weighting)

- Comparison of the impact of the pairs with individual 5 and the pairs without

## Impact with individual weighting

Number	variant									
	standard		5 at 65%		5 at 35%		5 & 6 at 65%		5 & 6 at 35%	
	Eigen value	Proportion	Eigen value	Proportion	Eigen value	Proportion	Eigen value	Proportion	Eigen value	Proportion
1	18.08	47.0%	18.01	49.0%	17.98	50.2%	17.20	49.1%	16.78	50.5%
2	5.09	13.3%	5.06	13.8%	5.05	14.1%	4.79	13.7%	4.67	14.1%
3	3.49	9.1%	3.01	8.2%	2.99	8.3%	3.01	8.6%	2.98	9.0%
4	2.85	7.4%	2.52	6.9%	2.47	6.9%	2.42	6.9%	2.30	6.9%
5	2.35	6.1%	2.03	5.5%	1.97	5.5%	1.84	5.3%	1.81	5.4%
6	1.91	5.0%	1.79	4.9%	1.64	4.6%	1.65	4.7%	1.40	4.2%
7	1.46	3.8%	1.42	3.9%	1.40	3.9%	1.40	4.0%	1.23	3.7%
8	1.25	3.3%	0.98	2.7%	0.81	2.3%	0.91	2.6%	0.79	2.4%

# Impact on the Q digital protocol 1

Numerous software and on-line services

## On-line

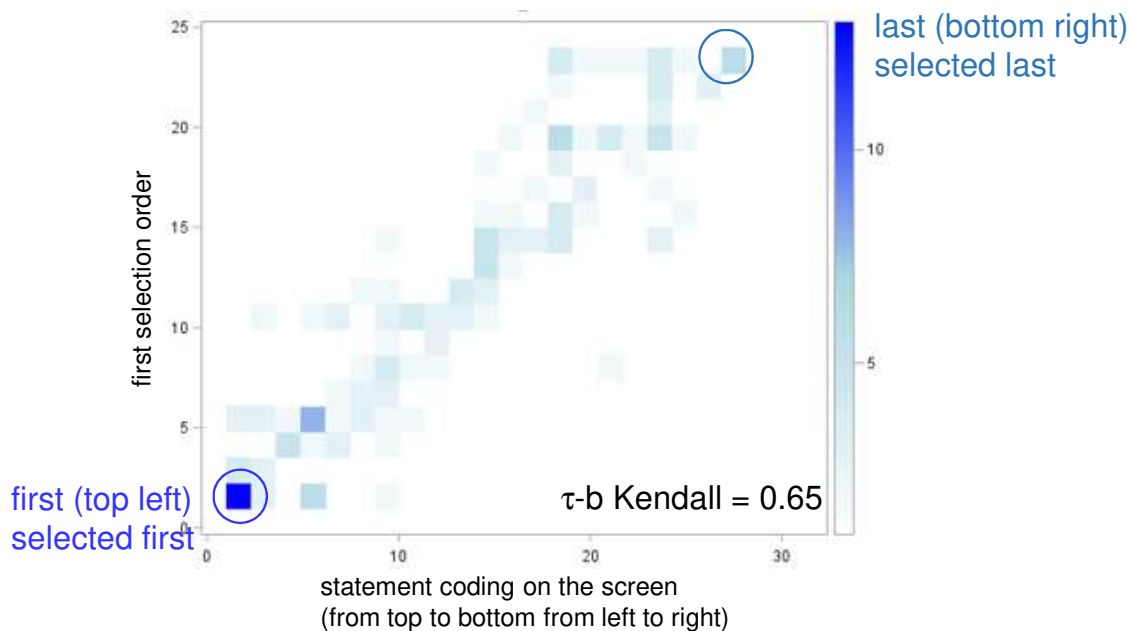
- HTML-Q  
<https://github.com/aproxima/htmlq>
- qsoftware  
<http://www.qsoftware.net>
- Ken-Q Data  
<https://shawnbanasick.github.io/ken-q-data/index.html#section1>
- Easy-HtmlQ  
[https://docs.google.com/presentation/d/1fOYxQOo2XpgR1IZ4gyGO\\_dRi9Ehh6-0TN98us2xPEPs/edit#slide=id.p](https://docs.google.com/presentation/d/1fOYxQOo2XpgR1IZ4gyGO_dRi9Ehh6-0TN98us2xPEPs/edit#slide=id.p)
- Ken-Q Analysis  
<https://shawnbanasick.github.io/ken-q-analysis/>
- Qsortouch  
<https://qsortouch.com/>
- vqmethod  
<https://www.vqmethod.com/Home>
- webQ  
<http://schmolck.org/qmethod/webq/>

## Off-line

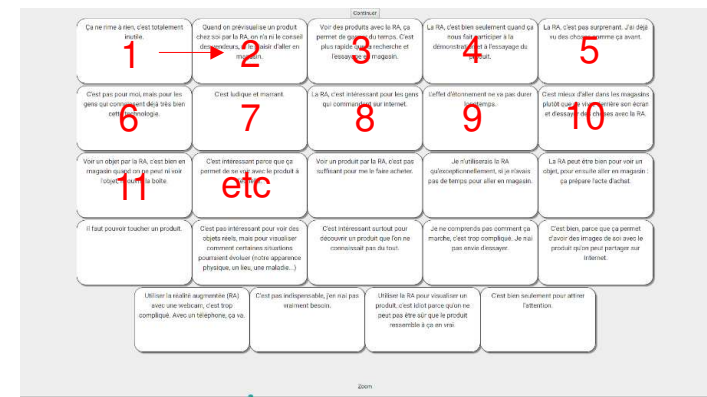
- PQMethod  
<http://schmolck.org/qmethod/#PQMethod>
- PCQ  
<http://www.pcqsoft.com/>
- Qmethod  
<https://cran.r-project.org/web/packages/qmethod/index.html>
- FlashQ  
<http://www.hackert.biz/flashq/home/>
- QFACTOR  
<https://ideas.repec.org/c/boc/bocode/s458326.html>
- QCONECRT  
<https://ideas.repec.org/c/boc/bocode/s458325.html>
- Attachment Q-Sorter  
<http://www.aqs.stoneclearing.net/>
- Lloyd's Q Sort Tool  
<http://www.nowhereroad.com/qsort/>
- rap  
<http://rap.ucr.edu/qsorter/>

# Impact on the Q digital protocol 2

**Question:** what is impact on the results of the statement order on the screen ?



**stage 1**



*Hypothesis:* no holistic vision and consequently no significant difference between a global presentation and a statement by statement presentation

# Next future ...



New survey and new traces



Understanding the behaviors



Integrating the new informations in the Q process and go beyond ...