

# A cognitive process of responding to a spatial survey

Textual analysis versus network analysis and spatial analysis  
for analysing an international corpus  
of co-citations about countries

C. GRASLAND, F. GUÉRIN-PACE,  
S. DE RUFFRAY, T. GIRAUD



# Introduction : a framework for the analysis of ordered list of places



# A very general problem ...



- Analysis of **list of items** given as answer to an **open question** is a very general problem that has been developed intensively by statisticians (Benzecri & al.), specialists of textual analysis, psychologists, sociologists, ...

*Which words would you associate to « environment » ...*

- **Ranking list of items** is also a well known problem that has been the object of many research, in the field for example of marketing, advertisement, ...

*In which order of preferences would you classify the following marks of automobiles ...*

... but with specificities when answers are ordered list of geographical objects »



- « What are the most important **cities** in global economy »
- « What are the most powerful **states** at the beginning of the XXIth century »
- « Which **touristic places** of Europe have you visited ... »

Answer of type : **Place1-Place2-Place3-Place4-...**

- **Can be located on maps**
- **Can be associated to distances**
- **Can be related to space-time process**





1. List of places in surveys : the result of a mental process ?

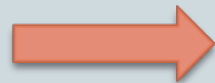


# Simple list of places

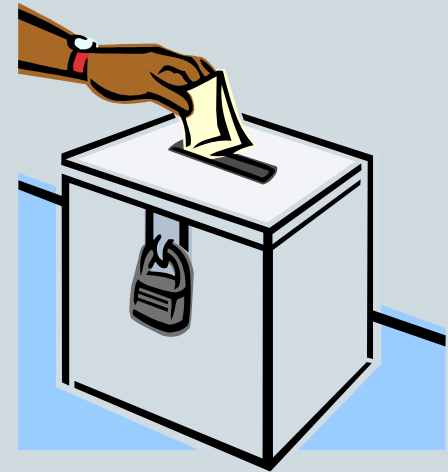
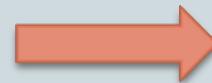
« List up 3 countries where you would like to live in the near future ?



Individual



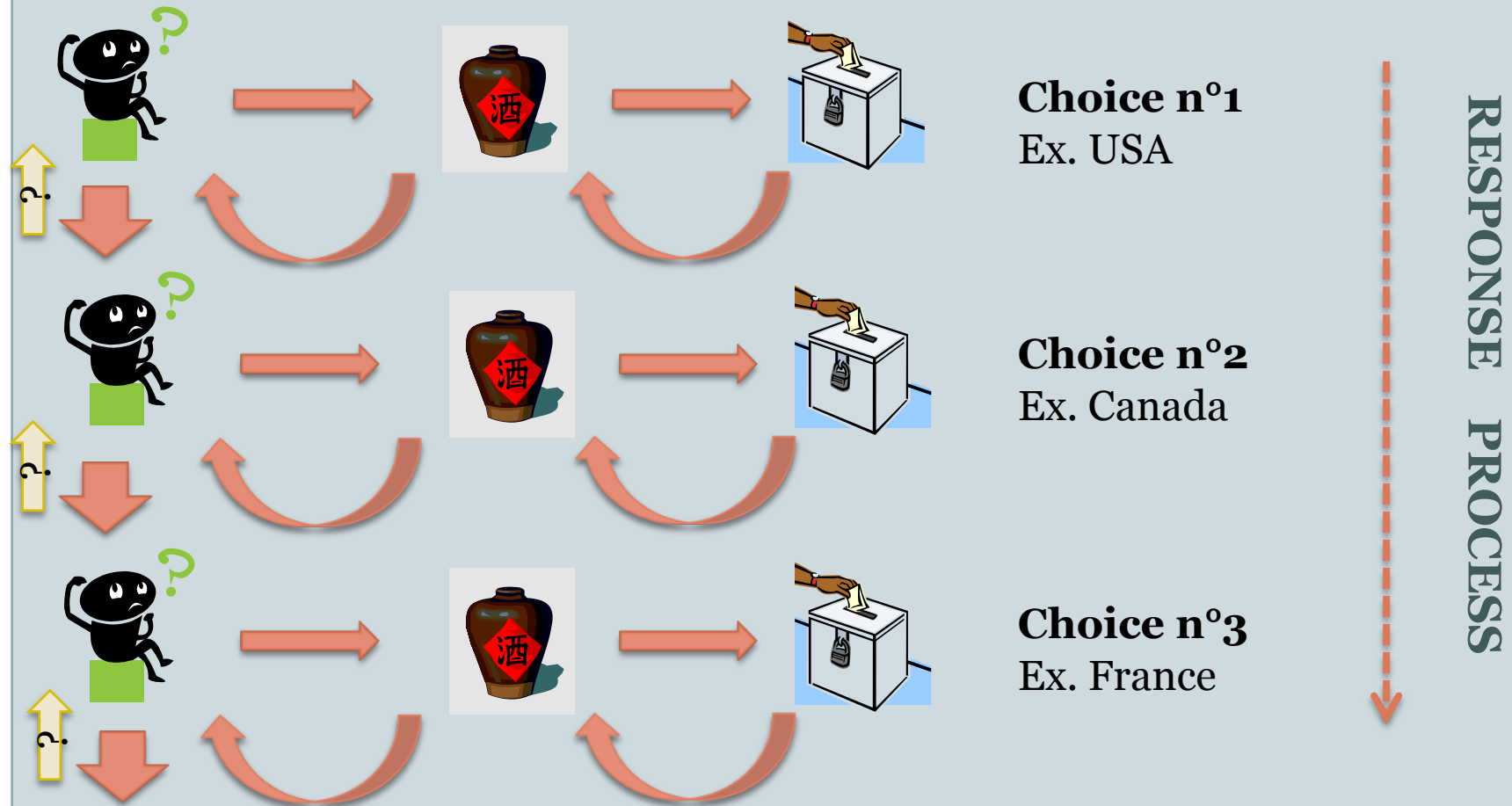
Set of possibilities



Choice

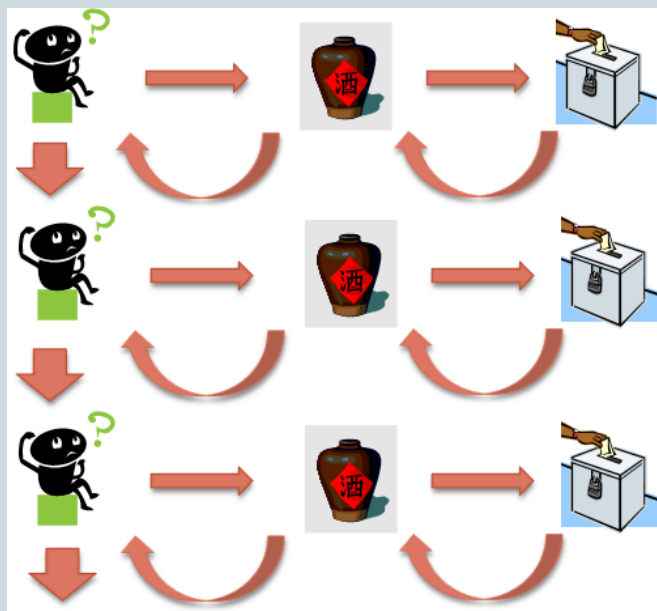


# Simple list of places



# Twin lists of places

« List up 3 countries where you would like to live in the near future and 3 others where you would not like to live »



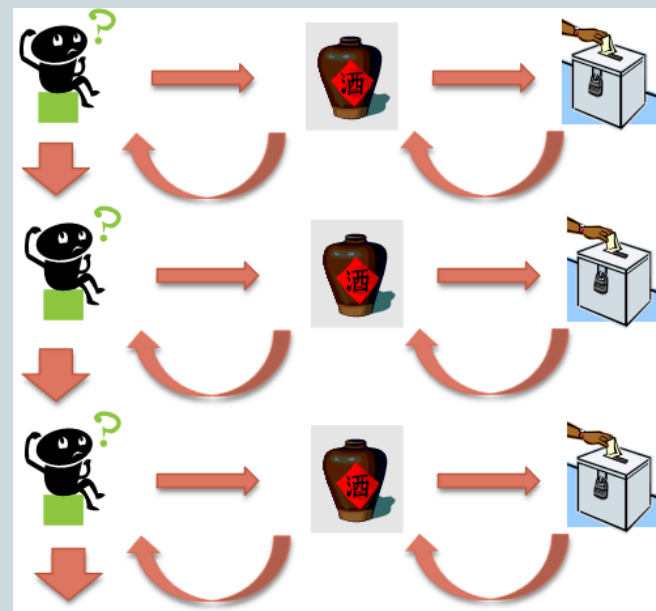
**Choice 1.1**  
Ex. USA



**Choice 1.2**  
Ex. Canada



**Choice 1.3**  
Ex. France



**Choice 2.1**  
Ex. Iran

**Choice 2.2**  
Ex. Iraq

**Choice 2.3**  
Ex. Soudan



**RESPONSE PROCESS (?)**

# Multiple list of places



## YOUR VISION OF THE WORLD

**B.1.a) With the exception of the cities of the country, or countries, where you currently have citizenship, list up to 5 cities for each of the following questions ...**

....where you would like to live in the near future?

- 1 \_\_\_\_\_
- 2 \_\_\_\_\_
- 3 \_\_\_\_\_
- 4 \_\_\_\_\_
- 5 \_\_\_\_\_

....where you would **NOT** like to live in the near future?

- 1 \_\_\_\_\_
- 2 \_\_\_\_\_
- 3 \_\_\_\_\_
- 4 \_\_\_\_\_
- 5 \_\_\_\_\_

**B.1.b) With the exception of the country or country(ies) where you currently have citizenship, list up to 5 country(ies) for each of the following questions ...**

....where you would like to live in the near future?

- 1 \_\_\_\_\_
- 2 \_\_\_\_\_
- 3 \_\_\_\_\_
- 4 \_\_\_\_\_
- 5 \_\_\_\_\_

....where you would **NOT** like to live in the near future?

- 1 \_\_\_\_\_
- 2 \_\_\_\_\_
- 3 \_\_\_\_\_
- 4 \_\_\_\_\_
- 5 \_\_\_\_\_

# Example of answers



code	city	stud	sexe	age	lik1	lik2	lik3	lik4	lik5	unl1	unl2	unl3	unl4	unl5
1	BUA	ART	F	20	USA	FRA	GBR	DEU	ITA	IRQ	PAK	NGA	MAR	CHN
2	BUA	ART	F	22	AUS	BEL	TUN	GRC	DNK	DEU	FRA	CHN	IRN	PAK
3	BUA	ART	F	23	USA	CAN	FRA	SWE	EGY	IRQ	AFG	DEU	ZAF	NGA
4	BUA	ART	F	24	USA	ZAF	GBR	SWE	NLD	FRA	CHN	NGA	GHA	RUS
5	BUA	ART	F	26	USA	IRL	DNK	SWE	BEL	IRQ	AFG	CHN	MNG	RUS
6	BUA	ART	F	30	ZAF	EGY	JPN	GHA	PAK	SAU	ETH	SDN	NGA	KWT
7	BUA	ART	M	21	USA	GBR	NGA	CAN	CHE	EGY	FRA	CHN	RUS	DEU
8	BUA	ART	M	21	USA	GBR	DNK	FRA	CAN	IRQ	PAK	AFG	IRN	VEN
9	BUA	ART	M	22	FRA	USA	GBR	ESP	ZAF	EGY	SAU	SDN	MRT	ITA
10	BUA	ART	M	23	GBR	USA	ZAF	NGA	CAN	FRA	CIV	TCD	IRQ	CHN
11	BUA	ART	M	23	NOR	GBR	USA	GHA	ESP	IRQ	FRA	CHN	ETH	EGY
12	BUA	ART	M	24	GHA	CAN	USA	JPN	GBR	FRA	TCD	IRN	IRQ	SDN

## Background

e.g. Survey realized in Buea (CMR), student in Arts, Woman, 20 years old.

## List of pull countries

e.g. USA, France, Great Britain, Germany, Italy

## List of push countries

e.g. Irak, Pakistan, Nigeria, Morroco, China

# Which types of analysis can be applied ?



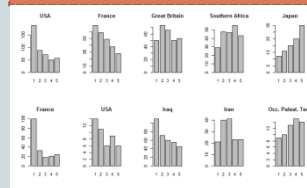
- **Statistical analysis** : list of items is a very specific type of answer where the *order of answers* does eventually matter.
- **Network Analysis** : list of items can be considered as nodes related by links where the *connexity of answers* does eventually matter.
- **Textual Analysis** : list of items can be considered as *words and sentences in a text* for which the *order of words* does matter
- **Spatial analysis** : *list of places* defines a geographical path where *distance between answers* does eventually matter.

# Outlook of the presentation

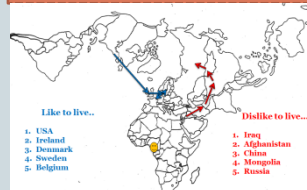
1. List of places as the result of a mental process



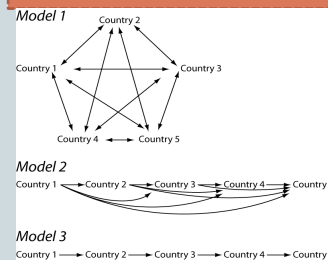
2. Testing the effect of ranks in the lists



3. Testing the effect of distances in space



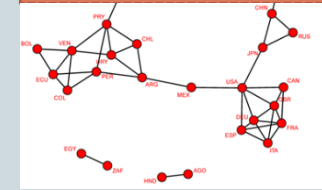
4. Selection of relevant segments



5. Textual Analysis of segments



6. Network Analysis of segments





## 2. Testing the effect of ranks on the lists of answers



# Why rank does matter ?

code	city	stud	sexe	age	lik1	lik2	lik3	lik4	lik5	unl1	unl2	unl3	unl4	unl5
1	BUA	ART	F	20	USA	FRA	GBR	DEU	ITA	IRQ	PAK	NGA	MAR	CHN
2	BUA	ART	F	22	AUS	BEL	TUN	GRC	DNK	DEU	FRA	CHN	IRN	PAK
3	BUA	ART	F	23	USA	CAN	FRA	SWE	EGY	IRQ	AFG	DEU	ZAF	NGA
4	BUA	ART	F	24	USA	ZAF	GBR	SWE	NLD	FRA	CHN	NGA	GHA	RUS
5	BUA	ART	F	26	USA	IRL	DNK	SWE	BEL	IRQ	AFG	CHN	MNG	RUS
6	BUA	ART	F	30	ZAF	EGY	JPN	GHA	PAK	SAU	ETH	SDN	NGA	KWT
7	BUA	ART	M	21	USA	GBR	NGA	CAN	CHE	EGY	FRA	CHN	RUS	DEU
8	BUA	ART	M	21	USA	GBR	DNK	FRA	CAN	IRQ	PAK	AFG	IRN	VEN
9	BUA	ART	M	22	FRA	USA	GBR	ESP	ZAF	EGY	SAU	SDN	MRT	ITA
10	BUA	ART	M	23	GBR	USA	ZAF	NGA	CAN	FRA	CIV	TCD	IRQ	CHN
11	BUA	ART	M	23	NOR	GBR	USA	GHA	ESP	IRQ	FRA	CHN	ETH	EGY
12	BUA	ART	M	24	GHA	CAN	USA	JPN	GBR	FRA	TCD	IRN	IRQ	SDN

**USA** are more frequently chosen than Sweden and *often at the first places*

**China** is chosen more frequently than France but *not in the first choices*.

# Testing random distribution of ranks



- **P** , set of students **1...i...N** each with answers provided  $R_{i,1} \dots R_{i,5}$ .
- **Q**, set of countries **1..j...k** proposed as a response by at least one student
- **F<sub>j,k</sub>** , the proportion of students who provided the answer **j** to the rank **k**.

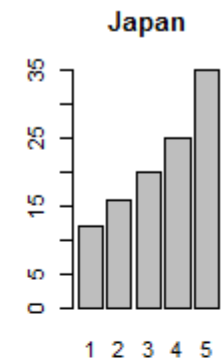
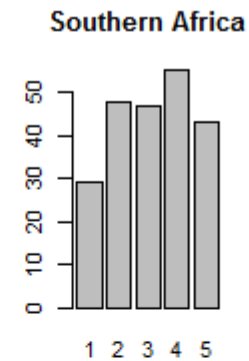
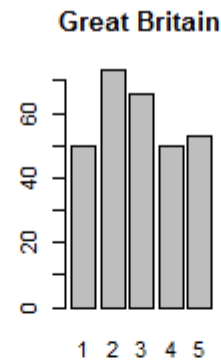
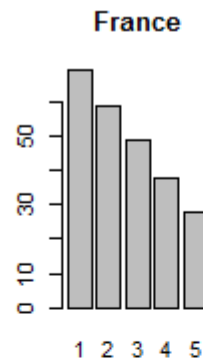
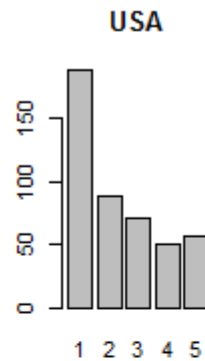
Then for any country **j** member of **Q** it is assumed that :

$$\mathbf{F_{1,k} = F_{2,k} = F_{3,k} = F_{4,k} = F_{5,k} = F_{tot, k}}$$

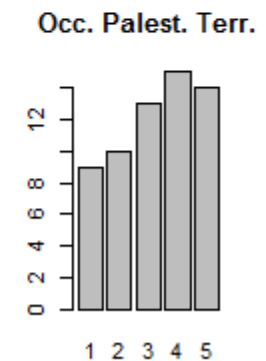
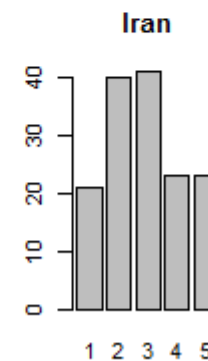
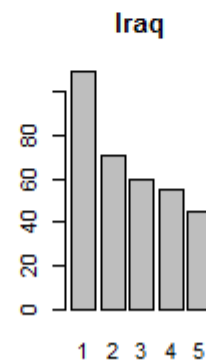
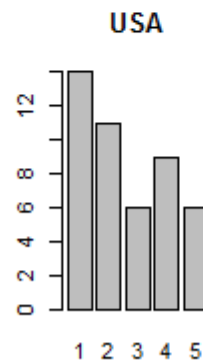
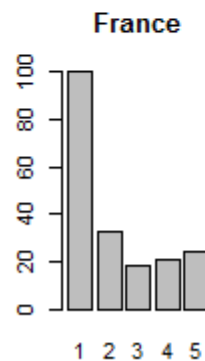
# A test on the sample of Cameroon Students ...



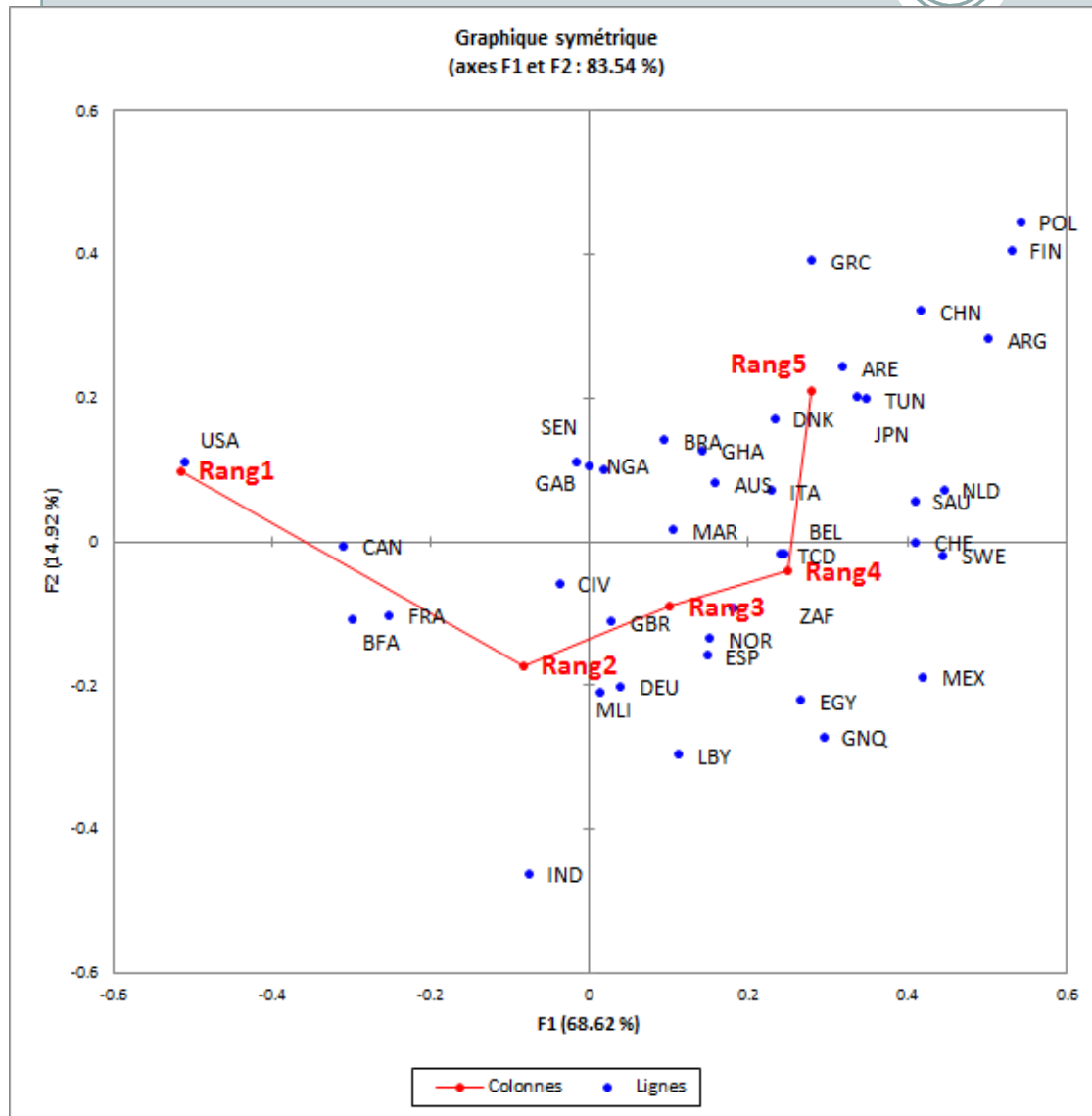
**Would  
like to  
live in ...**



**Would  
Not  
like to  
live in ...**



... does not support the hypothesis of random allocation of ranks.

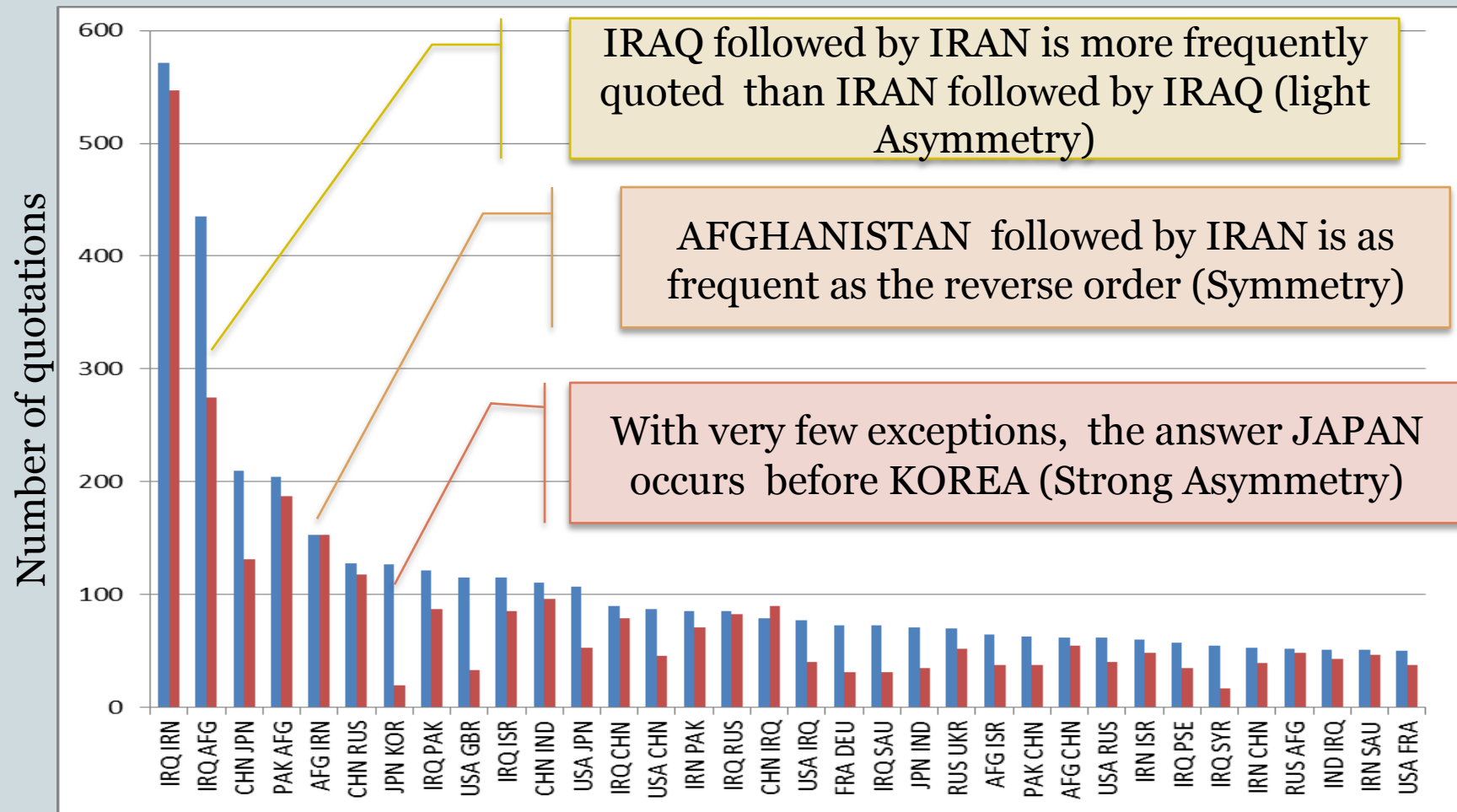


- **Null hypothesis is rejected**

(Chi-2 = 417.7, d.l.=152,  $p < 0.001$ )

- **A hierarchical organization appears** (high frequency of quotations is associated with first ranks).
- **But with a lot of specific exceptions**

# Benchmarking of couples of following places in the list is another way to explore the asymmetry of ranks

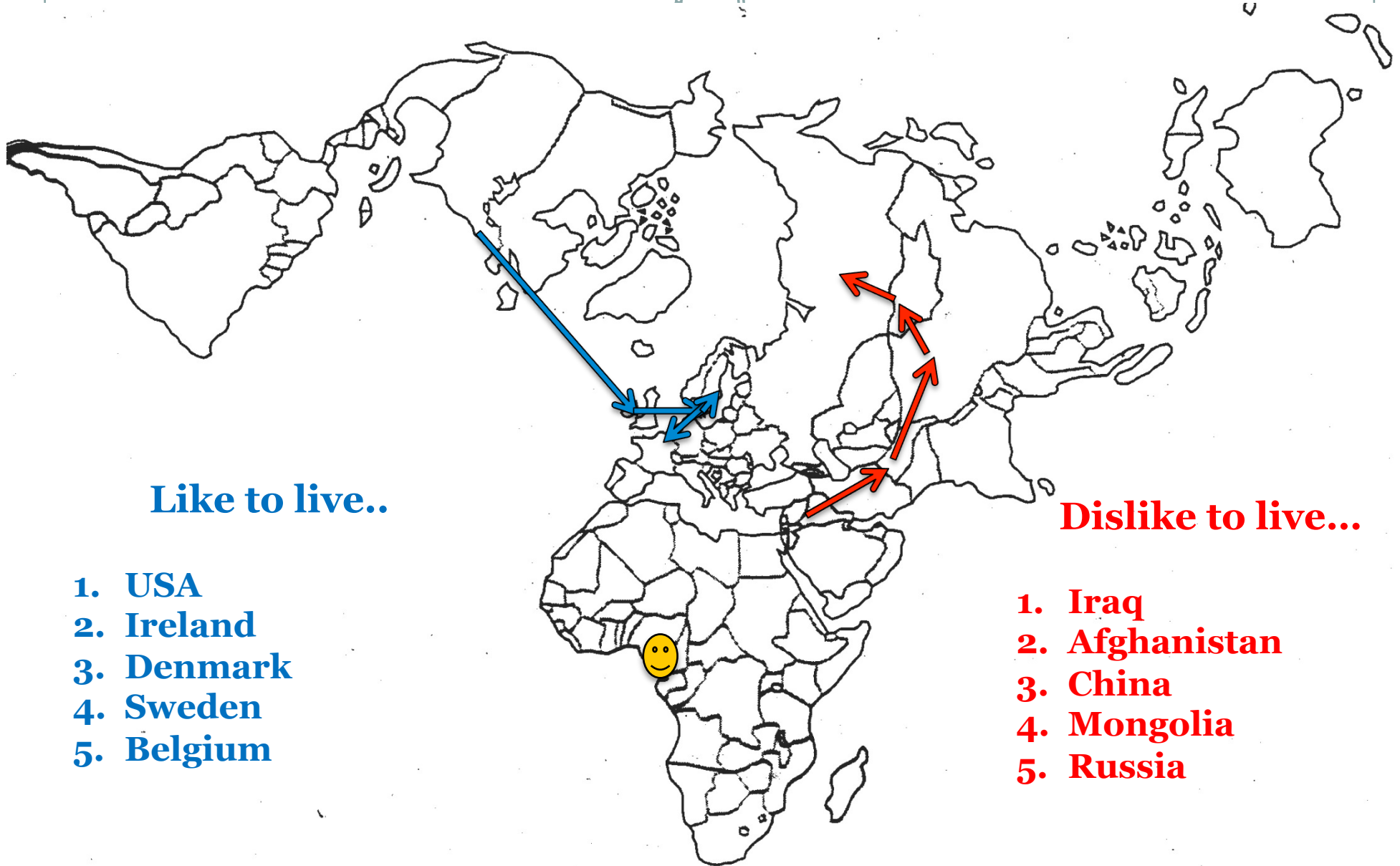


### 3. Testing the existence of spatial autocorrelation in the path of answers ?



# Example 1

## (Buea – Cameroun Anglophone)



**Like to live..**

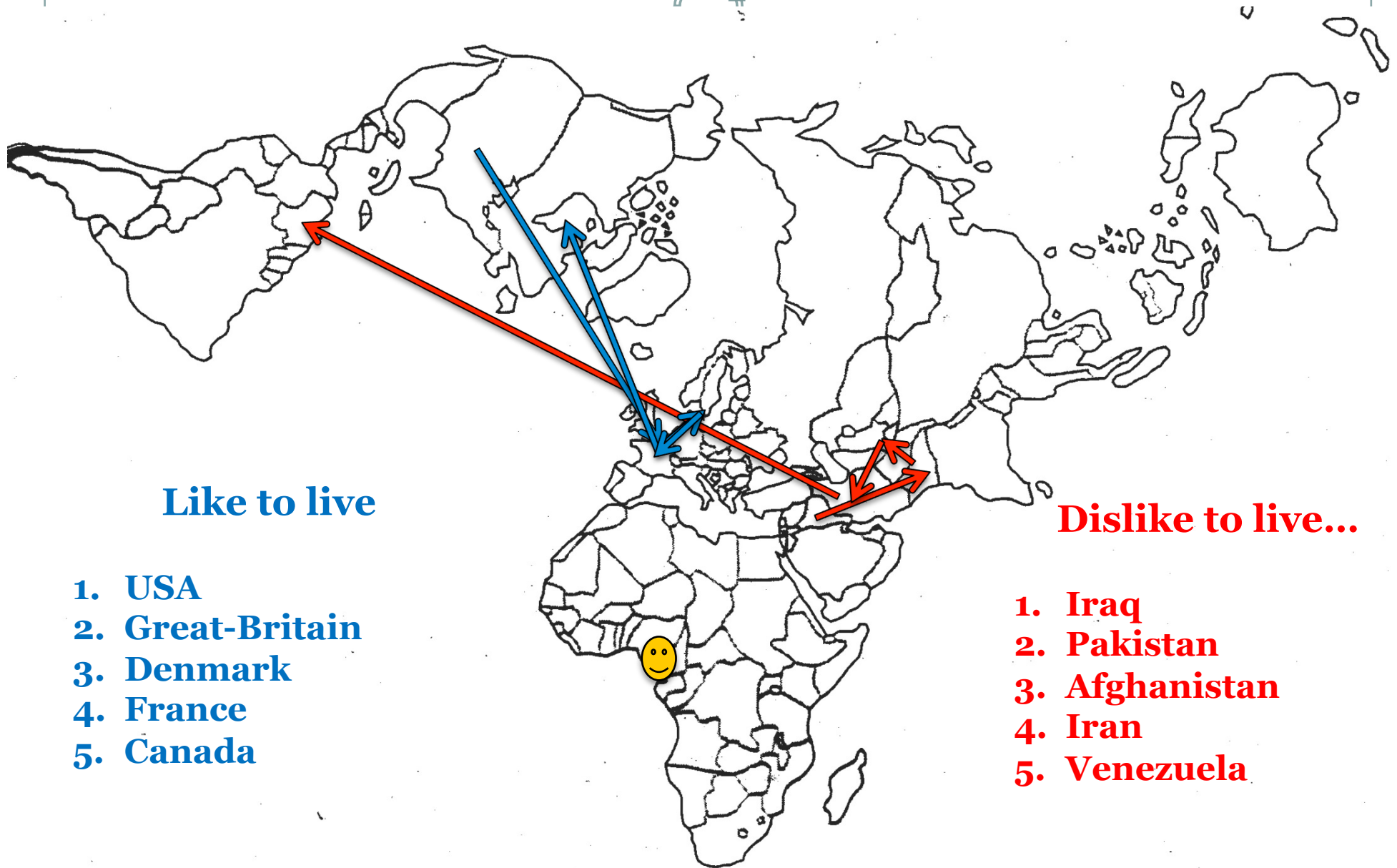
1. USA
2. Ireland
3. Denmark
4. Sweden
5. Belgium

**Dislike to live...**

1. Iraq
2. Afghanistan
3. China
4. Mongolia
5. Russia



## Example 2



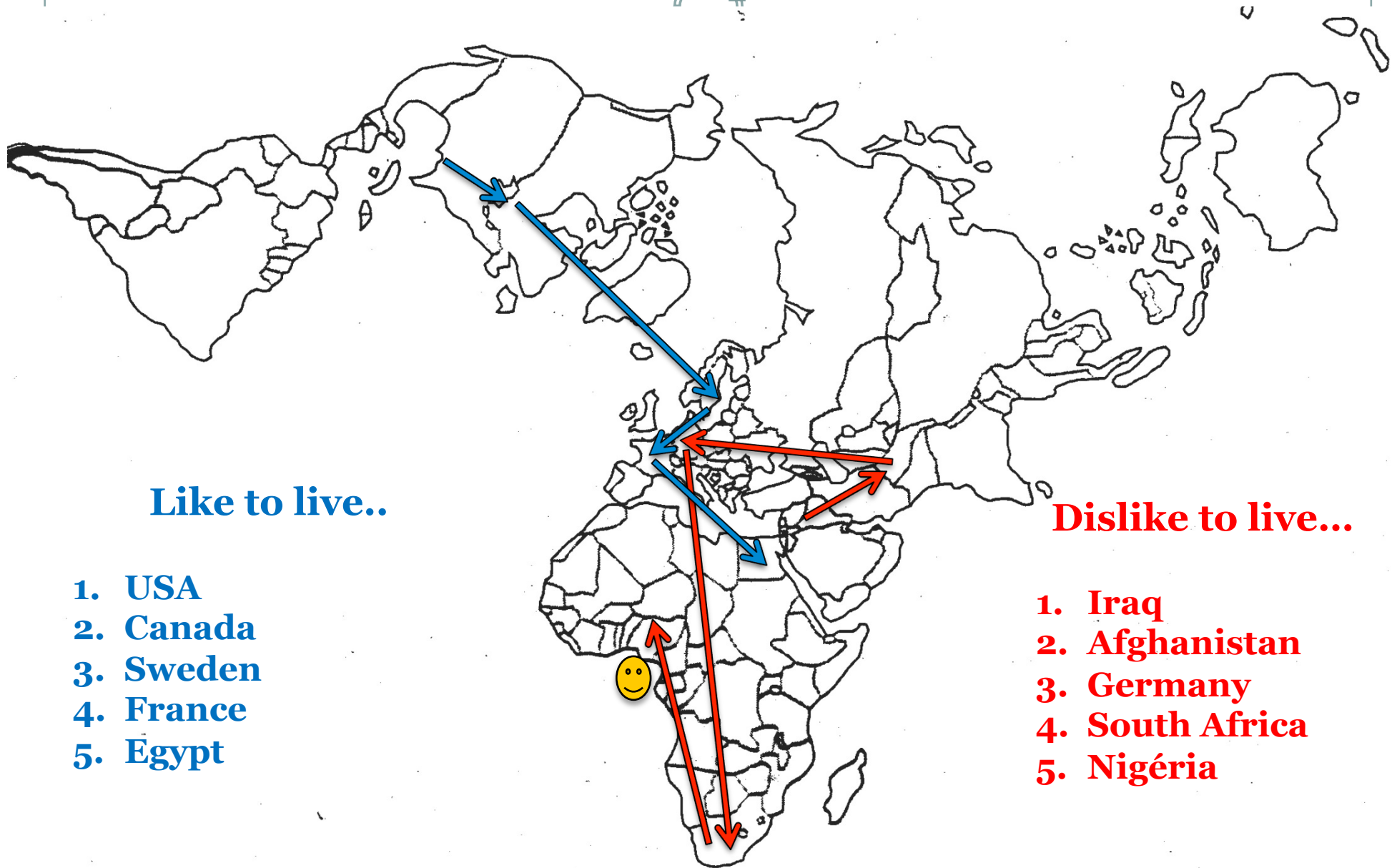
**Like to live**

1. USA
2. Great-Britain
3. Denmark
4. France
5. Canada

**Dislike to live...**

1. Iraq
2. Pakistan
3. Afghanistan
4. Iran
5. Venezuela

## Example 3



# Testing spatial autocorrelation



- **L** : set of lists of countries composed with  $k$  elements :  $R_1, R_2, R_3, R_4, R_5$ .
- **P** : set of couples of countries which appear in the same list.
- For any couple of countries  $(i, j)$ , we compute two distances, respectively the distance between the items ranks in the list ( $D_1$ ) and the countries distance in the geographical area ( $D_2$ ) .

For a given definitions of  $d_1$  and  $d_2$  are, it is assumed

- **$H_0$  : there is no relation between the order of quotation ( $D_1$ ) and the distance between quoted countries ( $D_2$ )**

Assumption is rejected : distance in the list is not independent from distance in geographical space

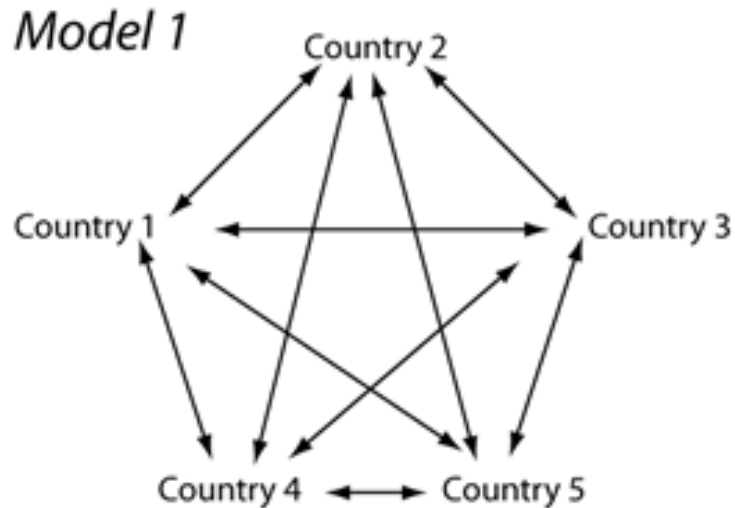
Rank of answers			Like to live		Dislike to live	
Rank i	Rank j	Dif. Rank	Distance	Contiguity	Distance	Contiguity
<b>1</b>	<b>2</b>	<b>1</b>	<b>5054</b>	<b>18,4%</b>	<b>4330</b>	<b>25,1%</b>
1	3	2	5768	12,5%	5267	13,5%
1	4	3	6222	10,4%	5680	10,2%
<b>1</b>	<b>5</b>	<b>4</b>	<b>6235</b>	<b>10,1%</b>	<b>5770</b>	<b>9,5%</b>
<b>2</b>	<b>3</b>	<b>1</b>	<b>5415</b>	<b>14,8%</b>	<b>4551</b>	<b>19,0%</b>
2	4	2	6054	10,2%	5455	10,7%
2	5	3	6104	10,1%	5645	9,5%
<b>3</b>	<b>4</b>	<b>1</b>	<b>5751</b>	<b>12,8%</b>	<b>4836</b>	<b>18,1%</b>
3	5	2	6073	9,8%	5581	9,9%
<b>4</b>	<b>5</b>	<b>1</b>	<b>5918</b>	<b>12,1%</b>	<b>4985</b>	<b>16,5%</b>
<b>Whole</b>			<b>5837</b>	<b>12,3%</b>	<b>5182</b>	<b>14,5%</b>

N.B. Analysis of variance lead in all case to the rejection of the null hypothesis of equal distance or contiguity between countries located in different rank in the list with  $p < 0.0001$

## 4. Hypothesis on the links between answers



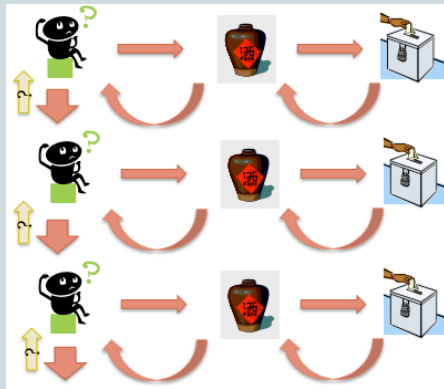
# Model 1 : Full linkage between answers



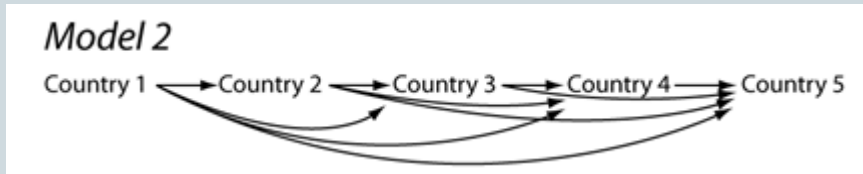
**Assumption** : the order of answers in the list is randomly distributed = each list of  $K$  answers produce  $K(K-1)$  links.

e.g. (A,B,C,D)

=> AB, AC, AD,  
BA,BC,BD, CA,CB,CD



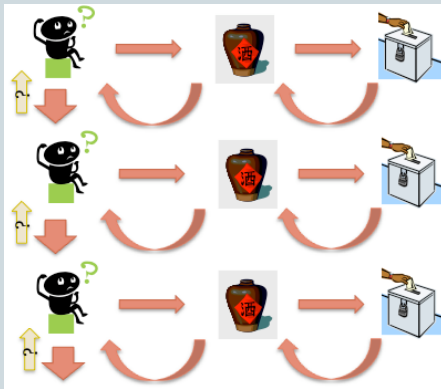
## Model 2 : Any follower linkage



**Assumption** : the order of answers is random but not the connexity. A list of  $k$  answers produces  $K(K-1)/2$  links

e.g. (A,B,C,D)

=> AB, AC, AD, BBC, BD, CD



# Model 3 : Strict follower linkage

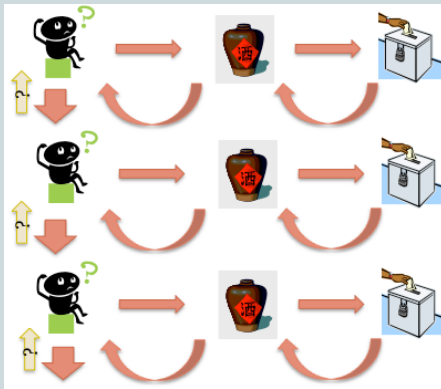


*Model 3*

Country 1 → Country 2 → Country 3 → Country 4 → Country 5

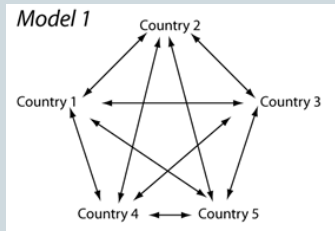
**Assumption** : the order of answers is random as well as the connexity. A list of k answers will produce (K-1) links

e.g. (A,B,C,D)  
=> AB, BC, CD



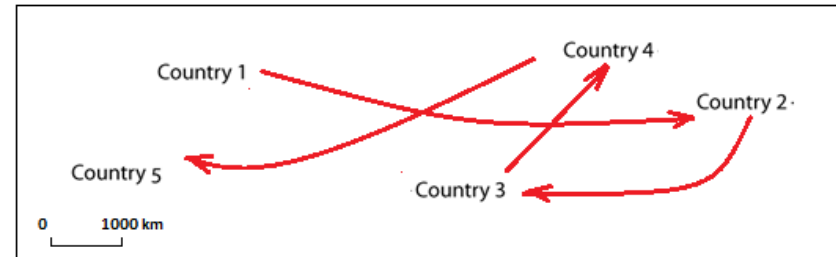


# Geographical interpretation of the 3 models



## Model 1

Random distribution of answers in geographical space : the geographical distance is not related to the rank of countries in the list.

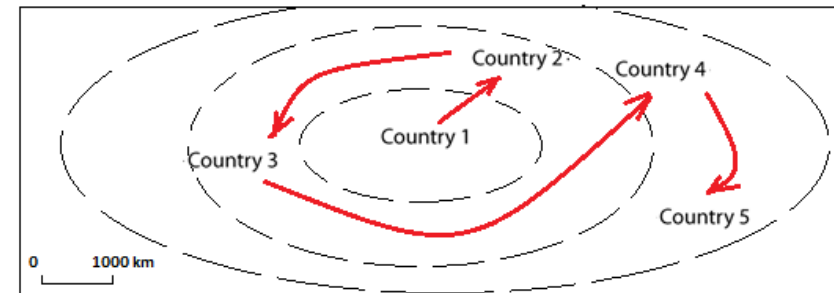


## Model 2

Country 1 → Country 2 → Country 3 → Country 4 → Country 5

## Model 2

Organisation of answers according to a gradient model : the geographical distance are increasing between the first answers and the final answers (Spiral)

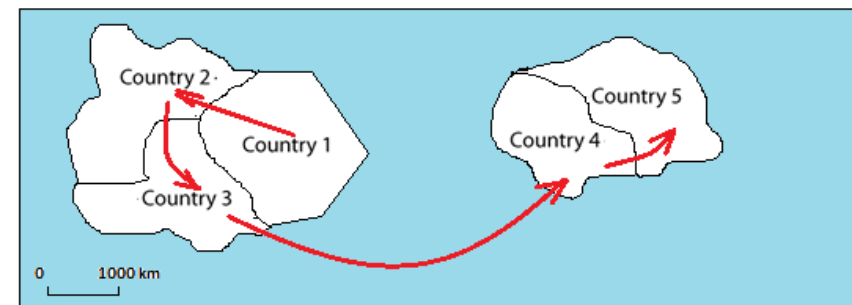


## Model 3

Country 1 → Country 2 → Country 3 → Country 4 → Country 5

## Model 3

Organisation of answers according to a minimal geographical path : the choice of successive answers is influenced by euclidean distance, common borders, belonging to a same continent, etc.



## 5. Application of textual analysis to linkages



The link USA-RUSSIA is not associated to the same links than RUSSIA-USA

**countries for students from Izm**

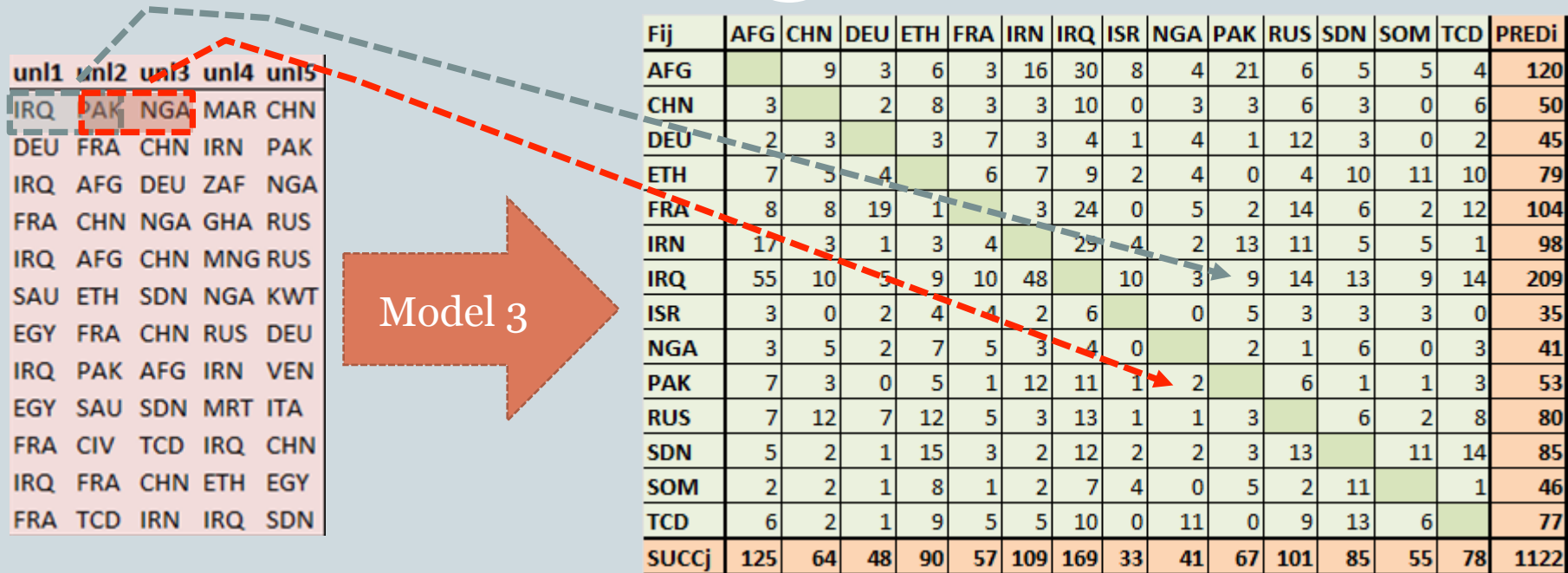
**ISRAEL is mainly associated to USA & « West » for students from Erzurum**

Axe 1 - 15.22 %

## 6. Application of interaction model to linkages



# Step 1 : building linkage matrix (ex. Cameroon)



The segment of countries are introduced in a matrix of linkage according to different possible assumption (**here: Model 3**). The sum of row defines the number of time a country has been put before another one ( $PREDi$ ), the sum of raw, the number of time it has been put after another one ( $SUCCj$ )

## Step 2 : Double Constraint Model

Fij	AFG	CHN	DEU	ETH	FRA	IRN	IRQ	ISR	NGA	PAK	RUS	SDN	SOM	TCD	PREDi
AFG		9	3	6	3	16	30	8	4	21	6	5	5	4	120
CHN	3		2	8	3	3	10	0	3	3	6	3	0	6	50
DEU	2	3		3	7	3	4	1	4	1	12	3	0	2	45
ETH	7	5	4		6	7	9	2	4	0	4	10	11	10	79
FRA	8	8	19	1		3	24	0	5	2	14	6	2	12	104
IRN	17	3	1	3	4		29	4	2	13	11	5	5	1	98
IRQ	55	10	5	9	10	48		10	3	9	14	13	9	14	209
ISR	3	0	2	4	4	2	6		0	5	3	3	3	0	35
NGA	3	5	2	7	5	3	4	0		2	1	6	0	3	41
PAK	7	3	0	5	1	12	11	1	2		6	1	1	3	53
RUS	7	12	7	12	5	3	13	1	1	3		6	2	8	80
SDN	5	2	1	15	3	2	12	2	2	3	13		11	14	85
SOM	2	2	1	8	1	2	7	4	0	5	2	11		1	46
TCD	6	2	1	9	5	5	10	0	11	0	9	13	6		77
SUCCj	125	64	48	90	57	109	169	33	41	67	101	85	55	78	1122

F*ij	AFG	CHN	DEU	ETH	FRA	IRN	IRQ	ISR	NGA	PAK	RUS	SDN	SOM	TCD	PREDi
AFG		7.3	5.5	10.6	6.9	13.1	23.3	3.7	4.7	7.7	11.9	10.1	6.3	9.2	120
CHN	6.0		2.1	4.1	2.7	5.1	9.1	1.5	1.8	3.0	4.7	3.9	2.5	3.6	50
DEU	5.3	2.5		3.7	2.4	4.5	8.1	1.3	1.6	2.7	4.1	3.5	2.2	3.2	45
ETH	9.7	4.6	3.5		4.3	8.3	14.7	2.4	2.9	4.9	7.5	6.4	4.0	5.8	79
FRA	12.4	5.9	4.4	8.6		10.6	18.8	3.0	3.8	6.2	9.6	8.1	5.1	7.4	104
IRN	12.3	5.9	4.4	8.5	5.5		18.6	3.0	3.7	6.2	9.5	8.1	5.0	7.3	98
IRQ	28.7	13.7	10.2	19.7	12.8	24.4		6.9	8.7	14.3	22.2	18.7	11.7	17.1	209
ISR	4.1	2.0	1.5	2.8	1.8	3.5	6.2		1.2	2.0	3.2	2.7	1.7	2.4	35
NGA	4.8	2.3	1.7	3.3	2.2	4.1	7.3	1.2		2.4	3.7	3.2	2.0	2.9	41
PAK	6.4	3.0	2.3	4.4	2.8	5.4	9.7	1.5	1.9		4.9	4.2	2.6	3.8	53
RUS	10.0	4.7	3.5	6.9	4.4	8.5	15.1	2.4	3.0	5.0		6.5	4.1	5.9	80
SDN	10.4	5.0	3.7	7.2	4.7	8.9	15.8	2.5	3.2	5.2	8.1		4.3	6.2	85
SOM	5.5	2.6	2.0	3.8	2.4	4.7	8.3	1.3	1.7	2.7	4.2	3.6		3.3	46
TCD	9.4	4.5	3.3	6.5	4.2	8.0	14.2	2.3	2.8	4.7	7.3	6.1	3.8		77
SUCCj	125	64	48	90	57	109	169	33	41	67	101	85	55	78	1122

$$F^*_{ij} = a_i \cdot \text{PRED}_i \cdot b_j \cdot \text{SUCC}_j$$

(Double constraint model)

The estimated values ( $F^*_{ij}$ ) of linkage can be obtained through the application of a double constraint model based on the marginal sums of the matrix (taking into account the asymmetry and the empty diagonal).

# Step 3 : Analysis of Residuals



Fij	AFG	CHN	DEU	ETH	FRA	IRN	IRQ	ISR	NGA	PAK	RUS	SDN	SOM	TCD	PREDi
AFG		9	3	6	3	16	30	8	4	21	6	5	5	4	120
CHN	3		2	8	3	3	10	0	3	3	6	3	0	6	50
DEU	2	3		3	7	3	4	1	4	1	12	3	0	2	45
ETH	7	5	4		6	7	9	2	4	0	4	10	11	10	79
FRA	8	8	19	1		3	24	0	5	2	14	6	2	12	104
IRN	17	3	1	3	4		29	4	2	13	11	5	5	1	98
IRQ	55	10	5	9	10	48		10	3	9	14	13	9	14	209
ISR	3	0	2	4	4	2	6		0	5	3	3	3	0	35
NGA	3	5	2	7	5	3	4	0		2	1	6	0	3	41
PAK	7	3	0	5	1	12	11	1	2		6	1	1	3	53
RUS	7	12	7	12	5	3	13	1	1	3		6	2	8	80
SDN	5	2	1	15	3	2	12	2	2	3	13		11	14	85
SOM	2	2	1	8	1	2	7	4	0	5	2	11		1	46
TCD	6	2	1	9	5	5	10	0	11	0	9	13	6		77
SUCCj	125	64	48	90	57	109	169	33	41	67	101	85	55	78	1122



$$\text{CHI}^2_{ij} = \frac{(F_{ij} - F^*_{ij})^2}{F^*_{ij}}$$



CHI2ij	AFG	CHN	DEU	ETH	FRA	IRN	IRQ	ISR	NGA	PAK	RUS	SDN	SOM	TCD
AFG		0.6	-1.1	-1.4	-1.5	0.8	1.4	2.2	-0.3	4.8	-1.7	-1.6	-0.5	-1.7
CHN	-1.2		-0.1	1.9	0.2	-0.9	0.3	-1.2	0.9	0.0	0.6	-0.5	-1.6	1.3
DEU	-1.4	0.3		-0.4	3.0	-0.7	-1.4	-0.3	1.9	-1.0	3.9	-0.3	-1.5	-0.7
ETH	-0.9	0.2	0.3		0.8	-0.4	-1.5	-0.2	0.6	-2.2	-1.3	1.4	3.5	1.8
FRA	-1.3	0.9	6.9	-2.6		-2.3	1.2	-1.7	0.6	-1.7	1.4	-0.8	-1.4	1.7
IRN	1.3	-1.2	-1.6	-1.9	-0.6		2.4	0.6	-0.9	2.8	0.5	-1.1	0.0	-2.3
IRQ	4.9	-1.0	-1.6	-2.4	-0.8	4.8		1.2	-1.9	-1.4	-1.7	-1.3	-0.8	-0.7
ISR	-0.5	-1.4	0.5	0.7	1.6	-0.8	-0.1		-1.1	2.1	-0.1	0.2	1.0	-1.6
NGA	-0.8	1.8	0.2	2.0	1.9	-0.5	-1.2	-1.1		-0.3	-1.4	1.6	-1.4	0.1
PAK	0.3	0.0	-1.5	0.3	-1.1	2.8	0.4	-0.4	0.1		0.5	-1.6	-1.0	-0.4
RUS	-0.9	3.3	1.8	2.0	0.3	-1.9	-0.5	-0.9	-1.2	-0.9		-0.2	-1.0	0.9
SDN	-1.7	-1.3	-1.4	2.9	-0.8	-2.3	-1.0	-0.3	-0.7	-1.0	1.7		3.3	3.1
SOM	-1.5	-0.4	-0.7	2.2	-0.9	-1.2	-0.5	2.3	-1.3	1.4	-1.1	3.9		-1.3
TCD	-1.1	-1.2	-1.3	1.0	0.4	-1.1	-1.1	-1.5	4.8	-2.2	0.7	2.8	1.1	

F*ij	AFG	CHN	DEU	ETH	FRA	IRN	IRQ	ISR	NGA	PAK	RUS	SDN	SOM	TCD	PREDi
AFG		7.3	5.5	10.6	6.9	13.1	23.3	3.7	4.7	7.7	11.9	10.1	6.3	9.2	120
CHN	6.0		2.1	4.1	2.7	5.1	9.1	1.5	1.8	3.0	4.7	3.9	2.5	3.6	50
DEU	5.3	2.5		3.7	2.4	4.5	8.1	1.3	1.6	2.7	4.1	3.5	2.2	3.2	45
ETH	9.7	4.6	3.5		4.3	8.3	14.7	2.4	2.9	4.9	7.5	6.4	4.0	5.8	79
FRA	12.4	5.9	4.4	8.6		10.6	18.8	3.0	3.8	6.2	9.6	8.1	5.1	7.4	104
IRN	12.3	5.9	4.4	8.5	5.5		18.6	3.0	3.7	6.2	9.5	8.1	5.0	7.3	98
IRQ	28.7	13.7	10.2	19.7	12.8	24.4		6.9	8.7	14.3	22.2	18.7	11.7	17.1	209
ISR	4.1	2.0	1.5	2.8	1.8	3.5	6.2		1.2	2.0	3.2	2.7	1.7	2.4	35
NGA	4.8	2.3	1.7	3.3	2.2	4.1	7.3	1.2		2.4	3.7	3.2	2.0	2.9	41
PAK	6.4	3.0	2.3	4.4	2.8	5.4	9.7	1.5	1.9		4.9	4.2	2.6	3.8	53
RUS	10.0	4.7	3.5	6.9	4.4	8.5	15.1	2.4	3.0	5.0		6.5	4.1	5.9	80
SDN	10.4	5.0	3.7	7.2	4.7	8.9	15.8	2.5	3.2	5.2	8.1		4.3	6.2	85
SOM	5.5	2.6	2.0	3.8	2.4	4.7	8.3	1.3	1.7	2.7	4.2	3.6		3.3	46
TCD	9.4	4.5	3.3	6.5	4.2	8.0	14.2	2.3	2.8	4.7	7.3	6.1	3.8		77
SUCCj	125	64	48	90	57	109	169	33	41	67	101	85	55	78	1122

The Chi-2 residual make possible to test the linkages that are significantly more or less important as that was expected according to the marginal distribution of  $PRED_i$  and  $SUCC_j$

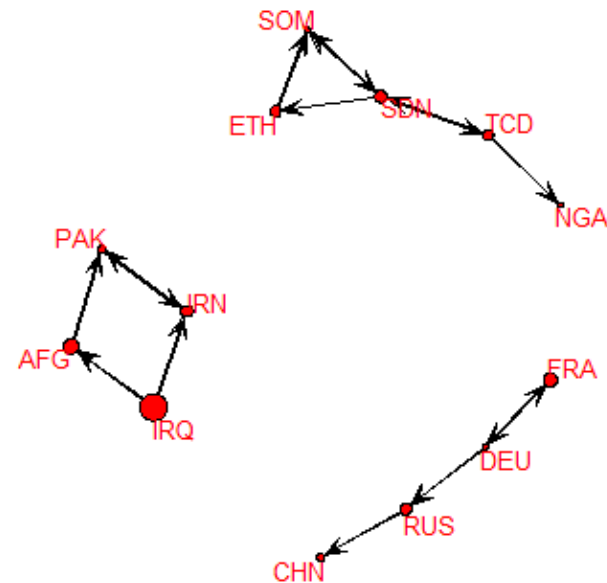
# Step 4 Network analysis of residuals

Fij	AFG	CHN	DEU	ETH	FRA	IRN	IRQ	ISR	NGA	PAK	RUS	SDN	SOM	TCD	PREDI
AFG		9	3	6	3	16	30	8	4	21	6	5	5	4	120
CHN	3		2	8	3	3	10	0	3	3	6	3	0	6	50
DEU	2	3		3	7	3	4	1	4	1	12	3	0	2	45
ETH	7	5	4		6	7	9	2	4	0	4	10	11	10	79
FRA	8	8	19	1		3	24	0	5	2	14	6	2	12	104
IRN	17	3	1	3	4		29	4	2	13	11	5	5	1	98
IRQ	55	10	5	9	10	48		10	3	9	14	13	9	14	209
ISR	3	0	2	4	4	2	6		0	5	3	3	3	0	35
NGA	3	5	2	7	5	3	4	0		2	1	6	0	3	41
PAK	7	3	0	5	1	12	11	1	2		6	1	1	3	53
RUS	7	12	7	12	5	3	13	1	1	3		6	2	8	80
SDN	5	2	1	15	3	2	12	2	2	3	13		11	14	85
SOM	2	2	1	8	1	2	7	4	0	5	2	11		1	46
TCD	6	2	1	9	5	5	10	0	11	0	9	13	6		77
SUCCJ	125	64	48	90	57	109	169	33	41	67	101	85	55	78	1122

$R_{ij} = 1$  if  
 (a)  $F_{ij} > n$   
 (b)  $CHI^2_{ij} > CHI^2(1, \alpha)$

CHI2ij	AFG	CHN	DEU	ETH	FRA	IRN	IRQ	ISR	NGA	PAK	RUS	SDN	SOM	TCD
AFG		0.6	-1.1	-1.4	-1.5	0.8	1.4	2.2	-0.3	4.8	-1.7	-1.6	-0.5	-1.7
CHN	-1.2		-0.1	1.9	0.2	-0.9	0.3	-1.2	0.9	0.0	0.6	-0.5	-1.6	1.3
DEU	-1.4	0.3		-0.4	3.0	-0.7	-1.4	-0.3	1.9	-1.0	3.9	-0.3	-1.5	-0.7
ETH	-0.9	0.2	0.3		0.8	-0.4	-1.5	-0.2	0.6	-2.2	-1.3	1.4	3.5	1.8
FRA	-1.3	0.9	6.9	-2.6		-2.3	1.2	-1.7	0.6	-1.7	1.4	-0.8	-1.4	1.7
IRN	1.3	-1.2	-1.6	-1.9	-0.6		2.4	0.6	-0.9	2.8	0.5	-1.1	0.0	-2.3
IRQ	4.9	-1.0	-1.6	-2.4	-0.8	4.8		1.2	-1.9	-1.4	-1.7	-1.3	-0.8	-0.7
ISR	-0.5	-1.4	0.5	0.7	1.6	-0.8	-0.1		-1.1	2.1	-0.1	0.2	1.0	-1.6
NGA	-0.8	1.8	0.2	2.0	1.9	-0.5	-1.2	-1.1		-0.3	-1.4	1.6	-1.4	0.1
PAK	0.3	0.0	-1.5	0.3	-1.1	2.8	0.4	-0.4	0.1		0.5	-1.6	-1.0	-0.4
RUS	-0.9	3.3	1.8	2.0	0.3	-1.9	-0.5	-0.9	-1.2	-0.9		-0.2	-1.0	0.9
SDN	-1.7	-1.3	-1.4	2.9	-0.8	-2.3	-1.0	-0.3	-0.7	-1.0	1.7		3.3	3.1
SOM	-1.5	-0.4	-0.7	2.2	-0.9	-1.2	-0.5	2.3	-1.3	1.4	-1.1	3.9		-1.3
TCD	-1.1	-1.2	-1.3	1.0	0.4	-1.1	-1.1	-1.5	4.8	-2.2	0.7	2.8	1.1	

**Cameroon**  
 Marginal Model on disliked countries  
 Oriented couples of places

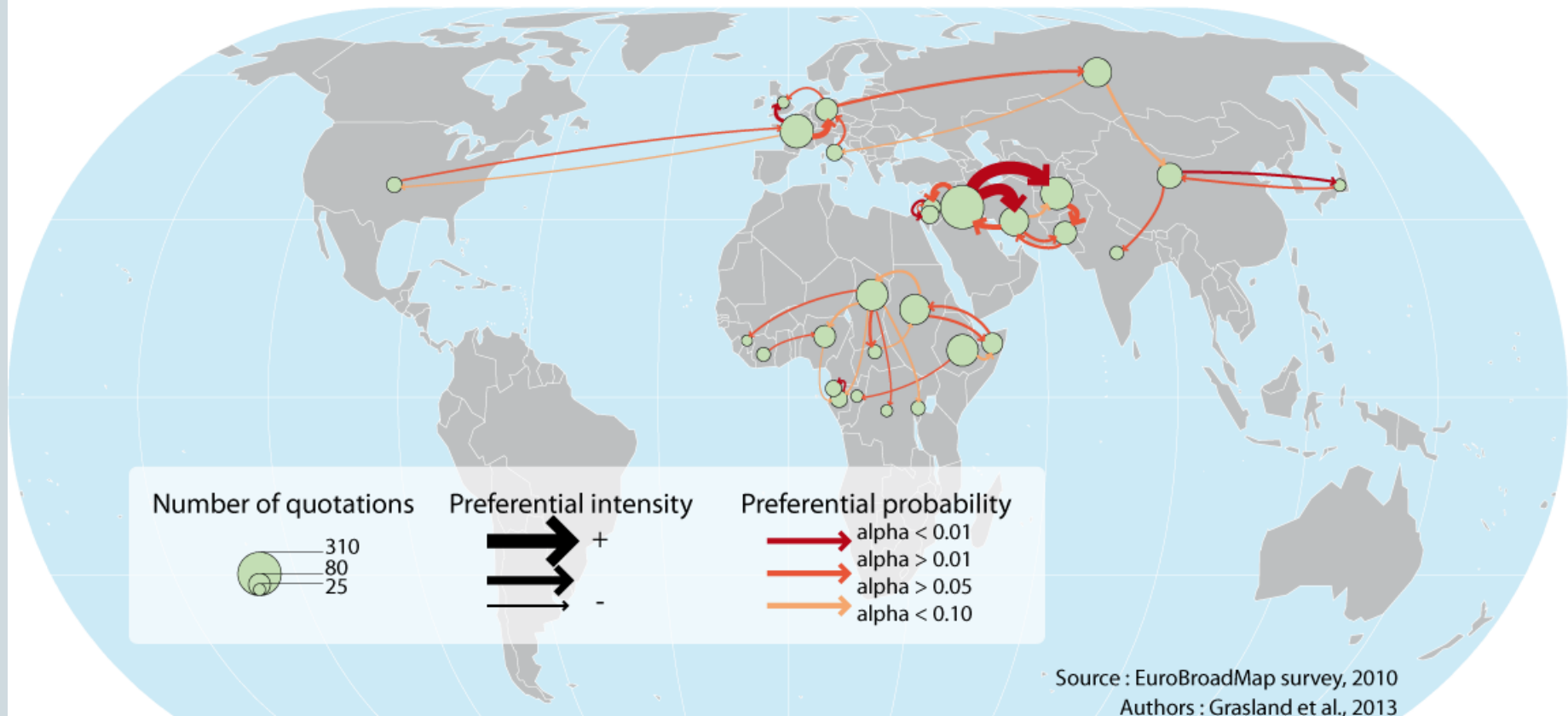


fij=5 - alpha=0.10 - r2= 0.53



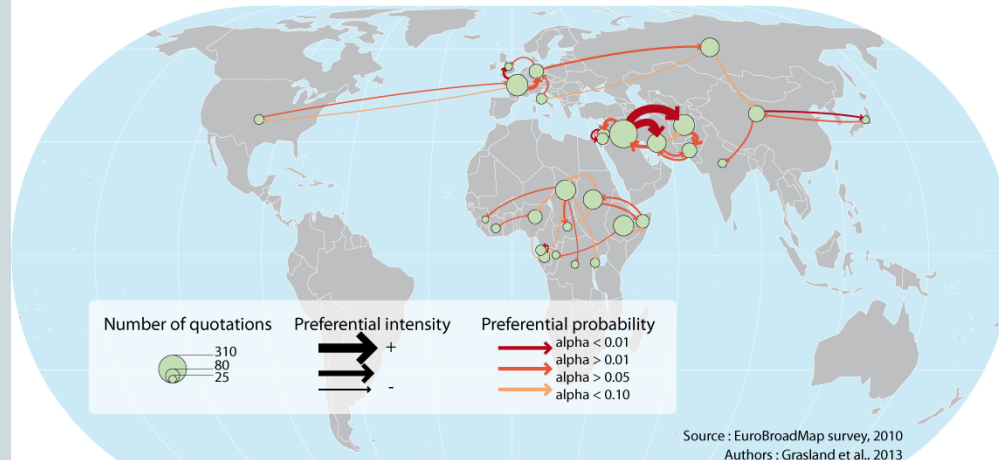
# Step 5 : Cartographic analysis of residuals

Preferential strict followers in the list of countries where Cameroonian students would not like to live

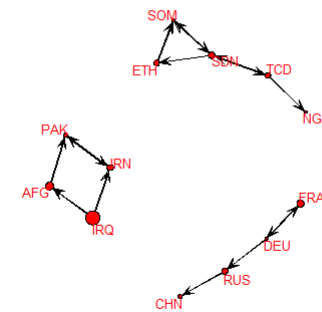


# Comparison of samples

Preferential strict followers in the list of countries where Camerooneese students would not like to live

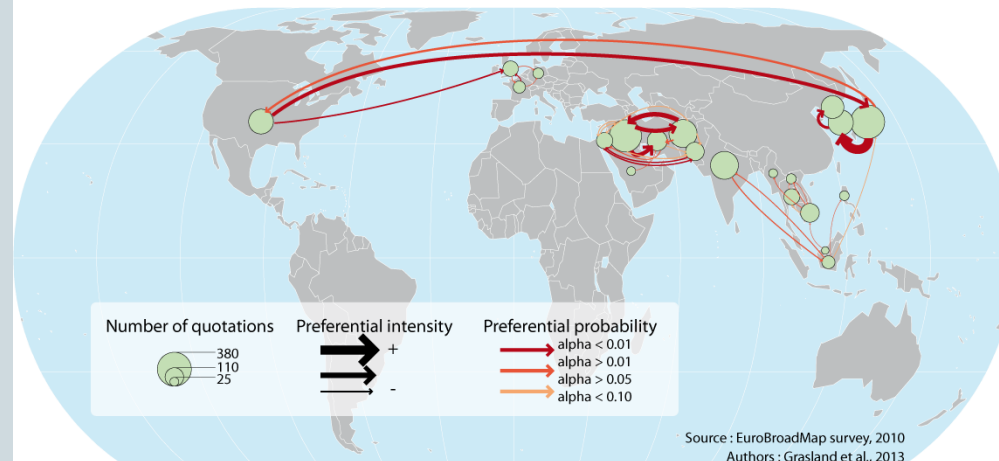


**Cameroon**  
Marginal Model on disliked countries  
Oriented couples of places

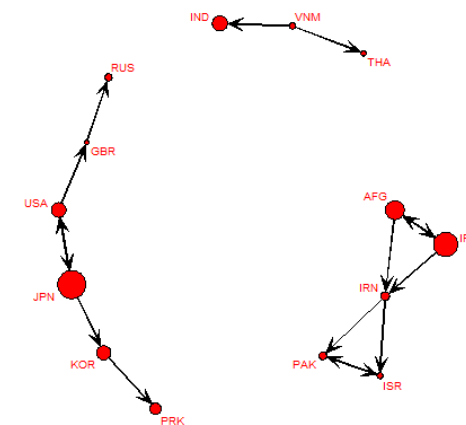


fij=5 -  $\alpha=0.10$  -  $r^2=0.53$

Preferential strict followers in the list of countries where Chinese students would not like to live



**China**  
Marginal Model on disliked countries  
Oriented couples of places



fij=5 -  $\alpha=0.10$  -  $r^2=0.56$

# CONCLUSION AND PERSPECTIVES



# (1) Application of spatial interaction models to modelize spatial dependency

$$F^*_{ij} = a_i \cdot \text{PRED}_i \cdot b_j \cdot \text{SUCC}_j \cdot (\text{DIST}_{ij})^\alpha \cdot \gamma^{\text{CONT}_{ij}}$$

(Double constraint model of spatial interaction)

	model 1: any order of the answers				model 2: relative followers				model 3: strict follower			
Gravity model	like		Dislike		like		Dislike		like		Dislike	
	Distance	Contiguity	Distance	Contiguity	Distance	Contiguity	Distance	Contiguity	Distance	Contiguity	Distance	Contiguity
BRAZIL	-0,22	NS	-0,37	1,15	-0,2	NS	-0,34	1,16	-0,31	NS	-0,65	1,25
CAMEROON	-0,3	0,82	-0,58	NS	-0,27	0,85	-0,6	NS	-0,34	NS	-0,91	NS
CHINA	-0,12	NS	-0,38	NS	-0,12	NS	-0,39	NS	-0,26	NS	-0,7	NS
RUSSIA	-0,34	NS	-0,65	1,14	-0,35	NS	-0,67	NS	-0,48	NS	-0,97	1,48

- Distance effect systematially significant but not contiguity
- Spatial dependance higher for repulsive than attractive places

## (2) Research of explanation for the spatial autocorrelation of answers



- **Structural explanation:** the criteria chosen by students to order countries are non-spatial (war, poverty, language, religion, ...) but the distribution of countries of same type is spatially autocorrelated.  
(ex. *Iran-Iraq-Afghanistan-Pakistan*)
- **Mental Map explanation :** Once the student has chosen a first country, his mind is focused on a particular area and the next choice is influenced by the proximity (real/perceived) of countries.  
(ex. *Chad-Sudan-Ethiopia* or *Japan-Korea-Russia*)

### (3) Application to other kind of lists of places



- **Countries** mentionned in the title of **media news** (does order matter ? )
- **Cities** of location of researchers in **scientific publications** (order does certainly matter... but in which direction ?)
- **Biographical data** : the order of places in the life on an individual path reflect migratory strategies

THANK YOU FOR YOUR ATTENTION

