Conductome

Top down versus bottom up? Using Machine Learning in psychometrics to construct cultural profiles

Christopher R. Stephens

C3 – Centro de Ciencias de la Complejidad y Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México

IACCP 2024 5th-9th August, Bali, Indonesia

Work in collaboration with: Dagmara Wrzecionkowska, Gabriela Navarro, Sofia Michaelian

Universidad Nacional Autónoma de México
Coordinación de la Investigación Científica UNAM
C3 Centro de Ciencias de la Complejidad
Facultad de Psicología
FES Zaragoza
Instituto de Investigaciones Jurídicas unam
IBERO CIUDAD DE MÉXICO ®
UNIVERSIDAD DE GUANAJUATO
IMSS

What is a cultural “profile”?

Cross-cultural psychology (CCP) examines similarities and differences in psychological functioning in different cultures. Triandis 2001

A set of features, $\mathbf{X} = (X_1, X_2, \ldots, X_N)$ that can be used to describe a group/population, $\boldsymbol{C}$, and distinguish it from other groups/populations, $\overline{\boldsymbol{C}}$.

Two questions:

i) Just what count as relevant cultural features?

a. Languages and idioms?
b. Beliefs and attitudes?
c. Customs?
d. Arts and sciences?
e. Socio-demographic and socio-economic factors?
f. Psychological constructs? *
g. Behaviours?
h. Health states – mental and physical?
i. Others?

- Hofstede´s cultural dimensions vs. individual based constructs such as personality, negative affect, delayed gratification, locus of control, auto-efficacy, etc.

- See also the talk of Dagmara Wrzecionkowska and the poster of Gabriela Navarro

ii) How do we quantify the differences in profiles $\mathbf{X}(C)$ versus $\mathbf{X}(\overline{C})$?

a. Distance functions?
b. Cultural niches?
c. Others?

* Do we want them to be valid across cultures, e.g., for job applications, exams etc. or reflect cultural differences?

The dimensions of culture space

9 psychological instruments:

- Delayed gratification (frequency) – N = 379; 35 items; 3 factors, $\alpha = 0.838$
- Delayed gratification (intensity) – N = 379; 35 items; 4 factors; $\alpha = 0.852$
- Locus of control (food) – N = 296; 66 items; 4 factors; $\alpha = 0.918$
- Locus of control (exercise) – N = 296; 43 items; 3 factors; $\alpha = 0.960$
- Auto-efficacy (food) – N = 297; 30 items; 2 factors; $\alpha = 0.944$
- Auto-efficacy (exercise) – N = 297; 18 items; 1 factor; $\alpha = 0.939$
- Three-factor Eating questionaire – N =232; 18 items; 2 factors ; $\alpha = 0.839$
- Beliefs (obesity) – N =297; ; 6 items; 1 factor; $\alpha = 0.844$
- Beliefs (exercise) – N = 297; 10 items; 2 factors; $\alpha = 0.754$

- Instruments and associated data can be viewed and downloaded at: chilam.c3.unam.mx/proy-42/datos-proy42

PSYCHOMETRICS

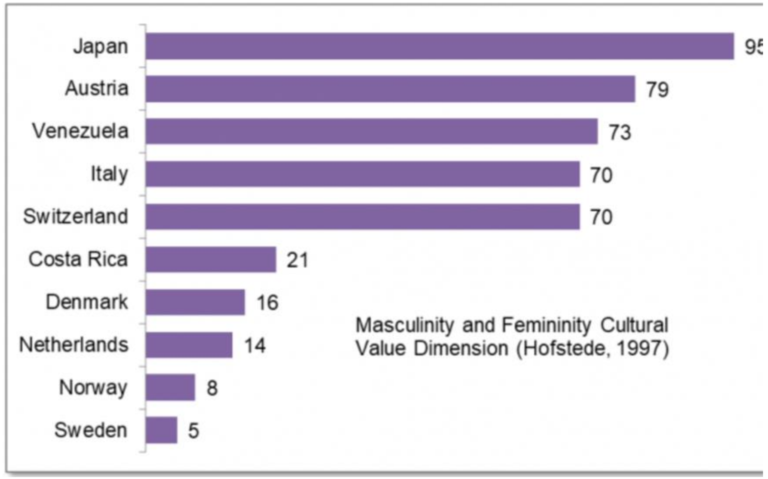
Possible cultural spaces:

X = (X_1, X_2,…, X_9) 9 instruments (all items or just validated items) = 9 dimensions
X = (X_1, X_2,…, X_{22}) 22 factors = 22 dimensions
X = (X_1, X_2,…, X_{360}) 261 items = 261 dimensions
In principle, we can also combine these dimensions

How do we judge the value of a cultural dimension?

i) Dimensional reduction: more "explainability"?
ii) Discrimination between populations?
iii) Predictive power?

A cultural profile $\mathbf{X}(\boldsymbol{C})$ of a population $\boldsymbol{C}$ is a vector in one of these spaces



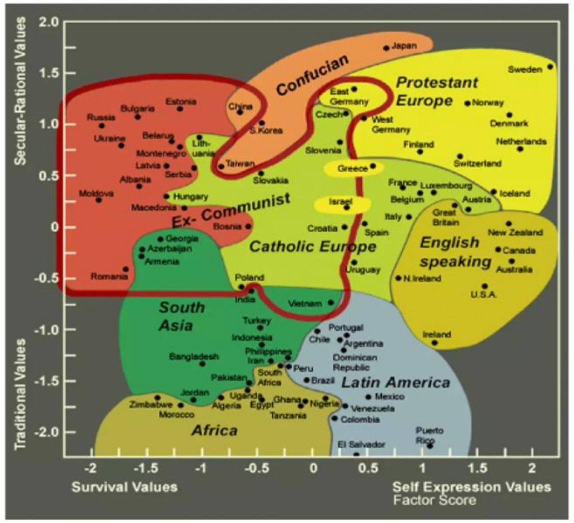
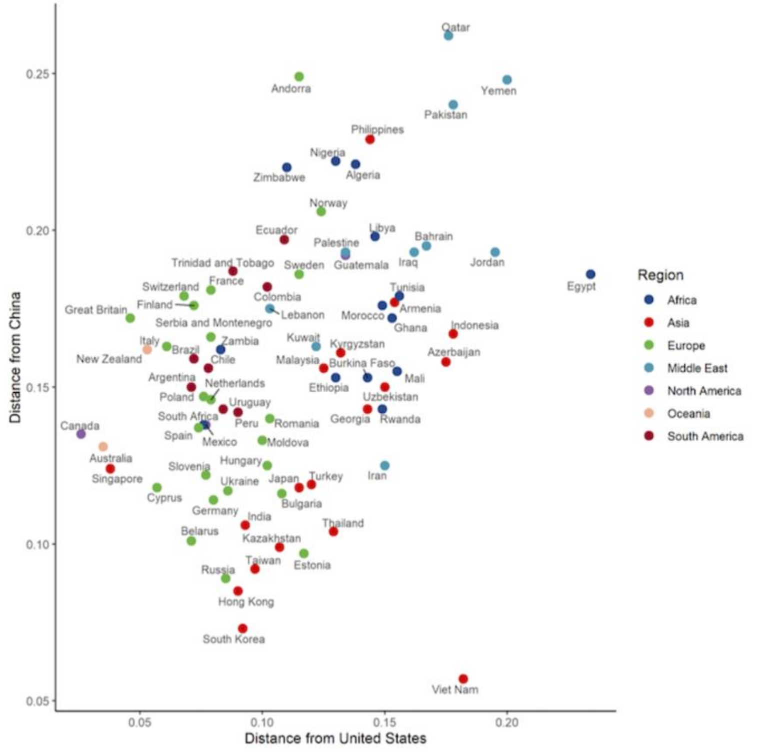


Figure 2. Map of countries, based on the Inglehart's World Values Survey [Inglehart & Welzel (2005, p.63)]

Two-dimensional perspectives



Courtesy of Michael Muthukrishna

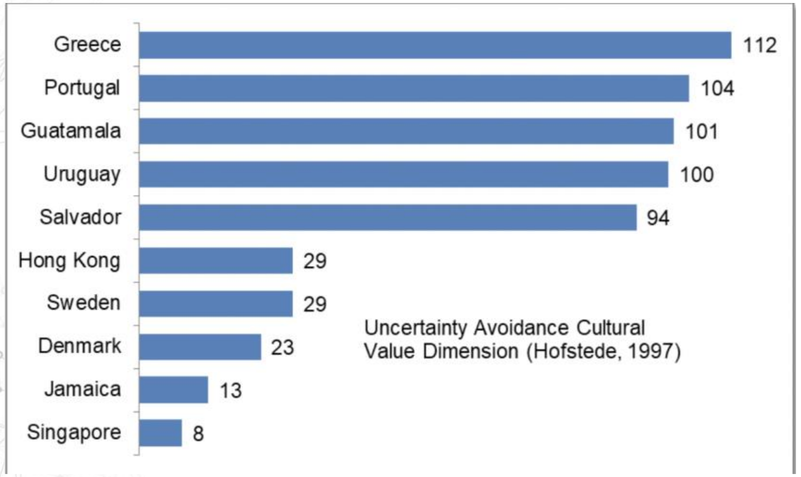
One dimensional perspective –

Distance is the average population distance

There are many distance metrics that can be used in culture space:

1. Euclidean distance
2. Kogut and Singh Index
3. Fixation index

Existence of a distance function depends on the profile dimensions being metric.



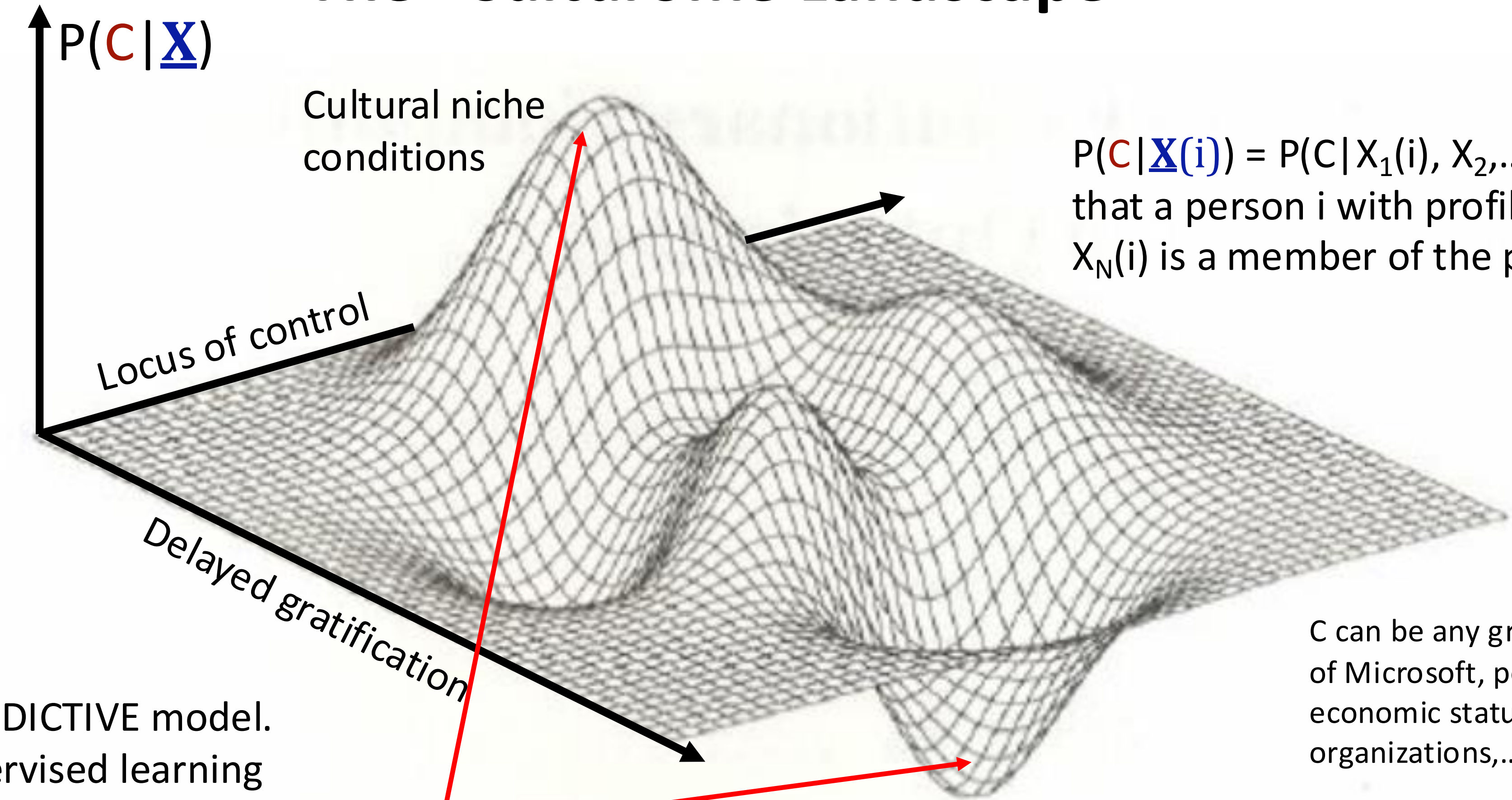
This is a **descriptive** not **predictive** approach

2. Kogut, B., & Singh, H. 1988. The effect of national culture on the choice of entry mode. Journal of International Business Studies, 19(3): 411–432.

3. Muthukrishna, Michael and Bell, Adrian and Henrich, Joseph and Curtin, Cameron and Gedranovich, Alexander and McInerney, Jason and Thue, Braden, Beyond WEIRD Psychology: Measuring and Mapping Scales of Cultural and Psychological Distance (January 15, 2020). Available at SSRN: https://ssrn.com/abstract=3259613 or http://dx.doi.org/10.2139/ssrn.3259613

Characterizing Cultural Niches

The “Culturome Landscape”



$P(C|\mathbf{X}(i)) = P(C|X_1(i), X_2,\ldots, X_N)$ is the probability that a person i with profile $\mathbf{X}(i) = X_1(i), X_2(i),\ldots, X_N(i)$ is a member of the population C.

C can be any group – Mexicans, employees of Microsoft, people with different socio-economic status, members of different organizations,…

Now we have a PREDICTIVE model. We can train a supervised learning model and use it to characterize cultural profiles and predict who most fits the profile.

High/low probability to be a member of the population C

Cultural anti-niche conditions

The Machine Learning Perspective

P(C|$\underline{\mathbf{X}}$(i)) is a **classifier.**

- We use it to predict who is likely to be in the population C.
- We train the model on a set of subjects where we know if they are members of the population C or not.
- We can then apply the model, if we choose, to persons whose status is “unknown”.
- We evaluate the predictive power of the model using standard metrics: e.g., confusion matrix, ROC curve, etc.
- We can judge the different cultural dimensions through their **predictability** – i.e., how they contribute to the performance of the classifier model
- Feature selection linked to maximizing classification performance will determine if the construct, its factors, the individual items, or a particular combination of them are more predictive

The Machine Learning Perspective

$P(C|\underline{\mathbf{X}}) = N(C\,\underline{\mathbf{X}})/N(\underline{\mathbf{X}})$ cannot be computed directly in frequentist terms , $N(\underline{\mathbf{X}}) = 0, 1$

$P(C|\underline{\mathbf{X}}) = P(\underline{\mathbf{X}}|C)P(C)/P(\underline{\mathbf{X}})$ Bayes Theorem

$P(\underline{X}|C) = \prod_{i=1}^{N} P(X_i|C)$ Naïve Bayes approximation

$P(C|\underline{X}) \rightarrow S(C|\,\underline{X}) = \ln(P(C|\underline{X})/P(\bar{C}|\underline{X})) = \sum_{i=1}^{N} s(X_i) + \ln(P(C)/P(\bar{C})$

where $s(X_i) = \ln(P(X_i|C)/P(X_i|\bar{C}))$

If $s(X_i) > 0$/$P(C|X_i) > P(C)$ then X_i is a cultural niche factor and
if $s(X_i) < 0$/$P(C|X_i) < P(C)$ then X_i is a cultural anti-niche factor

We use a binomial test to determine the statistical significance of the feature X_i

$$\varepsilon = \frac{N(X_i)(P(C|X_i) - P(C))}{\sqrt{N(X_i)P(C)(1 - P(C))}}$$

$P(C)$ is the null hypothesis. $|\varepsilon| > 1.96$ represents the 95% CI that the observed $P(C|X_i)$ is inconsistent with the null hypothesis.

The Populations

Two populations of Mexico City-based undergraduate university students:

Universidad Iberoamericana (Class C to be characterized)
Private university – Pri (https://ibero.mx/)

“Entrusted to the Society of Jesus, more than 80 years after its creation, IBERO has dedicated itself to forming the men and women that Mexico needs. Its educational leadership has gone beyond the classroom and has placed it as one of the best private universities nationally and internationally. Although a large part of its members are animated by Christian principles, IBERO respects and accommodates all ideology, and does not intend to use its academic activities for the benefit of any particular doctrine.”

Universidad Nacional Autónoma de México
Public university - Pub (https://www.unam.mx/)

“The National Autonomous University of Mexico was founded on September 21, 1551 with the name of the Royal and Pontifical University of Mexico. It is the largest and most important university in Mexico and Latin America. Its primary purpose is to be at the service of the country and humanity, to train professionals useful to society, to organize and carry out research, mainly on national conditions and problems, and to extend the benefits of culture as widely as possible.”

N = 380; P(C) ~ 0.23

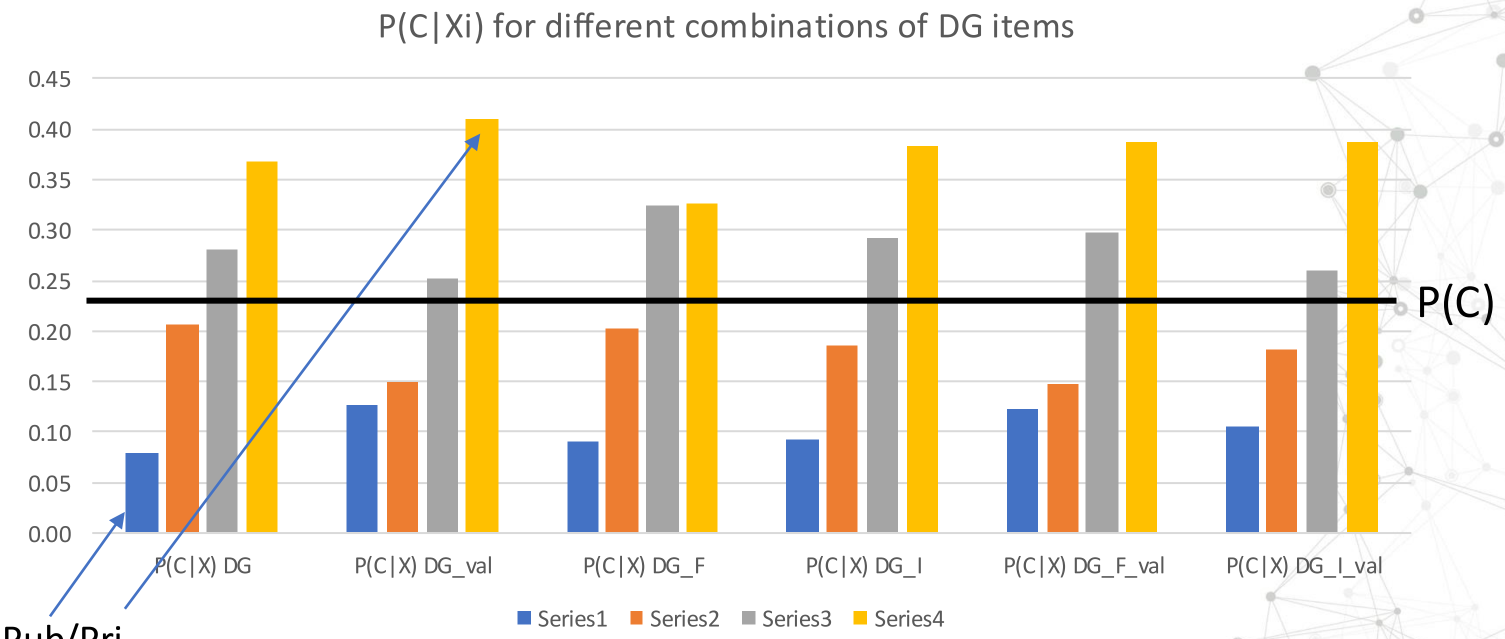
Top down versus bottom up…

What is more predictive? Constructs, factors of constructs, items,…?

Constructs: Multiple constructs? only items validated by exploratory factor analysis in the population,…?
Example: DG Frequency scale – assume 5-point Likert scale is interval and map to 1, 2, 3, 4, 5. For each participant sum the scores. Typically this scale as independent variable can then be regressed against a dependent variable.
In our machine learning classifier approach we divide the scale into M intervals, X_i, each with approximately the same number of participants to obtain the same statistical weight $N(X_i)$ for each interval.

Variable	Interval values	Bin	N(Xi)	N(CXi)	P(C|Xi) DG
DG	(155.999, 242.0]	1	61	5	0.08
DG	(242.0, 264.0]	2	58	12	0.21
DG	(264.0, 284.0]	3	60	17	0.28
DG	(284.0, 334.0]	4	59	22	0.37

Each version of the constructs DG frequency and intensity or their combination, using all items or only those involved in a factor, have slightly different predictability properties.

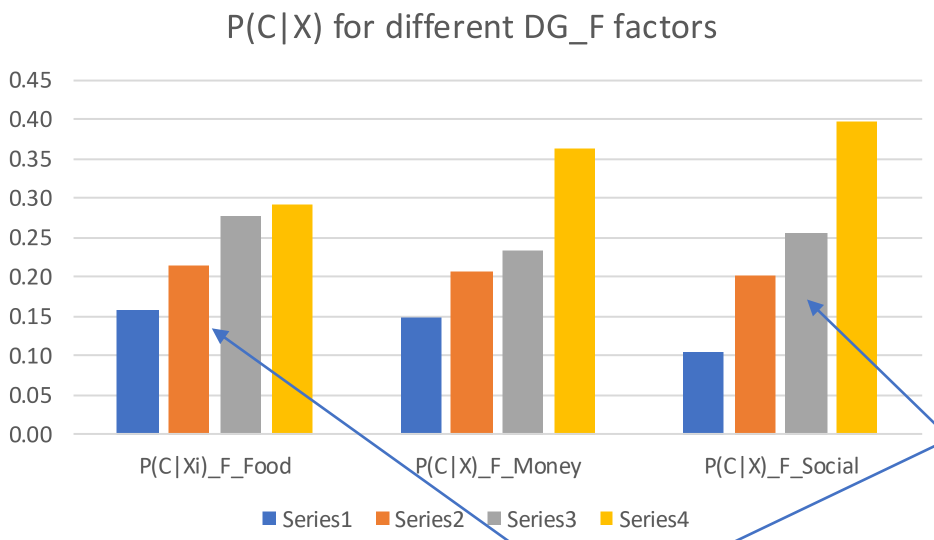


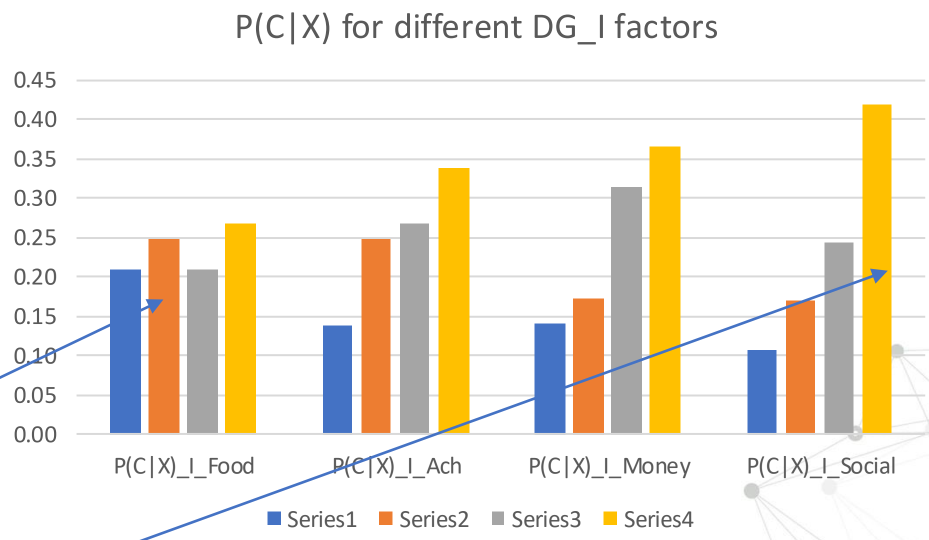
Predicts better who is Pub/Pri

Approximately linear

Top down versus bottom up…

Factors?





Less predictive factor

More predictive factor

Variable	Interval values	Epsilon	N(Xi)	N(Cxi)	P(C|Xi)
DG_F_Food	(2.999, 8.0]	-1.36	61	10	0.16
DG_F_Food	(8.0, 10.0]	-0.34	67	14	0.21
DG_F_Food	(10.0, 12.0]	0.81	58	16	0.28
DG_F_Food	(12.0, 15.0]	1.00	52	15	0.29
DG_F_Money	(5.999, 21.0]	-1.70	75	11	0.15
DG_F_Money	(21.0, 24.0]	-0.49	66	14	0.21
DG_F_Money	(24.0, 26.0]	0.03	38	9	0.23
DG_F_Money	(26.0, 30.0]	2.40	60	22	0.36
DG_F_Social	(3.999, 11.0]	-2.38	62	6	0.10
DG_F_Social	(11.0, 12.0]	-0.56	63	13	0.20
DG_F_Social	(12.0, 14.0]	0.44	63	16	0.26
DG_F_Social	(14.0, 15.0]	2.77	50	20	0.40

Less predictive values of the factor

More predictive values of the factor

Still approximately linear

In a regression approach you include in the entire variable

Top down versus bottom up…

What is more explicative?

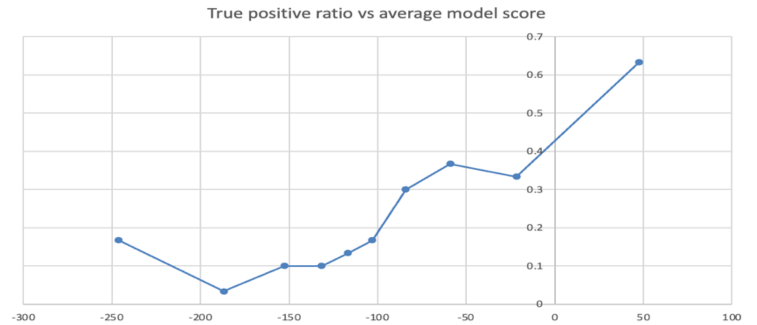
The 10 most significant item-based niche and anti-niche factors

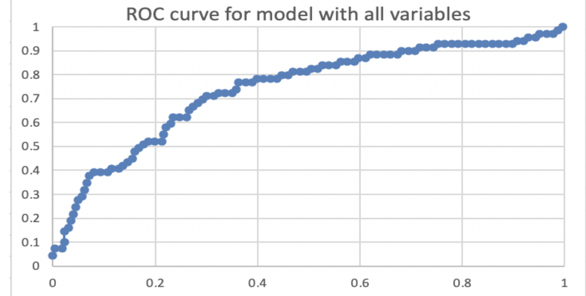
These are the most significant DG_I factors in the cultural profile of our Pri population relative to the Pub population.
The top 10 niche factors and the top 10 anti-niche factors. Note that they involve all four factors as well as non-validated items.

Item	Response	Validated	Factor	P(C|X)	Epsilon	Score
I manage my money well	Totally Agree	Y	Money	0.41	4.06	1.22
I have abandoned physical pleasure and comfort in order to achieve my goals	Totally disagree	N		0.40	2.52	2.02
I value the needs of the people around me	Totally Agree	Y	Social	0.39	5.42	0.40
I try to spend my money wisely	Totally Agree	Y	Money	0.39	4.73	0.66
Sometimes I eat until I feel bad	Neither agree/disgree	Y	Food	0.39	3.19	1.41
I have a hard time maintaining a healthy/special diet	Totally disagree	Y	Food	0.37	2.77	1.34
I try to consider how my actions will affect others in the long run	Totally Agree	Y	Social	0.36	4.15	0.34
I have a hard time maintaining a healthy/special diet	disgree	Y	Food	0.35	3.30	0.59
I generally try to consider how my actions affect others	Totally Agree	Y	Social	0.34	3.93	0.04
I never think about how my behavior affects others	Totally disagree	Y	Social	0.34	4.01	-0.02
I can resist eating junk food whenever I want	Totally disagree	Y	Food	0.09	-2.24	0.00
I value the needs of the people around me	Neither agree/disgree	Y	Social	0.07	-3.59	-0.95
I've tried to work hard at school/work so I can have a better future	Neither agree/disgree	N		0.05	-3.20	-1.94
I generally try to consider how my actions affect others	Neither agree/disgree	Y	Social	0.05	-3.78	-1.20
There's no point in considering how my actions affect other people	Agree	Y	Social	0.00	-1.97	-5.62
At school/work, I always try to take the easy way out	Totally Agree	Y	Achievement	0.00	-1.99	-5.62
Generally I try to consider how my actions affect others	disgree	Y	Social	0.00	-2.20	-5.84
I enjoy spending my money the moment I receive it	Totally Agree	Y	Money	0.00	-2.47	-6.08
I'd rather take the easy way out in life to get ahead	Agree	Y	Achievement	0.00	-3.01	-6.48
When someone gives me money, I like to spend it right away	Agree	Y	Money	0.00	-3.21	-6.60

vs

All significant at the $p < 0.05$ level





Conclusions

- The construction of cultural profiles **X**(C) for a population C is a way to achieve the goal of CCP to "examine similarities and differences in psychological functioning in different cultures."
- Psychometrics can be used to define cultural dimensions associated with important cultural constructs to try and quantify similarities and differences.
- These similarities and differences can be measured by constructing probabilities P(C|**X**) to determine which "cultural dimensions" X_i best distinguish the population C from other populations
- P(C|**X**) is a **prediction** model and can be visualized as a "Culturome Landscape". Dimensions where $P(C|X_i) > P(C)$ represent the cultural niche of the population, and anti-niche in the case $P(C|X_i) < P(C)$.
- P(C|**X**) can be constructed using machine learning techniques. Bayesian classifiers are particularly useful and appropriate.
- These techniques use ***predictability*** as an **objective** criterion for feature selection, to be able to choose the most predictive cultural dimensions – constructs and factors (latent variables), items (observables) - for quantifying and understanding the cultural niche of a population and its cultural profile.
- Understandability/explainability is a **subjective** criterion for feature selection.

chilam.c3.unam.mx

Top down versus bottom up...

Items?