"Exploring the nuances of democratic ideals: Insights from a semantic clustering and valence analysis in Chile"

- Sergio Chaigneau & Ricardo González

JAMES S. McDONNELL FOUNDATION

CSCN
CENTRO DE NEUROCIENCIA SOCIAL Y COGNITIVA
ESCUELA DE PSICOLOGÍA - UAI

LEAS
LABORATORIO DE ENCUESTAS Y ANÁLISIS SOCIAL

UAI
UNIVERSIDAD ADOLFO IBÁÑEZ

“Exploring the nuances of democratic ideals: Insights from a semantic clustering and valence analysis in Chile”

Collaborating Team:

- Nicolás Marchant (postdoc @CSCN UAI)
- Kevin Blanco (DNSC UAI)
- Rodrigo Lagos (DNSC UAI)
- Enrique Canessa (Engineering and Sciences UAI)
- Sebastián Moreno (Engineering and Sciences UAI)
- Diego Ramos (Engineering and Sciences UAI)
- Cristóbal Heredia (Engineering and Sciences UAI)

JAMES S. McDONNELL FOUNDATION

CSCN
CENTRO DE NEUROCIENCIA SOCIAL Y COGNITIVA
ESCUELA DE PSICOLOGÍA - UAI

LEAS
LABORATORIO DE ENCUESTAS Y ANÁLISIS SOCIAL

UAI
UNIVERSIDAD ADOLFO IBÁÑEZ

Motivation

- Recent trends in established democracies across North America and Western Europe have highlighted a concerning shift in perceptions of democratic values, especially among younger generations.
- Foa and Mounk (2016) claim there is a broad decline in the appreciation for democracy, extending beyond mere dissatisfaction with political leaders to fundamental questions about the political system itself.
- This generational divide reveals millennials' markedly lower enthusiasm for democracy compared to older generations.

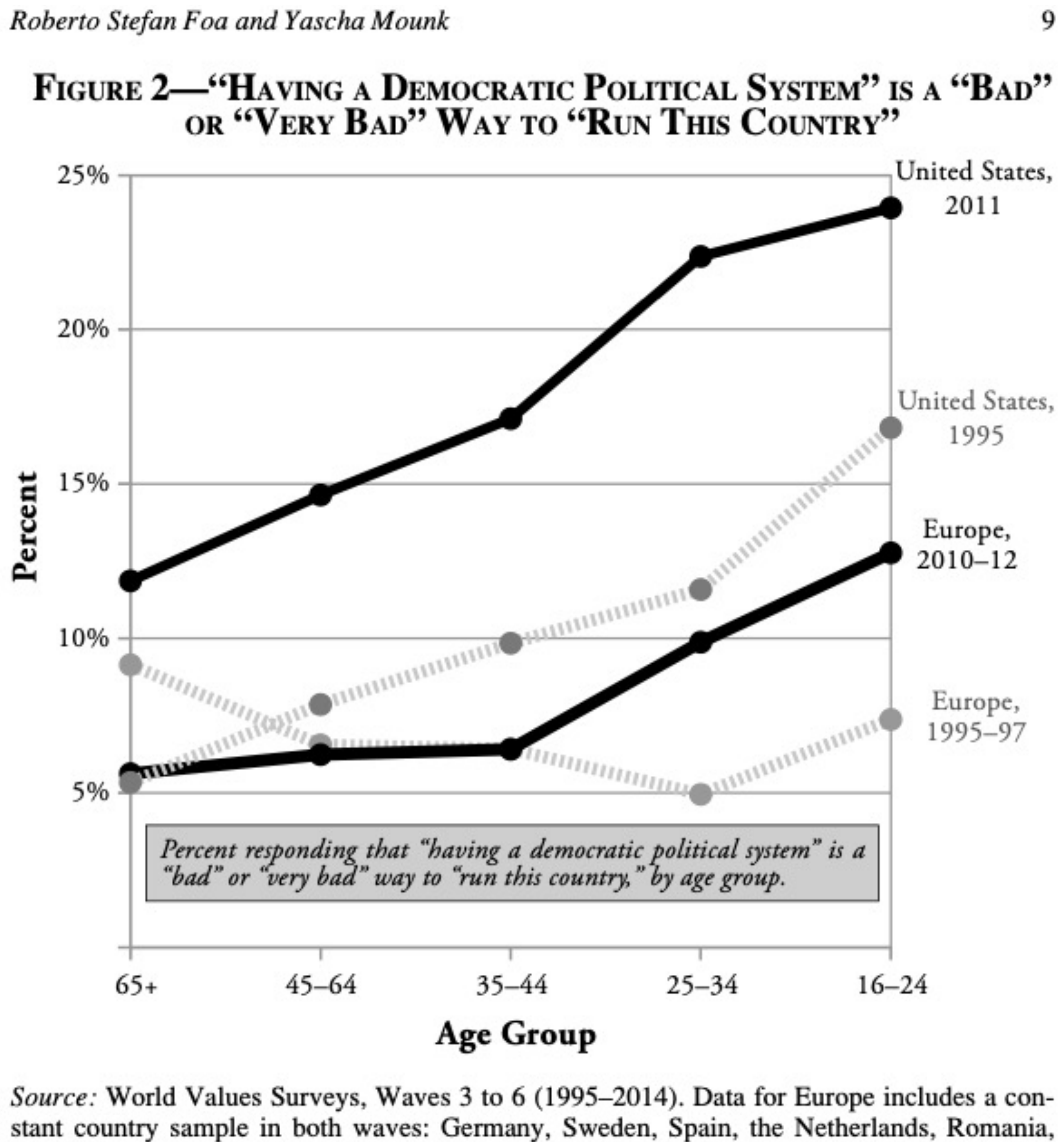
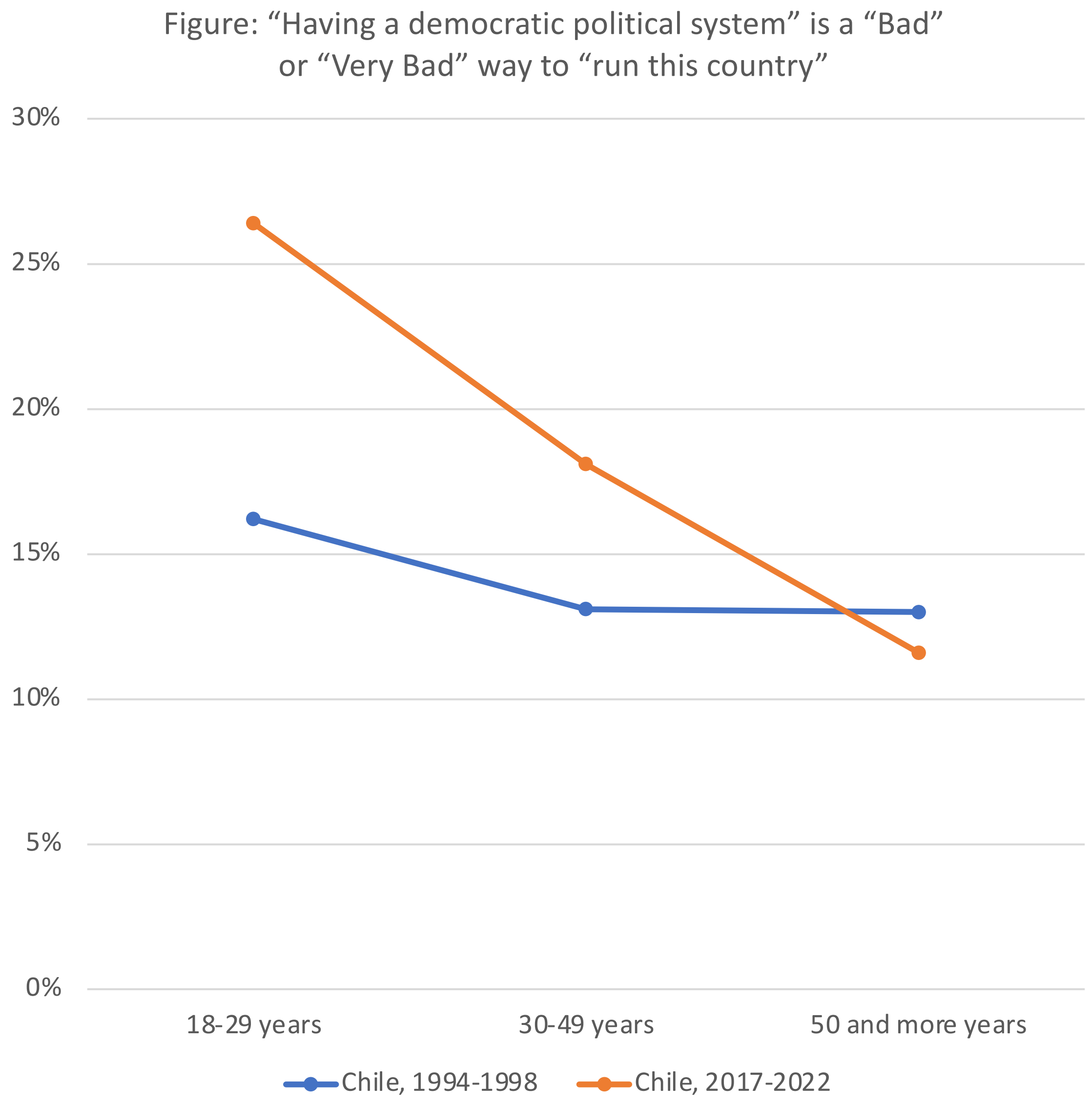


FIGURE 2—"HAVING A DEMOCRATIC POLITICAL SYSTEM" IS A "BAD" OR "VERY BAD" WAY TO "RUN THIS COUNTRY"

Source: World Values Surveys, Waves 3 to 6 (1995–2014). Data for Europe includes a constant country sample in both waves: Germany, Sweden, Spain, the Netherlands, Romania,

Motivation

- The questioning of democracy has been notably mirrored in Chile.
- This generational divide reveals millennials' markedly lower enthusiasm for democracy compared to older generations in Chile as well.



Source: World Values Surveys, Waves 3 and 7.

Motivation

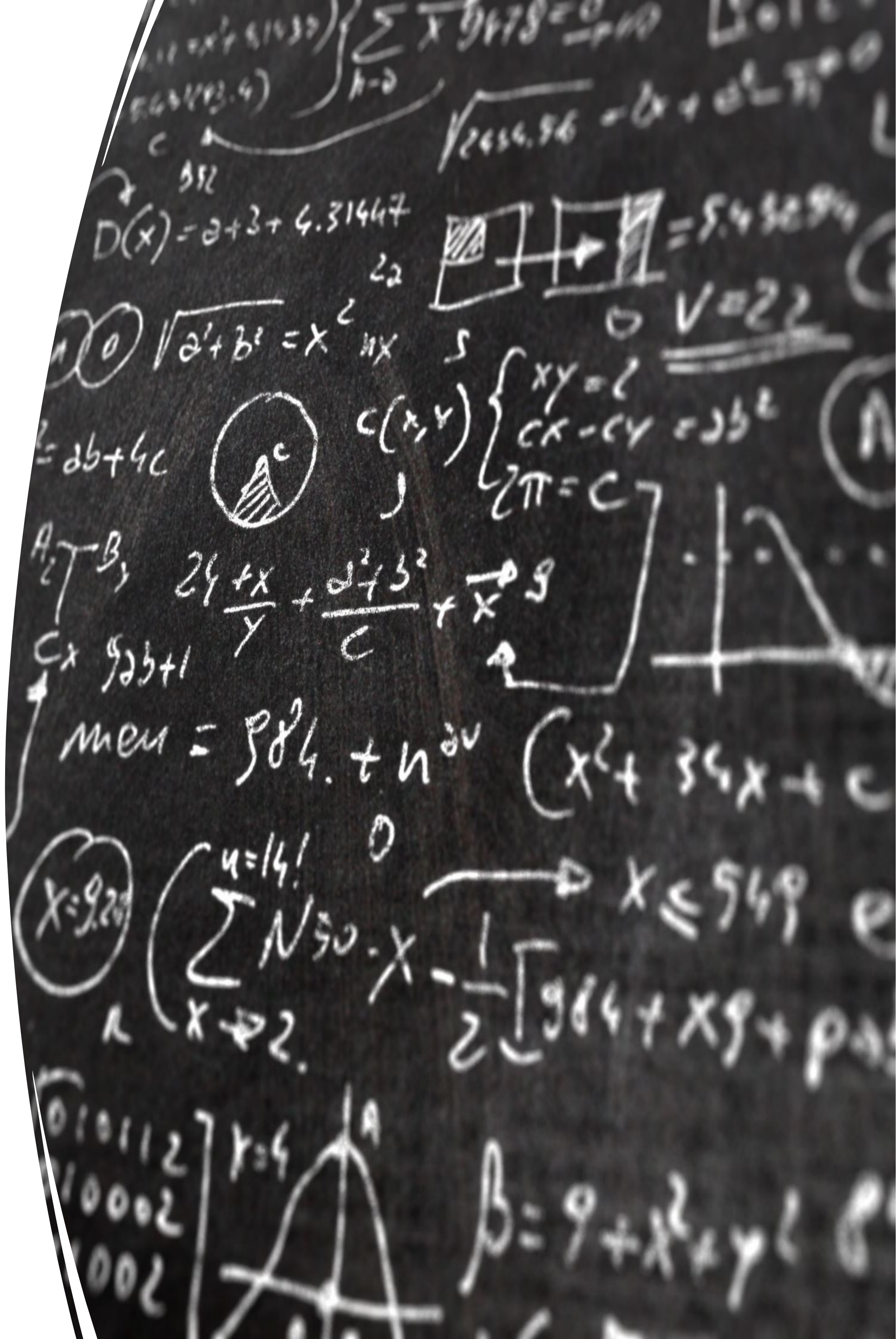
- Democratic backsliding? What does this mean? We need to know what people mean by democracy.
- A recent systematic review shows that the most common approach to this question is to predefine the conceptions of democracy to be measured; 81 out of 98 studies do this (König et al., 2023).
- Potential problems:
- Survey measures reflect assumptions about what political principles, values, and practices define democracy and therefore implicitly presume that citizens elsewhere share such assumptions (not obvious).
- Number of conceptions of democracy covered per article range between 1.5 and 1.7 concepts per study (not wide enough) (König et al., 2023).

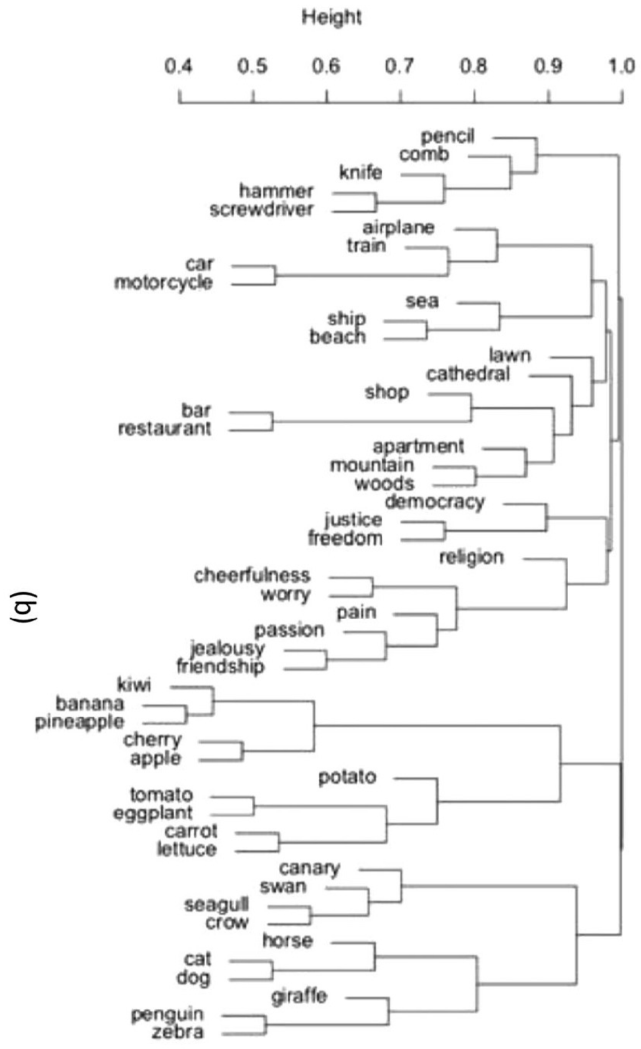
Motivation

- 17 out of 98 articles use open-ended questions to explore citizens views on democracy (König et al., 2023).
- Responses to these questions are categorized either through inductive grouping (Canache, 2012a; Doorenspleet, 2015; Miller et al., 1997) or pre-defined constructs (Baviskar & Malone, 2004; Canache, 2012b; Dalton et al., 2007; Tianjian Shi & Jie Lu, 2010).
- Potential problems
- Categorizing responses usually means imposing a structure on citizens' preferences and limiting them to specific categories.
- Reporting these responses as the number or percentage of mentions (Chu et al., 2008; Miller et al., 1997) may not capture associations between different views on democracy, potentially missing different conceptions of democracy.

Plan for today

- To gain a deeper understanding of what this democratic backsliding of democratic values means, we use the Word Association Task (WAT), a method adapted from experimental cognitive psychology to study social groups.
- We show how the WAT can be used to explore semantic memory for social concepts like *democracy* and *nature*, with the underlying assumption that semantic content matters for behavior.
- We explore groups based on age in a Chilean sample.
- We offer an overview of assumptions and methods of analysis.
- We discuss preliminary results exploring *democracy* and *nature* (used here as a contrast category), limitations and future work.
- This work is still preliminary and exploratory.





Shared semantic memory

- **Do people share semantic memory?**
- **Does semantic memory correlate with behavior?**
- Yes
  - For concrete concepts (e.g., objects)
  - For abstract concepts (e.g., stereotypes → prejudiced behavior)

Conditions for semantic memory to be shared

- **Stable structuring factors**
- Common perceptual experience.
- Common linguistic/cultural environment.
- Generally assumed to be true of concrete concepts.

- **Can these factors operate on abstract concepts?**
- Yes.
  - Stereotypes.
  - Why not for other abstract concepts, such as *nature*, *democracy*, *climate change*?
  - Just as stereotypes, these seem to be influenced by social group, current cultural dynamics, have cognitive and emotional components and correlate with behavior.
  - It is known, however, that abstractions pose particular challenges even if studied in the lab.

Summary of assumptions

I. We assume that when people search Semantic Memory (SM) to produce their lists, words that occur nearby in their list come from the same area of SM.

II. To the extent that people belong to a true social group (i.e., shared environment), then their lists should be able to be combined into “representative” lists.

Word Association Task (WAT)

WAT

Participants are asked to list words or word combinations (e.g., Department Head) that come to mind when thinking of a target concept.

Advantages

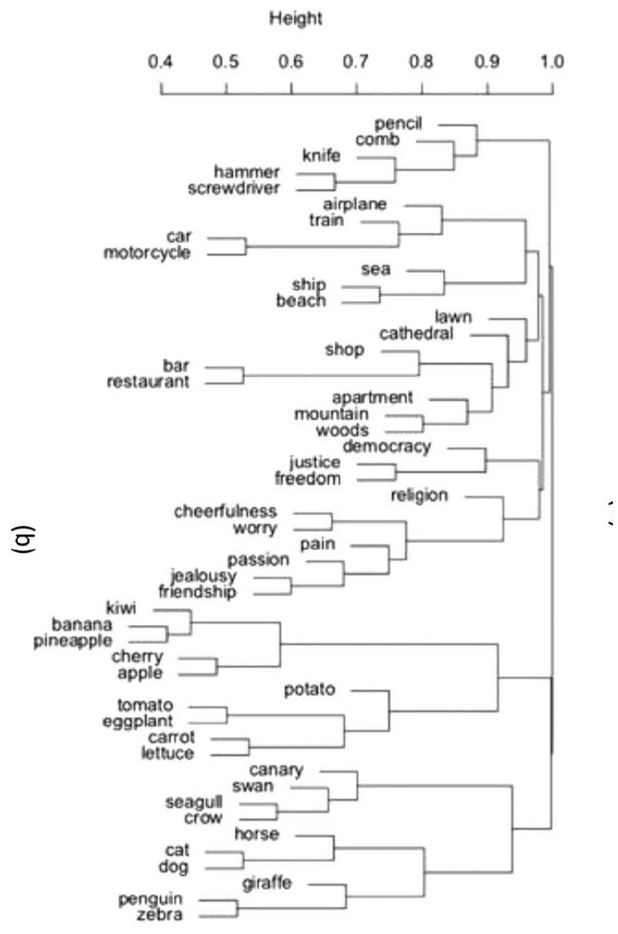
Ease of Collection: simple instructions, fast to collect.

Naturalness: we impose minimal structure by the way questions are framed.

Innovative Analysis Methods: offer both qualitative and quantitative insights.

Sample and Procedures

- Participants freely provided a variable length list of things they associated with *democracy* and *nature*.
- Data was collected through a self-administered web-based survey programmed in Qualtrics, targeting Chilean adults aged 18 and older.
- Respondents were selected using a non-probability sampling method, applying quotas for age, education level, and geographic regions to ensure a demographically diverse sample. This data collection took place between August 31 and September 12, 2023.
- For analyses we report here, we focus on a N = 419 sample (young = 268 (18 – 34 yrs.), old = 151 (55 and over))



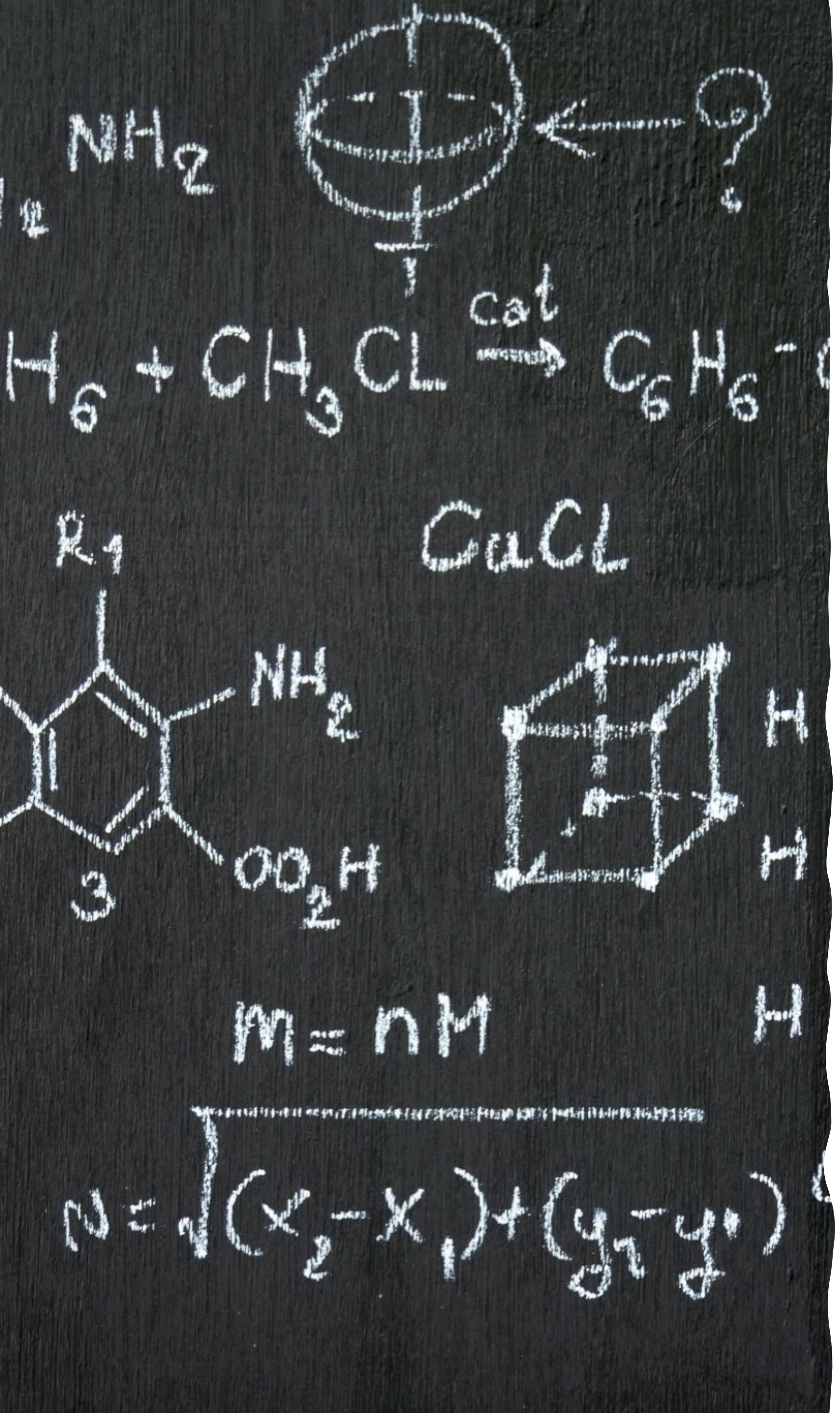
Tools for analyses

Coverage: loosely speaking, percent of SM (shared + ideosyncratic) that has been sampled. This is our interpretation of "representativeness".

Valence: averaged subjective ratings of positive or negative emotional valence on a 9-point scale (1 = unhappy, 9 = happy).

List distance: a similarity measure that takes distances between words in lists as index of semantic distance.

Heat maps and clustering: reflect association strength based on the list similarity measure.

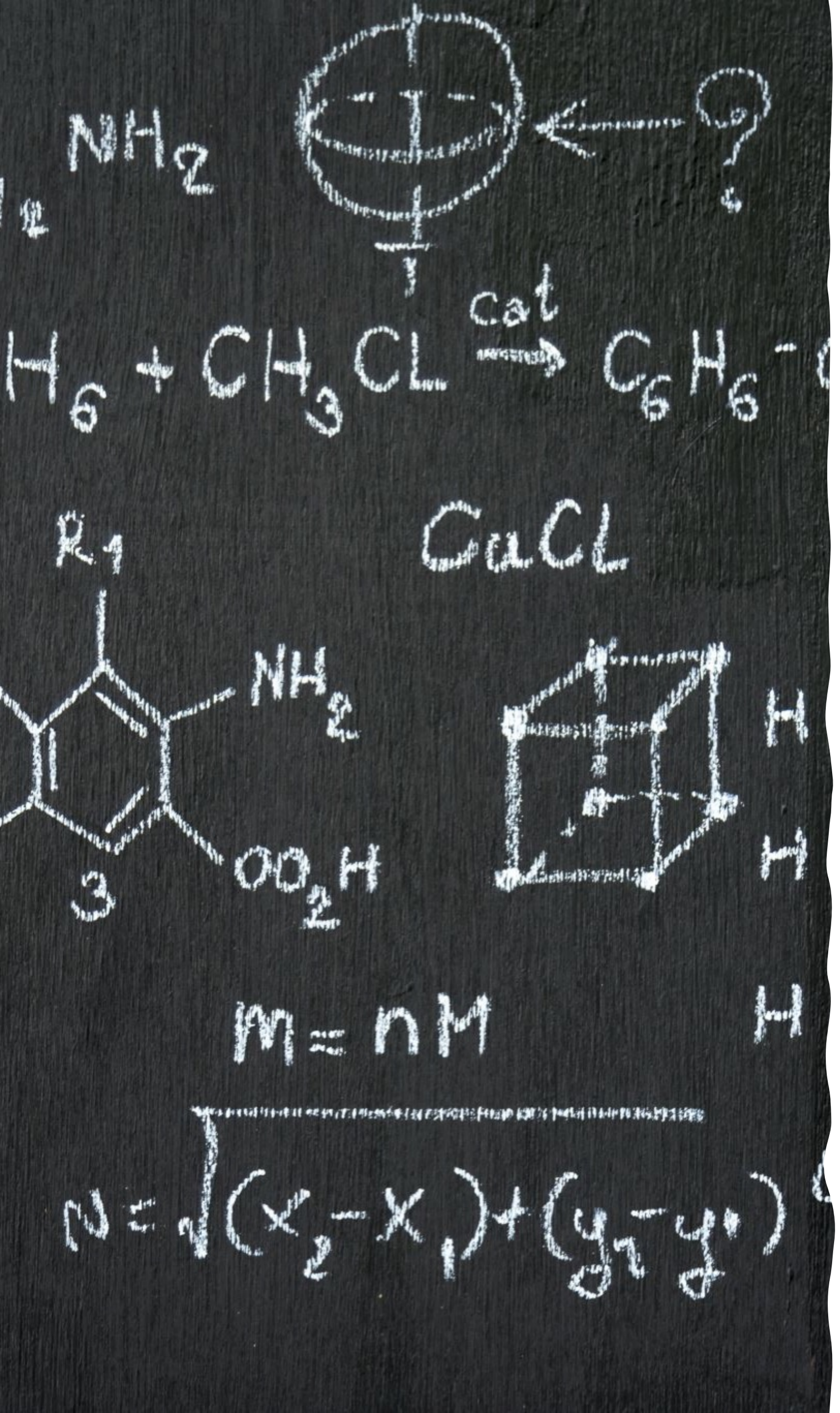


Coverage

- When collecting lists, have we “covered” sufficiently possible responses to the target word?
- Coverage is an estimate of the percent of total available responses in the relevant population that has been captured in the current sample. “Completeness” is another way of thinking about this.
- The corresponding intuition is that once words start to repeat across lists (Q1 approaches zero), then full coverage is approached

$$\hat{C}(T) = 1 - \frac{Q_1}{\mathcal{U}}\left[\frac{Q_1\,(T-1)}{Q_1\,(T-1) + 2Q_2}\right] \quad (7)$$

, where $\mathcal{U} = \sum_{k=1}^{T} kQ_k = \sum_{i=1}^{S_{obs}} Y_i$ (i.e., total number of properties listed by all participants for a concept).



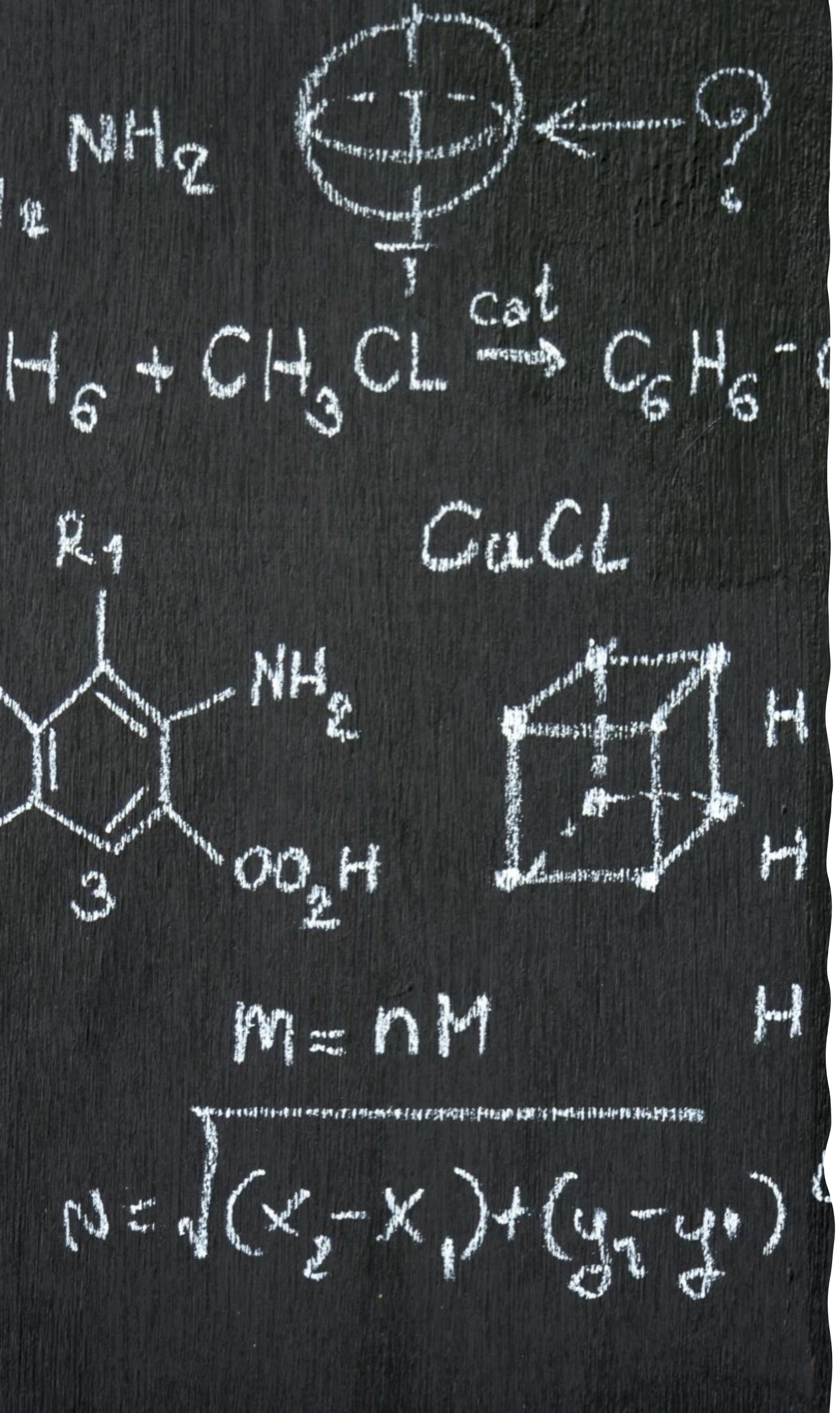
Coverage

- Importantly, to compare groups, similar and relatively high coverages ensure that there is not a large proportion of Q1 singletons, which would affect the list similarity measure by increasing zero values in the similarity matrix (as discussed next).

$$\hat{C}(T) = 1 - \frac{Q_1}{\mathcal{U}}\left[\frac{Q_1\,(T-1)}{Q_1\,(T-1) + 2Q_2}\right] \tag{7}$$

, where $\mathcal{U} = \sum_{k=1}^{T} kQ_k = \sum_{i=1}^{Sobs} Y_i$ (i.e., total number of properties listed by all participants for a concept).

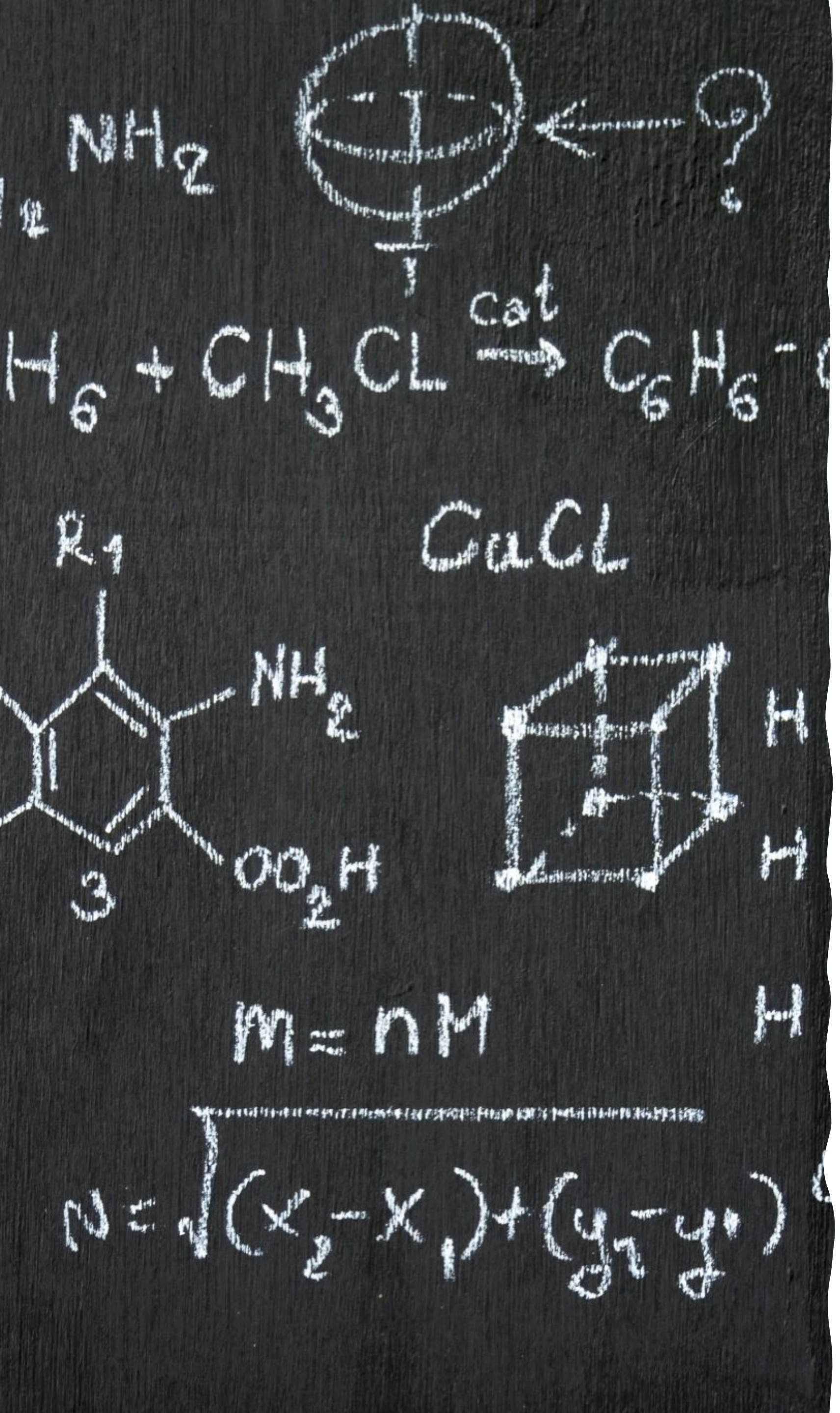
Chao, A., & Jost, L. (2012). Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology*, *93*(12), 2533–2547. https://doi.org/10.1890/11-1952.1



Valence

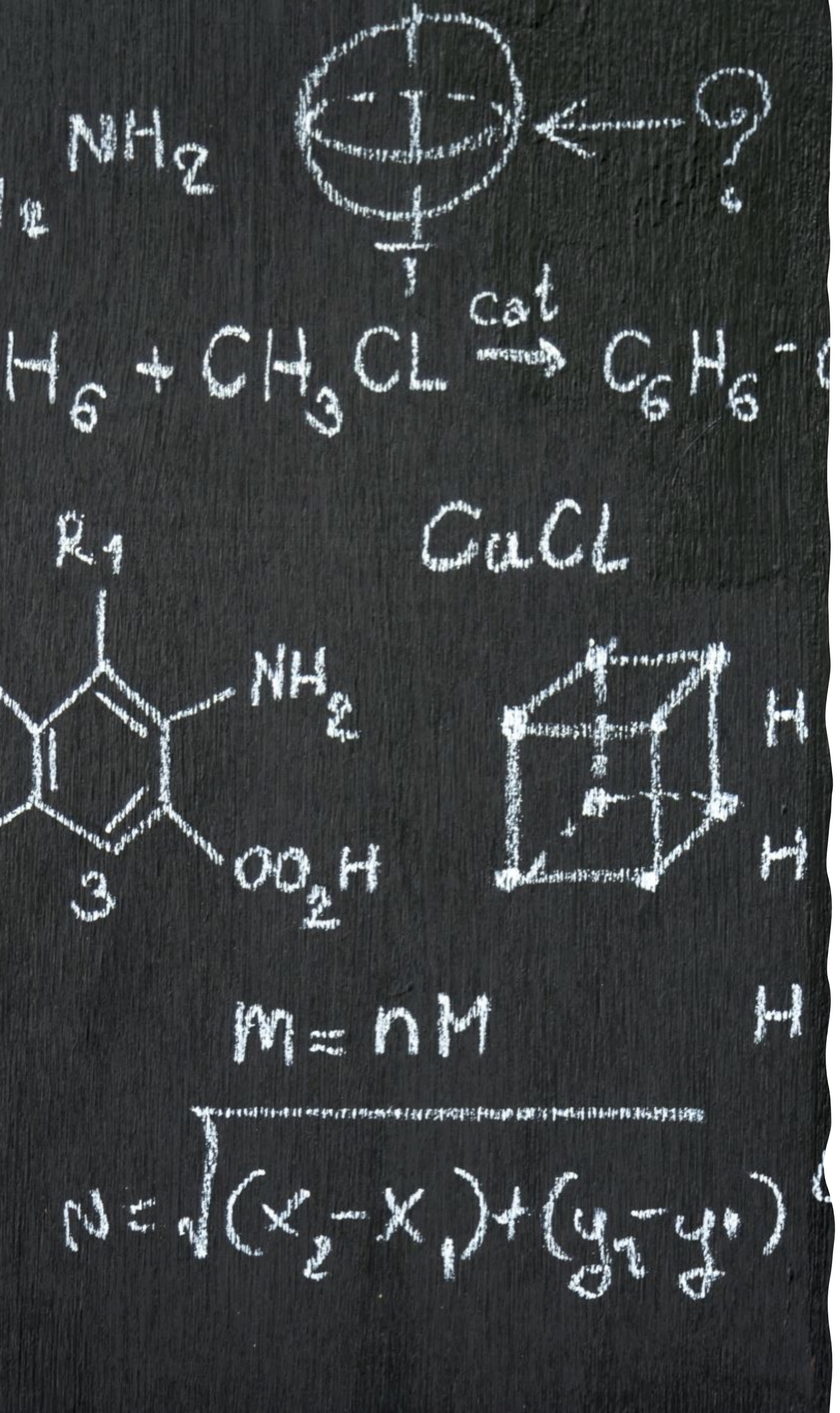
- Averaged subjective ratings of positive or negative emotional valence on a 9-point scale (1 = unhappy, 9 = happy).
- Collected in Stadthagen-Gonzalez et al. (2017).

Stadthagen-Gonzalez, H., Imbault, C., Pérez Sánchez, M.A. *et al.* (2017), Norms of valence and arousal for 14,031 Spanish words. *Behav Res* **49**, 111–123 https://doi.org/10.3758/s13428-015-0700-2



List similarity

- For each subject, similarities are computed such that words that occur close to each other have a maximal similarity, and words that are farther apart have a lower similarity. By averaging similarities across participants, a similarity matrix can be obtained.
- 0 ≤ list similarity ≤ 1, with 0 meaning that the two words were never mentioned in the same list, and 1 meaning that both words were always mentioned in direct proximity.
- Though this method perhaps produces lower fidelity similarity data relative to other methods (e.g., direct pairwise similarity ratings), it has the advantage of allowing data collection outside the lab.



List similarity

- For example, if we were to obtain direct similarity judgments for all pairs out of just 20 words, each participant would need to provide judgments for 20*19/2 = 190 pairs.
- Assuming that each judgment requires 20 seconds, each participant would require 20*190 ≈ 63 minutes. This, not even considering instructions, set-up and rest periods.
- In contrast, the word association task typically takes 1 minute or less per participant.

Analysis Pipeline

People produce lists of words associated with concept X.

Collect lists for analysis.

Analyze the Coverage of the initial lists.

- If Coverage is insufficient, collect more lists.
- Relatively high and similar coverages avoid a source of ambiguity in posterior analyses.

Compute average list valence for each subject.

Compute List Distance to generate a distance/similarity matrix (representing the group's Semantic Memory structure).

Generate heatmaps and dendrograms based on the computed distances and clusters.

Select "Representative" or "closely knit" words that best represent the different groups.

Coverage for samples of different age levels listing for democracy

	Q_1	Q_2	T	S_{obs}	U	S_{hat}	$C(T)$	T^*	S_{hat}^*
Age groups									
Young	242	51	130	380	1267	949.74	**0.81**	198	651.08
Adults	241	50	145	397	1317	973.80	**0.82**	210	658.59
Older	133	38	71	235	615	464.47	**0.79**	94	357.61

- *Note: young = 18 to 34 yrs, adult = 35 to 54 yrs., older = 55 and over*

$$\widehat{C}(T) = 1 - \frac{Q_1}{U}\left[\frac{Q_1\,(T-1)}{Q_1\,(T-1) + 2Q_2}\right] \qquad (7)$$

, where $U = \sum_{k=1}^{T} kQ_k = \sum_{i=1}^{Sobs} Y_i$ (i.e., total number of properties listed by all participants for a concept).

Coverage for samples of different age levels listing for nature

	Q_1	Q_2	T	S_{obs}	U	S_{hat}	**$C(T)$**	T^*	S_{hat}^*
					Grupos etarios				
Jóvenes	202	61	138	359	1356	691.04	**.85**	90	467.55
Adultos	253	35	153	405	1427	1313.4	**.82**	306	792.7
Mayores	170	39	80	271	741	636.88	**.77**	143	477.04

- *Note: young = 18 to 34 yrs, adult = 35 to 54 yrs., older = 55 and over*

$$\hat{C}(T) = 1 - \frac{Q_1}{\mathcal{U}}\left[\frac{Q_1\,(T-1)}{Q_1\,(T-1) + 2Q_2}\right] \tag{7}$$

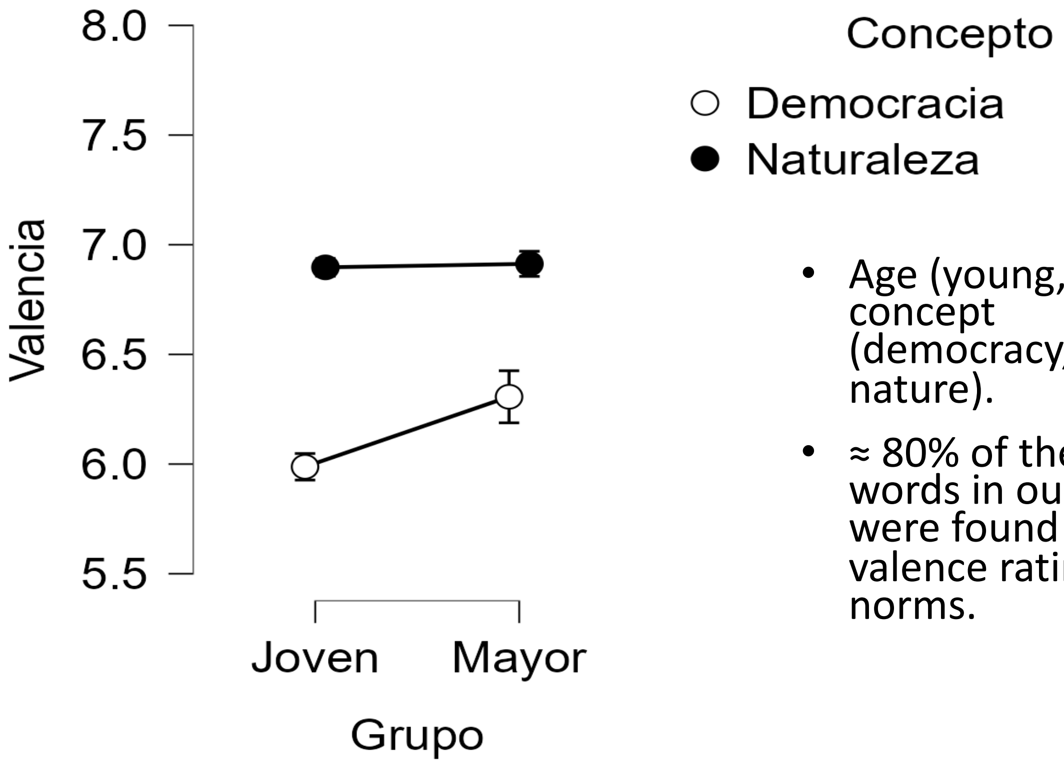
, where $\mathcal{U} = \sum_{k=1}^{T} kQ_k = \sum_{i=1}^{Sobs} Y_i$ (i.e., total number of properties listed by all participants for a concept).

Valence results

ANOVA - Valencia

Cases	Sum of Squares	df	Mean Square	F	p	η²
Grupo	2.707	1	2.707	6.134	0.014	0.011
Concepto	55.305	1	55.305	125.326	< .001	0.217
Grupo ✻ Concepto	2.213	1	2.213	5.016	0.026	0.009
Residuals	183.134	415	0.441			

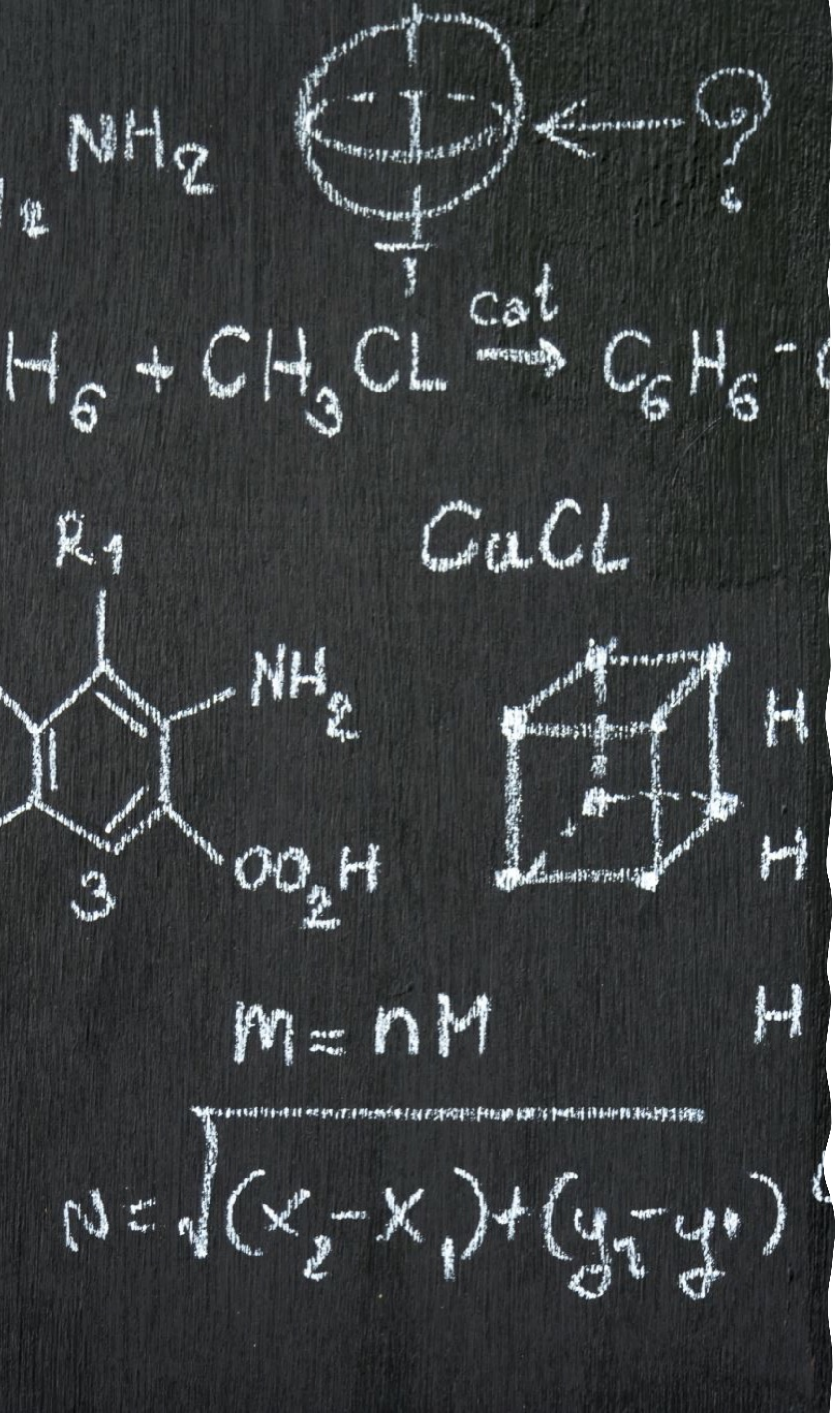
Note. Type III Sum of Squares



- Age (young,old) x concept (democracy, nature).
- ≈ 80% of the words in our lists were found in the valence ratings norms.

Data cleaning for similarity estimates

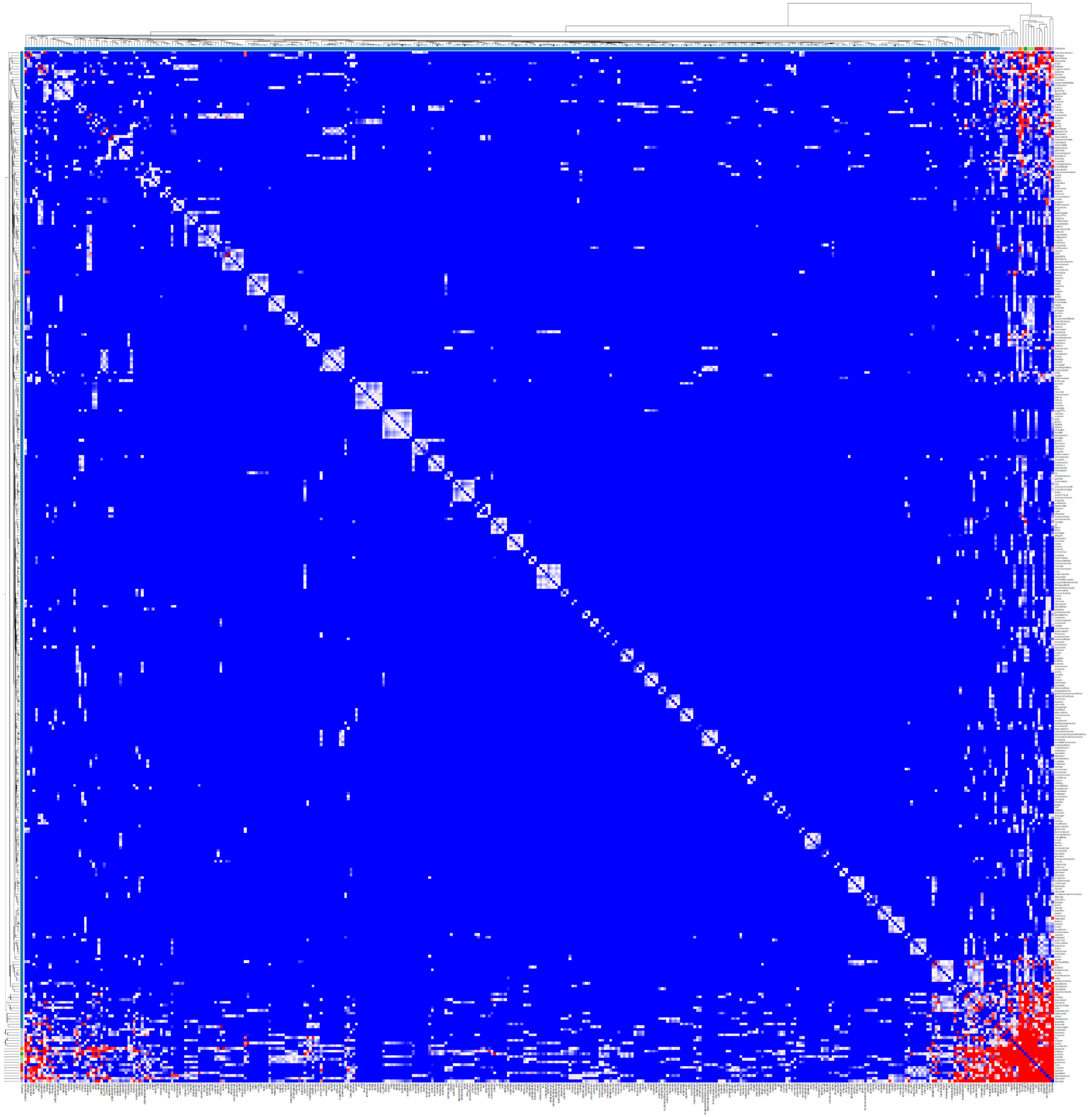
- Subjects who provided 2 or less word associates were excluded from this analysis.
- Unequal numbers across age groups
  - Young-nature = 139; adults-nature = 153; old-nature = 80
  - Young-democracy = 139; adults-democracy = 153; old-democracy = 80
- Analyses here were made by associates words rather tan by participant.
  - Democracy N = 345
  - Nature N = 372
- Recall that 0 ≤ list similarity ≤ 1, with 0 meaning that the two words were never mentioned in the same list, and 1 meaning that both words were always mentioned in direct proximity.



Heat-maps

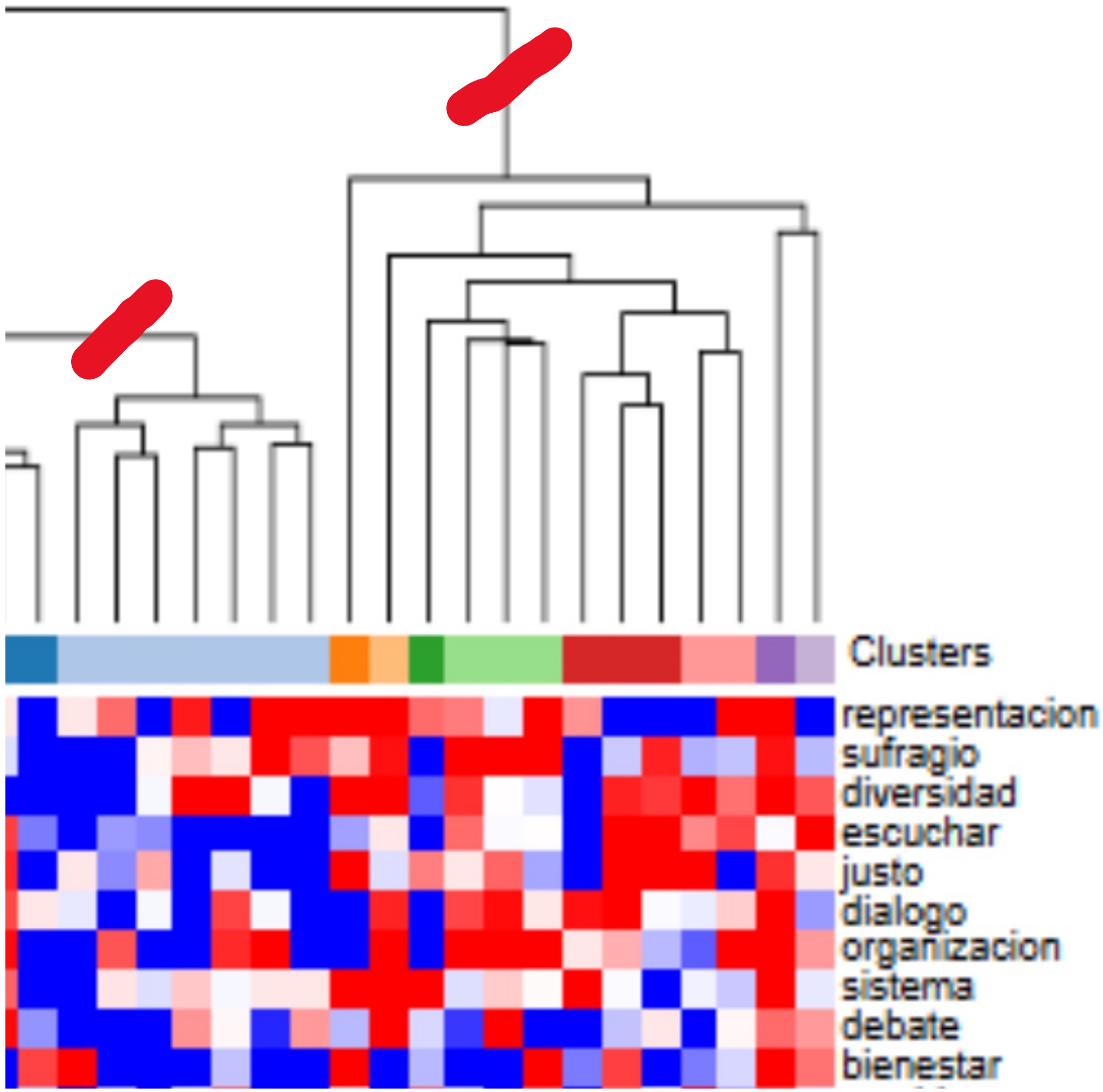
- A way of inspecting structure in our similarity matrices is dendrograms and heat-maps. This is a standard way of processing distance/similarity data in cognition research.
- For well defined groups with shared SM, they are known to capture shared high-level semantic structure (e.g., for concrete objects).
- We used standard hierarchical clustering with Euclidean distance and complete linkage.

Heat-map for young participants listing for democracy



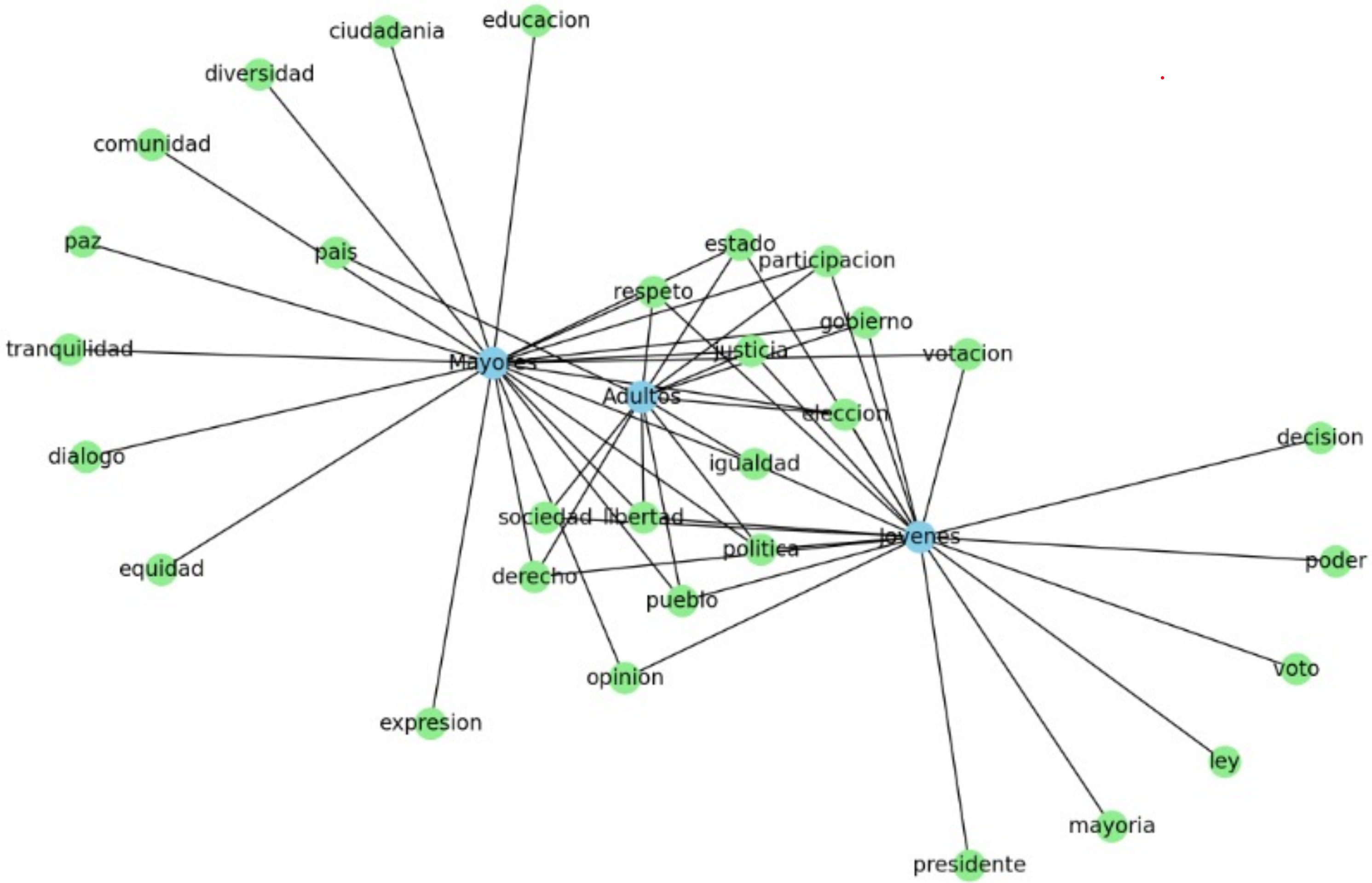
Shows area of dense shared semantic space (in red) versus more idiosyncratic content (noise).

Heat-map for young participants listing for democracy

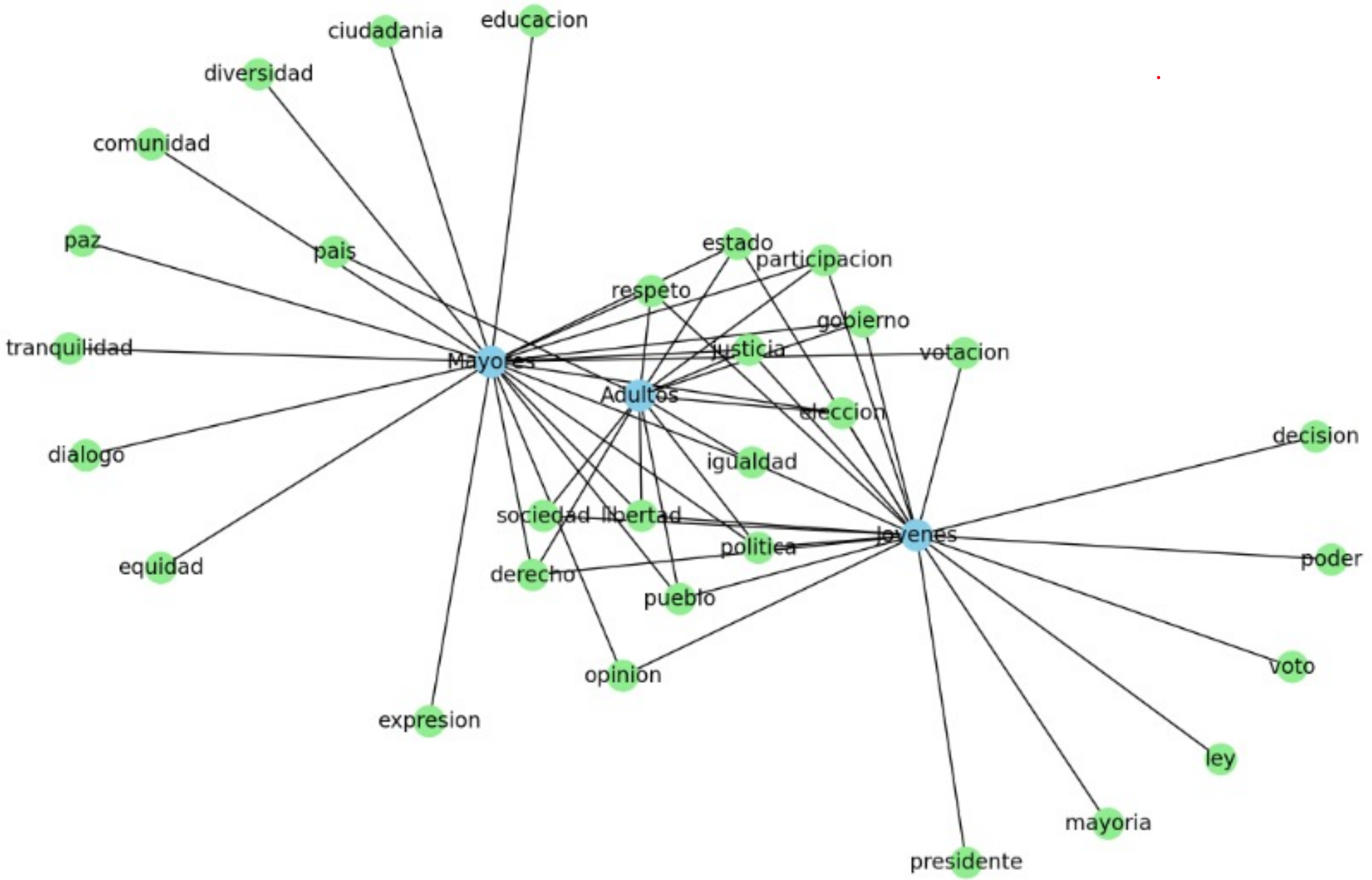


- Cuts are visually selected to obtain those clusters of words that are closer among themselves than they are to words in alternative clusters (i.e., coherent clusters).
- These correspond to high-density areas.

Semantic map comparing age level groups based on their selected "representative" words (i.e., words that tend to come in clusters for a certain group) for democracy.

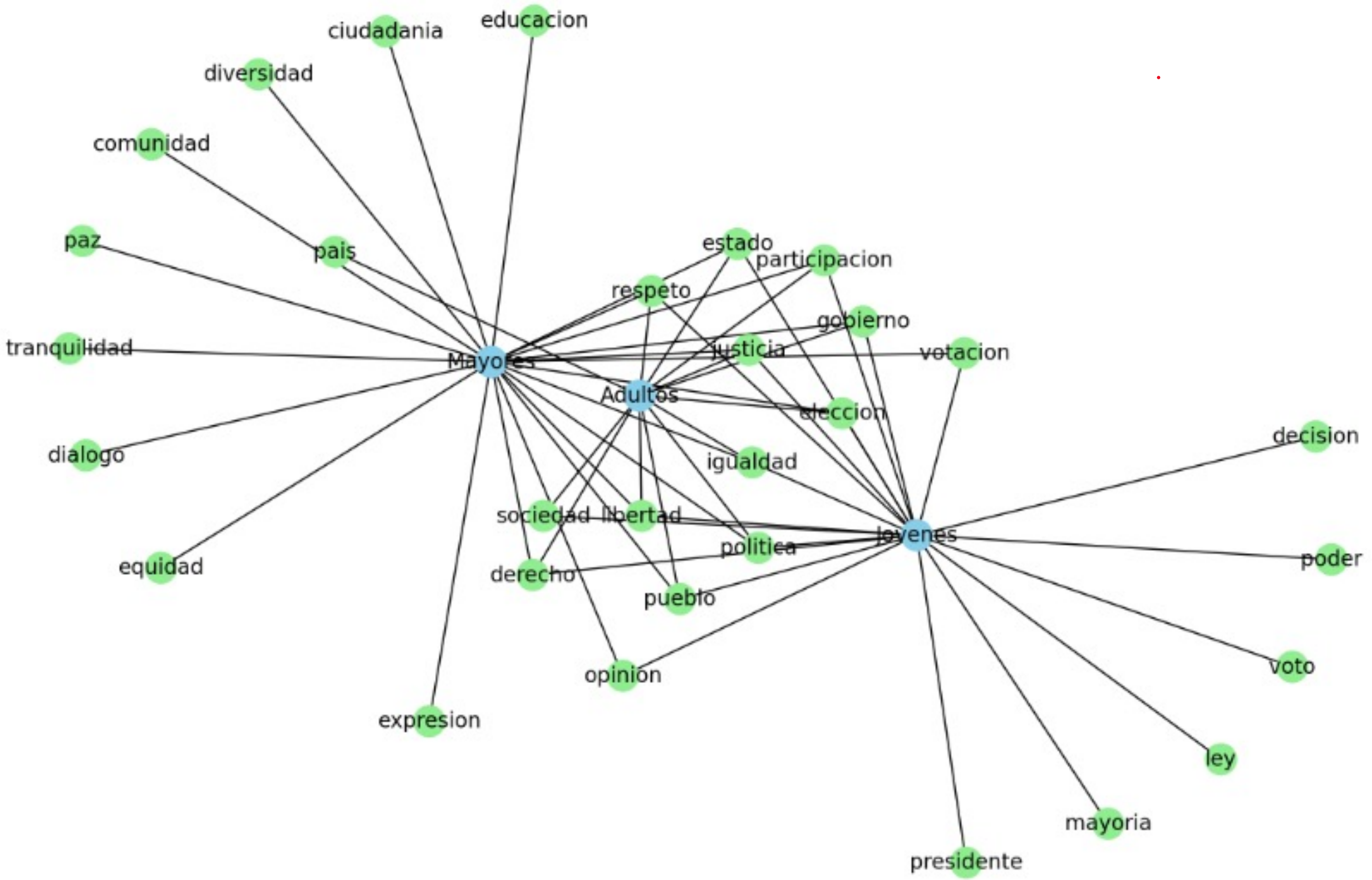


This analysis suggests that older subjects and younger subjects have a greater variety of closely knit associations relative to adults



And, these closely knit associations appear qualitatively different:

Older	Young
peace	power
diversity	majority
dialogue	decision
education	law
citizen	vote
community	president



Interim Conclusions

Our data show that it is posible to discriminate young vs old at the group level in terms of the emotional valence they attach to the concept *democracy.*

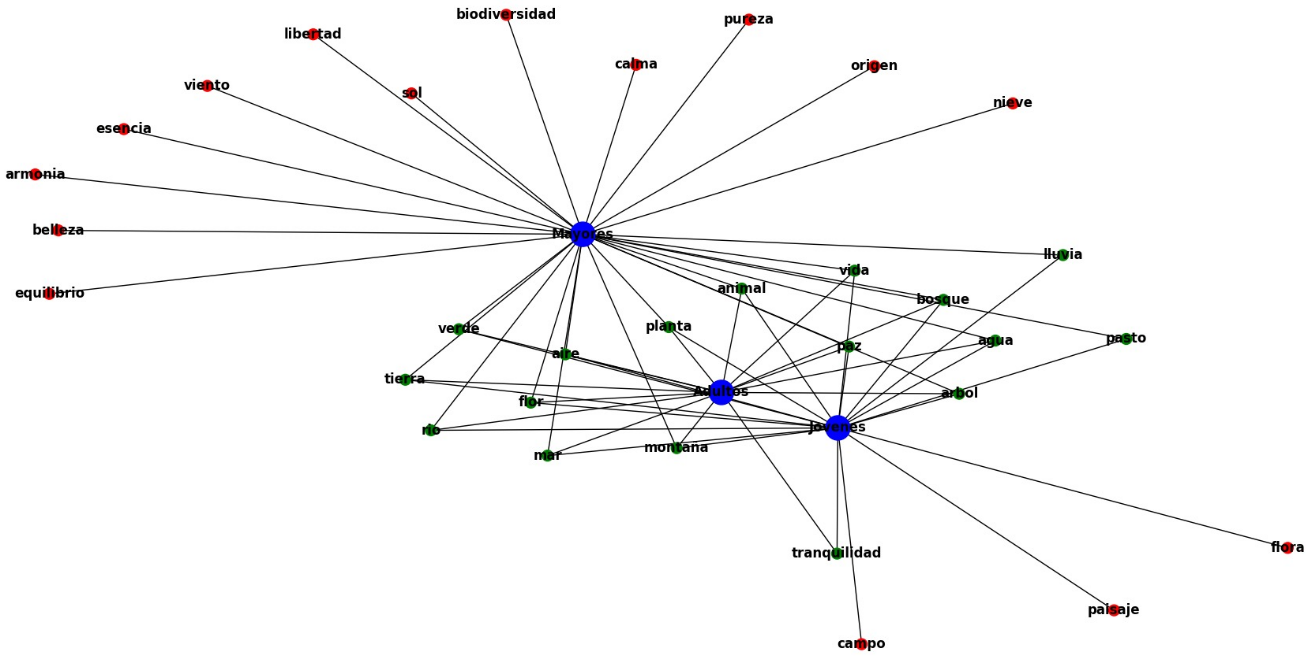
Due to factors we can only speculate about, it would seem that the concept of *democracy* changes with age (developmental or cohort effects?).

If we take content differences at face value, the concept of *democracy* for older participants incorporates the idea of *community.*

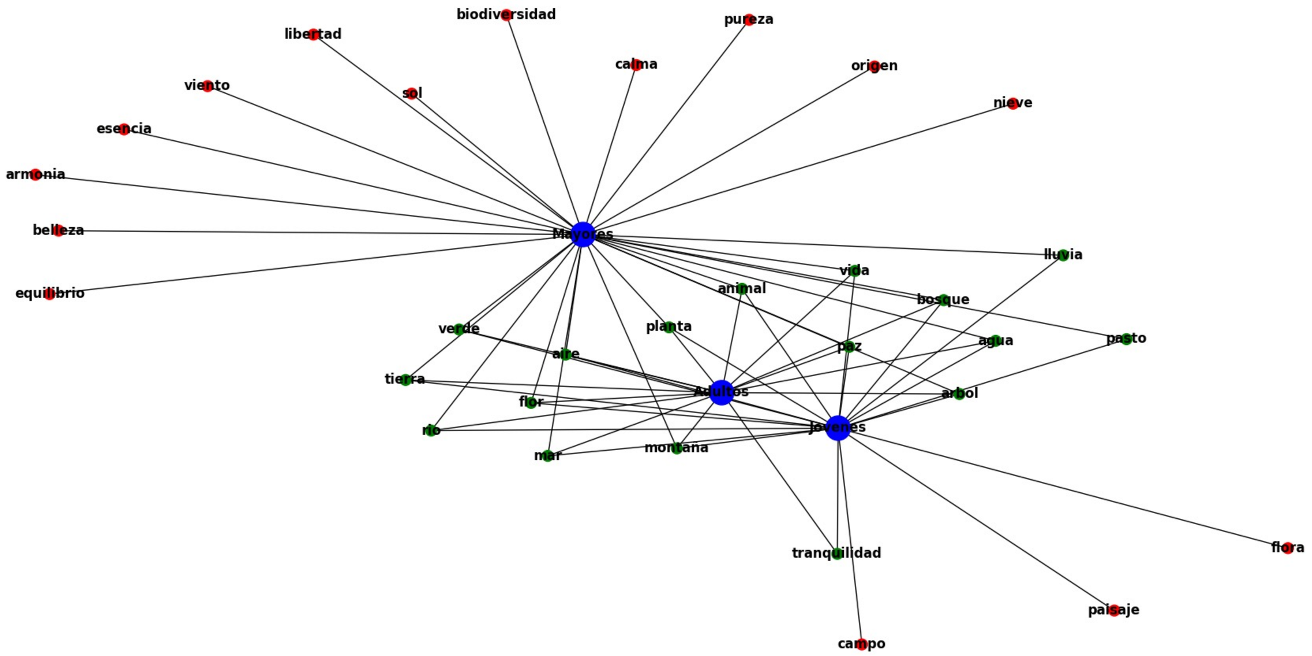
It is possible that such values account for the more positive valence in the older sample.

We now apply these same procedures to the *Nature* concept, which should provide us with a suitable contrast.

After the hierarchical clustering and homogenous clusters selection, semantic map comparing age level groups based on their selected word associates ("representative" words) for nature

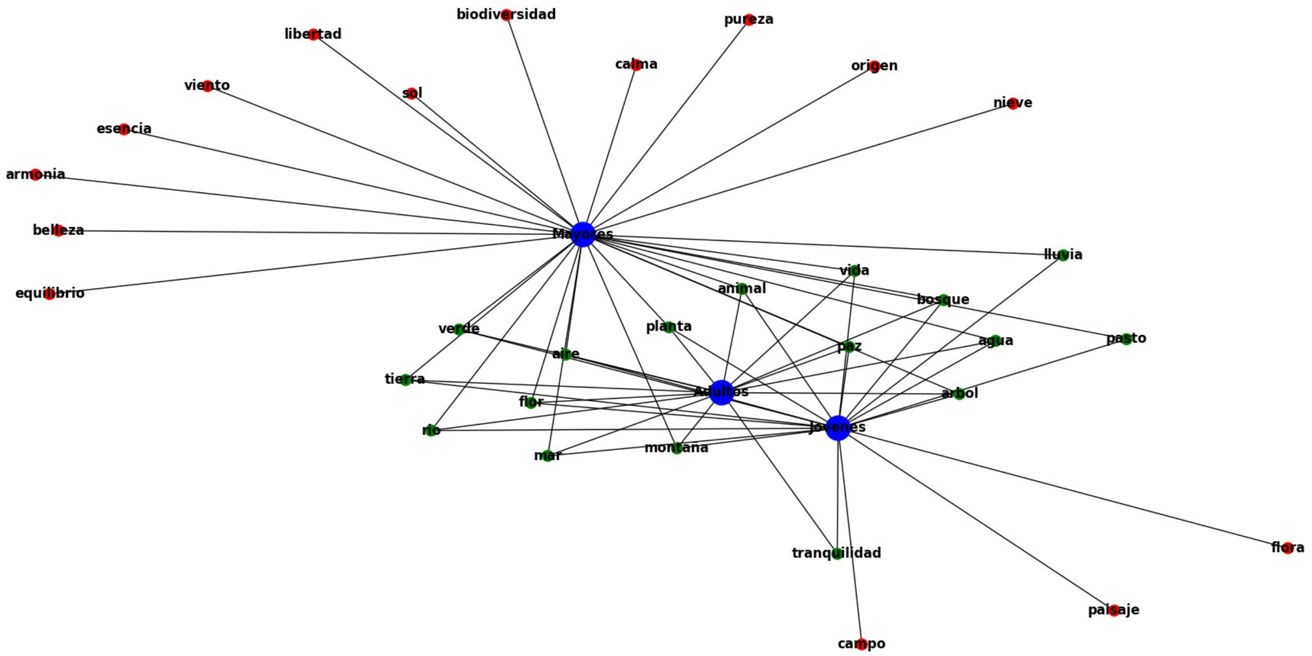


Consistently with previous analyses, older subjects seem to have a particular variety of closely knit associations that contrast with those of adults and young participants



Some of these closely knit associations include:

harmony, beauty, biodiversity, purity, origin, essence, equilibrium, calmness



These associations can be qualitatively described as "humanistic" and "aesthetic", according to Lumber et al.'s (2017) classification of pathways to nature connectedness.

harmony, beauty, biodiversity, purity, origin, essence, equilibrium, calmness

For younger and adult participants, associations seem more "utilitarian" and "naturalistic" (e.g., farmland, forest, mountain, green, water).

- Lumber, R., Richardson, M., & Sheffield, D. (2017). Beyond knowing nature: Contact, emotion, compassion, meaning, and beauty are pathways to nature connection. *PloS one*, *12*(5), e0177186. https://doi.org/10.1371/journal.pone.0177186

Conclusions for valence

Our statistical analyses on valence ratings suggest that our older sample has more positive associations with the concept of *democracy* than our younger sample.

Might this explain why opinion polls in Chile have shown a decrease among younger population of belief in democracy as a form of government?

They also show that *nature* has an overall more positive associations than *democracy*.

Conclusions for similarity

Our data suggest that it is posible to qualitatively discriminate young vs old participants in terms of their associations to the concepts of democracy and nature.

It is interesting that older participants show content differences suggesting that their idea of democracy incorporates the idea of community.

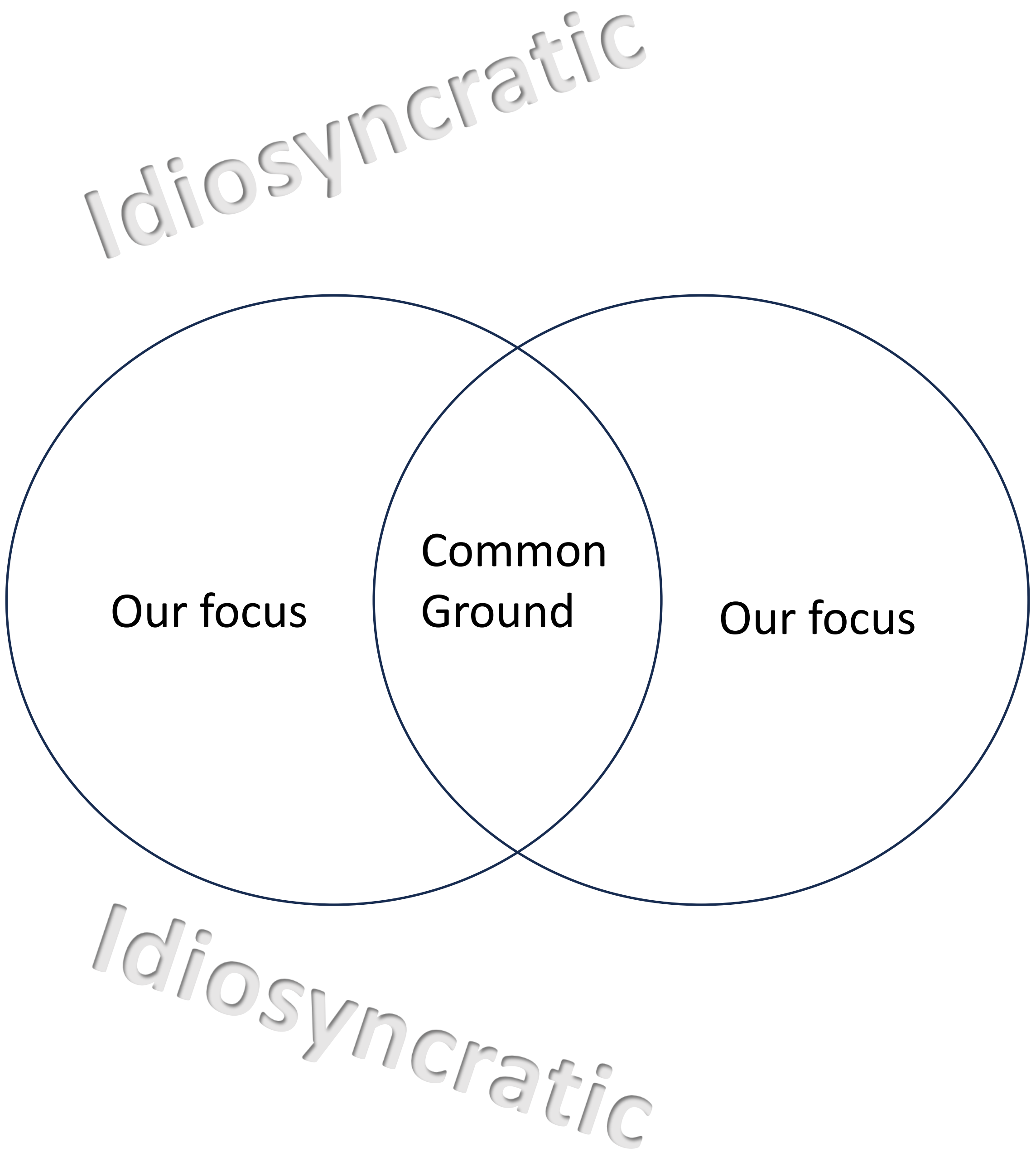
It is also interesting that older participants seem to incorporate more aesthetic and humanistic associations for the concept of nature.

We do not have a definitive explanation for these differences (cohort, developmental?).

We speculate that results from our similarity analyses relate to those of the valence analysis (i.e., values have a more positive valence).

Limitations

- Do these differences imply differences in opinion, attitude, behavior?
- Consider that our analyses take advantage of those associations that are not common across our groups, so the large common ground that our groups have may wash out the differences our data suggest.



Limitations

- Are our characterizations generalizable outside our sample?
    - Our coverage measure attempts to solve this issue. However, we cannot substitute traditional representativeness considerations.
    - We envision using a two-pronged approach, where samples need to show population representativeness and content completeness to allow proper generalizations and comparisons across populations.
- Our similarity analyses are mainly at the group level. Can they be extended to our individual subjects? (an issue of intersubject variability).

Unknowns

- We don't know when we have a group in terms of semantic memory. We are currently simply assuming that some demographic (e.g., age, educational level) also defines a social group in terms of shared Semantic Memory.
    - Our analyses suggest that this assumption is founded, because our groups tend to differ in valence for the word associates that we extracted from their lists.
    - However, we do not have criteria for predicting when a certain sociodemographic grouping should result in a distinctive SM structure. This might ultimately be an empirical issue.

Unknowns

- We have been assuming that what we get in the listing task are the most automatic or memory-available associations, and these associations should have consequences for cognitive processing, e.g., is it easier to process messages associated with this information. Our valence analyses suggest this might be true, however, more direct evidence is needed. To provide this evidence, we could:
    - Study processing at the individual level, with the hypothesis that if someone belongs to a target group, then the semantic structures being uncovered should have an effect in cognitive processing. For example, we could explore if the uncovered preferred associations make information more acceptable (e.g., argument acceptability) or memorable (e.g., that older subjects remember better or pay more attention to information about nature's aesthetic aspects)
    - We could also use our data to generate hypotheses, e.g., that younger participants agree more with claims about root problems of democracy being structural (in line with a procedural view of democracy), whereas older participants agree more with the claim that root problems involve lack of engagement (in agreement with a deeper value assigned to democracy).

"Exploring the nuances of democratic ideals: Insights from a semantic clustering and valence analysis in Chile"

- Sergio Chaigneau & Ricardo González

Thanks!!

JAMES S. McDONNELL FOUNDATION

CSCN
CENTRO DE NEUROCIENCIA SOCIAL Y COGNITIVA
ESCUELA DE PSICOLOGÍA - UAI

LEAS
LABORATORIO DE ENCUESTAS Y ANÁLISIS SOCIAL

UAI
UNIVERSIDAD ADOLFO IBÁÑEZ

References on WAT

- Ramos, D., Moreno, S., Canessa, E. *et al.* (2023). AC-PLT: An algorithm for computer-assisted coding of semantic property listing data. *Behav Res*. https://doi-org.uai.idm.oclc.org/10.3758/s13428-023-02260-9
- Marchant, N., & Chaigneau, S. E. (2022). Computational Cognitive models of Categorization: Predictions under Conditions of Classification Uncertainty. *Psykhe*. https://doi.org/10.7764/psykhe.2021.37971
- Canessa, E., Chaigneau, S. E., Moreno, S., & Lagos, R. (2022). CPN Coverage Analysis: An R package for parameter estimation in conceptual properties norming studies. *Behavior Research Methods*, doi:10.3758/s13428-022-01811-w
- Canessa, E., Chaigneau, S. E., Lagos, R., & Medina, F. A. (2021). How to carry out conceptual properties norming studies as parameter estimation studies: Lessons from ecology. *Behavior Research Methods, 53(*1), 354-370. doi:10.3758/s13428-020-01439-8
- Canessa, E., Chaigneau, S. E., & Moreno, S. (2021). Language processing differences between blind and sighted individuals and the abstract versus concrete concept difference. *Cognitive Science, 45*(10) doi:10.1111/cogs.13044
- Chaigneau, S.E., Canessa, E., Lenci, A., & Devereux, B. (2020). Eliciting semantic properties: methods and applications. *Cognitive Processing, 21*, 583–586. https://doi.org/10.1007/s10339-020-00999-z
- Canessa, E., & Chaigneau, S. E. (2020). Mathematical regularities of data from the property listing task. *Journal of Mathematical Psychology, 97* doi:10.1016/j.jmp.2020.102376
- Canessa, E., Chaigneau, S. E., Moreno, S., & Lagos, R. (2020). Informational content of cosine and other similarities calculated from high-dimensional conceptual property norm data. *Cognitive Processing,* doi:10.1007/s10339-020-00985-5
- Chaigneau, S. E., Canessa, E., & Gaete, J. (2012). Conceptual agreement theory. *New Ideas in Psychology, 30*(2), 179-189.