

Spatial representations of similarity between countries

Nimra Nadeem

Abstract

In this study, I use similarity judgment data from a diverse set of subject pools in order to construct a spatial representation of the similarity between countries, using Shepard's non-metric multidimensional scaling algorithm. My goal was to discover the underlying structure of mental representations of countries, and specifically to compare the variation in this structure across subject pools from different parts of the world. Despite limitations in the data collection method, the study reveals fascinating variation in the spatial representation of country similarity as the background of the subject pool varies. Particularly, similarity judgments for countries appear to be informed significantly by the political perception of given countries in the specific country of the subject pool.

Keywords: similarity; multidimensional scaling; non-metric

Introduction

Background

In his 1962 paper, Roger Shepard described the non-metric multidimensional scaling (MDS) algorithm, which uses non-metric similarity information between each pair of a set points to reconstruct the metric configuration of a set of points in 2D space. ((Shepard, 1962)) In his 1980 paper, Shepard applied this non-metric MDS algorithm to data on similarity judgments between colors collected in a study by G. Ekman in 1954. ((Shepard, 1980; Ekman, 1954)) His result had a fascinating circular configuration, emulating the image laid down in Newton's color wheel.

Shepard's non-metric MDS algorithm provides a useful technique for inferring the structure of mental representations based on similarity judgment data from people. The application of this MDS method has been mostly done on data from a predominantly western subject pool, i.e. mostly North American or West European subject pools. For me, collecting similarity judgment data from a diverse subject pool would be a fascinating way to discover the variation in structures of mental representation across the world.

Research Question

For this study, my goal was to discover the structure of mental representations of countries, and specifically to compare the variation in this structure across subject pools from different parts of the world.

One possible hypothesis was that this structure would emulate geographical relationships between countries, i.e. neighboring countries would be closer together while countries on

different continents would be represented as further apart. However, I thought that many other factors such as politics, culture, language, religion, familiarity, etc. could skew this representation in interesting ways depending on the exposure of the specific subject pool.

Method

The subjects were asked to take an online similarity test, built using TypeForm (credits). I used this test to collect similarity judgments on 14 different countries. The test started by asking the subjects to be seated in a quiet room without any distractions or interruptions. They were then told that there are no right or wrong answers, and the purpose of the test to get data on their individual instinctive judgments.

The test had a total of 91 questions, one for each pair of countries. I divided the test into four sections to make it less daunting for the subjects. Each section had around 25 questions. All questions on the test had the same format. For each question, I gave the subjects the names of two countries, and asked to rate on a scale from 1-10 how similar they feel the countries are, 1 meaning not similar at all and 10 meaning identical. I asked them to answer each question as quickly as they could, preferably in less than 30 seconds, because I wanted to gauge the first immediate judgment the individuals had.

Before starting I asked the subjects to rate on a scale from 1-10 how well they have understood the instructions. I then had them answer a sample question to make them familiar with the format, with two countries that weren't on my actual list of countries. Finally, I asked them to take a moment to think about the 1-10 scale and what they might consider to be a case of 1 (not similar at all), or 5 (average case) or 10 (identical). I asked them to give an example of a country pair for each of the 3 categories to make sure they actually thought about the scale before beginning.

The subjects were 20-50 years old, and all of them had grown up and spent most of their adult life in the country of choice. I collected data from 133 subjects. Unfortunately, I could not use all these responses in my eventual analysis because some countries did not have enough subjects.

My aim was to collect data from at least 10 subjects from each of at least 2 countries from each continent, to get the chance to compare across a diverse subject pool. Due to limited resources, however, I could not get the exact number of

subjects I needed from each region, so instead I used whatever responses I got. I am aware of the statistical error that the non-uniformity of subject pools would introduce, so in an extension of this study I would focus mainly on more refined data collection.

Results

Spatial representation of similarity between countries - North American Subject Pool (USA)

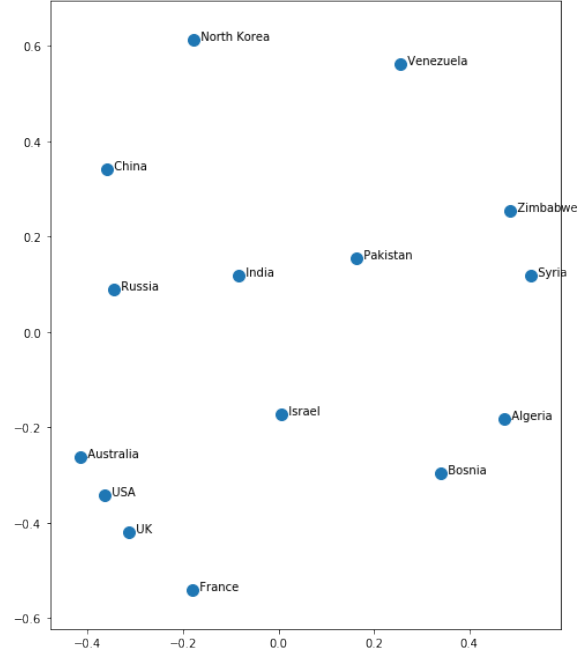


Figure 2: Results from subjects belonging to the US. The pool included a total of 20 subjects.

Spatial representation of similarity between countries - Pakistani Subject Pool

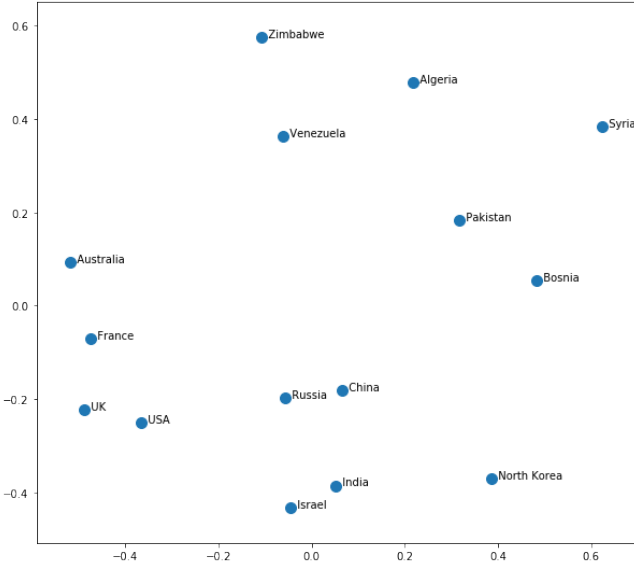


Figure 1: Results from subjects belonging to Pakistan. The pool included a total of 46 subjects.

Spatial representation of similarity between countries - Nepali Subject Pool

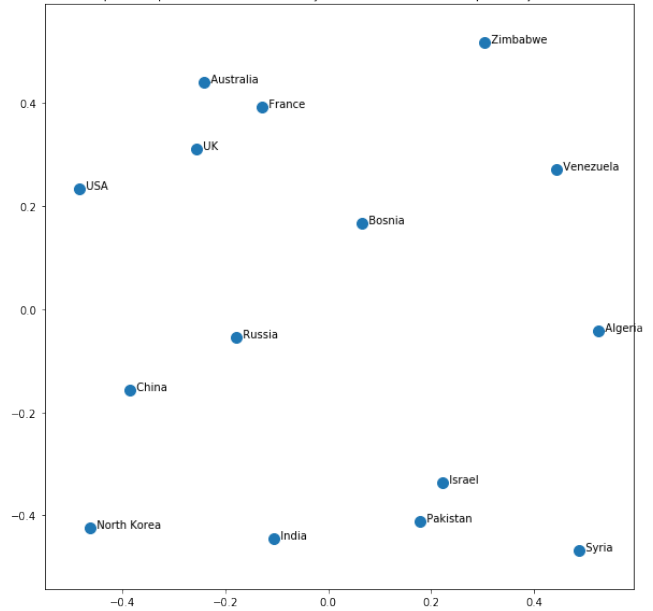


Figure 3: Results from subjects belonging to Nepal. The pool included a total of 11 subjects.

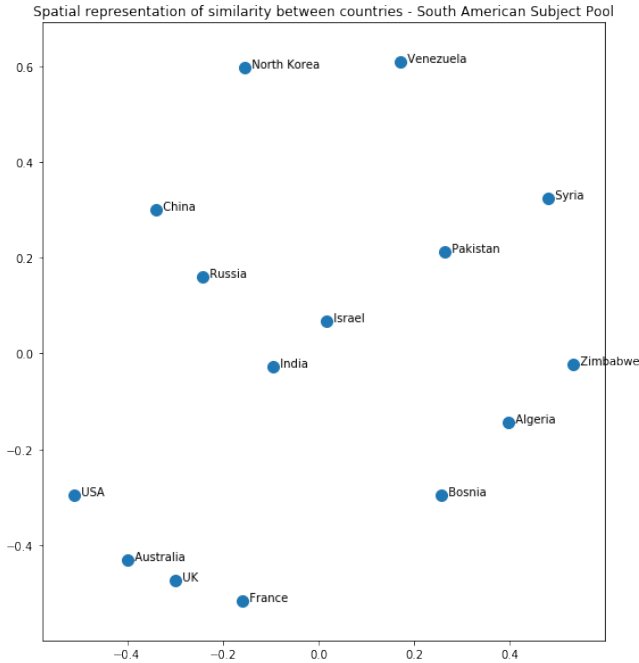


Figure 4: Results from subjects belonging to the South American region, with the following exact number of subjects from each country. Costa Rica: 6, Brazil: 2, Bolivia: 1, Guatemala: 1, Peru: 1, Uruguay: 1

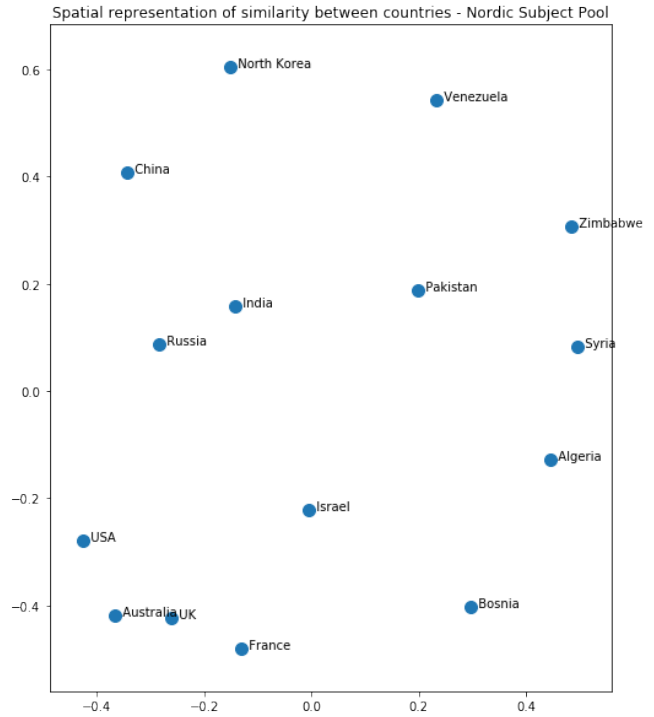


Figure 6: Results from subjects belonging to the Nordic region, with the following exact number of subjects from each country. Iceland: 8, Denmark: 4, Norway: 3, Sweden: 1



Figure 5: Results from subjects belonging to Romania. The pool included a total of 8 subjects.

Evaluation

Fascinatingly, the predominant characteristic dictating the structure of these mental representations appears to be politics. China, Russia and North Korea are grouped closer together by almost all subject pools. I did not manage to collect enough data from subjects belonging to any of these 3 countries, or a country politically and culturally close to them. The countries I did collect data from have a dominant conception of these countries being politically aligned, often as the block opposed to the US, so that is probably why they are grouped together. Of course, China and Russia are also geographically close, which is all the more reason for people to perceive them to be similar.

In a similar vein, France, UK, USA and Australia are grouped together by almost all subject pools as well. This grouping is clearly not geographical, as this group contains countries from 3 different continents. These 4 do appear to be politically aligned, often perceived to be in the pro-US block. They are also developed countries with strong economies and a dominant imperial influence across the globe. USA, UK and Australia also share the same language as a result of British colonization, which makes them even more similar. Across diverse subject pools, these similarity judgments do not seem to vary much.

The most fascinating results are on the more politically unstable countries, with more complicated bilateral relations. Notice that almost all subject pools perceive India and Pak-

istan as close together, because of the obvious cultural, linguistic, geographical, historical and religious similarities. However, the Pakistani subject pool perceives India as further away because of the dominant nationalistic sentiment in the country, and the tendency to reinforce themselves as unique from their significant political rival. In fact, the Pakistani subject pool perceives Israel and India to be closer together than any other subject pool, because the general public resentment against Israel in Pakistan is incredibly strong, so to group it closer to the most significant political rival appears more natural.

Most subject pools group Pakistan and Syria relatively close, possibly because of their common religion, and political instability. Israel being an extremely controversial country, has a widely varying relative position in the mental representations. Venezuela, Zimbabwe and Algeria are perceived by the Pakistani subject pool as relatively unfamiliar countries, many people don't hear a lot about any of these. So they tend to group them as further away from all the others.

The subject pool from Nordic countries is mostly politically neutral, educated and globally aware individuals, so the locations of Bosnia and Algeria appear closer to their geographically closer countries. The Nepali subject pool perceives India as close to communist North Korea, the two forces of political and economic oppression of the country. They also group the politically unstable countries caught in the Islam-centered conflicts (Israel, Pakistan and Syria) as closer together. Similar to the Pakistani subject pool, Zimbabwe, Venezuela and Algeria are perceived as further from all the others, possibly because of unfamiliarity.

Limitations

There were several limitations of the method employed in this study. Most significantly, the variable subject pool sizes introduce a significant discrepancy in the statistical quality of the data, and clearly I did not manage to collect data for enough subjects from enough countries to have a more complete picture of the variation in the structure of mental representations around the globe. Also, while collecting the data through an online form makes it possible to reach a larger set of people, it makes the set up less controlled and makes it harder to ensure that everyone understands and follows the instructions properly. I also got feedback from most people that they felt the test was too long. While I couldn't change the number of questions if I wanted to consider 14 countries, I think the length of the test appeared more daunting because of the impersonal nature of the test. I feel that if the subjects were asked to answer questions in a more interactive second, it would feel less lengthy.

The most important step in fixing the problems with this study would be to improve the data collection method, conducting it in a more uniform and controlled manner.

Conclusion

Even with all the above limitations, however, the data reveals patterns in the structure of mental representations that make

a lot of sense given the demographic of the respective subject pool. The results give a fascinating insight into a potentially novel method of visualizing the variation in public perception of the global political scenario.

Acknowledgments

Thank you to the United World Colleges Alumni network for helping me collect data from multiple countries. Thank you to Rachit Dubey and Tom Griffiths for advising me on the topic of this paper.

References

- Ekman, G. (1954). Dimensions of color vision. *The Journal of Psychology*, 38(2), 467–474.
- Shepard, R. N. (1962). The analysis of proximities: multidimensional scaling with an unknown distance function. i. *Psychometrika*, 27(2), 125–140.
- Shepard, R. N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science*, 210(4468), 390–398.