

BCS_RSA_final_stimuli_list

Norming Study Results

This document produces a table merging the results from the BCS Noun Norming and the Color Norming studies. This table is meant to serve as a master list of potential stimuli. Stimuli are included only if they met the inclusion criteria of both the noun and color norming studies. The criteria are given below:

Noun Norming Inclusion Criteria:

1. After accounting for spelling errors and spelling/pronunciation differences across dialects, more than 70% of responses answered with a single label
2. There was no label that accounted for 70% of responses. However the combination of the top 2-3 labels did account for 70% of the responses, and they were all of the same gender.

Color Norming Inclusion Criteria:

1. Stimuli were excluded if responses to at least one of the object's four colors did not reach a 70% threshold. Multiple color responses (e.g. "blue-grey") and modified colors (e.g. "light blue") were counted as separate from the single color responses (e.g. "blue").

Below I reproduce the analysis from both of these studies, but only for the items that met the thresholds for both studies.

The final table is given below:

##	item	colorScheme	gender	label	NounLabelAccuracy	color1
## 1	bracelet	1	F	narukvica	0.7200000	1.0000000
## 2	shirt	1	F	majica	0.7600000	1.0000000
## 3	bucket	1	F	kanta	0.7733333	1.0000000
## 4	armchair	1	F	fotelja	0.8000000	0.7333333
## 5	lamp	1	F	lampa	0.8000000	1.0000000
## 6	shoe	1	F	majority	0.8266667	1.0000000
## 7	stapler	1	F	majority	0.8266667	1.0000000
## 8	guitar	1	F	gitara	0.8666667	0.9600000
## 9	mug	1	F	majority	0.8666667	1.0000000
## 10	mask	1	F	maska	0.8800000	1.0000000
## 11	iron	1	F	pegla	0.9066667	0.9444444
## 12	chair	1	F	stolica	0.9066667	1.0000000
## 13	fence	1	F	ograda	0.9066667	1.0000000
## 14	book	1	F	knjiga	0.9200000	1.0000000
## 15	candle	1	F	sveca	0.9200000	1.0000000
## 16	pencil	1	F	olovka	0.9600000	1.0000000
## 17	dress	1	F	haljina	0.9733333	1.0000000
## 18	crown	1	F	kruna	0.9866667	1.0000000
## 19	door	1	F	vrata	1.0000000	1.0000000
## 20	scarf	1	M	sal	0.8266667	1.0000000
## 21	calculator	1	M	majority	0.8800000	0.9583333
## 22	belt	1	M	majority	0.8800000	1.0000000
## 23	cushion	1	M	jastuk	0.9333333	1.0000000
## 24	comb	1	M	cesalj	0.9333333	1.0000000
## 25	ring	1	M	prsten	0.9466667	1.0000000

## 26	flower	1	M	cvet	0.9733333	1.0000000
## 27	wallet	1	M	novcanik	0.9733333	1.0000000
## 28	bed	1	M	krevet	0.9866667	0.9130435
## 29	butterfly	1	M	leptir	1.0000000	0.9130435
## 30	balloon	1	M	balon	1.0000000	1.0000000
## 31	phone	1	M	telefon	1.0000000	1.0000000
## 32	purse	2	F	torba	0.7200000	0.9333333
## 33	vase	2	F	vaza	0.7333333	1.0000000
## 34	shovel	2	F	lopata	0.7733333	0.9615385
## 35	slipper	2	F	papuca	0.7866667	0.9523810
## 36	whistle	2	F	majority	0.7866667	0.9200000
## 37	necklace	2	F	ogrlica	0.8266667	0.9523810
## 38	bowl	2	F	majority	0.8800000	0.9333333
## 39	ribbon	2	F	masna	0.8800000	1.0000000
## 40	sock	2	F	carapa	0.8933333	1.0000000
## 41	shell	2	F	skoljka	0.9066667	0.9375000
## 42	fork	2	F	majority	0.9200000	1.0000000
## 43	glove	2	F	rukavica	0.9200000	1.0000000
## 44	tie	2	F	kravata	0.9466667	0.9615385
## 45	basket	2	F	majority	0.9600000	0.9444444
## 46	fish	2	F	riba	0.9733333	1.0000000
## 47	die	2	F	kocka	0.9866667	0.8695652
## 48	bandaid	2	M	flaster	0.7333333	1.0000000
## 49	knife	2	M	noz	0.7866667	0.9473684
## 50	lipstick	2	M	majority	0.7866667	0.9130435
## 51	plate	2	M	majority	0.7866667	1.0000000
## 52	lock	2	M	majority	0.8000000	0.9130435
## 53	screwdriver	2	M	majority	0.8000000	1.0000000
## 54	luggage	2	M	kofer	0.8400000	0.9444444
## 55	radio	2	M	majority	0.8666667	0.9130435
## 56	robot	2	M	robot	0.8800000	0.9500000
## 57	helicopter	2	M	helikopter	0.8800000	0.9500000
## 58	drum	2	M	bubanj	0.8933333	0.8333333
## 59	binoculars	2	M	majority	0.9066667	1.0000000
## 60	microscope	2	M	mikroskop	0.9600000	0.8947368
## 61	truck	2	M	kamion	0.9733333	1.0000000
## 62	calendar	2	M	kalendar	0.9866667	0.7272727
## 63	hammer	2	M	cekic	0.9866667	0.9090909
##	color2	color3	color4			
## 1	1.0000000	1.0000000	0.9523810			
## 2	0.9583333	1.0000000	1.0000000			
## 3	1.0000000	1.0000000	0.9545455			
## 4	1.0000000	0.9090909	0.7692308			
## 5	1.0000000	1.0000000	1.0000000			
## 6	1.0000000	0.9523810	0.7647059			
## 7	1.0000000	1.0000000	1.0000000			
## 8	0.9333333	1.0000000	0.9500000			
## 9	1.0000000	1.0000000	1.0000000			
## 10	0.9333333	1.0000000	1.0000000			
## 11	0.8823529	0.9200000	0.9000000			
## 12	0.8333333	0.9615385	1.0000000			
## 13	0.9500000	1.0000000	1.0000000			
## 14	0.9677419	1.0000000	0.7333333			
## 15	1.0000000	1.0000000	1.0000000			

```

## 16 1.0000000 1.0000000 0.9583333
## 17 0.9629630 0.9375000 0.9166667
## 18 0.9500000 0.7500000 0.8888889
## 19 0.9500000 0.9600000 1.0000000
## 20 0.9473684 1.0000000 0.8260870
## 21 0.9473684 0.9333333 0.8636364
## 22 1.0000000 1.0000000 1.0000000
## 23 1.0000000 1.0000000 0.7619048
## 24 0.9230769 1.0000000 1.0000000
## 25 1.0000000 0.9523810 0.7272727
## 26 1.0000000 1.0000000 1.0000000
## 27 0.8787879 1.0000000 1.0000000
## 28 0.9333333 1.0000000 0.9473684
## 29 0.8260870 1.0000000 0.8235294
## 30 0.9000000 0.8750000 0.8333333
## 31 0.9047619 1.0000000 0.9523810
## 32 1.0000000 0.7857143 0.9310345
## 33 1.0000000 0.9166667 1.0000000
## 34 1.0000000 0.8125000 0.9473684
## 35 1.0000000 0.8947368 0.8636364
## 36 1.0000000 0.7272727 1.0000000
## 37 1.0000000 0.8333333 1.0000000
## 38 1.0000000 0.9230769 0.8695652
## 39 1.0000000 0.8125000 1.0000000
## 40 1.0000000 1.0000000 0.9411765
## 41 1.0000000 0.7894737 1.0000000
## 42 1.0000000 0.8421053 1.0000000
## 43 1.0000000 0.9583333 1.0000000
## 44 1.0000000 0.7647059 0.8823529
## 45 1.0000000 0.9047619 1.0000000
## 46 0.9545455 0.8333333 0.8421053
## 47 0.8421053 0.7000000 0.8888889
## 48 0.9545455 0.7272727 1.0000000
## 49 1.0000000 0.8235294 0.9583333
## 50 0.9523810 0.8571429 1.0000000
## 51 1.0000000 0.9375000 1.0000000
## 52 0.9523810 0.7368421 0.9411765
## 53 1.0000000 0.7058824 0.9444444
## 54 1.0000000 0.9615385 0.9523810
## 55 1.0000000 1.0000000 0.9565217
## 56 1.0000000 0.7500000 0.8823529
## 57 1.0000000 0.7222222 1.0000000
## 58 0.8000000 0.7142857 0.8076923
## 59 1.0000000 0.8400000 0.9047619
## 60 0.9285714 0.7142857 1.0000000
## 61 0.8947368 0.8888889 1.0000000
## 62 0.9545455 0.9047619 0.9230769
## 63 0.9411765 0.7500000 0.9523810

```

Items are then split up from this table based on color scheme and gender:

Color Scheme 1 Feminine

##	item	colorScheme	gender	label	NounLabelAccuracy	color1	color2
----	------	-------------	--------	-------	-------------------	--------	--------

## 1	door	1	F	vrata	1.0000000	1.0000000	0.9500000
## 2	crown	1	F	kruna	0.9866667	1.0000000	0.9500000
## 3	dress	1	F	haljina	0.9733333	1.0000000	0.9629630
## 4	pencil	1	F	olovka	0.9600000	1.0000000	1.0000000
## 5	book	1	F	knjiga	0.9200000	1.0000000	0.9677419
## 6	candle	1	F	sveca	0.9200000	1.0000000	1.0000000
## 7	iron	1	F	pegla	0.9066667	0.9444444	0.8823529
## 8	chair	1	F	stolica	0.9066667	1.0000000	0.8333333
## 9	fence	1	F	ograda	0.9066667	1.0000000	0.9500000
## 10	mask	1	F	maska	0.8800000	1.0000000	0.9333333
## 11	guitar	1	F	gitara	0.8666667	0.9600000	0.9333333
## 12	mug	1	F	majority	0.8666667	1.0000000	1.0000000
## 13	shoe	1	F	majority	0.8266667	1.0000000	1.0000000
## 14	stapler	1	F	majority	0.8266667	1.0000000	1.0000000
## 15	armchair	1	F	fotelja	0.8000000	0.7333333	1.0000000
## 16	lamp	1	F	lampa	0.8000000	1.0000000	1.0000000
## 17	bucket	1	F	kanta	0.7733333	1.0000000	1.0000000
## 18	shirt	1	F	majica	0.7600000	1.0000000	0.9583333
## 19	bracelet	1	F	narukvica	0.7200000	1.0000000	1.0000000
##	color3	color4					
## 1	0.9600000	1.0000000					
## 2	0.7500000	0.8888889					
## 3	0.9375000	0.9166667					
## 4	1.0000000	0.9583333					
## 5	1.0000000	0.7333333					
## 6	1.0000000	1.0000000					
## 7	0.9200000	0.9000000					
## 8	0.9615385	1.0000000					
## 9	1.0000000	1.0000000					
## 10	1.0000000	1.0000000					
## 11	1.0000000	0.9500000					
## 12	1.0000000	1.0000000					
## 13	0.9523810	0.7647059					
## 14	1.0000000	1.0000000					
## 15	0.9090909	0.7692308					
## 16	1.0000000	1.0000000					
## 17	1.0000000	0.9545455					
## 18	1.0000000	1.0000000					
## 19	1.0000000	0.9523810					

There are a total of 19 stimuli that are in color scheme 1 and are feminine.

If we are to choose the top 12, they are:

- door
- crown
- dress
- mug
- pencil
- book
- candle
- iron
- chair
- fence
- mask
- guitar

Of these, the only one with multiple answers is “mug” {šolja, šalica}. Both are feminine and this is a dialectal difference. So not too much variability.

We should consider if we want to include “fence” or another object, since that is the only larger object out of all of these.

In terms of color, white book has a 73% and yellow crown has a 75%..

If we don’t like fence, white book, or yellow crown, possible replacements are: - stapler, 3 different feminine labels, but not used interchangeably, colors are super clear - lamp: the only reason it does not reach 100% for noun nameability is because some people answered “desk lamp” as a compound

Any other replacements, one of the criterion fall below 80%.

Color Scheme 1 Masculine

##	item	colorScheme	gender	label	NounLabelAccuracy	color1	color2
## 1	butterfly	1	M	leptir	1.0000000	0.9130435	0.8260870
## 2	balloon	1	M	balon	1.0000000	1.0000000	0.9000000
## 3	phone	1	M	telefon	1.0000000	1.0000000	0.9047619
## 4	bed	1	M	krevet	0.9866667	0.9130435	0.9333333
## 5	flower	1	M	cvet	0.9733333	1.0000000	1.0000000
## 6	wallet	1	M	novcanik	0.9733333	1.0000000	0.8787879
## 7	ring	1	M	prsten	0.9466667	1.0000000	1.0000000
## 8	cushion	1	M	jastuk	0.9333333	1.0000000	1.0000000
## 9	comb	1	M	cesalj	0.9333333	1.0000000	0.9230769
## 10	calculator	1	M	majority	0.8800000	0.9583333	0.9473684
## 11	belt	1	M	majority	0.8800000	1.0000000	1.0000000
## 12	scarf	1	M	sal	0.8266667	1.0000000	0.9473684
##	color3	color4					
## 1	1.0000000	0.8235294					
## 2	0.8750000	0.8333333					
## 3	1.0000000	0.9523810					
## 4	1.0000000	0.9473684					
## 5	1.0000000	1.0000000					
## 6	1.0000000	1.0000000					
## 7	0.9523810	0.7272727					
## 8	1.0000000	0.7619048					
## 9	1.0000000	1.0000000					
## 10	0.9333333	0.8636364					
## 11	1.0000000	1.0000000					
## 12	1.0000000	0.8260870					

There are a total of 12 stimuli that are in color scheme 1 and are masculine

This is the limiting case, we should include all of them.

Color Scheme 2 Feminine

##	item	colorScheme	gender	label	NounLabelAccuracy	color1	color2
## 1	die	2	F	kocka	0.9866667	0.8695652	0.8421053
## 2	fish	2	F	riba	0.9733333	1.0000000	0.9545455
## 3	basket	2	F	majority	0.9600000	0.9444444	1.0000000
## 4	tie	2	F	kravata	0.9466667	0.9615385	1.0000000
## 5	fork	2	F	majority	0.9200000	1.0000000	1.0000000

## 6	glove	2	F	rukavica	0.9200000	1.0000000	1.0000000
## 7	shell	2	F	skoljka	0.9066667	0.9375000	1.0000000
## 8	sock	2	F	carapa	0.8933333	1.0000000	1.0000000
## 9	bowl	2	F	majority	0.8800000	0.9333333	1.0000000
## 10	ribbon	2	F	masna	0.8800000	1.0000000	1.0000000
## 11	necklace	2	F	ogrlica	0.8266667	0.9523810	1.0000000
## 12	slipper	2	F	papuca	0.7866667	0.9523810	1.0000000
## 13	whistle	2	F	majority	0.7866667	0.9200000	1.0000000
## 14	shovel	2	F	lopata	0.7733333	0.9615385	1.0000000
## 15	vase	2	F	vaza	0.7333333	1.0000000	1.0000000
## 16	purse	2	F	torba	0.7200000	0.9333333	1.0000000
##	color3	color4					
## 1	0.7000000	0.8888889					
## 2	0.8333333	0.8421053					
## 3	0.9047619	1.0000000					
## 4	0.7647059	0.8823529					
## 5	0.8421053	1.0000000					
## 6	0.9583333	1.0000000					
## 7	0.7894737	1.0000000					
## 8	1.0000000	0.9411765					
## 9	0.9230769	0.8695652					
## 10	0.8125000	1.0000000					
## 11	0.8333333	1.0000000					
## 12	0.8947368	0.8636364					
## 13	0.7272727	1.0000000					
## 14	0.8125000	0.9473684					
## 15	0.9166667	1.0000000					
## 16	0.7857143	0.9310345					

There are a total of 16 stimuli that are in color scheme 2 and are feminine.

Just take the top 12, I have no opinions about them.

Color Scheme 2 Masculine

##	item	colorScheme	gender	label	NounLabelAccuracy	color1
## 1	calendar	2	M	kalendar	0.9866667	0.7272727
## 2	hammer	2	M	cekic	0.9866667	0.9090909
## 3	truck	2	M	kamion	0.9733333	1.0000000
## 4	microscope	2	M	mikroskop	0.9600000	0.8947368
## 5	binoculars	2	M	majority	0.9066667	1.0000000
## 6	drum	2	M	bubanj	0.8933333	0.8333333
## 7	robot	2	M	robot	0.8800000	0.9500000
## 8	helicopter	2	M	helikopter	0.8800000	0.9500000
## 9	radio	2	M	majority	0.8666667	0.9130435
## 10	luggage	2	M	kofer	0.8400000	0.9444444
## 11	lock	2	M	majority	0.8000000	0.9130435
## 12	screwdriver	2	M	majority	0.8000000	1.0000000
## 13	knife	2	M	noz	0.7866667	0.9473684
## 14	lipstick	2	M	majority	0.7866667	0.9130435
## 15	plate	2	M	majority	0.7866667	1.0000000
## 16	bandaid	2	M	flaster	0.7333333	1.0000000
##	color2	color3	color4			
## 1	0.9545455	0.9047619	0.9230769			

```
## 2  0.9411765 0.7500000 0.9523810
## 3  0.8947368 0.8888889 1.0000000
## 4  0.9285714 0.7142857 1.0000000
## 5  1.0000000 0.8400000 0.9047619
## 6  0.8000000 0.7142857 0.8076923
## 7  1.0000000 0.7500000 0.8823529
## 8  1.0000000 0.7222222 1.0000000
## 9  1.0000000 1.0000000 0.9565217
## 10 1.0000000 0.9615385 0.9523810
## 11 0.9523810 0.7368421 0.9411765
## 12 1.0000000 0.7058824 0.9444444
## 13 1.0000000 0.8235294 0.9583333
## 14 0.9523810 0.8571429 1.0000000
## 15 1.0000000 0.9375000 1.0000000
## 16 0.9545455 0.7272727 1.0000000
```

There are a total of 16 stimuli that are in color scheme 2 and are masculine

Of the top 12 there are 3 that are potentially problematic: - truck/helicopter -> only larger objects - radio -> radio versus kasetefon (radio versus cassette player)

I say we replace radio with knife. The only reason knife is not higher is due to compounds, but I believe a lot of them are because “noz” is a 3 letter word, and the program didn’t allow people to enter 3 letter words. There is only a single response that doesn’t have the word “noz” in it, and all the colors come across very clearly.

I think truck/helicopter is better than the other words. So we should keep them and fence.

The stimuli:

Majority combinations we chose are:

- mug
- calculator
- belt
- basket
- fork
- bowl
- binoculars
- lock
- screwdriver

Bowl, binoculars, and screwdriver fall slightly below the 70% threshold for the dialectal measure. That is, if we split it up by BIH/HR/SRB (country), for BIH and SRB one label comes out over the 70% threshold. For HR speakers it does not. However, I still count this as dialectal as the other label is primarily driven by HR speakers.

Calculator is not dialectal.

Our Options

Option 1: Randomize

Per participant: 4 scenarios 2 genders 6 items

Let's go through an example to show how this works. So say we have scenario 4 for a masculine target. We need to build a scenario that looks like:

[Target: object 1, gender a, color x],
[competitor 1: object 1, gender a, color y],
[competitor 2: object 2, gender b, color y]

We randomly pick a color scheme: it is color scheme 1. We then randomly pick an object from cs1 that is masculine: balloon. We randomly pick a color from cs1: blue.

[Target: object 1 = balloon, gender a = masculine, color x = blue]

We now move onto competitor 1. We already know object 1 and gender a. So we just have to pick a different color in cs1. We pick: yellow.

[competitor 1: object 1 = balloon, gender a = masculine, color y = yellow]

We finally pick competitor 2. There are two genders, and we want the opposite of masculine. So we have to pick a feminine object.

We randomly pick an object from cs1 feminine objects: book. The color likewise has to be the same as before. So we are already done.

[competitor 2: object 2 = book, gender b = feminine, color y = yellow]