UNIVERSITY OF PADOVA

Department of General Psychology

Bachelor's Degree Course in Psychological Science

Final dissertation

The enigma of 5 vs. 7 in number comparison tasks

Supervisor: **Professor Mariagrazia Ranzini**
Co-supervisor: **Dr. Samuel Lepoittevin,**
Institute of Neuroscience, Louvain-la-Neuve, Belgium

Candidate: **Chiara Capozzo**
Student ID number: **2067429**

Academic Year 2024-2025

TABLE OF CONTENTS

ABSTRACT…………………………………………………………………..4

1. INTRODUCTION
    1.1- The Concept of Numerosity……………………….…………...…5
    1.2- Language and Digits…………………………….……………...7
    1.3- Number Comparison Task……………………………………….8
    1.4- Current Study……………………………………………………..9

2. METHODS
    2.1- Sample……………………………………………………………11
    2.2- Setting……………………………………………………………11
    2.3- Procedure………………………………………….…………..12

3. ANALYSIS AND DISCUSSION
    3.1- Analyses and Results……………………………………………14
    3.2- Discussion and Conclusions……………………………………..17

REFERENCES…………………………………………………………..20

ABSTRACT

The pair 5 vs. 7 seems to be particularly difficult to process in number comparison tasks. Previous studies have reported this difficulty in samples of French-speaking individuals. In this thesis I report a study aiming at understanding whether the difficulty with 5 vs. 7 is a peculiarity of French-speakers, possibly bound to language, or whether it affects Italian speakers too. The experiment consists of a number comparison task, administered to 40 Italian-speaking University students. The results show a problematic processing of the pair 5 vs. 7, with increased response times and poorer accuracy. The fact that the pair 5 vs. 7, and especially 7 itself, is related to slower and poorer performance in a group of Italian speakers suggests that this difficult processing is not related to language. Further studies are needed to understand the specific causes of this phenomenon.

1- INTRODUCTION

1.1 – THE CONCEPT OF NUMEROSITY

Numbers constitute an important and pervasive aspect of our everyday lives: working, studying, going to the grocery store, using our phones, are all activities during which we interact extensively with numbers. Numbers are written, cited, exposed in so many different settings. As humans, we managed to structure and organize them into the domain of Mathematics, enabling us to use numbers as tools to develop more and more complex instruments and machines. But how were we able to recognize the potential of numbers?

The field of research on numerical cognition investigates how our mind interacts with the perceived world through numbers, that come in different forms and different names, ranging from digits, to words, to actual quantities. The Triple Code System (Dehaene,1992), a shared and common numerical cognition model, tries to hypothesize how we internally structure numerical representations. This system, as the name recalls, is divided into three mechanisms: the Analog Magnitude Representation, the Auditory Verbal Word Frame and the Visual Arabic Number Form. The interaction of these three processes is what constitutes our ability to process different forms of numbers and the model states that even the symbolic representations (Verbal Word Frame and Visual Arabic Number Form) are tightly related to Analog Magnitude Representation. In fact, according to this model, our analog system enables us to grasp a real sense of number by getting activated each time we interact with a number word or digit. Symbols that humans derived from this magnitude representation system and that form the other two elements of the Triple Code Model, would feel empty without any kind of tangible connection to perceived reality and numerosity.

We are used to thinking that only humans have acquired a special sensitivity to numbers, but if we look at other animal species, we realize that such ability is common in the natural world. As a matter of fact, the sense of number is wired in the animal brain (Dehaene, 1997): many studies suggest that there is an underlying nonverbal system that relies on a continuous representation of quantity, or, more specifically, *numerosity,* the one described in the Triple Code Model as Analog Magnitude Representation. This nonverbal system appears to be part of natural equipment shared by different animals, ranging from crows to monkeys. From an evolutionary perspective, the ability to estimate the number of items we are interacting with, being it food or possible predators, is precious and useful. This discerning numerosity process seems to be organized as an accumulator. The act of counting itself is discrete and specific, since each number has a clear-cut meaning and correspondence to reality. Our internal system is, however, built on a continuous mechanism. Dehaene (1997) helps us to visualize it by using a metaphor, comparing our accumulator to a bowl of water that gets filled little by little, giving us the idea of numerosity, but with little precision. Similar to adding water, a variable and noisy action, adding items leads to imperfect and inconstant representations of numerical quantities. Another major model of numerical cognition is the Mental Line, which theorizes that numbers are arranged on a compressed line. This theory has been supported by different empirical data, suggesting that number representation relies on space to create an organic numerical structure (Hubbard *et al.*, 2005).

However, there seems to be differences in the speed and accuracy of numerical processing, based on the quantity of items presented. For smaller quantities there is a specific mechanism that can be observed in individuals among countries and even among species: subitizing. Subitizing is the process by which individuals recognize quantities with cardinal value lower than 4 and it appears to be much more effortless and rapid than value estimation, which involves larger numbers recognition (Trick, Pylyshyn, 1994). On the other hand, large numbers seem to be tightly connected to the accumulator mechanism cited above, going through a process of pure analog estimation. As said before, this estimation is variable, and, more specifically, its noise increases with the increment of items' quantity. This phenomenon could be explained by Weber's law, an important psychophysical law that illustrates how the perceived magnitude of quantities depends on their relative, and not absolute, value. It has been noted that our mind processes numbers like other physical magnitudes, e.g. length or volume (Moyer, Landauer,1967).

Clearly, our relationship with mathematics is more complex than this. Our numerical cognition relies on a third internal mechanism, counting, which is different from subitizing or estimating, and involves cultural symbols associated with perceived quantities. Our ability to communicate using number words and digits is almost given for granted in our society, but it is the result of an important and specific higher order cognitive structuring, that occurred only in the human species.

1.2 – LANGUAGE AND DIGITS

The acquisition of language is important to link our innate and noisy accumulator with our ability to count, using discrete units and representing numbers with words. Numbers, as we now refer to them, are born the moment in which our language gives us the necessary instrumentation to symbolize and discretize the continuous magnitude represented in our mind. As a matter of fact, in cultures that are not equipped with words even for larger numbers, but only verbally identify quantities such as "one", "two", or "many", calculations are made only approximately (Göbel *et al*, 2011). This therefore suggests the presence of a nonverbal approximate system of numerosity perception, and at the same time supports the hypothesis of how symbolic structure is needed to perform exact calculation. Evidence of this important role of language is found in how different cultures and languages are related to different mathematical skills. The transparency of a verbal numerical system seems to be partially predictive of children mathematical proficiency (Dowker *et al*., 2008). Transparency is defined as presence and correspondence of spoken and written number words and the base and regularity of the number system. Some languages, such as Chinese, Korean, or even Welsh, appear to be more understandable and easier to associate with perceived quantities, even though the concept of numerical performance is way too broad to be only explained and caused by language inflexions.

Moreover, verbal number acquisition is the first step to achieving visual digits comprehension. Children encounter both quantities and words to describe them before having the ability to connect the former and the latter. The analog magnitude system is activated very soon after birth, and the cultural names given to numbers are part of everyday life vocabulary, therefore becoming rapidly part of the child's experience. In fact, children first learn, usually around two years of age, to recite numbers in the right order, and only later do they become aware of how these words are related to already perceived quantities, acquiring then the concept of cardinality of verbal numbers. Different studies suggest that children's arithmetic skills are predicted not by their ability to discern and compare quantities of dots, but by their performance in Arabic digits comparison tasks (Knudsen *et al*., 2015). In non-symbolic number comparison tasks, a higher cognitive organization, given by digital symbols and how much children have managed to spatially organize them, appears to be a much more relevant predictor of precision and proficiency (Sella *et al*., 2020).

1.3 NUMBER COMPARISON TASK

Number comparison tasks consist of presenting participants with a pair of numbers, being them quantities, words or digits, and asking them to identify and select the larger or smaller one between the two. This kind of paradigm has been extensively used in numerical cognition, as it has highlighted many different mechanisms used by our mind to manipulate and interact with numbers.

As mentioned above, the role of order in number perception is intrinsically important, partially because we are talking about quantities, but certainly because our mind has developed to use the space dimension as a support for number comprehension. One of the main effects that have been discovered and support this theory is the SNARC effect. Apparently, individuals are faster at comparing pairs of numbers when the larger one is presented on the right side of the computer screen and the smaller one on the left. This effect, although mildly mediated by cultural experiences, for example reading direction (Göbel *et al*, 2011), is however an outstanding representation of how numbers could be arranged in our mental line, by following a left-to-right sequence, like the centimeters we see on the ruler.

Two other effects are considered steadily robust and therefore essential to be mentioned: the distance effect and the size effect. Both effects apparently rely on Weber's law of perception and were first found by Moyer and Landauer (1967). The distance effect consists in the increased difficulty experienced when comparing two quantities that are close in number to one another. This impairment is often combined with the size effect. The latter involves linearly increasing difficulty in processing and comparison of greater numbers, since our accumulator's representation loses precision as numbers grow larger. The two effects often co-occur. For example, if we had to compare two numbers, our performance would be different depending on their ratio: 2 is double 1 and hence it is clearly greater, but 47 is just a bit larger than 46. Since the data we refer to is noisy, our ability to compare close numbers decreases according to their proximity and their size. According to ANS theory, that supports an automatic and natural transformation of symbolic discrete number units into continuous internal magnitude representation, these two effects should work both in magnitude comparison tasks and in digits comparison tasks. The real effect would be, as mentioned above, the ratio effect explained by Weber’s law, that is just articulated along two different dimensions, size and distance. However different models are now trying to justify the same results by theorizing different underlying mental processes (Krajcsi, 2017), (Fias, & Verguts, 2004).

1.4 CURRENT STUDY

The aim of this study is to investigate the performance of a group of healthy adults in a number comparison task. In a recent study by Lepoittevin and colleagues (Lepoittevin, Geers & Andres, *A re-examination of the sub-base-five effect in single digit comparison*, talk at the annual meeting “Advances in Numerical Cognition Research”, 2024) the authors described an effect of impairment when comparing 5 vs. 7, in a group of French speakers. One possible explanation was that this impaired processing could be related to language: indeed, 5 and 7, when pronounced out loud in French, have similar sounds (/sɛ̃k/ and /sɛt/). This close phonetic similarity could cause a similar activation of the two digits and result in the difficulties mentioned. The aim of this study was therefore to replicate this previous study on a group of Italian speakers. In Italian, 5 and 7 are verbally represented by very different phonetic codes (/*'ʧiŋkwe/* and */'sɛtte/*), making Italian speakers an acceptable comparison group for the Belgian study. The question to investigate is whether Italian speakers show the same difficulty in comparing 5 vs. 7 as French speakers.

If our data showed no problematic processing of the mentioned pair in Italian-speaking sample, then the linguistical explanation of French speakers impaired performance would be supported. On the contrary, if our findings confirmed the presence of this effect in the Italian sample too, further studies would be necessarily needed to untangle the question. In both cases, the Discrete Semantic System model (DSS) could be used to explain the observed behavioral performance (Krajcsi, 2017). This model theorizes that the two systems, symbolic and non-symbolic, are not so tightly related as ANS describes, but can conversely work as separate units. An effect that involves activation of specific linguistic or visual circuits, related to symbolic representation of numbers, would possibly support this model, that deviates from the Triple Code System model.

In the next chapter we will explain the methods of this experiment, which stems from a collaboration with the Ph.D. Samuel Lepoittevin and Michael Andres, working at the Catholic University of Louvain in Belgium.

2- METHODS

2.1 - SAMPLE

The tested participants were 40. The inclusion/exclusion criteria were the following: being an Italian native speaker, being between 18 and 25 years old and showing no history of psychiatric, neurological or learning disorder (dyscalculia or history of Mathematical Learning Disorder in particular). Complementary data were collected before and after the task; they included: age, gender, study level, Edinburgh Handedness Index score (Oldfield, 1971), mother tongue, other languages spoken, and task order.

Twenty-three of the participants identified as females and the remaining seventeen identified as males. No participant showed criteria to be excluded. The mean age of the dataset was 21 years and 3 months, with a Standard Deviation (SD) of one year and 3 months. The most spoken second language was English, mentioned by all participants. All the participants were University students, most of them studying in Padua, showing a study level with a minimum of 13 years and the mean equal to 16 years. The Edinburgh Handedness Inventory score was computed for every participant. Almost all participants showed themselves to be right-handed: however, three participants turned out to be ambidextrous and two turned out to be left-handed. The mean EHI score was 0.73 with SD equal to 0.45.

The participants were recruited either using Google Form platform or by direct contact. All the participants were given and signed a written consent form to take part in the experiment and let the experimenter consensually analyze their data.

2.2 - SETTING

The experimental setting was as neutral as possible: all participants were tested in the same laboratory room, with no windows, same artificial light and all sitting on the same comfortable chair.

The material required for the experiment consisted just in the computer, and participants had to answer using the keyboard. The keyboard used was an AZERTY keyboard, to maintain the similarity with the Belgian part of the study. The distance between participant and screen has been kept around 60 cm, following guidelines developed for the Belgian study.

2.3 - PROCEDURE

The task was developed using PsychoPy (v2021.2.3), by Michael Andres and Samuel Lepoittevin. The task consisted in the translated version of the one used for the Belgian part of the study, to maintain the experience of the two linguistically divergent samples as close and as similar as possible. The duration of the whole experiment was about 40 minutes, depending on the speed of each participant. The task was divided into two conditions: larger or smaller conditions.

The aim was to correctly select the larger or smaller number, depending on the condition, by choosing between the two options presented and pressing either the S or L key to select the chosen number. The participant was asked to be as fast and as accurate as possible. All participants were presented with both conditions, the larger and smaller version, but the order in which the tasks were administered was counterbalanced: half of the participants started with the larger (L) task and the other half started with the smaller (S) task.

If the target number appeared on the left side of the screen, the participant was instructed to press the S key, and to press the L key if the target number appeared on the right. Before starting the actual task, in both conditions a practice section with 12 trials was presented, during which both instructions and feedback on the performance were given.

The main structure of the task consisted in the repetition of 440 trials per condition, with smaller and longer pauses at the end of different loops. Every time the single trial started, the scheme worked as follows: Inter Trial Interval (ITI), fixation point, target. The Inter Trial Interval was initially shown for 500 ms, followed by the appearance of a fixation point at the center of the screen, for the duration of 500 ms. The last part of the trial was the target presentation: a pair of numbers, randomly selected by a given list, were shown for 500 ms, one on the left of the fixation point, and the other on the right. There was no time limit for the participant to press one of the two keys, but once that happened, the Trial loop was repeated. In **figure 2.1** the loop here described can be seen more accurately.

The pair shown during the Trial loop was randomly selected from a list containing 22 pairs: 11 possible pairs and their mirrored version (ex: 1 vs 3 and 3 vs 1). The frequency at which pairs appeared was the same for all of them, only their order was randomized. Since the pair investigated was 5 vs. 7, to control the distance effect, all the other pairs used for comparison consisted in numbers, in the range from 1 to 9, two digits away from one another (ex: 1 vs. 3, 7 vs. 9). Four additional pairs involving 0 and 10 were added, to avoid 1 and 9 being automatically selected just because they were the smallest and biggest number of the set, respectively. 0 and 10 were labeled as fillers, were not considered in the data analysis and appeared in comparison with numbers both two and one digits away, differently from all the other pairs.

At the end of each condition task, the participant should have compared 440 pairs, and each pair should have appeared 22 times. Since all participants were tested on both conditions, 880 pairs were totally compared, and each pair was compared 44 times. The high number of repetitions was designed to let researchers analyze all data and consider it still valid, even in the eventuality of losing some trials due to software lag or participants' concentration.

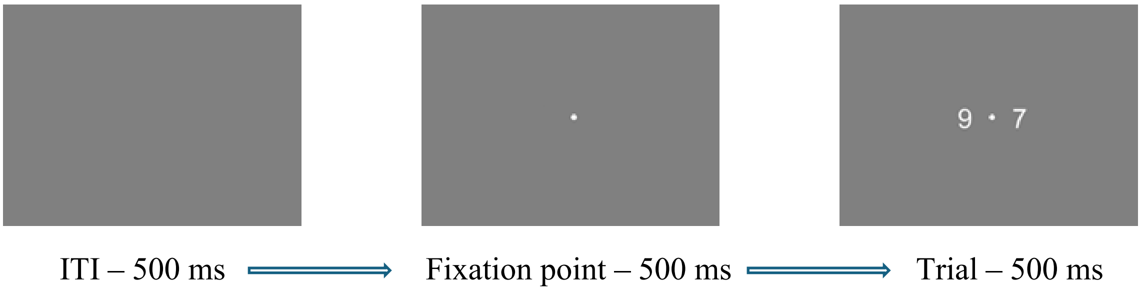


Figure 2.1 *– The scheme of Trial Loop worked as follows. Once the trial item disappeared, the screen went back to grey, waiting for the participant's response.*

3- ANALYSES AND DISCUSSION

We will now focus on the analysis conducted on our data collection and the discussion based on our findings.

3.1 – ANALYSES AND RESULTS

The data included at first 35200 trials, with 880 trials for each participant. The first step was filtering out all the errors, which were 8,37% of the total data. We then set thresholds for Reaction Time filtering (RT). The lower threshold was set at 250ms.

For the higher threshold, we filtered out outliers with normalized value over 3 SD from the mean, calculating it first on the whole group and then on each participant. We chose this double filtering because, during most experiments' administration, different software lags occurred, leading to registration of oddly slow RTs. The duration of these lags was long enough to impair the mean and SD calculation at the individual level, but by first filtering data from the whole group, these software errors were excluded, and we were then able to individuate participants' outlier due to performance and not system lags.

After filtering, we started the analysis of our data. During all the analysis, it is important to mention, we categorized the independent variable of compared number pairs using their calculated mean, e.g. 5 and 7 will from now on be indicated by the label "6", which is the mean between the two. Indeed, the previously mentioned "Fillers" were not included in the analysis, so the independent variable's categories will range from 2 to 8, meaning, from pair 1-3 to pair 7-9. This new naming process is shown more clearly in **fig. 3.1**.

Original Pair	1 vs. 3	2 vs. 4	3 vs. 5	4 vs. 6	5 vs. 7	6 vs. 8	7 vs. 9
Mean label	2	3	4	5	6	7	8

Figure 3.1 *– Table showing labeling process used in the analyses to summarize the pair with just one word*

We first conducted a regression analysis, scripted in R, to simulate the size effect that is usually found in these tasks. After calculating the RT mean of every participant for every condition (in this case, mean of the numbers' pair) and plotting the boxplot summarizing these values (**fig 3.2**), we computed the linear regression (**fig 3.3**) on the mean of all the different means, to control the different numerosity of trials analyzed across participants. The function of the regression line was $y = 0.011x + 0.420$. We then checked the linear fit, which confirmed the accuracy of our linear regression, verified with different tests (Multiple R-squared = 0.880), (Adjusted R-squared = 0.856), (F (1,5) = 36.69, p=0.002). The summary of the linear regression showed us residuals computed between the mean of each condition and the calculated linear regression point (**fig 3.4**).

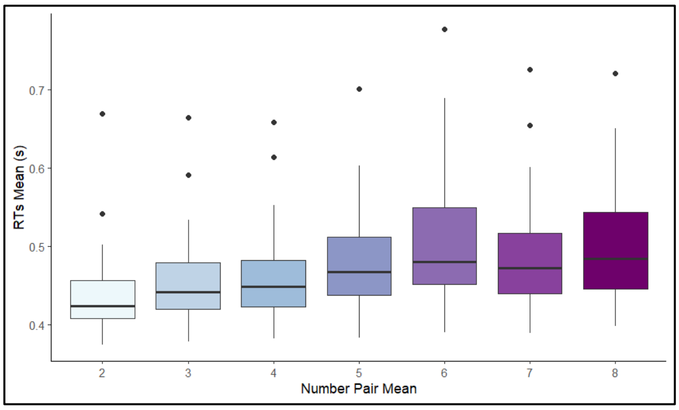
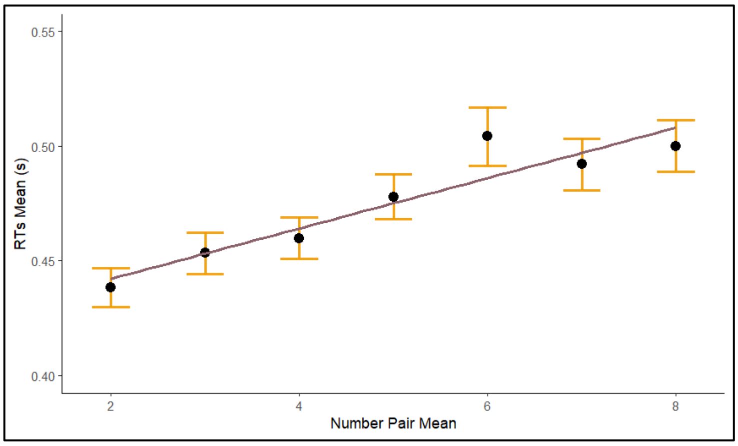


***Figure 3.2** (above right) – Boxplot with RT means per condition (Pair Mean, **fig 3.1**).*

***Figure 3.3** (below left) – Linear regression with RT means and standard error. Pair 6 (5 vs. 7) appears to be more deviant from regression line than others.*

***Figure 3.4** (below right) – Table showing residuals of each condition. Pair 6 (5 vs. 7) is 10 times more distant from regression line than all other pairs.*

Pair 2	-0.0038
Pair 3	0.0002
Pair 4	-0.0042
Pair 5	0.0029
Pair 6	0.0177
Pair 7	-0.0050
Pair 8	-0.0080



The second analysis was conducted, similarly to the study we aim to replicate, on residuals. We first computed linear regression for each participant and then considered the mean of residuals and the standard error across all participants (**fig. 3.5**). We analyzed data in two different ways: t-test for paired samples and repeated measures ANOVA. First, we chose a t-test for paired sample, comparing the mean of “pair 6” (our 5 vs. 7) with the mean between all other groups "no pair 6" (H0: mean of pair6 = mean of nopair6). The result showed to be statistically significant (t (39) = -7.20, $p < 0.0001$), so we can say that pair 6 processing deviates from the mean of other pairs' RT. We then computed a repeated measures ANOVA, (F (6) = 14.139, $p < 0.001$), using JASP software. Once we knew that the mean difference was statistically relevant, we decided to conduct a Post Hoc test. In this analysis (**fig. 3.6**), it can be seen how pair 6 differentiates from all other pairs, while pair 5 and pair 8 were the only two other pairs different enough from one another to reach statistical significance.

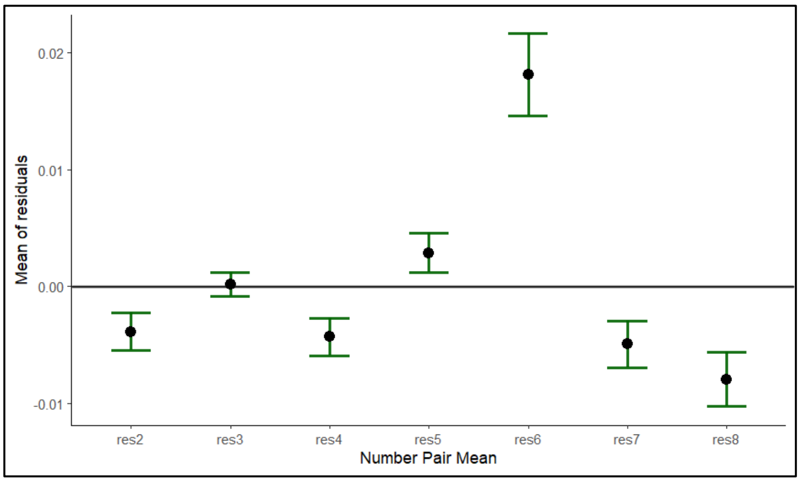


***Figure 3.5** – Residuals' means and standard errors computed across all participants. We can see how res6, residuals calculated on pair 5 vs. 7, is descriptively different from other pairs.*

		Mean Difference	SE	t	p_{bonf}	p_{holm}
2	3	−0.004	0.003	−1.239	1.000	1.000
	4	3.690×10^{-4}	0.003	0.115	1.000	1.000
	5	−0.007	0.003	−2.103	0.768	0.402
	6	−0.022	0.003	−6.697	< .001	< .001
	7	0.001	0.003	0.371	1.000	1.000
	8	0.004	0.003	1.197	1.000	1.000
3	4	0.004	0.003	1.354	1.000	1.000
	5	−0.003	0.003	−0.863	1.000	1.000
	6	−0.018	0.003	−5.458	< .001	< .001
	7	0.005	0.003	1.611	1.000	1.000
	8	0.008	0.003	2.436	0.328	0.203
4	5	−0.007	0.003	−2.217	0.579	0.331
	6	−0.022	0.003	−6.812	< .001	< .001
	7	8.274×10^{-4}	0.003	0.257	1.000	1.000
	8	0.003	0.003	1.082	1.000	1.000
5	6	−0.015	0.003	−4.594	< .001	< .001
	7	0.008	0.003	2.474	0.295	0.197
	8	0.011	0.003	3.299	0.024	0.017
6	7	0.023	0.003	7.069	< .001	< .001
	8	0.025	0.003	7.894	< .001	< .001
7	8	0.003	0.003	0.825	1.000	1.000

***Figure 3.6** – Table showing the post-hoc analysis computed on repeated measures ANOVA. These were filtered using both Bonferroni and Holm corrections. Label 6 corresponds to pair 5 vs. 7, while pairs 5 and 8 correspond respectively to 4 vs.6 and 7 vs. 9.*

The last analysis we decided to conduct was a purely descriptive one, focusing on participants' accuracy rather than RT. We simply computed the percentage of accuracy across all participants in all conditions and then calculated the mean for each condition (**fig 3.7**). Once again, we can see how pair 6 seems to deviate from the other pairs, considering the role of the size effect.

Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Pair 7	Pair 8
96,78%	94,66%	95,38%	90,31%	80,22%	86,16%	79,16%

Figure 3.7 *– No statistical analysis was conducted on the accuracy, but we decided to create a table to visualize differences in accuracy between conditions. Pair 6, corresponding to accuracy found when comparing 5 vs. 7, appears to deviate from the linearly decreasing accuracy.*

3.2 – DISCUSSION & CONCLUSION

The analysis conducted showed and confirmed the presence of the size effect, meaning that RT proportionally increases with number's magnitude. We found this effect by plotting the mean RTs and regression line (**fig.3.2**, **fig 3.3**) and seeing how participants' RTs become progressively slower as the pair mean increases. The presence of the size effect, which we expected to find, means that our data is in line with typical number comparison performances and can be considered reliable. The chosen pairs avoid manipulating the distance effect, and this allowed us to observe the magnitude effect even if the range was small, confirming that choosing pairs of numbers two digits apart can be a good control procedure for the distance effect's interference.

Our study showed an impaired performance in comparing the pair 5 vs. 7 in an Italian-speaking sample during a number comparison task. The participants were slower when comparing this pair, as we statistically demonstrated using t-test and ANOVA test analyses.

These results confirm the effect found in the French-speaking sample of the study we decided to replicate, therefore weakening the linguistical hypothesis of number perception that we considered in the introduction.

We will now try to discuss some possible implications of the effect found in this study.

Since the linguistical hypothesis is not accurate enough to support our findings, another possible explanation could consider the visual paths involved in number perception and comparison, given the close relation between digits and images (Fias *et al.*, 2001). During the test administration, we took note of participants' comments about their performance during the task. Many of them were aware of having issues with pair 5 vs. 7, or more specifically, 7. Some of them reported problems both with pair 5 vs. 7 and 7 vs. 9, even if the latter does not show performance impairments found in the analysis that are associated with the former. We are planning to investigate the correlation between participants' impressions and individual performance. However, focusing on the actual enigma, we would like to hypothesize a possible visual and perceptual effect regarding number 7, which many participants recognized as problematic.

It could be possible that, when digit 7 is presented in its Arabic digital form, at a perceptual level even the digit 1 is activated, given their visual similarity. This perceptual activation could then cascade and partially activate the semantic information related to 1, resulting in increased difficulty and impaired performance when number 7 is being compared, as this double activation leads to perceiving 7 as smaller than how it really is. Number 1 is, as a matter of fact, an incredibly relevant number. It is extremely represented in languages, and it appears in everyday life much more frequently than any other number. According to the DSS, the frequency at which digits are represented is extremely relevant and influences our number comparison performance. This is even strengthened by the shape that number 1 has. It shares features with number 7, especially in some fonts, but number 1's shape is common in different notation systems, e.g. Roman, as it represents the easiest symbol to think of. Even when we consider early representations of quantities, or maybe the ones that generally come more natural to us, we will realize that bar lines are usually the most used counting method. One bar line is almost identical to digit 1, making it extremely easy for our brain to activate magnitude 1 when presented with a similar shape. Even in our experiment, we did not see any problematic processing with the pair 7 vs. 9. This finding agrees with our hypothesis, since 1 is smaller than 9, just like 7, and therefore creates no difficulty in comparison tasks.

We, however, recognize some limitations in our study. For example, participants were all University students, restricting our sample to people who are still studying. Moreover, since the study was conducted in Padova, we think that the population coming from the Veneto region is more represented in this study, as it is where the University is located, making the background of our participants less variegated. Another issue we would like to mention is that in our task, the type of font we used might have increased the similarity between 1 and 7, since 7 showed no horizontal bar and neither did 1. This could strengthen even more the possible major role of the vertical bar in perceptual activation of numbers, which we talked about earlier.

Further studies need to be designed to check this effect. One possible paradigm could involve exploring the weakness related to font selection by replicating this study with additional conditions, based on text font. Fonts in which the shape differences between 1 and 7 increase or decrease the similarity between the two digits could be selected for further studies. This should let us see the effect of font and general digit shape on the performance during number comparison tasks, giving us an insight into how our perceptual and semantic representations of numbers work together.

REFERENCES

Dehaene S. (1992). Varieties of numerical abilities. *Cognition*, 44(1-2), 1–42. https://doi.org/10.1016/0010-0277(92)90049-n

Dehaene, S. (1997). Il pallino della matematica: Scoprire il genio che è in noi. Arnoldo Mondadori. Retrieved from: Biblioteca di Dipartimento di Matematica (PD)

Dowker, A., Bala, S., Lloyd, D. (2008). Linguistic Influences on Mathematical Development: How Important Is the Transparency of the Counting System? *Philosophical Psychology*, 21(4), 523 - 538. Retrieved from: http://dx.doi.org/10.1080/09515080802285511

Fias, W., Reynvoet, B., & Brysbaert, M. (2001). Are Arabic numerals processed as pictures in a Stroop interference task?. *Psychological research*, 65(4), 242–249. https://doi.org/10.1007/s004260100064

Fias, W., & Verguts, T. (2004). The mental number line: exact and approximate. *Trends in cognitive sciences*, 8(10), 447–449. Retrieved from: https://doi.org/10.1016/j.tics.2004.08.009)

Göbel, S. M., Shaki, S., & Fischer, M. H. (2011). The Cultural Number Line: A Review of Cultural and Linguistic Influences on the Development of Number Processing. *Journal of Cross-Cultural Psychology*, 42(4), 543-565. https://doi.org/10.1177/0022022111406251

Hubbard, E., Piazza, M., Pinel, P. *et al.* Interactions between number and space in parietal cortex. *Nat Rev Neurosci* **6**, 435–448 (2005). https://doi.org/10.1038/nrn1684

Knudsen, B., Fischer, M. H., Henning, A., & Aschersleben, G. (2015). The Development of Arabic Digit Knowledge in 4- to 7-Year-Old Children. *Journal of Numerical Cognition*, 1(1), 21-37. https://doi.org/10.5964/jnc.v1i1.4

Krajcsi, A. (2017). Numerical distance and size effects dissociate in Indo-Arabic number comparison. *Psychon Bull Rev* 24, 927–934. Retrieved from: https://doi.org/10.3758/s13423-016-1175-6

Lepoittevin, S., Geers, L. & Andres, M. (2024, May). *A re-examination of the sub-base-five effect in single digit comparison*. Talk Presented at the annual meeting "Advances in Numerical Cognition Research" (Numbers and the Brain 2024), Padua, Italy.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, 215(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Sella, F., Lucangeli, D., Cohen Kadosh, R., & Zorzi, M. (2020). Making Sense of Number Words and Arabic Digits: Does Order Count More?. *Child development*, 91(5), 1456–1470. https://doi.org/10.1111/cdev.13335

Oldfield R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113. Retrieved from: https://doi.org/10.1016/0028-3932(71)90067-4

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological review*, 101(1), 80–102. Retrieved from: https://doi.org/10.1037/0033-295x.101.1.80