UNIVERSITÀ DEGLI STUDI
DI TRENTO

CiMeC

upf.

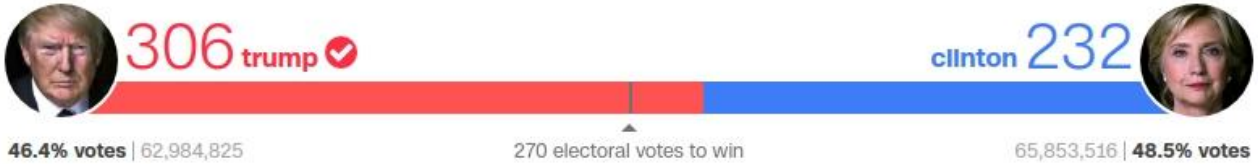
# Comparatives, Quantifiers, Proportions: A Multi-Task Model for the Learning of Quantities from Vision

***Pezzelle, S.**, §Sorodoc, I., *Bernardi, R.
*UniTN | §UPF, UniTN

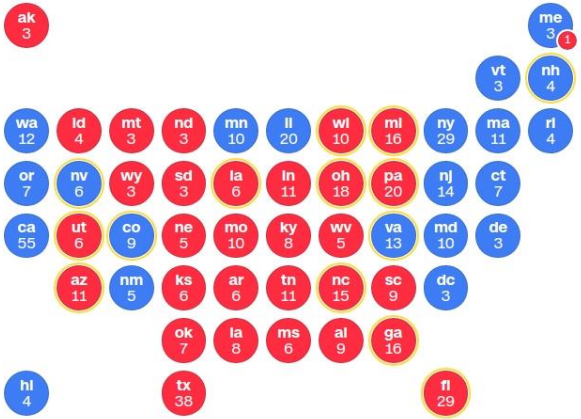
# What's *Quantification?*

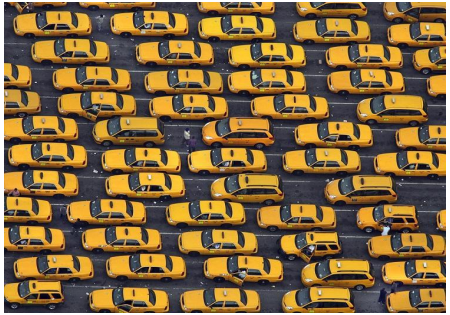
Operation of quantifying, i.e. expressing ***quantitative*** information

- "***More than half*** of the electoral votes were for Trump"
- "Indeed, he got ***306*** electoral votes out of ***538***"
- "In percentage, ***46.4%*** of Americans voted for him"
- "Though Clinton got ***more*** votes (48.5%), he was elected"



# Why *Vision*?

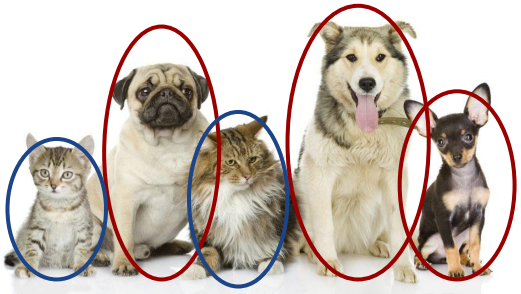

ak 3
me 3 1
vt 3
nh 4
wa 12
id 4
mt 3
nd 3
mn 10
il 20
wi 10
mi 16
ny 29
ma 11
ri 4
or 7
nv 6
wy 3
sd 3
ia 6
in 11
oh 18
pa 20
nj 14
ct 7
ca 55
ut 6
co 9
ne 5
mo 10
ky 8
wv 5
va 13
md 10
de 3
az 11
nm 5
ks 6
ar 6
tn 11
nc 15
sc 9
dc 3
ok 7
la 8
ms 6
al 9
ga 16
hi 4
tx 38
fl 29



# Different Ways to Quantify

- **Numbers**
  - *seven, 72, five, 123, etc.*

***count*** of exact, absolute cardinality of ***one set***



- **Comparatives**
  - *more, same, less*
- **Quantifiers**
  - *all, most, few, almost all, etc.*
- **Proportions**
  - *20%, 85%, thirty-three percent, etc.*

comparison or ***relation*** between ***two sets***

# ANS vs Counting

- Ability of ***comparing*** non-symbolic ***sets*** (a.k.a. ANS) reported in infants since youngest age, well before being able to count
  - [Piazza & Eger (2016), Xu & Spelke (2000), McCrink & Wynn (2004)]

- Proportional values extracted ***holistically***, i.e. w/out relying on the pre-computed cardinalities of sets
  - [Fabbri et al. (2012), Yang et al. (2015)]

- In language acquisition, Comparatives (~3.3 yrs) and Quantifiers (3.4-3.6 yrs) acquired ***before*** Numbers (3.5-)
  - [Odic et al. (2013), Halberda et al. (2008), Le Corre & Carey (2007)]

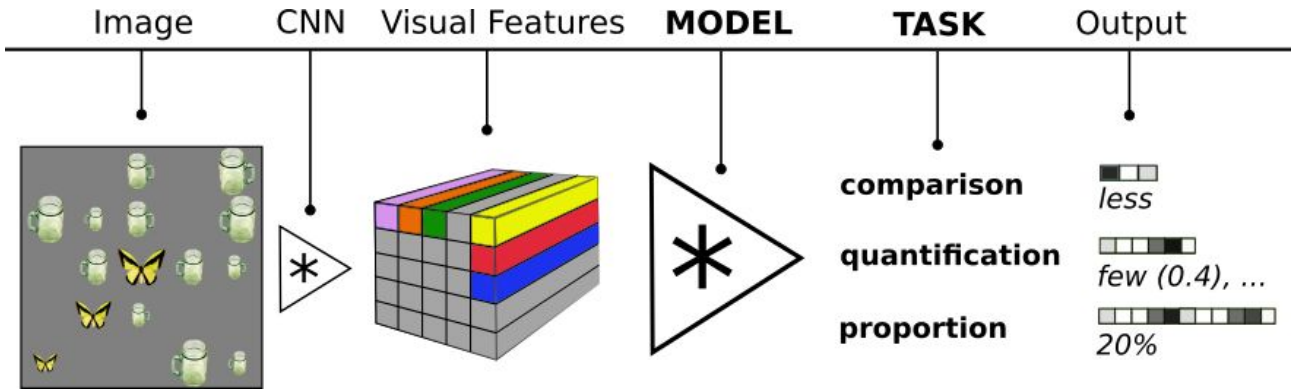
# Hypotheses

- ***Shared operation*** underlying Comparison, Vague Quantification, Proportion
  - counting not needed and perhaps **conflicting**

- Increasing-complexity ***hierarchy*** of relation-based mechanisms, shown by evidence from cognition and language acquisition:
  - 1. Comparison
  - 2. Vague Quantification
  - 3. Proportion

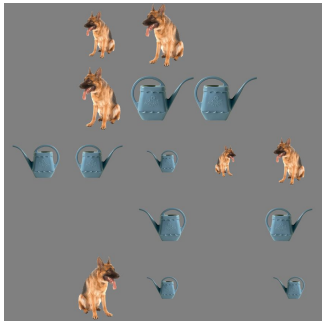
# Research Questions

- Can ANS-based tasks be learned by a single, **Multi-Task Learning** model?
- Are low-level tasks beneficial to high-level ones, and *vice versa*?



# Materials

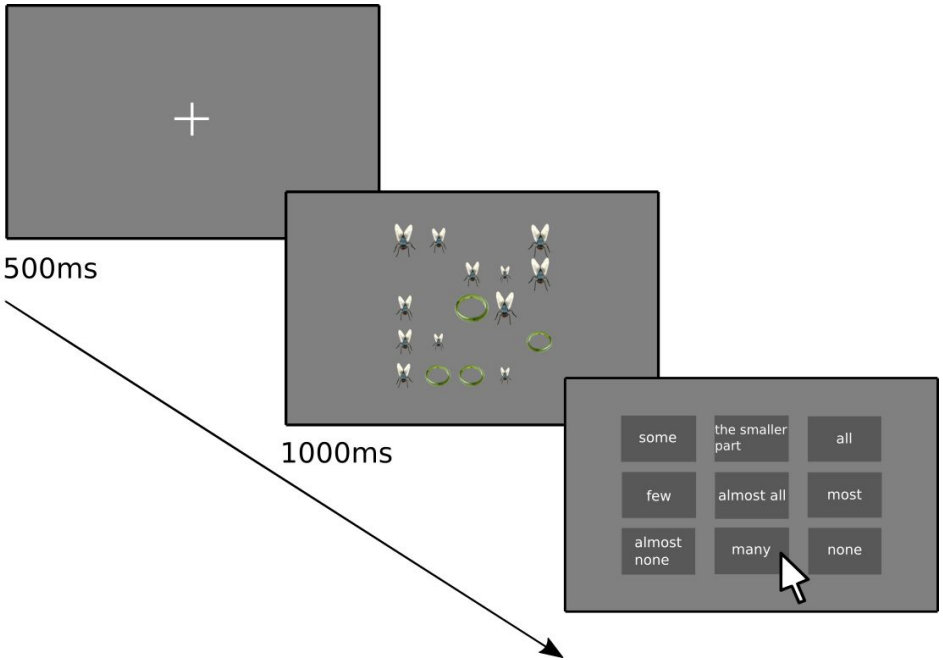
- Artificially built (11.9K train, 1.7K val, 3.4K test)
- 3-20 total objects (animals + artifacts) from [15] in the scene
- 17 *ratios*, i.e. proportions of animals (8 > 50%, 8 < 50%, 1 = 50%)
- Number cases balanced for ratio
- Size, position, orientation randomly varied



Comparative: *less*
Proportion: *40%*
Quantifier: ?

# How many of the Objects are *Animals*?

[Pezzelle, Bernardi, Piazza (*under review*). Cognition]



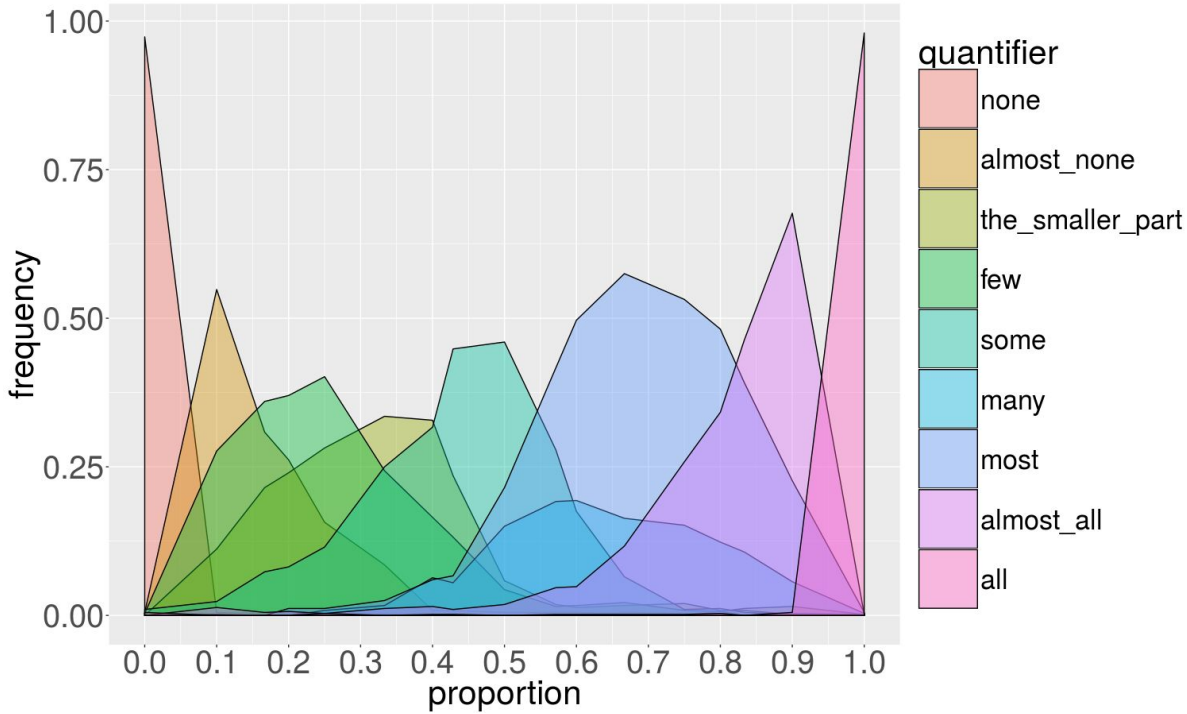

Behavioral experiment:

- 340 scenes (balanced ratios)
- 1,000ms exposure to scene
- 30 participants
- 10.2K responses

Analyses:

- *glmer* (6 main, 3 random)
- proportion ***best predictor***!

# Quantifiers and Proportions: Distribution

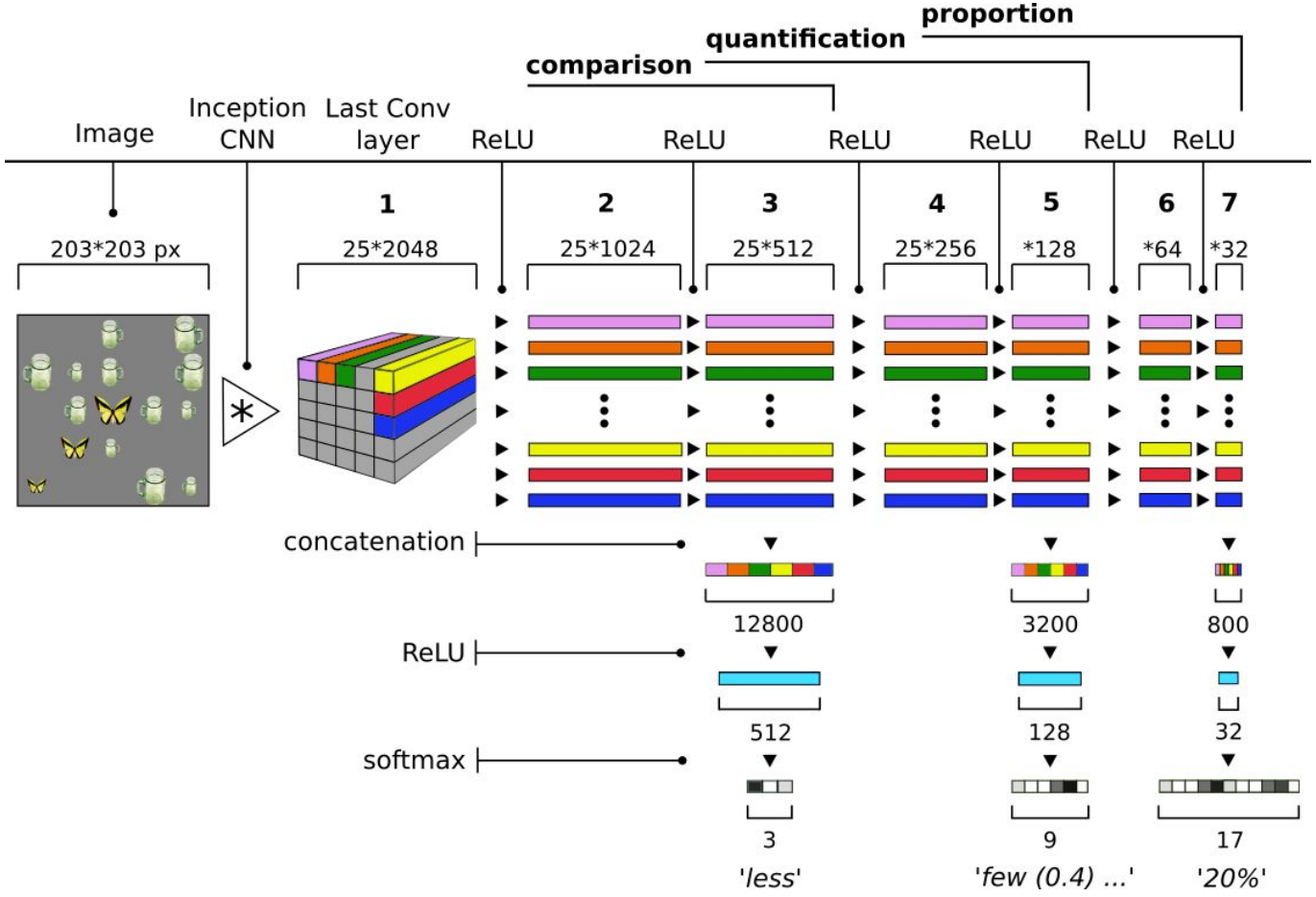


# One-Task Models

- **one-task-*frozen***
  - one-task models fed with ‘frozen’ visual features (average of last Conv layer of Inception v3 CNN pre-trained on ImageNet)

- **one-task-*end2end***
  - one-task models fed with raw images and embedding Inception v3 CNN

# MTL Model

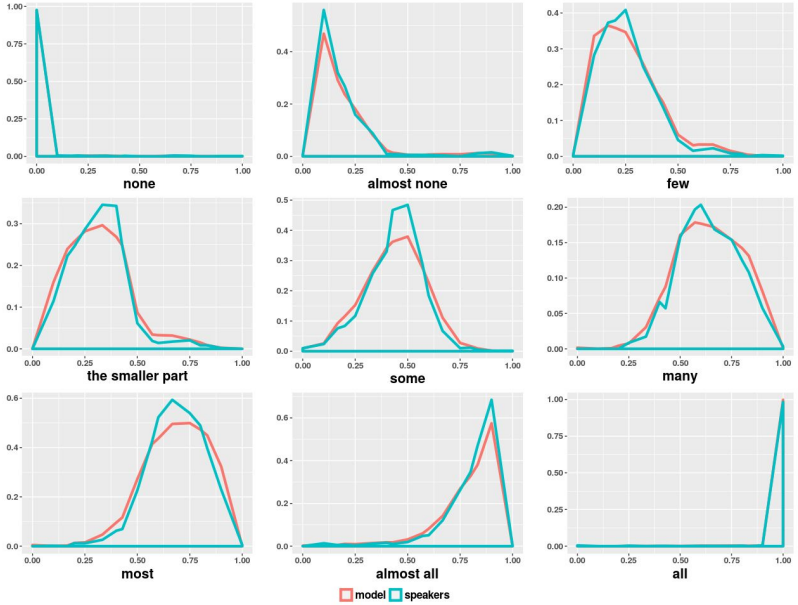

proportion
quantification
comparison
Image
Inception CNN
Last Conv layer
ReLU
ReLU
ReLU
ReLU
ReLU
ReLU
1
2
3
4
5
6
7
203*203 px
25*2048
25*1024
25*512
25*256
*128
*64
*32
concatenation
12800
3200
800
ReLU
512
128
32
softmax
3
9
17
'less'
'few (0.4) ...'
'20%'

# Results

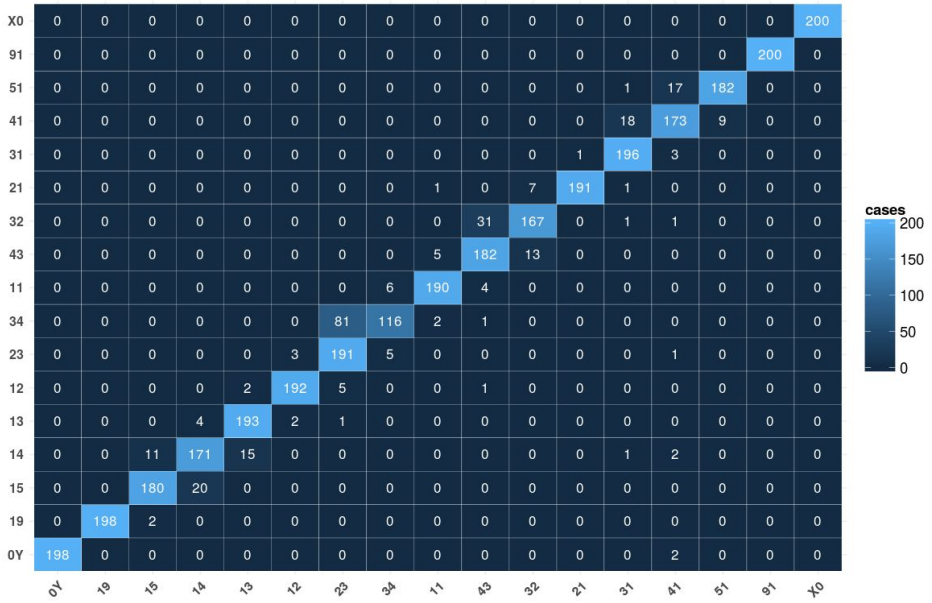
| **model** | **setComp** | **vagueQ** | **propTarg** | **nTarg** |
|---|---|---|---|---|
| | *accuracy* | *Pearson r* | *accuracy* | *accuracy* |
| *chance/majority* | 0.470 | 0.320 | 0.058 | 0.132 |
| one-task-frozen | 0.783 | 0.622 | 0.210 | 0.312 |
| one-task-end2end | 0.902 | 0.964 | 0.659 | **0.966** |
| multi-task-prop | **0.995** | **0.982** | **0.918** | – |
| multi-task-number | 0.854 | 0.807 | – | 0.478 |

# MTL Errors

## vagueQ

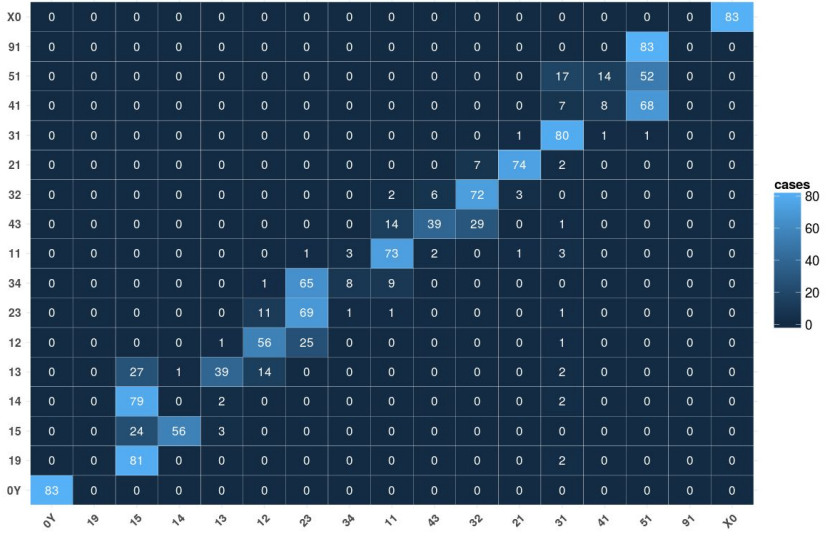



## propTarg



# Does it Generalize?

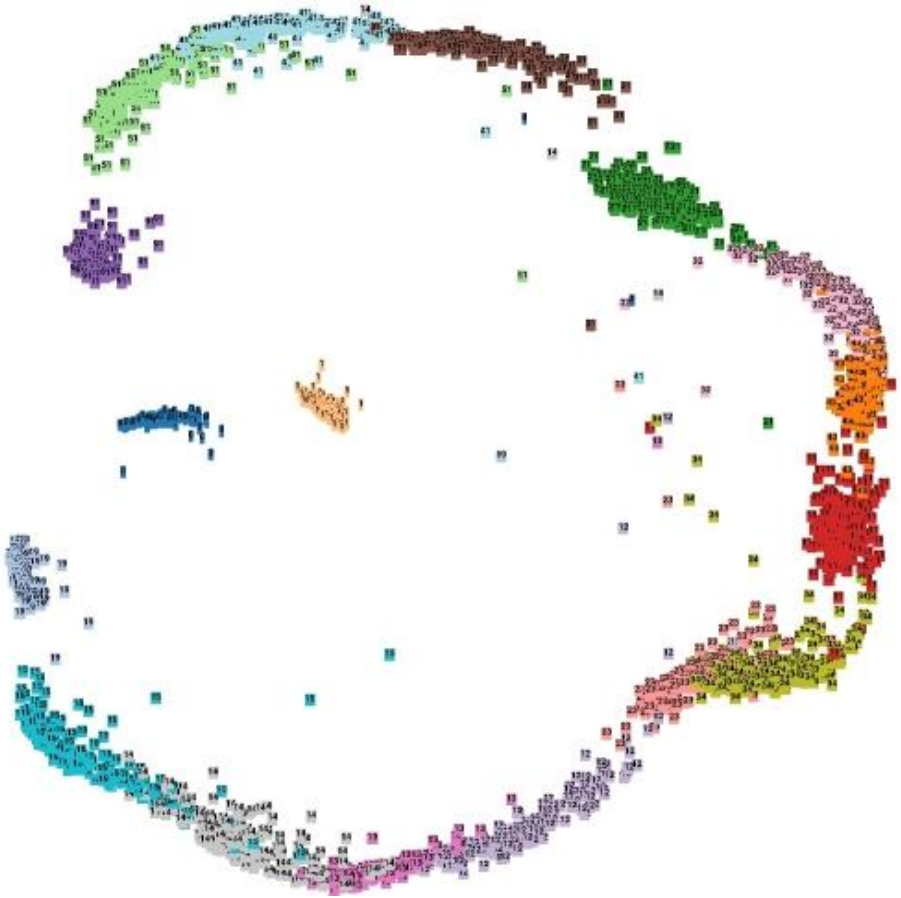
- MTL tested w/ unseen combinations
- *Plausible* errors, i.e. similar ratios




| model | setComp | vagueQ | propTarg |
|---|---|---|---|
| | *accuracy* | *Pearson r* | *accuracy* |
| *chance/majority* | 0.470 | 0.320 | 0.058 |
| one-task-frozen | 0.763 | 0.548 | 0.068 |
| one-task-end2end | 0.793 | 0.922 | 0.059 |
| multi-task-prop | **0.943** | **0.960** | **0.539** |

# Proportional Layer

- 2-dimensional PCA analysis on 32-d last layer of proportional task (before softmax)
- Proportions clearly clustered together and ordered *clockwise*



# Conclusions

- Sharing a **common core** boosts performance in all relation-based tasks, confirming they underlie same operation (relation between sets)
- Exact number is a different operation → **interference**
- MTL able to generalize to unseen combinations to some extent

# Ongoing Work

- Do the results hold when training-testing within other **modalities**?
- Is the core of the model (encoding quantities) **modality-independent**, and thus transferable?

# The Q-Team

Sandro Pezzelle
CIMeC, University of Trento

Ionut-Teodor Sorodoc
Universitat Pompeu Fabra, Barcelona

Aurelie Herbelot
CIMeC, University of Trento

Raffaella Bernardi
CIMeC, DISI, University of Trento

**QUANT**ities in **I**mages and **T**exts at **CLIC** lab (**QUANTIT-CLIC**)

https://quantit-clic.github.io/

# Thank you!

*Few* / *Some* / *Many* questions?

# References


1. Manuela Piazza. 2010. Neurocognitive start-up tools for symbolic number representations. *Trends in cognitive sciences* 14(12):542–551.
2. Manuela Piazza and Evelyn Eger. 2016. Neural foundations and functional specificity of number representations. *Neuropsychologia* 83:257–273.
3. Fei Xu and Elizabeth S Spelke. 2000. Large number discrimination in 6-month-old infants. *Cognition* 74(1):B1–B11.
4. Koleen McCrink and Karen Wynn. 2004. Large-number addition and subtraction by 9-month-old infants. *Psychological Science* 15(11):776–781.
5. Sara Fabbri, Sara Caviola, Joey Tang, Marco Zorzi, and Brian Butterworth. 2012. The role of numerosity in processing nonsymbolic proportions. *The Quarterly Journal of Experimental Psychology* 65(12):2435– 2446.
6. Ying Yang, Qingfen Hu, Di Wu, and Shuqi Yang. 2015.Children's and adults' automatic processing of proportion in a stroop-like task. *International Journal of Behavioral Development* 39(2):97–104.
7. Darko Odic, Paul Pietroski, Tim Hunter, Jeffrey Lidz, and Justin Halberda. 2013. Young children's understanding of "more" and discrimination of number and surface area. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 39(2):451.
8. Peter Bryant. 2017. *Perception and understanding in young children: An experimental approach*, volume 4. Routledge.
9. Felicia Hurewitz, Anna Papafragou, Lila Gleitman, and Rochel Gelman. 2006. Asymmetries in the acquisition of numbers and quantifiers. *Language learning and development* 2(2):77–96.
10. Justin Halberda, Len Taing, and Jeffrey Lidz. 2008. The development of 'most' comprehension and its potential dependence on counting ability in 852 preschoolers. *Language Learning and Development* 4(2):99–121.
11. Mathieu Le Corre and Susan Carey. 2007. One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition* 105(2):395–438.