

**FMSERA
2014**

A1

*research
in math
education*

**Sep 1
2014**

How to quantify the reliability of an error classification principle?

Hannu Tiitu, Jarmo Malinen, Antti Rasila
Aalto University

Birgitta Dahl, Tore Ståhl
Arcada University of Applied Sciences



Project MathProHealth



New learning environment



**Arcada
Sigma**

**Aalto
STACK**

**Earlier
experi-
ences**

***Guides
learning***

**Validation
&
certification**

Game-like

**Med.
calc.
LE**

AI

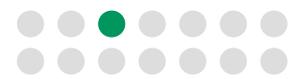
**More
realistic**

**More
authentic**

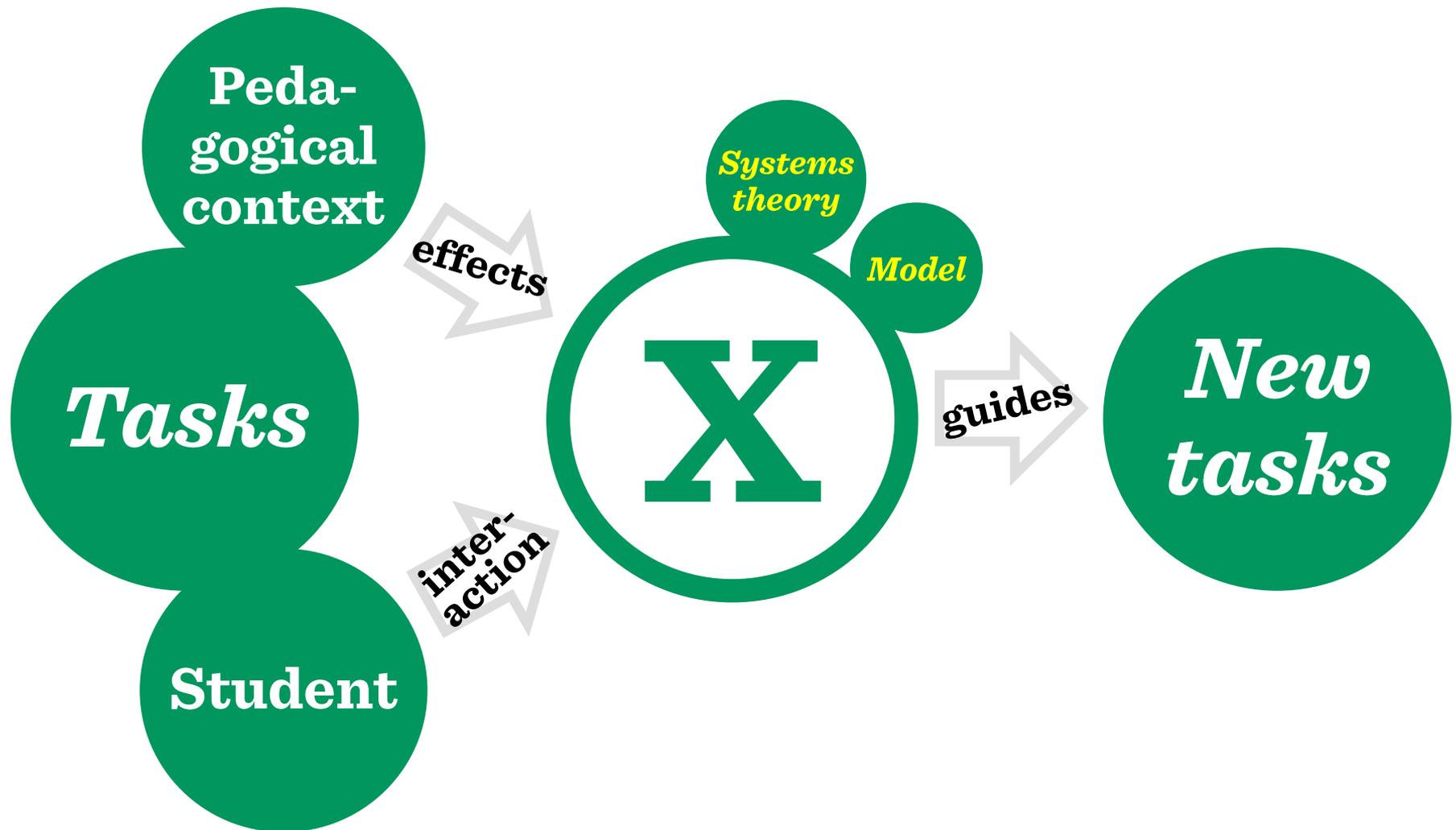
Motivating

HCI

**More
interactive**

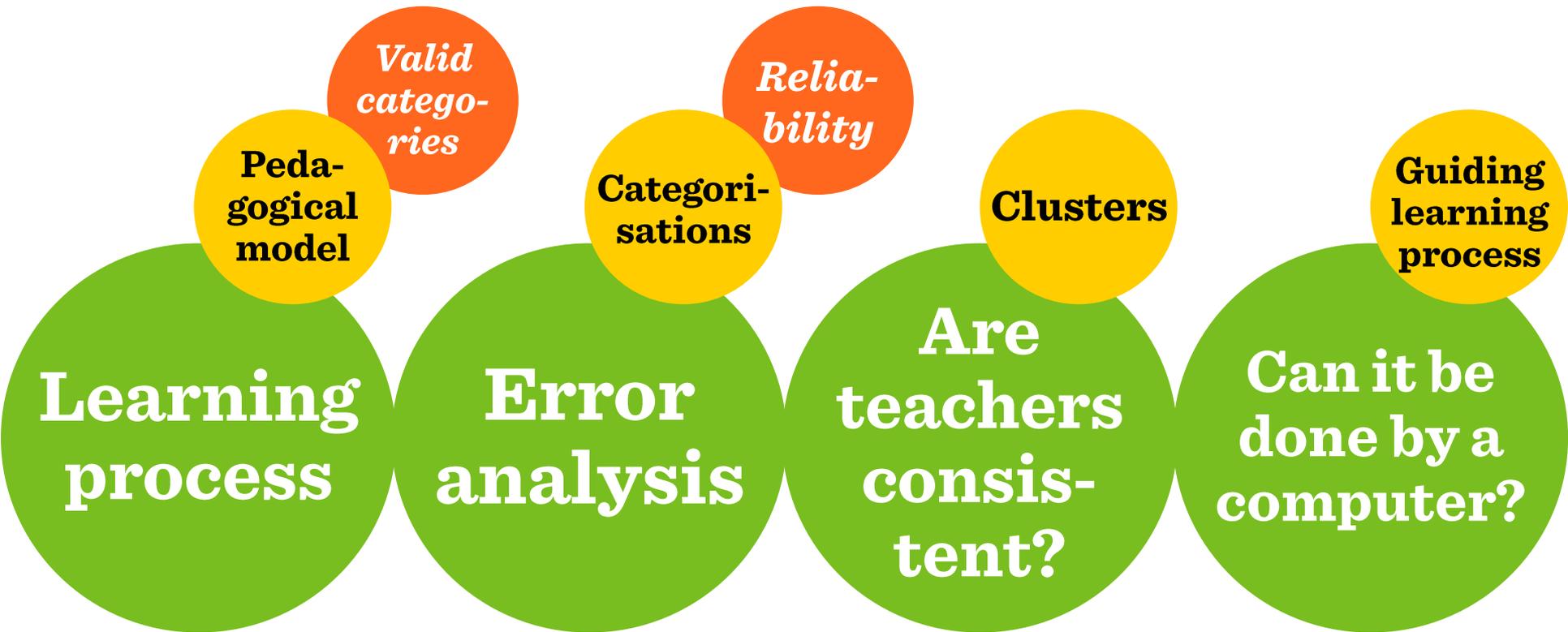


Motivation: *Design of intelligent LEs*





Designing the





**Error
analysis in
literature**

**Commonly:
reliability
not verified**

Needed

***Valid & reliable
error analysis***

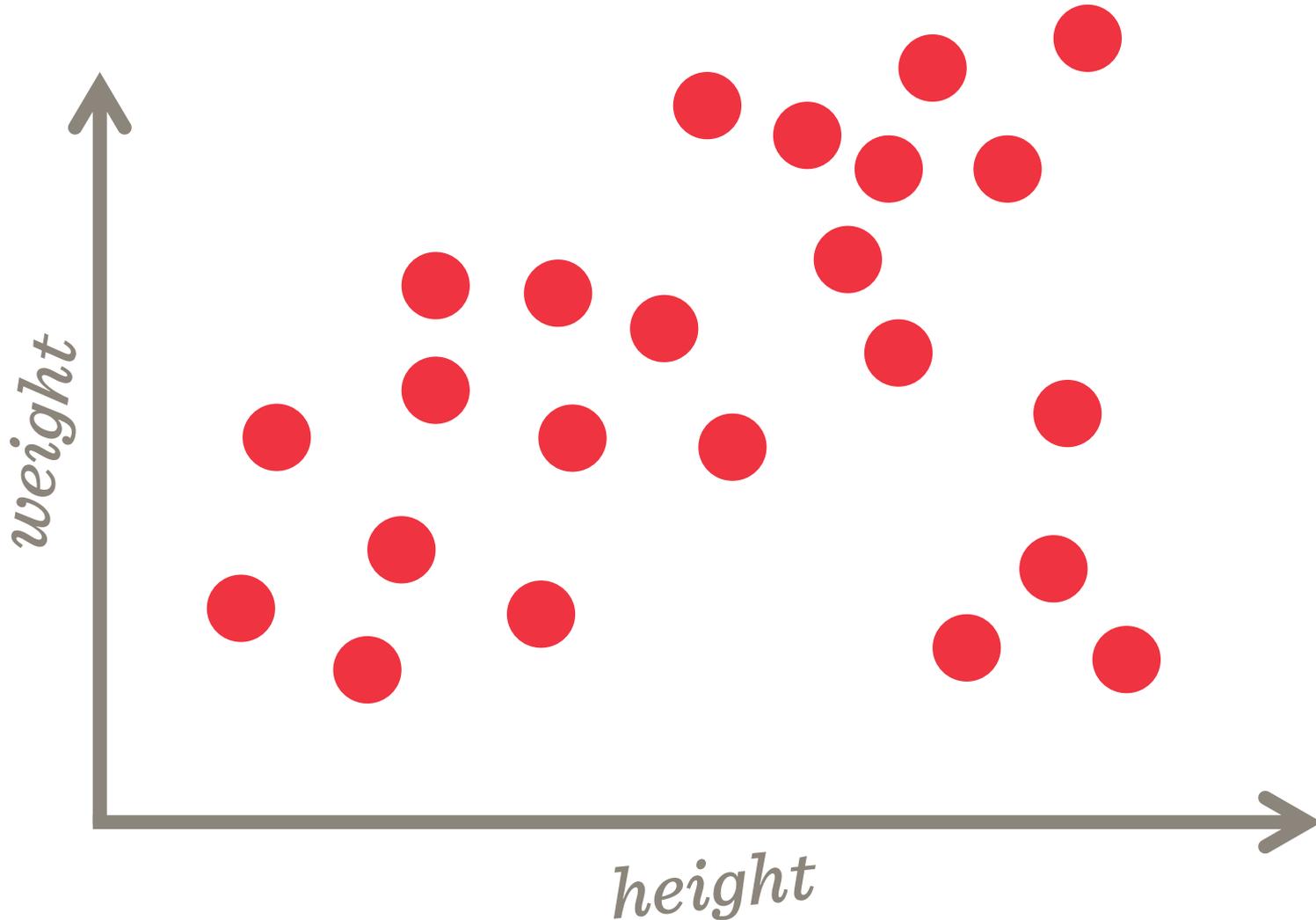


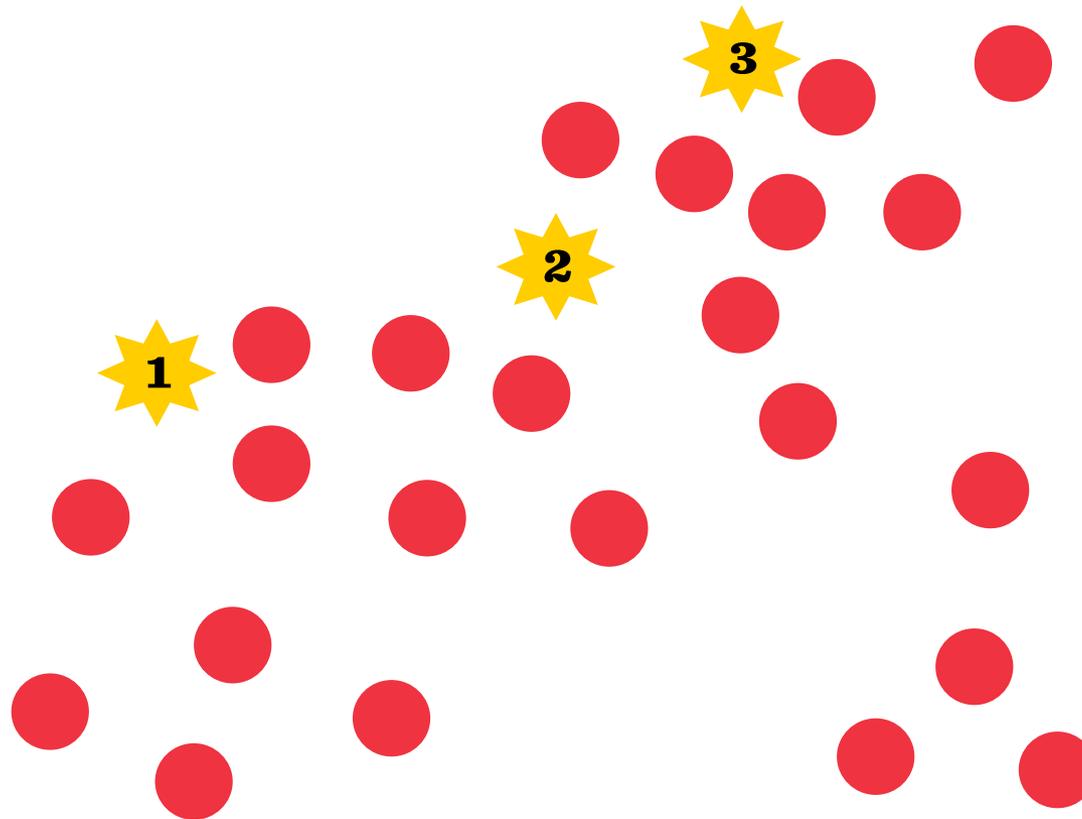
Our approach to analysing reliability



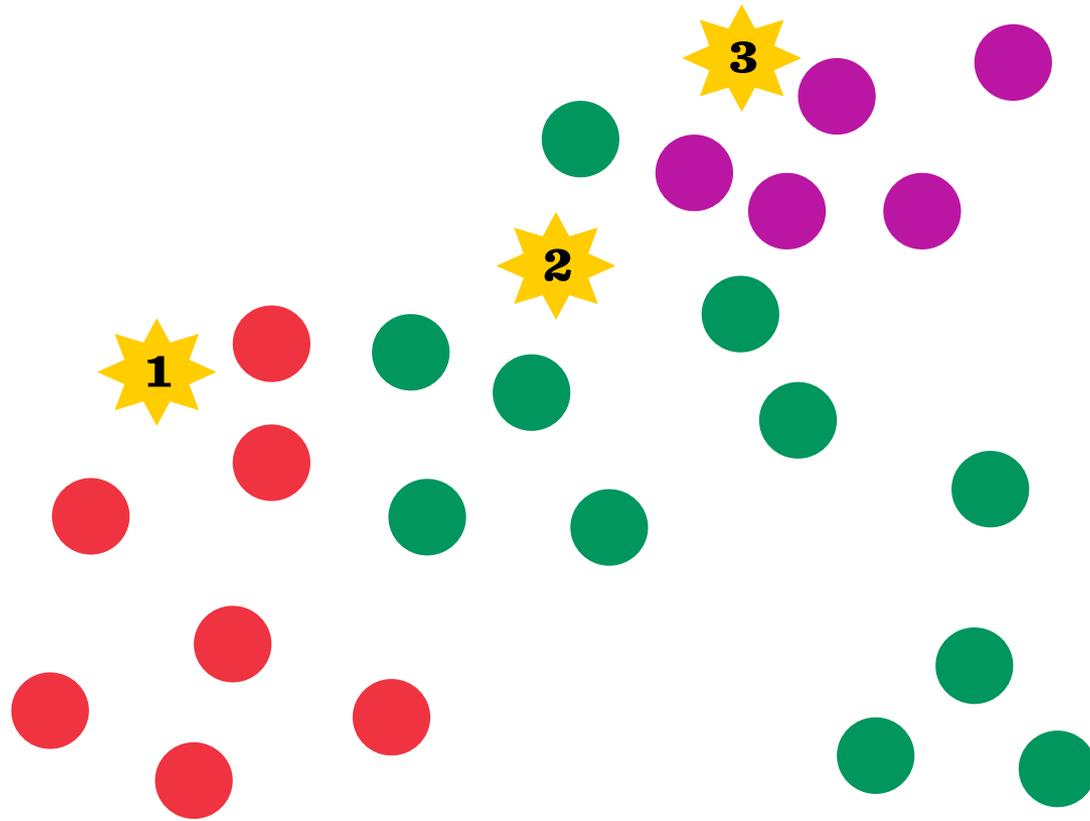


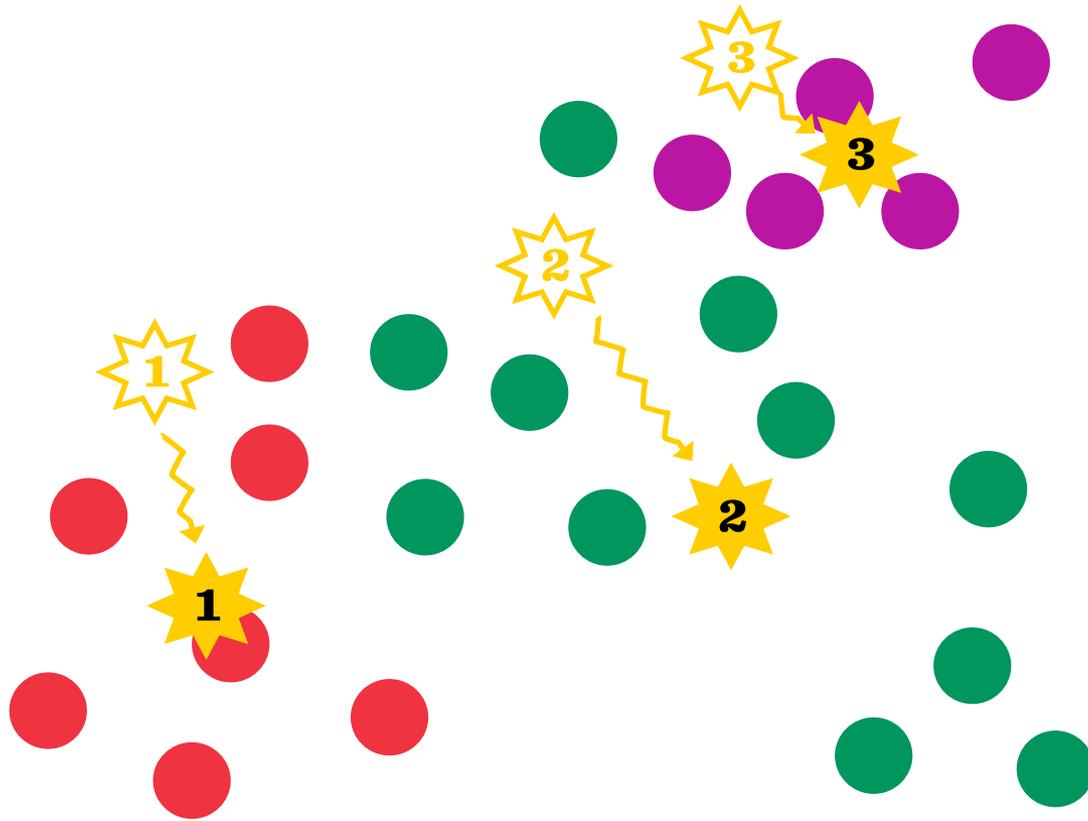
Example: *Clustering observations with two properties*

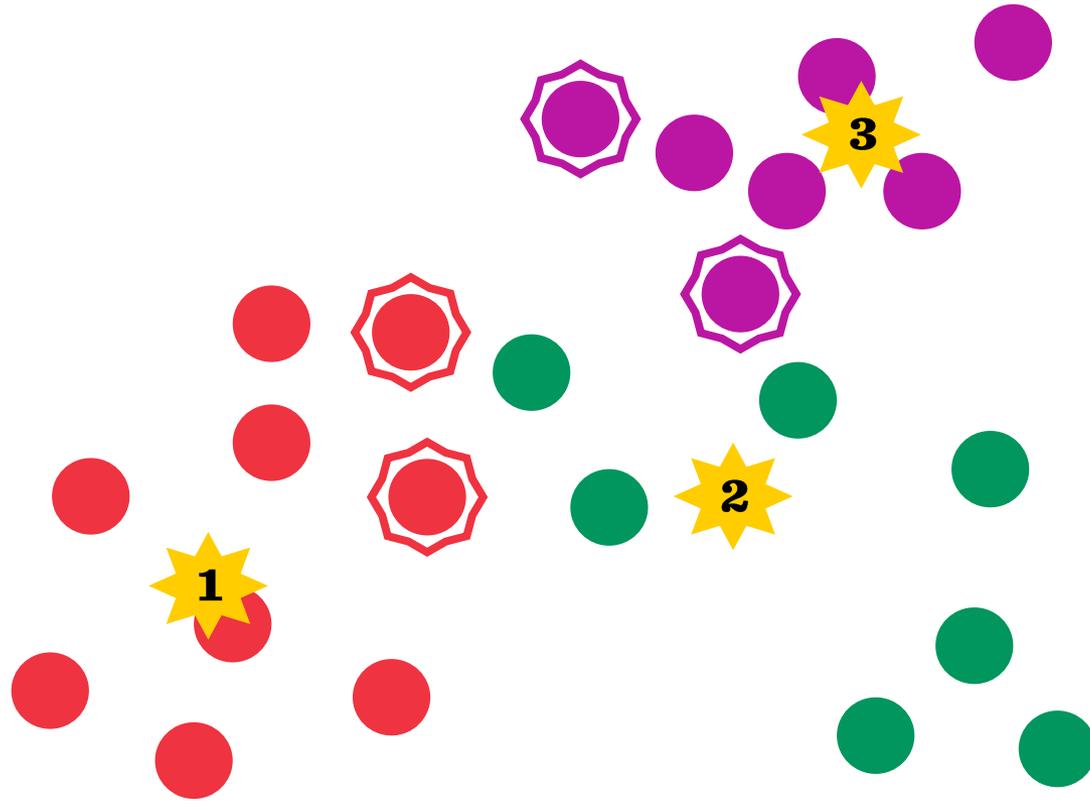


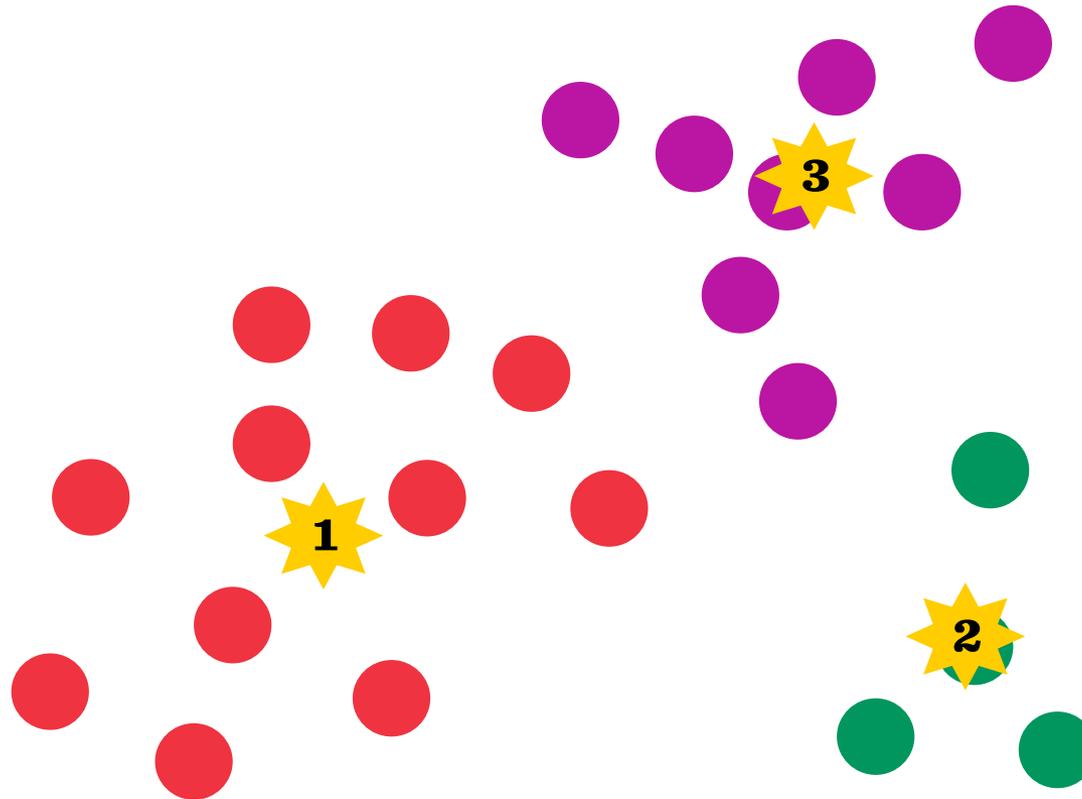


$k = 3$

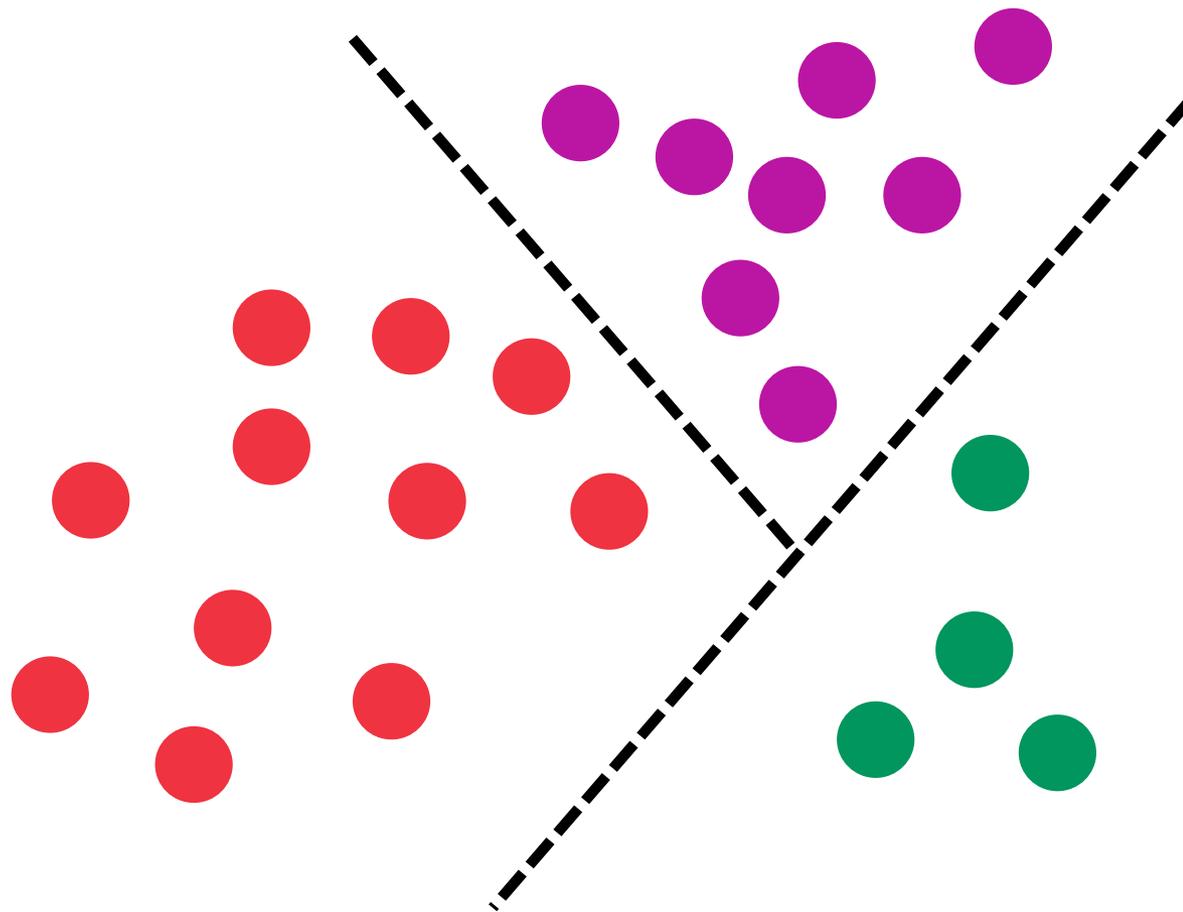








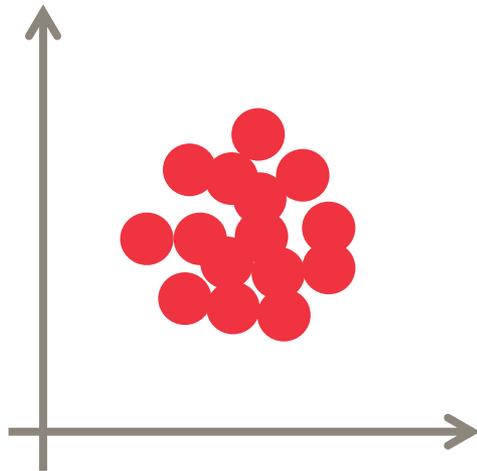
finally the *iteration ends*



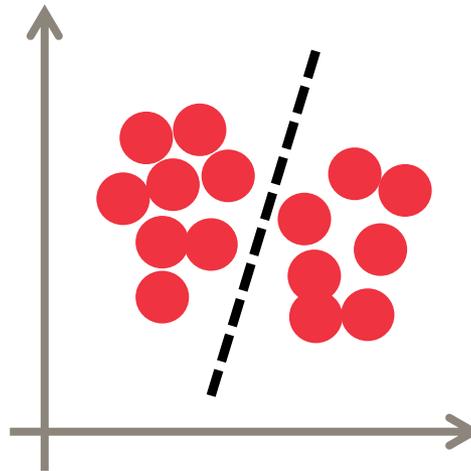
observations are *divided* into 3 clusters



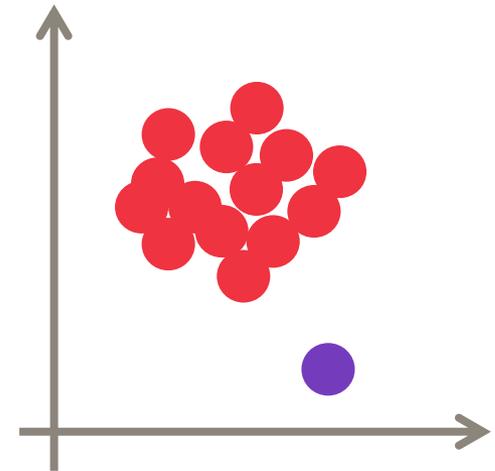
Outcomes and interpretations



**Single
consensus
cluster**



**Rivaling
consensus
clusters**



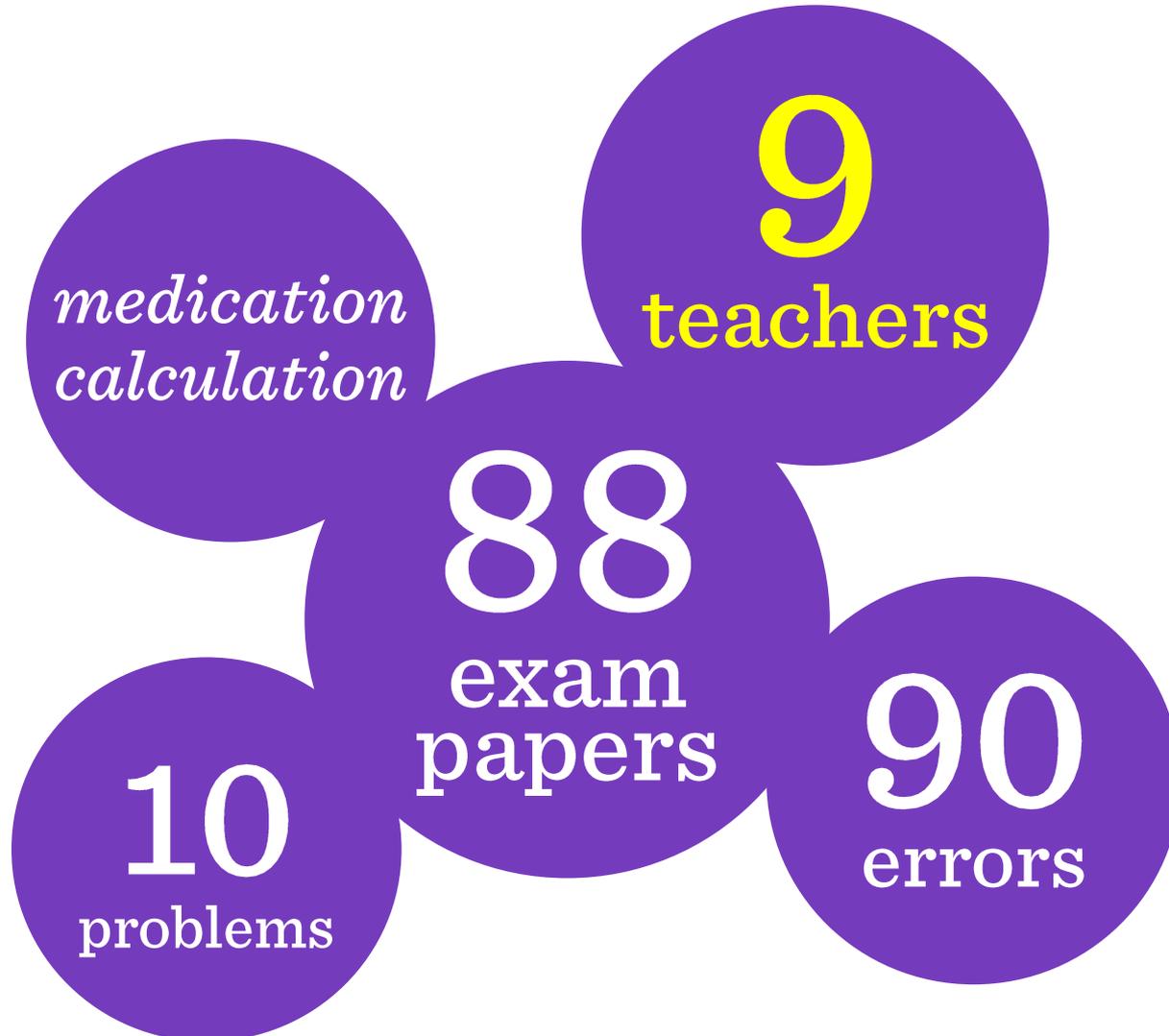
**Lone
classifiers**

Why?

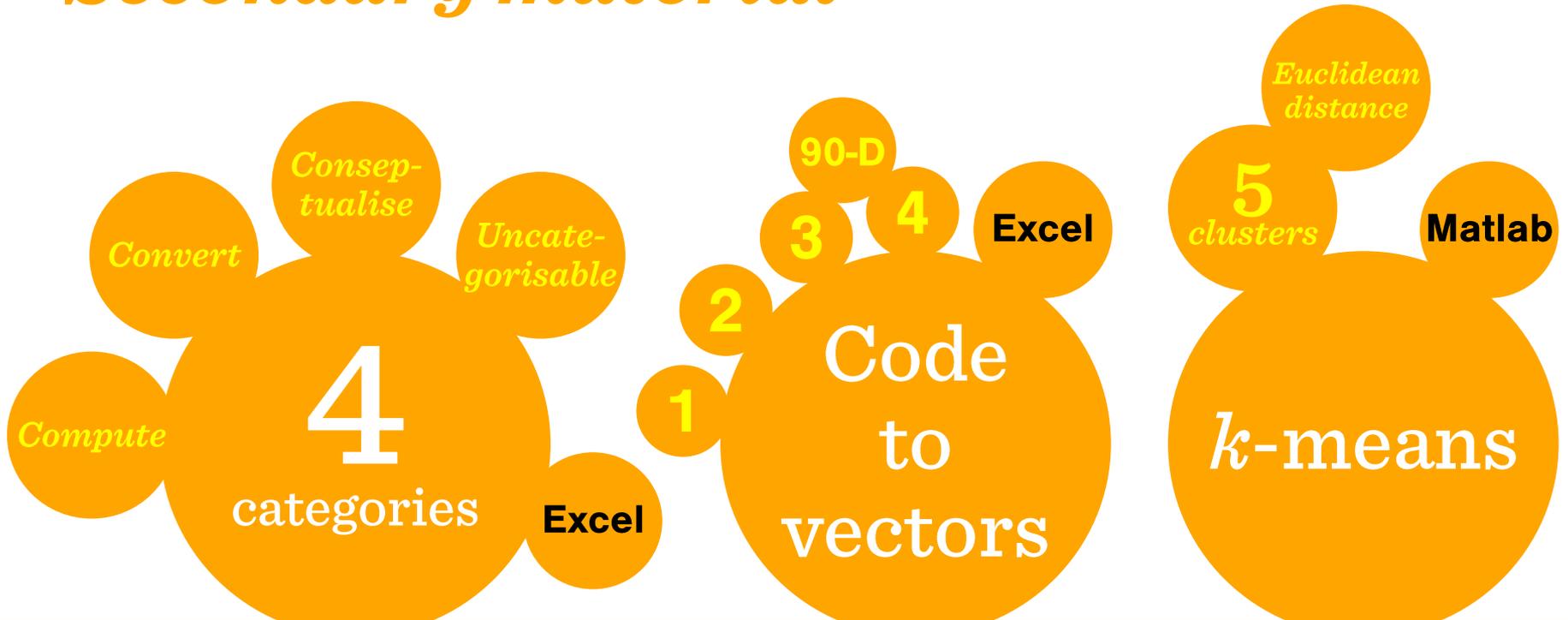
*Need for new
categories? Hidden
variables*



Primary material



Secondary material



| Exam | Problem | Error | Compute | Convert | Conceptualize | Critically evaluate | Uncategorizable | Error |
|------|---------|-------|---------|---------|---------------|---------------------|-----------------|-------|
| | | | C1 | C2 | C3 | C4 | X | |
| V412 | 7b | 28 | | | 1 | | | 28 |
| V412 | 6 | 29 | | 1 | | | | 29 |
| V412 | 6 | 30 | | 1 | | | | 30 |
| V412 | 5a | 31 | | | | | 1 | 31 |
| V412 | 1b | 32 | | | 1 | | | 32 |
| V412 | 5b | 33 | 1 | | | | | 33 |
| V412 | 5b | 34 | | | 1 | | | 34 |
| V412 | 3 | 35 | | | 1 | | | 35 |
| V412 | 6 | 36 | 1 | | | | | 36 |
| V412 | 4 | 37 | | | | | 1 | 37 |
| V412 | 6 | 38 | | 1 | | | | 38 |
| V411 | 3a | 39 | 1 | | | | | 39 |
| V411 | 3b | 40 | | | 1 | | | 40 |
| V411 | 5b | 41 | | | 1 | | | 41 |
| V411 | 6 | 42 | 1 | | | | | 42 |
| V411 | 5b | 43 | 1 | | | | | 43 |

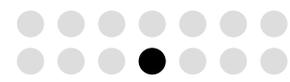
Data for K-means analysis

C1=1, C2=2, C3=3, X=4

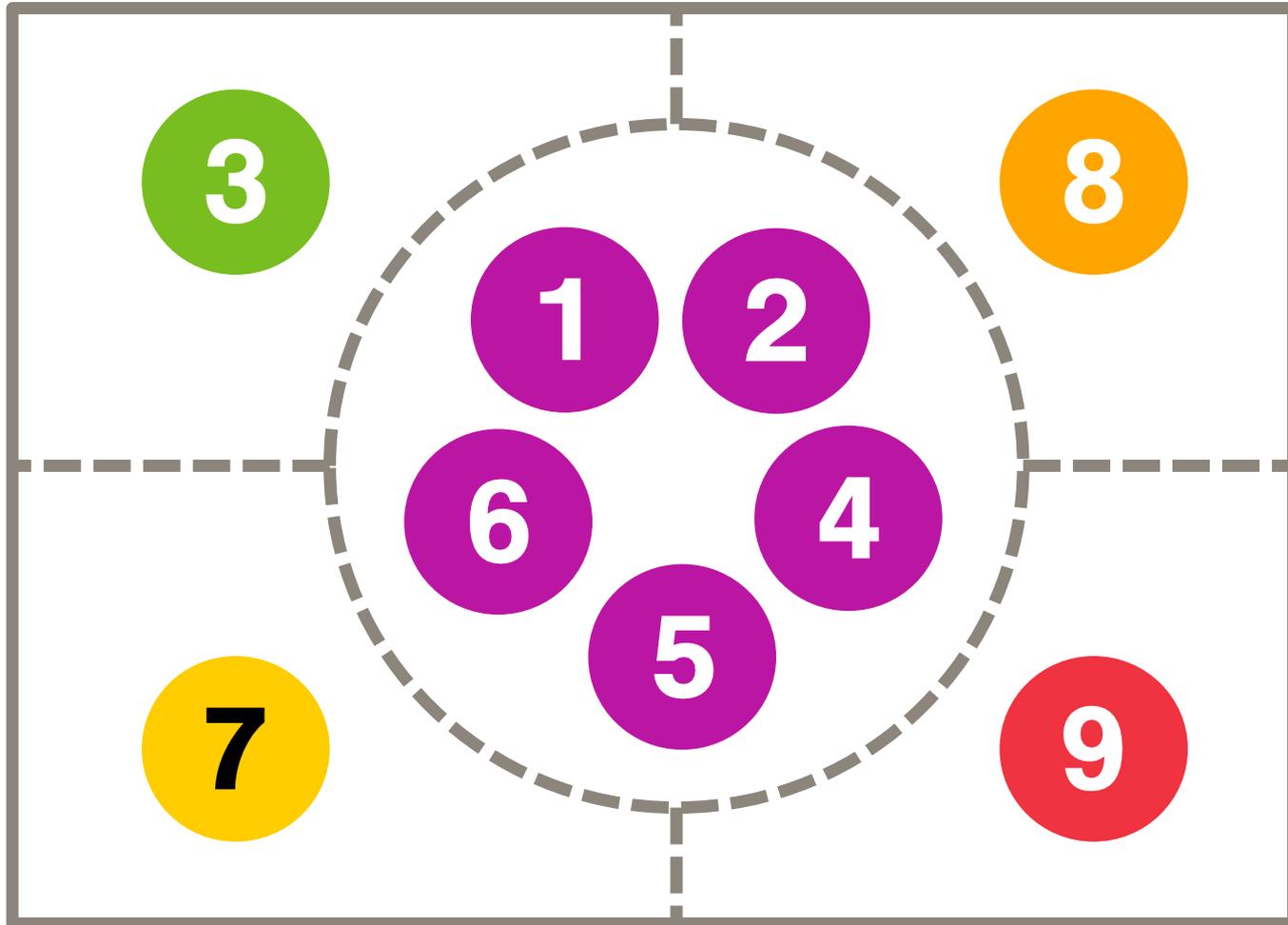
| | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 |
|----|----|----|----|----|----|----|----|----|----|
| 7 | 2 | 3 | 3 | 3 | 3 | 3 | 1 | 3 | 3 |
| 8 | 1 | 3 | 3 | 1 | 3 | 1 | 3 | 1 | 1 |
| 9 | 3 | 2 | 2 | 3 | 3 | 3 | 1 | 3 | 3 |
| 10 | 3 | 3 | 2 | 3 | 3 | 3 | 3 | 1 | 1 |
| 11 | 3 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 11 |
| 12 | 3 | 2 | 2 | 3 | 3 | 3 | 1 | 3 | 1 |
| 13 | 3 | 3 | 1 | 3 | 3 | 3 | 3 | 4 | 13 |
| 14 | 3 | 4 | 3 | 3 | 3 | 3 | 1 | 3 | 3 |
| 15 | 3 | 4 | 2 | 3 | 3 | 3 | 1 | 1 | 3 |
| 16 | 3 | 1 | 3 | 3 | 3 | 3 | 1 | 3 | 4 |
| 17 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 1 | 1 |
| 18 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 18 |
| 19 | 3 | 3 | 3 | 3 | 3 | 3 | 1 | 3 | 3 |
| 20 | 3 | 3 | 2 | 3 | 3 | 3 | 1 | 4 | 1 |
| 21 | 3 | 3 | 3 | 3 | 3 | 3 | 1 | 4 | 21 |
| 22 | 3 | 3 | 3 | 1 | 3 | 3 | 1 | 3 | 4 |
| 23 | 3 | 2 | 3 | 3 | 3 | 1 | 1 | 3 | 1 |
| 24 | 3 | 3 | 3 | 3 | 3 | 3 | 1 | 4 | 3 |

```

k = kmeans(X,5)
C = k(C1, C2, C3, X)
C
ans =
     2     2     5     2     2     2     3     1     4
     3     3     5     3     3     3     1     2     4
     1     5     4     1     1     1     2     3     3
     1     5     3     1     1     1     2     4     4
     4     4     2     4     4     4     1     5     3
     4     4     5     4     4     4     1     3     2
    
```

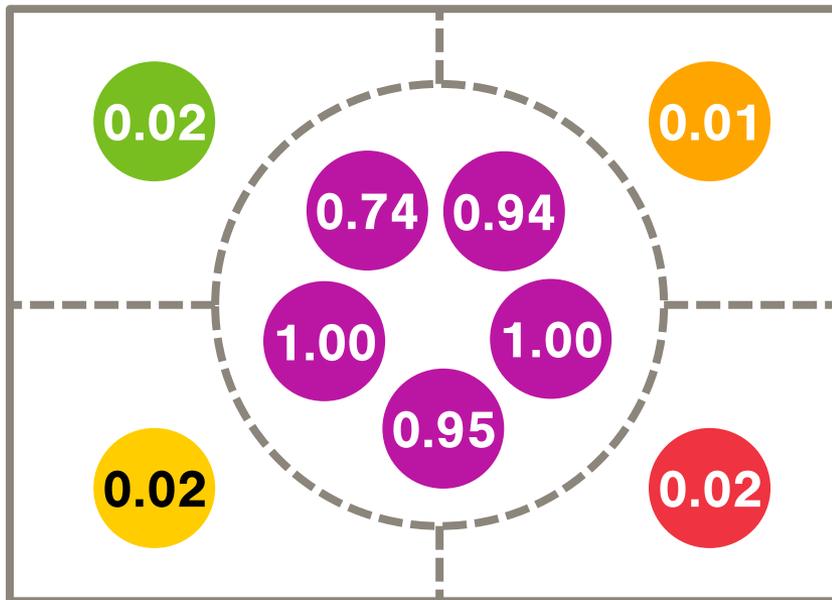


$k = 5, n = 90$

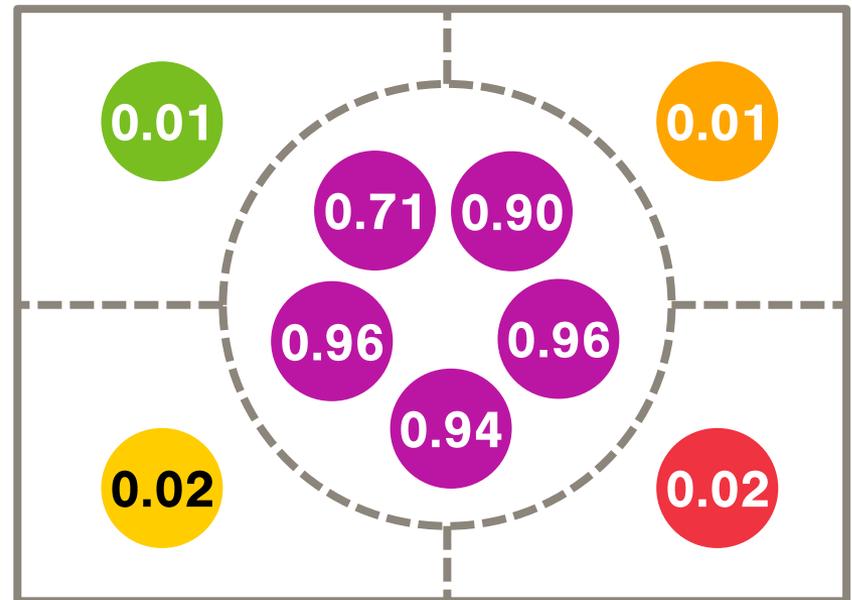




Robustness



**500 times randomly 5 % off ($n = 85$),
100 clustering for each**

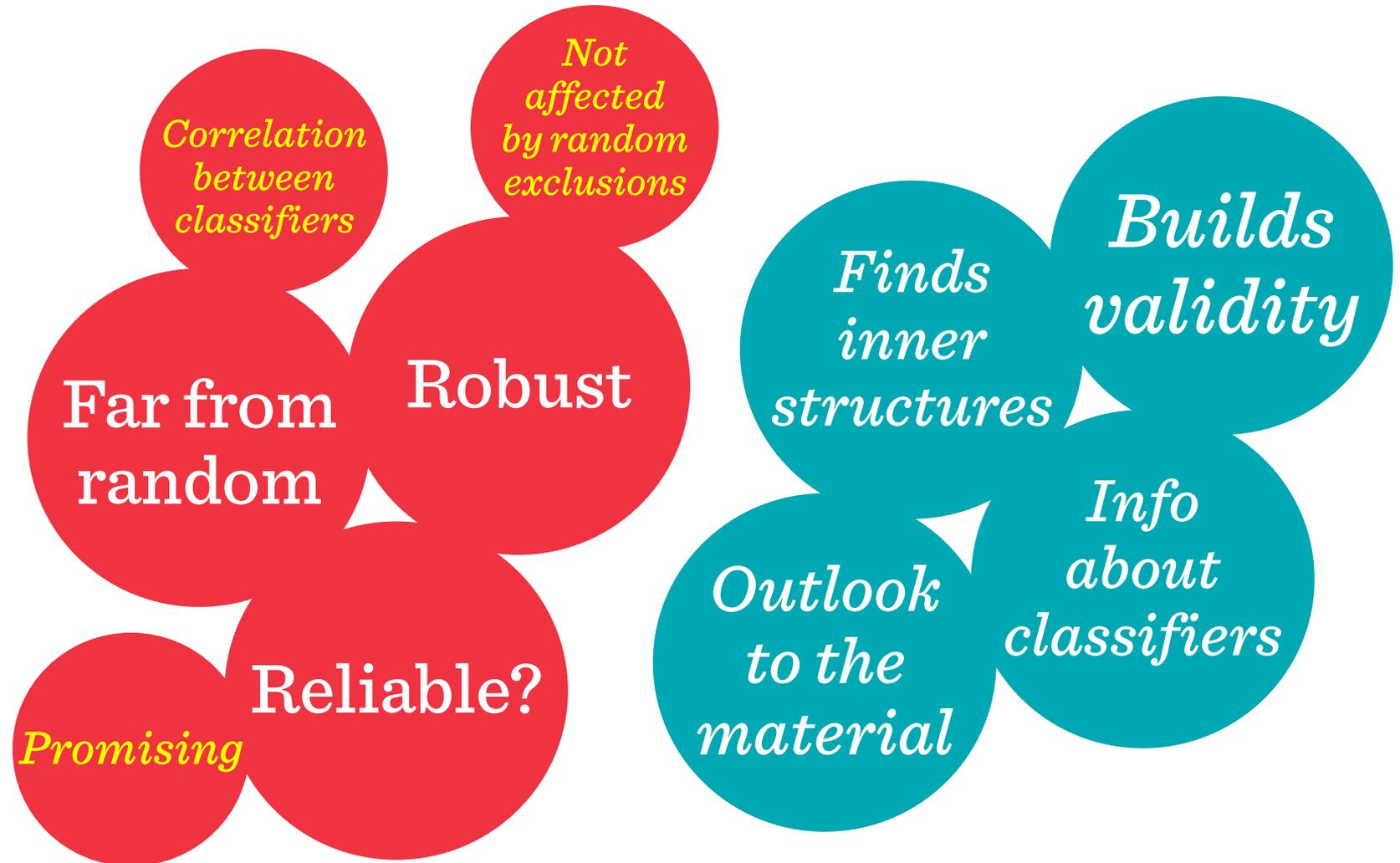


**500 times randomly 10 % off ($n = 81$),
100 clustering for each**



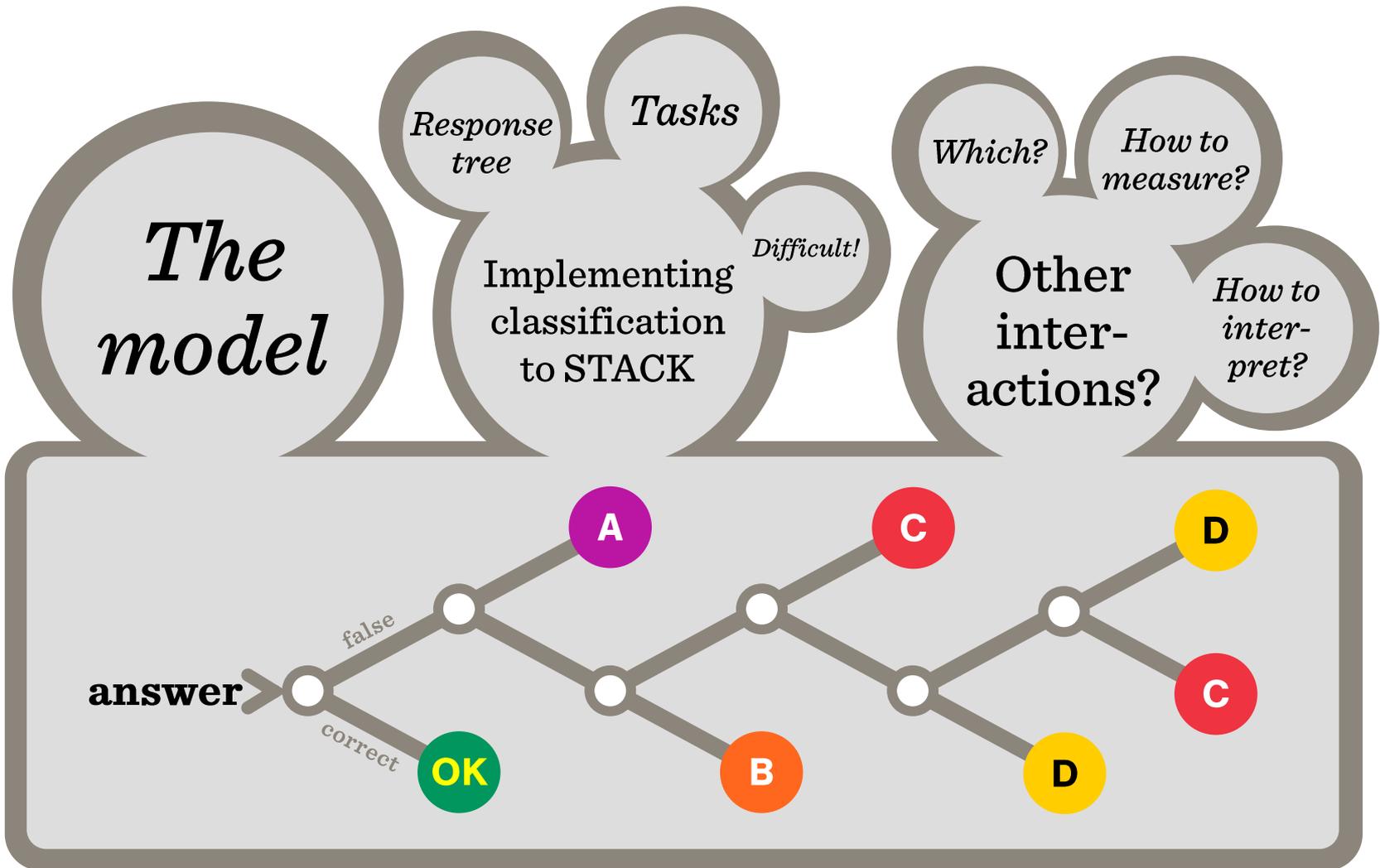
Observations on the data

Conclusions on the method





Future: *Building the AI*



Literature

- Bock, H. H. (2008). Origins and extensions of the means algorithm in cluster analysis. *Electronic Journ@l for History of Probability and Statistics*, 4(2). Retrieved from <http://www.jehps.net/Decembre2008/Bock.pdf>
- Dahl, B., Ståhl, T., Malinen, J., Rasila, A., & Tiitu, H. (2014). Diagnosing nursing students' errors in medication calculation. Designing a method based in the 4 Cs teaching model for analysing mathematical proficiency. In J. Viteli, & A. Östman (Eds.), *Tuovi 12: Interaktiivinen tekniikka koulutuksessa 2014 -konferenssin tutkijatapaamisen artikkelit* (TRIM Research Reports 12, pp. 82–92). Tampere: University of Tampere. Available from <http://urn.fi/URN:ISBN:978-951-44-9561-8>
- Leikas, S., Granberg, L., Ståhl, T., Kurko, T., Antikainen, O., Airaksinen, M., & Pohjanoksa-Mäntylä, M. (2012). Sigma- lääkelaskennan oppimisympäristö: kehittämistarpeet ja hyödyntämismahdollisuudet farmasian peruskoulutuksessa [Sigma: e-learning environment for learning medication calculation skills: pharmacy students' perceptions of its usefulness and development needs]. *Farmaseuttinen aikakauskirja Dosis*, 28(2), 106–117.
- Lloyd, S. P. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2), 129–137.
- MacQueen, J. B. (1967). Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Volume 1 (pp. 281–297). Berkeley, CA: University of California Press.
- Sangwin, C. (2013). *Computer Aided Assessment of Mathematics*. Oxford: Oxford University Press.
- Steinhaus, H. (1957). Sur la division des corps matériels en parties. *Bulletin de l'Académie Polonaise des Sciences*, 4(12), 801–804.