Rigging your Card Games: Differentiating Expert from Novice

Steven F. Wolf, D. P. Dougherty, Gerd Kortemeyer
Lyman Briggs College – Michigan State University

– Abstract –

A seminal study by Chi *et al.* firmly established the paradigm that novices categorize physics problems by "surface features" (e.g. "incline," "pendulum," "projectile motion," ...), while experts use "deep structure" (e.g. "energy conservation," "Newton 2," ...). Yet, efforts to replicate the study frequently fail, since the ability to distinguish experts from novices turns out to be highly sensitive to the problem set being used. Exactly what properties of problems are most important in problem sets that discriminate experts from novices in a measurable way? To answer this question, we studied the categorizations by known physics experts and novices using a large, diverse set of problems. We found that the number of questions required to accurately classify experts and novices can be surprisingly small so long as the problem set is carefully crafted to be composed of problems with particular pedagogical and contextual features. Finally, we found that not only was *what* you ask (deep structure) important, but also *how* you ask it (problem context).

– Context –

- Nobody has straightforwardly replicated Chi et. al.'s [1] study
  - Chi's questions are lost
  - Chi's exact analysis method is lost
- Conclusion bears examining considering above difficulty
- Similar [3–5] have used different methods

–Why Rigging?–

- Many studies have been attempted which do not get Chi's result [4]
- Successful studies [3-5] select problems carefully
- Epistemological resources are activated in a context dependent manner [7,8]
- Categorization result (*winning*) must be sensitive to problem selection (*rigging*)

–Current Study–

Questions for study:

1. What properties of problems best distinguish experts from novices? (How can we "rig" the game to "win"?)
2. How can we determine how well our method distinguishes expert from novice?

Study Method (Developed in [2]):

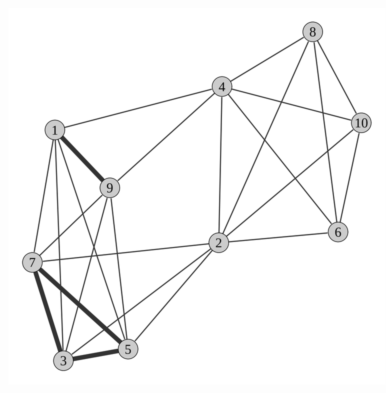
Step 1 - Convert a sorter's categorization into a graph
Step 2 - Compare each graph using distance metric
Step 3 - Visualize sorters using PCA

–Categorizations as graphs–

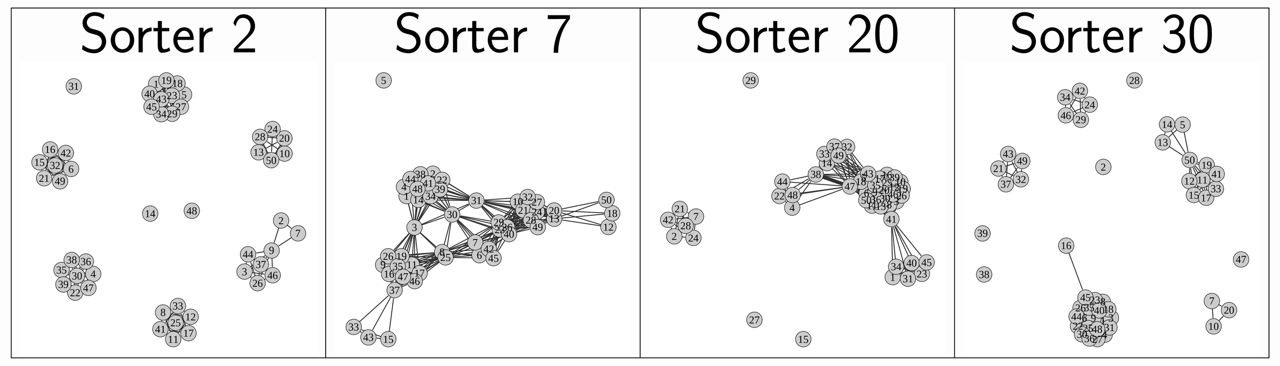
Toy example: Sample categorization of 10 problem set

Categories and problems:

- Newton's second Law $\{1, 3, 5, 7, 9\}$
- Conservation of Energy $\{2, 4, 6, 8, 10\}$
- Conservation of momentum $\{2, 3, 5, 7\}$
- Rotation $\{1, 4, 9\}$



–Typical Sorters–



- # 2 and 7 are *experts*
- # 20 and 30 are *novices*
- # 7 and 20 are *stackers*
- # 2 and 30 are *spreaders*

–Subset Properties–

Problem properties

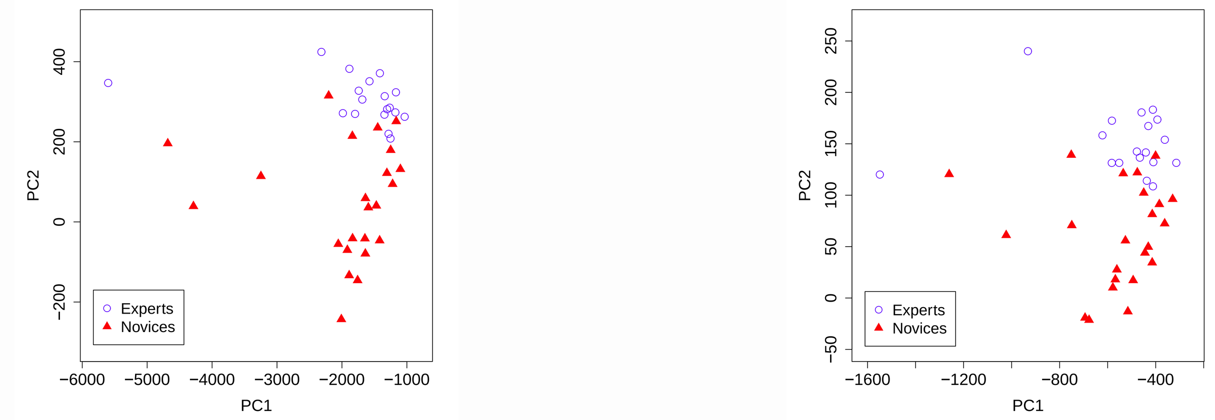
- Taxonomy of introductory physics problems (TIPP) [6]
- Problem difficulty (# of "dots")
- Chapter

Group discrimination properties

- Cramer test statistic
- Hotellings T^2 statistic
- Average Rate of Correct Classification (ARCC)

–Visualizing Sorters–

(a) Entire problem set (b) Problems from Singh's study [5]



- Principal components (PC1 and PC2) – largest sources of variation in categorization data
- PC1 puts stackers on the left and spreaders on the right
- PC2 puts experts to the top and novices to the bottom

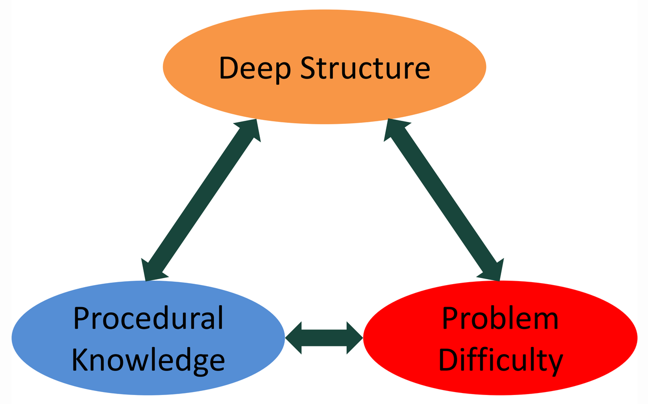
–Subset analysis results–

Problem variable group	Percent variability explained†
TIPP-Declarative	5.4
TIPP-Procedural	30.4
Difficulty	22.8
Chapter	41.4

† Percentage is of the variability explained by the study.

–Source of discrimination–

- Deep structure
  – Problem Chapter
- Procedural knowledge
  – "Make a flow chart"
- Problem difficulty
  – Easier is better



–Rigging plans and challenges–

- Monte Carlo technique: Analyze 5 and 10 problem subsets of large dataset
- Simulated annealing technique: Analyze 25 problem subsets of large dataset
- Study expert/novice discrimination dependence on problem characteristics of subsets
- Challenge: The search space is *huge*!

$$\begin{aligned} \text{Num 5 problem subsets} &= 2,118,760 \\ \text{Num 10 problem subsets} &= 10,272,278,170 \\ \text{Num 25 problem subsets} &= 1.264 \times 10^{14} \end{aligned}$$

- Solution: Run task on a supercomputer (High Performance Computing Cluster)

–Conclusions–

- We can differentiate expert from novice based on their categorizations
- Problem properties only account for 43% of the variability between expert and novice
- "Rigged" sets have problems with these properties:
  1. Variety of chapters
  2. Variety of tasks
  3. Focus on simple problems

–References–

[1] M. T. H. Chi, P. J. Feltovich, and R. Glaser, *Cognitive Science* **5**, 121 – 152 (1981).
[2] S. F. Wolf, D. P. Dougherty, G. Kortemeyer, *Phys. Rev. ST-PER* **8** 010124 (2012).
[3] A. Mason and C. Singh, *Phys. Rev. ST-PER* **7**, 020110 (2011).
[4] G. H. Veldhuis, *Science Education* **74**, 105 – 118 (1990).
[5] C. Singh, *American Journal of Physics* **77**, 73–80 (2009).
[6] R. Teodorescu, et. al. , *PERC Proceedings* **1064**, 203–206 (2008).
[7] B. W. Frank, et. al. , *Phys. Rev. ST-PER* **4**, 020102 (2008).
[8] D. Hammer, *American Journal of Physics* **68**, S52S59 (2000).