

Ganesa Thandavam Ponnuraj¹, Minsu Ha², Ross H. Nehm² (PI)

¹Department of Computer Science, SUNY Stony Brook

²Center for Science and Mathematics Education, SUNY Stony Brook

Assessing evolutionary explanations

Although evolution and natural selection are core ideas in the biological sciences (*Vision and Change*), undergraduate students have great difficulty understanding the causes of evolution and constructing explanations free of naïve ideas or misconceptions. EvoGrader is an assessment tool for grading written (typed) evolutionary explanations.

Evolutionary ideas that EvoGrader can detect:

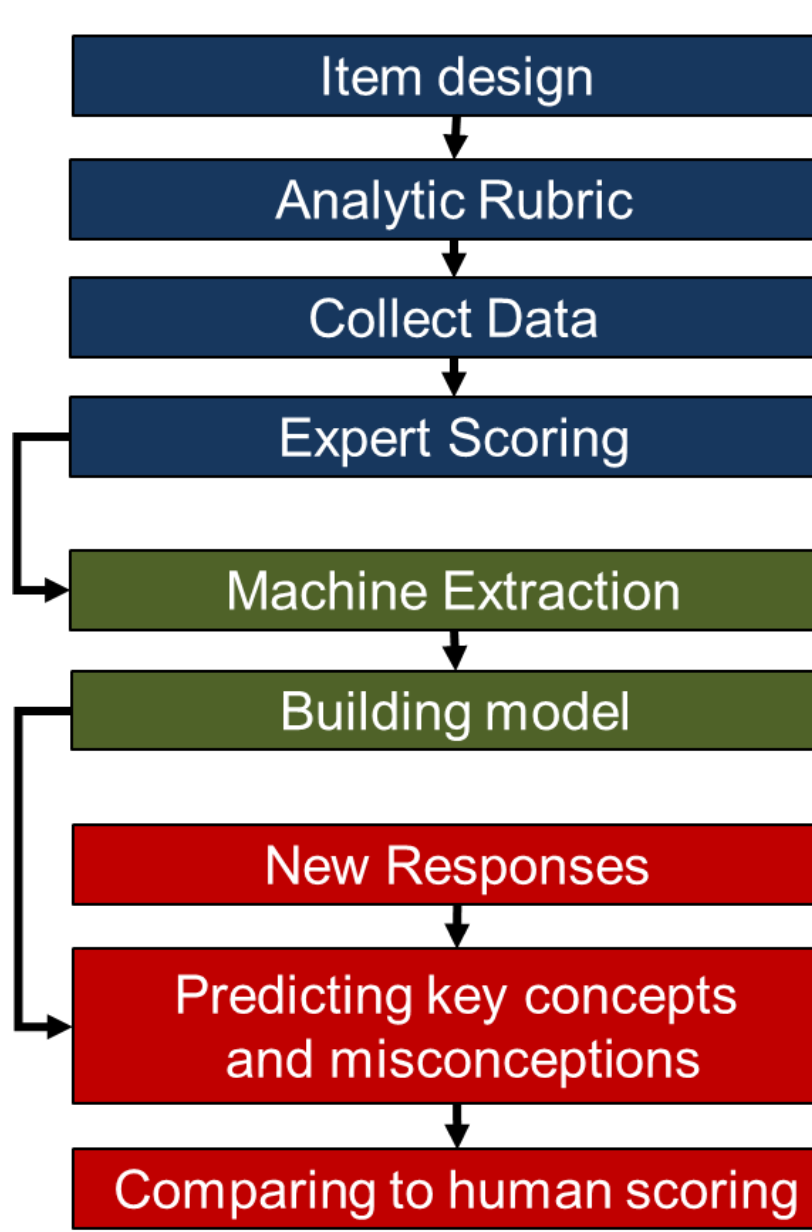
- Normative scientific ideas: Variation, Heritability, Competition, Limited resources, Differential survival/reproduction, and Non-adaptive idea
- Non-normative naïve ideas: Need/goal, Use/disuse, and Adapt/acclimation

Evolutionary reasoning patterns that EvoGrader can detect:

- Normative scientific model (Explanation composed exclusively of scientific ideas)
- Mixed model (Explanation composed of a mixture of non-normative naïve and scientific ideas)
- Non-normative naïve model (Explanation composed exclusively of non-normative naïve ideas)

Overview of the web-based system

- A special purpose interface was built using the LightSIDE program
- New models were built to detect evolutionary concepts
- An online platform was developed in J2EE
- The site is hosted by Amazon Elastic Cloud (reliable response time for concurrent queries)
- Designed in an Easy-To-Use fashion



Publications and more information

- REESE-0909999 (PI Nehm)
- TUES-1322872 (PI Nehm)
- Nehm lab: <http://www.nehmlab.org>
- <http://evolutionassessment.org>
- <http://evograder.org>
- Related publications

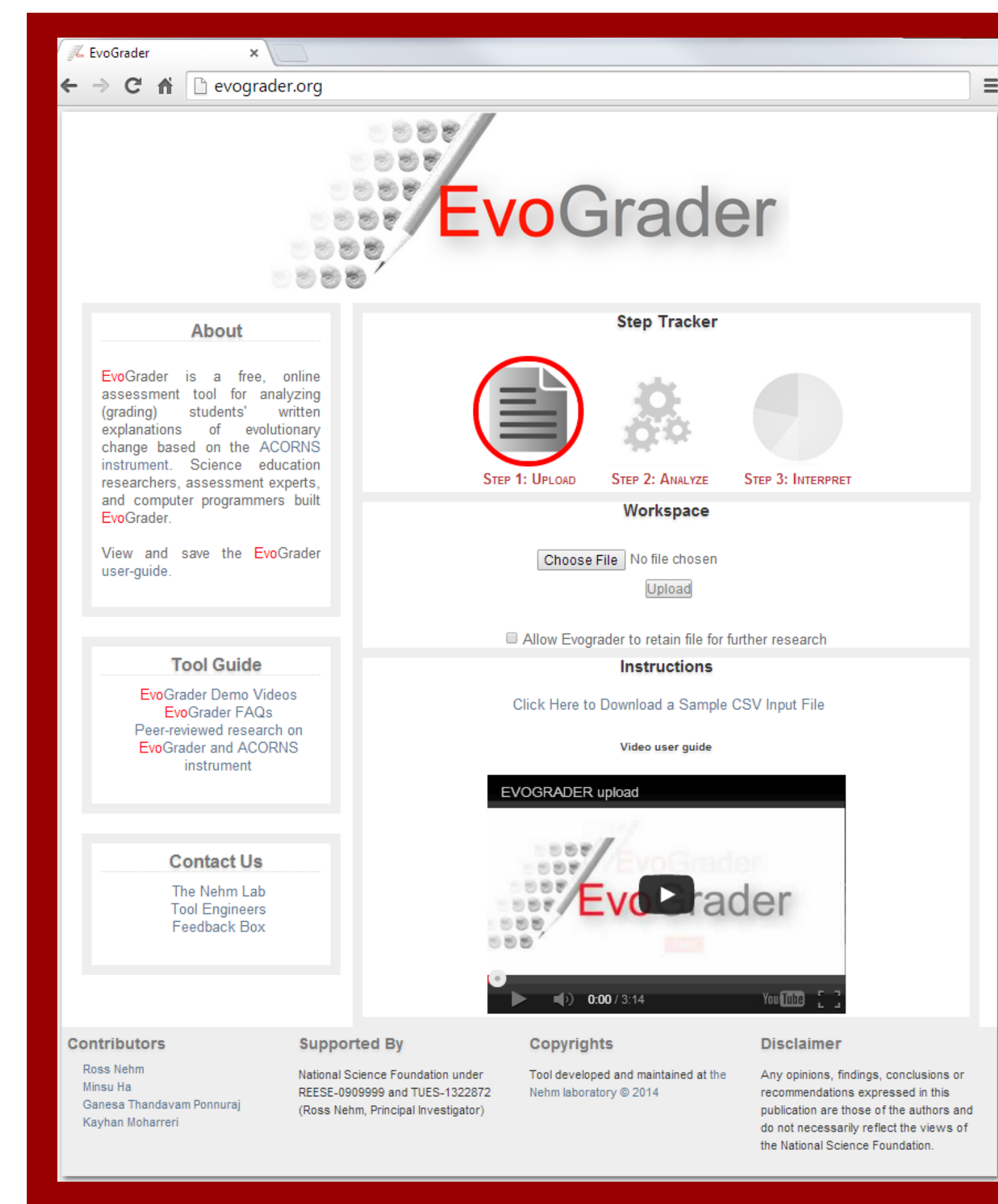


- Beggrow, E. P., Ha, M., Nehm, R. H., Pearl, D., & Boone, W. J. (2014). Assessing scientific practices using machine-learning methods: How closely do they match clinical interview performance? *Journal of Science Education and Technology*, 23(1), 160-182.
- Opfer, J. E., Nehm, R. H., & Ha, M. (2012). Cognitive foundations for science assessment design: Knowing what students know about evolution. *Journal of Research in Science Teaching*, 49(6), 744-777.
- Nehm, R. H., Ha, M., & Mayfield, E. (2012). Transforming biology assessment with machine learning: Automated scoring of written evolutionary explanations. *Journal of Science Education and Technology*, 21(1), 183-196.
- Ha, M., Nehm, R. H., Urban-Lurain, M., & Merrill, J. E. (2011). Applying computerized scoring models of written biological explanations across courses and colleges: Prospects and limitations. *CBE-Life Science Education*, 10, 379-393.

What is EvoGrader?

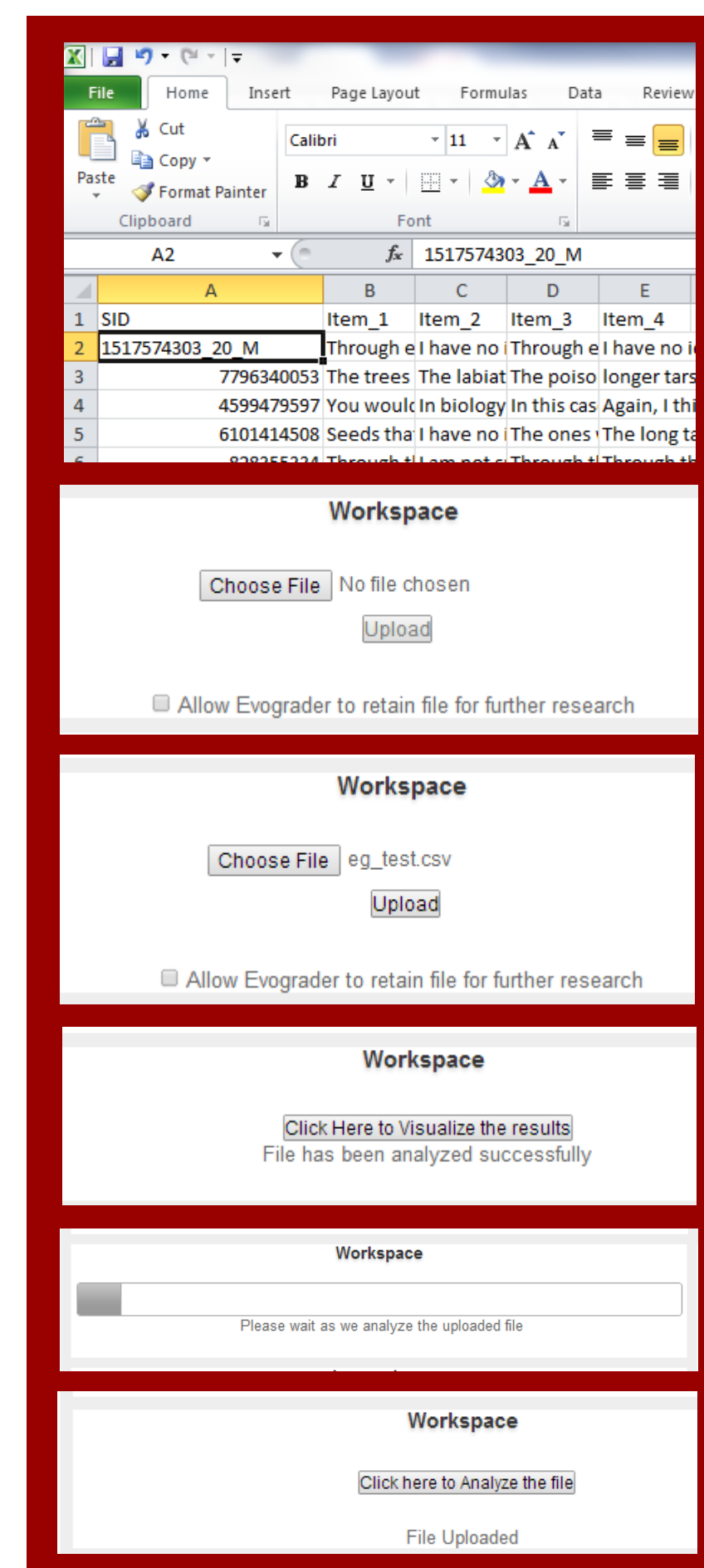
- EvoGrader is a free, online assessment tool for analyzing (grading) students' written explanations of evolutionary change.
- EvoGrader uses machine-learning methods to analyze student responses. A corpus of more than 10,000 human-graded written responses to many different types of questions was used to build the models that score student responses.

Overview of EvoGrader (<http://www.evograder.org>)



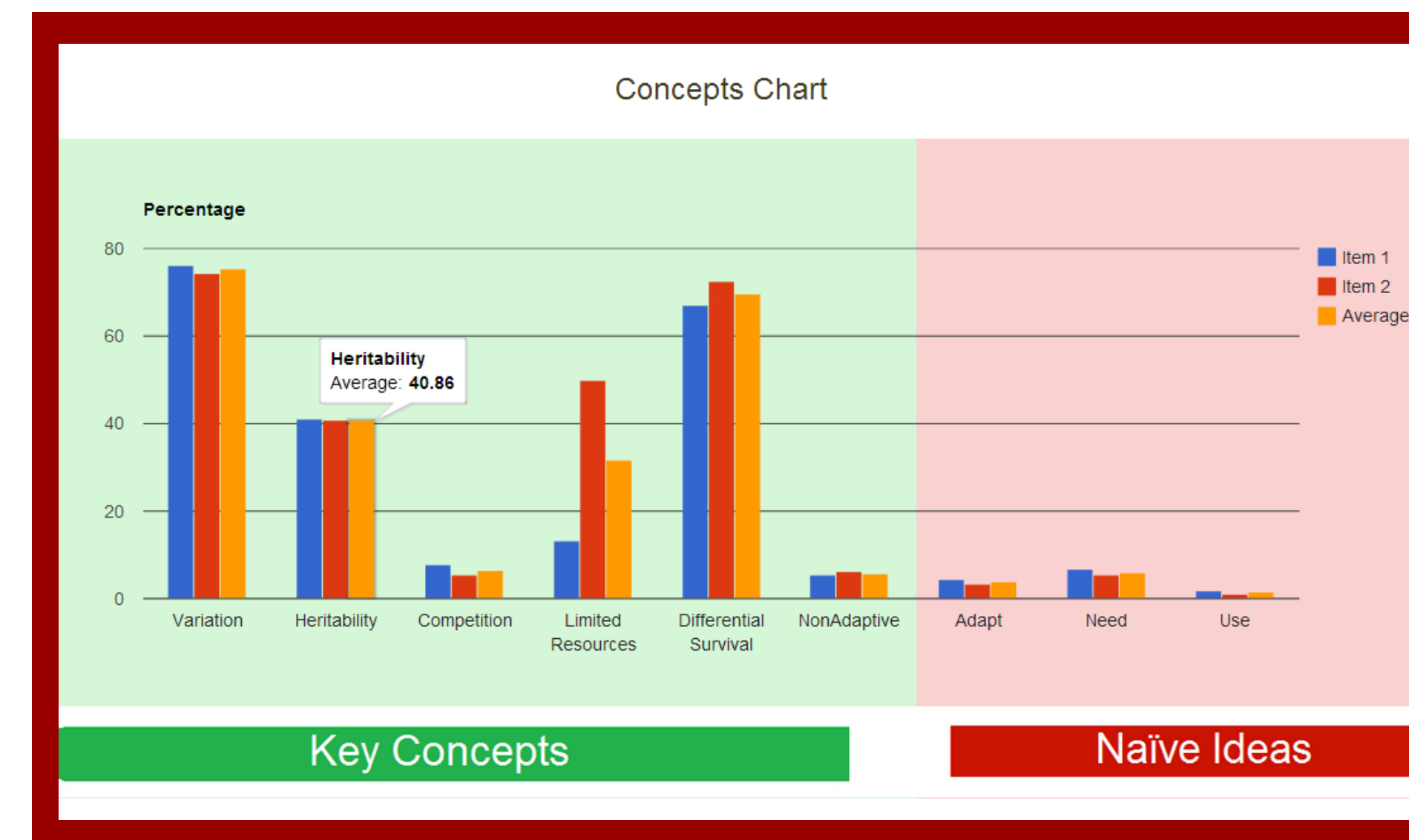
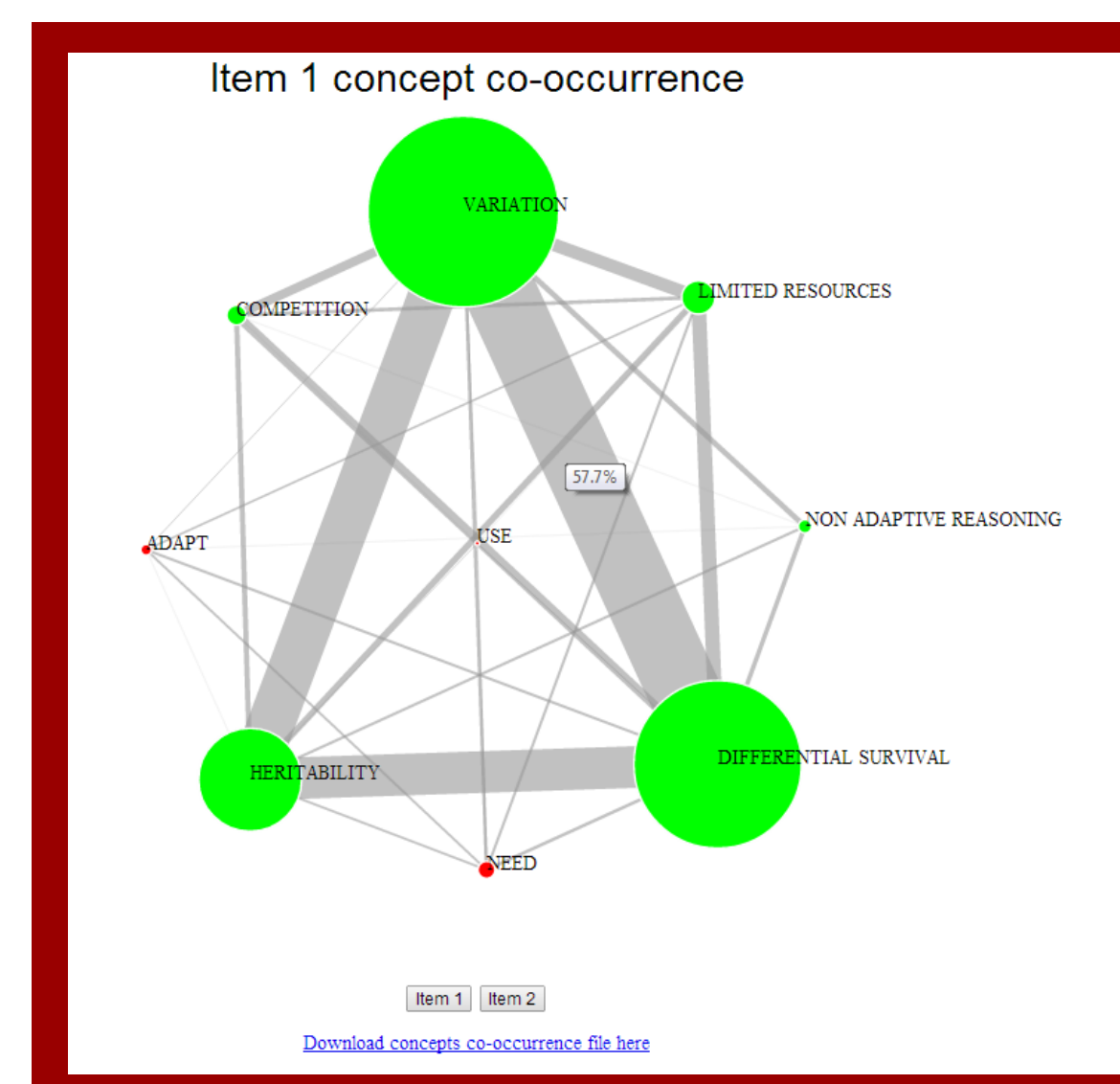
How to use EvoGrader

- It's easy to use EvoGrader! All you need to do is upload an Excel file (.csv format) with students' written responses, and EvoGrader analyzes the responses for a variety of concepts--both scientific and naïve--and provides data about how the students are performing.
- The first cell of the first column must contain the heading SID; you can add students' identification in the first column.
- The first cell of the second column must contain the heading 'Item' underscore '1'.
- The item order should always begin with 1 and extend up to 8. 8 is maximum number of item that the system can process in a single analysis.

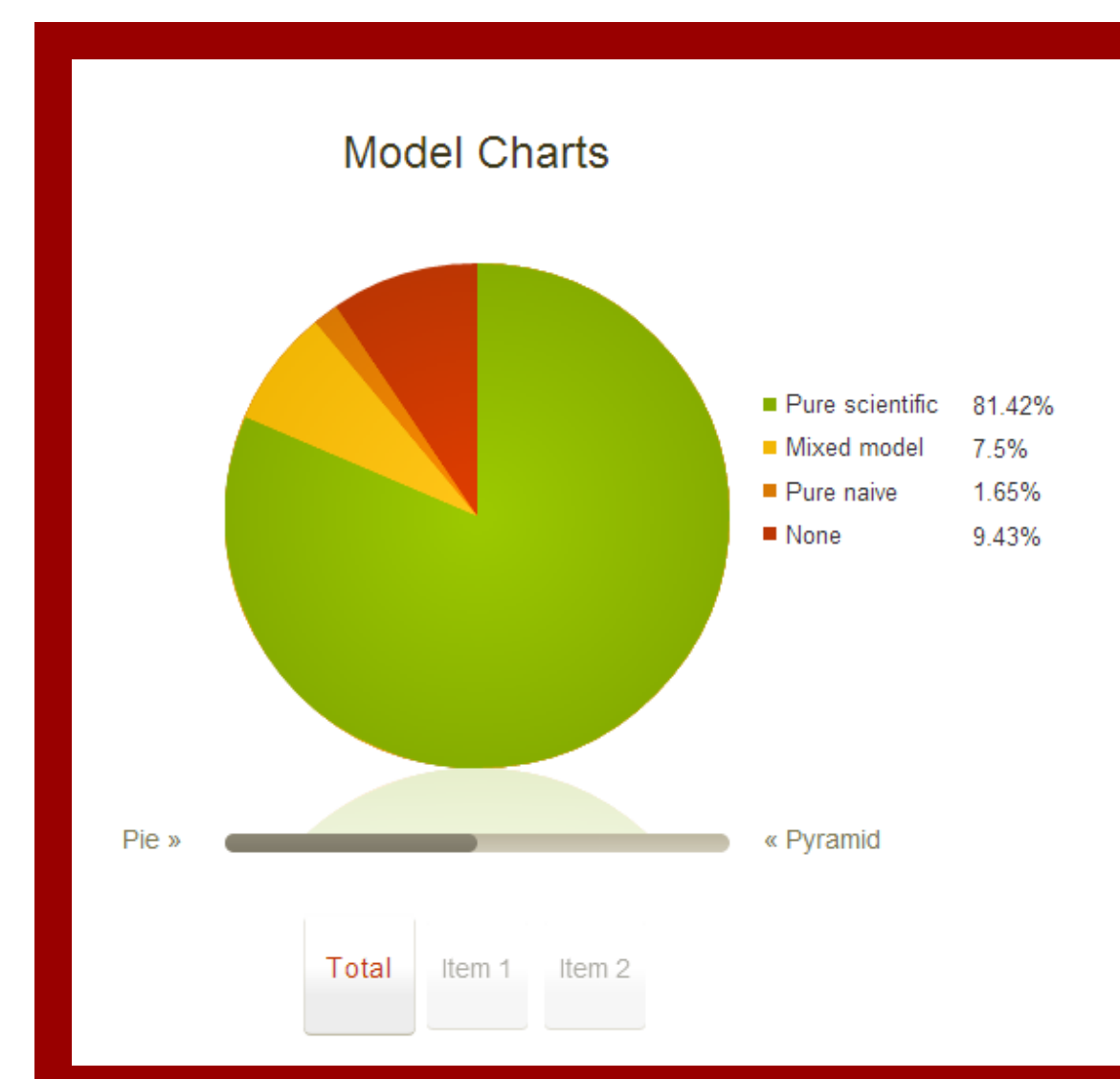


EvoGrader Results

- Bubble chart for concept co-occurrence
- Bar graph for concept frequency
- Table for concept frequency
- Pie chart for reasoning models
- EvoGrader scored file in excel



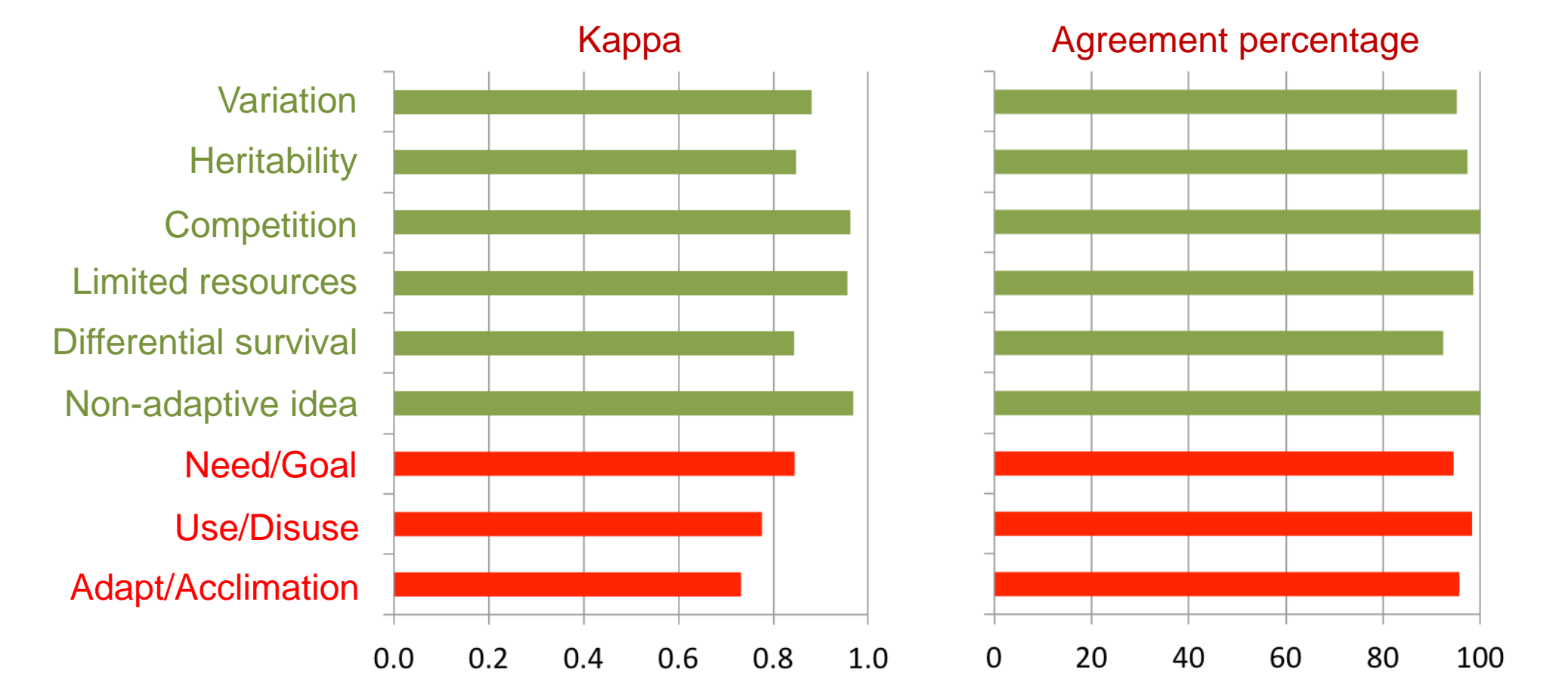
Reasoning Model	Percentage
Pure scientific	81.42%
Mixed model	7.5%
Pure naïve	1.65%
None	9.43%



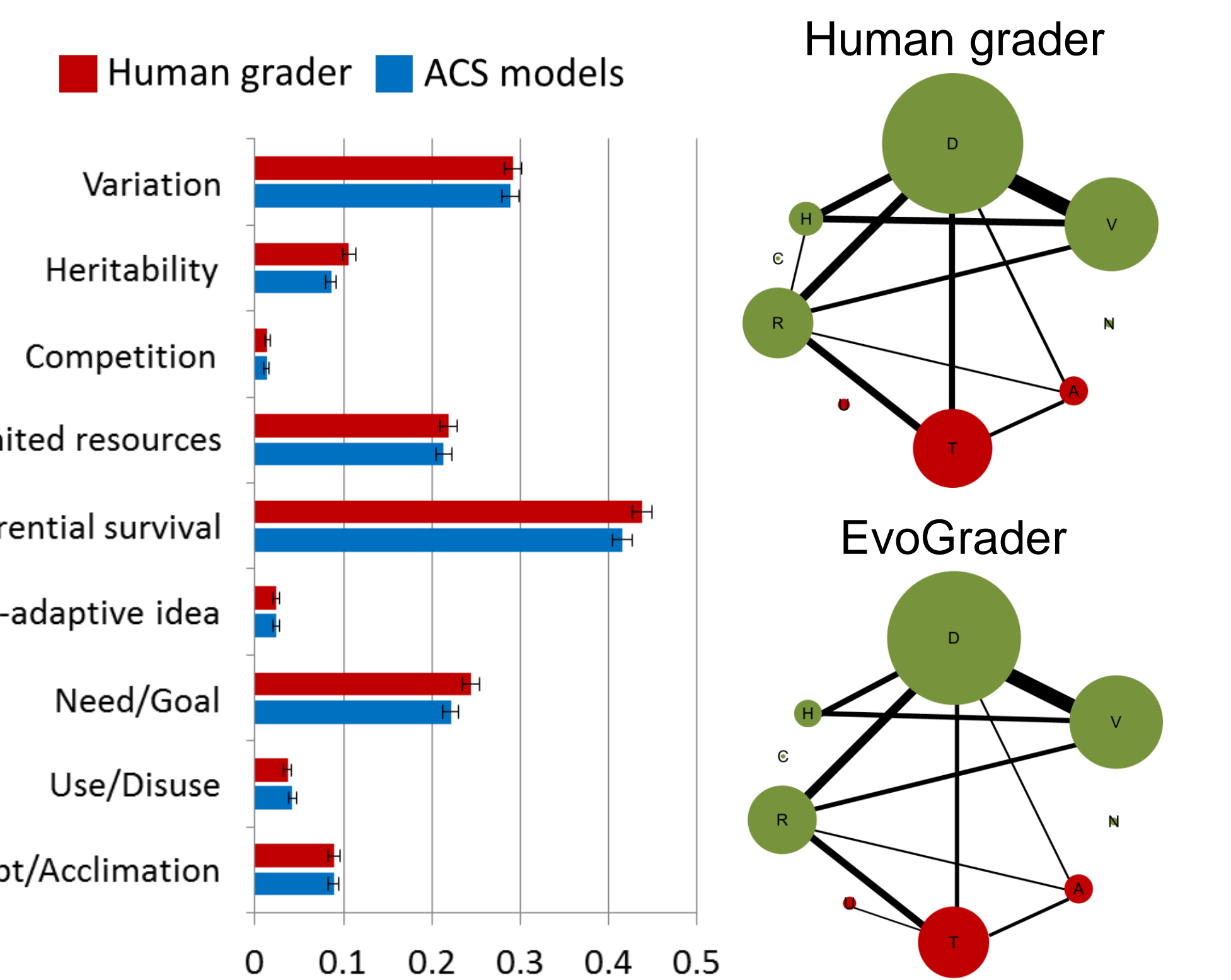
SID	Item_1	KC_Variation	KC_Heritability	KC_Competition	KC_LimitedResources	KC_DifferentialSurvival	KC_NonAdaptive	KC_Adapt	KC_Need	KC_Use	KC_Score	Model_Type
1	1317574303_20_M	1	1	0	0	1	0	0	0	0	3	0
2	7.8E+09	1	1	0	0	1	0	0	0	0	3	0
3	4.6E+09	1	0	0	1	1	0	0	0	0	3	0
4	6.1E+09	1	1	0	0	1	0	0	0	0	3	0
5	8.2E+09	1	1	0	0	1	0	0	0	0	3	0
6	8.0E+09	1	0	0	0	1	0	0	0	0	2	0
7	3.5E+09	1	0	0	0	1	0	0	0	0	2	0
8	7.0E+09	0	0	0	0	0	0	0	0	0	0	4
9	1.84E+09	1	1	1	1	1	1	1	1	1	5	0
10	1.97E+09	0	0	0	0	0	0	0	0	0	0	4
11	6.27E+09	0	0	0	0	0	0	0	0	0	0	4
12	5.84E+09	0	0	0	0	0	0	0	0	0	0	4
13	8.19E+09	1	0	1	0	1	0	0	0	0	3	0
14	8.17E+09	1	0	0	0	1	0	0	0	0	2	0

Evidence for the Efficacy of EvoGrader

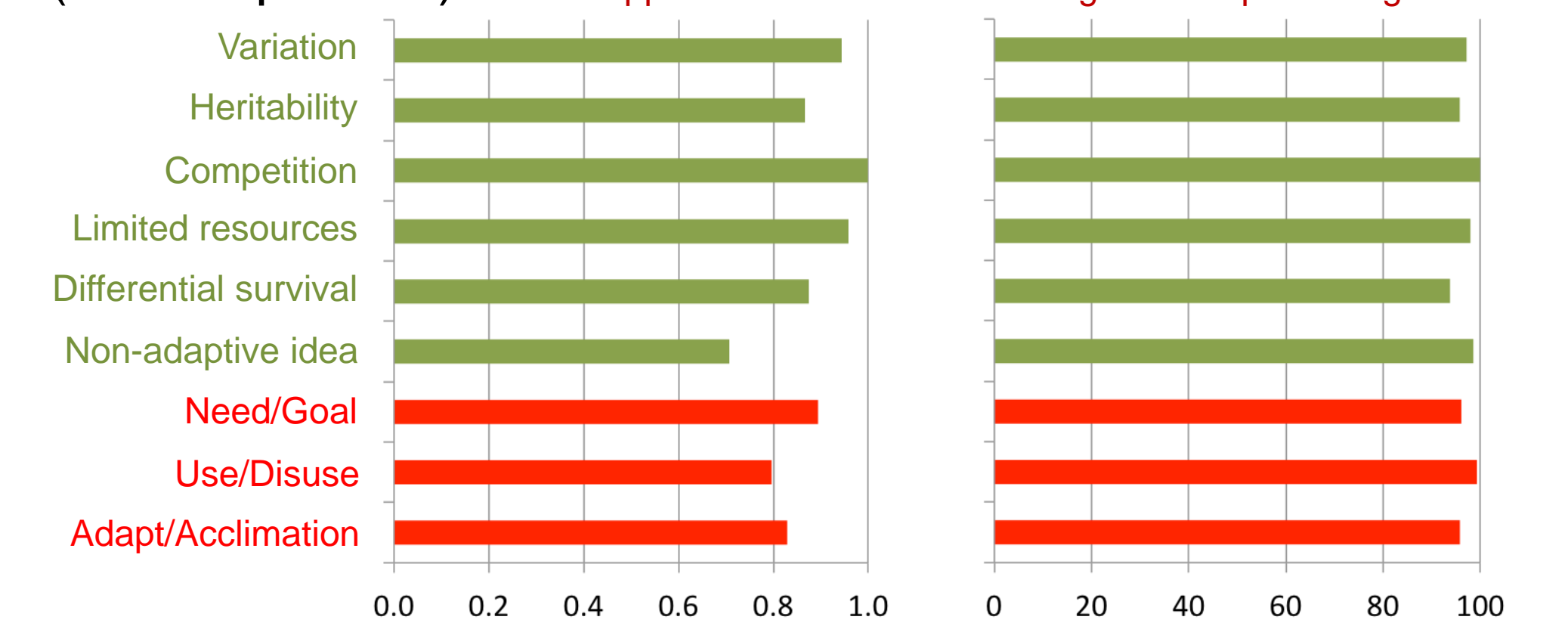
Human-EvoGrader correspondence (2000 responses)



Correlation: Scientific ideas ($r = 0.96, p < 0.0001$)
Naïve idea ($r = 0.90, p < 0.0001$)

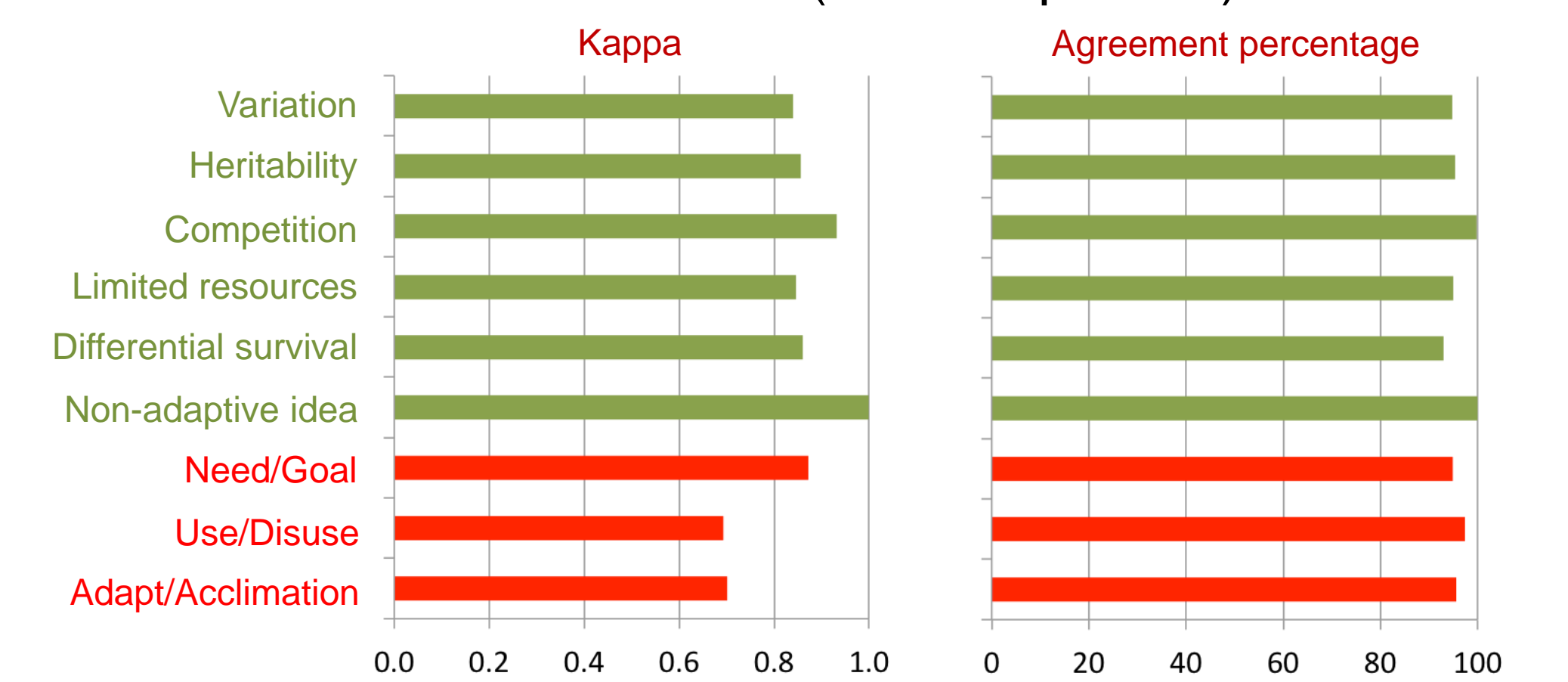


How well does EvoGrader work with non-ACORNs items? (285 responses)



Correlation: Scientific ideas ($r = 0.94, p < 0.0001$)
Naïve idea ($r = 0.88, p < 0.0001$)

Non-ACORNs: Orchid-leaves (1112 responses)



Correlation: Scientific ideas ($r = 0.90, p < 0.0001$)
Naïve idea ($r = 0.85, p < 0.0001$)