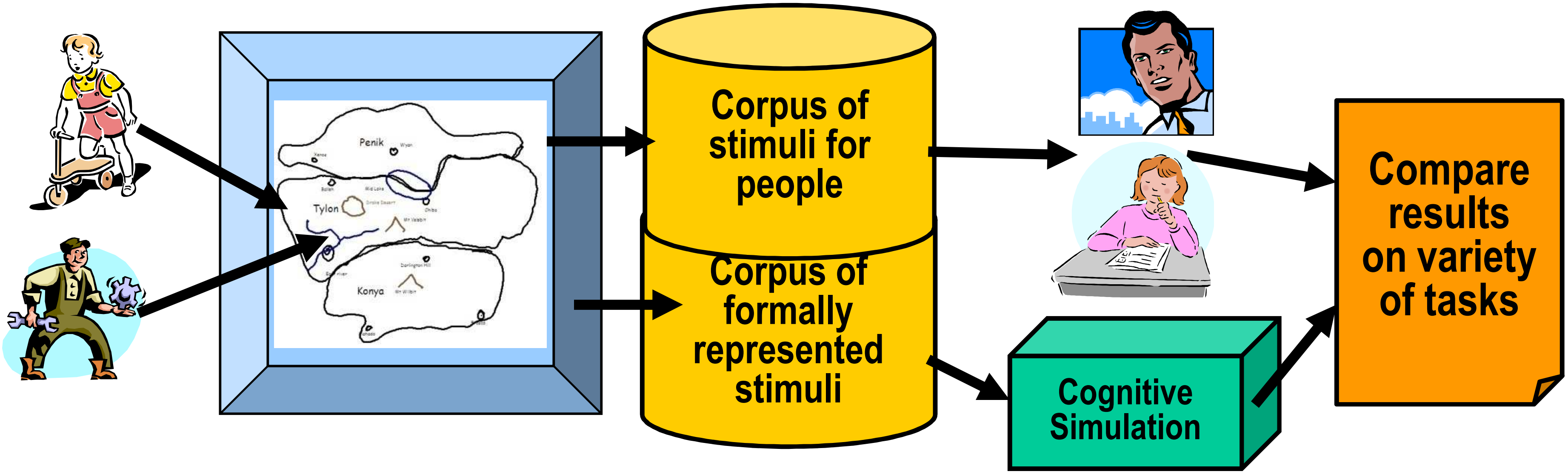
Support for Cognitive Science Experiments

CogSketch as Research Instrument



Gathering and modeling data in laboratory and classroom experiments

- Model existing psychological experiments
- Collect human data via sketching
- As visual/spatial processing calibrated, provide automatic data analysis facilities

Two Roles in Experiments

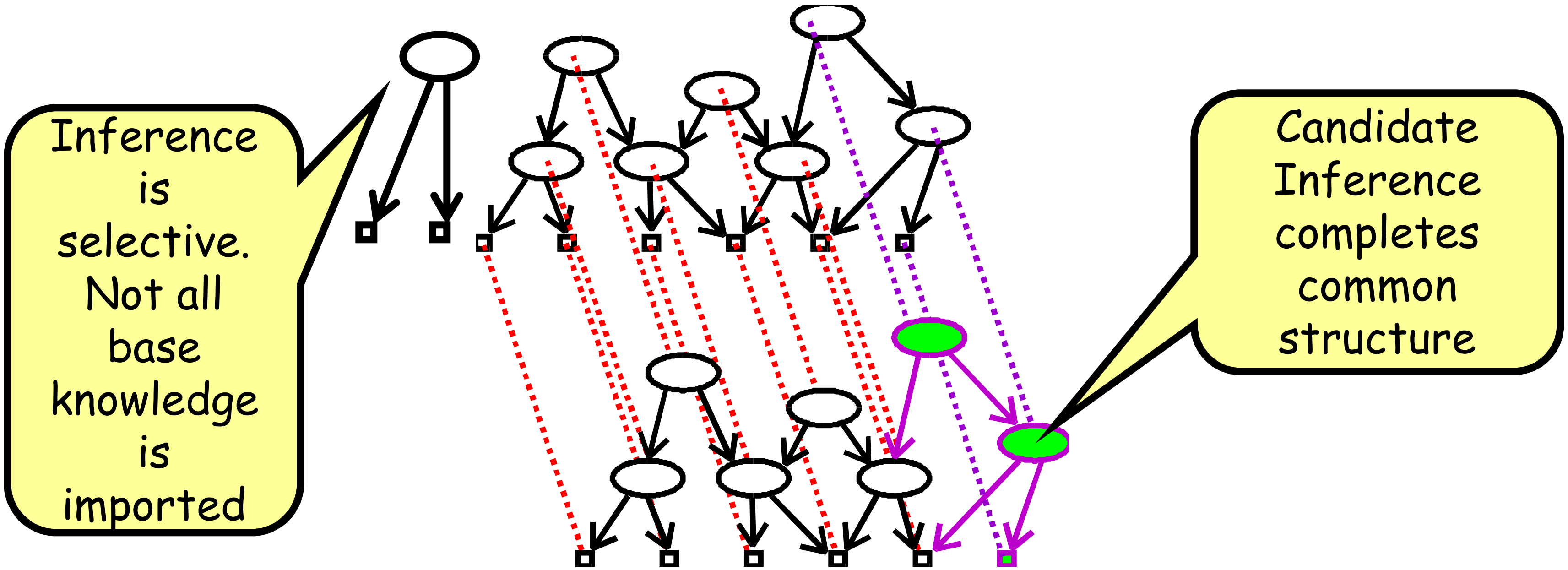
- Cognitive simulation platform
  - Including Evans analogy examples
- Gathering & analyzing human data
  - Exporting ink data
  - Interface simplifications

Two Hypotheses about Human Cognition

- Analogy is a central mechanism of reasoning and learning
  - cf. Gentner's *Why we're so smart* (2003)
- Common sense reasoning primarily relies on within-domain analogical reasoning and generalization, organized around qualitative representations
  - Forbus & Gentner, 1997
- Implications:
  - Symbolic, relational representations essential
  - Matching, not chaining

Structure-Mapping Theory (Gentner, 1983)

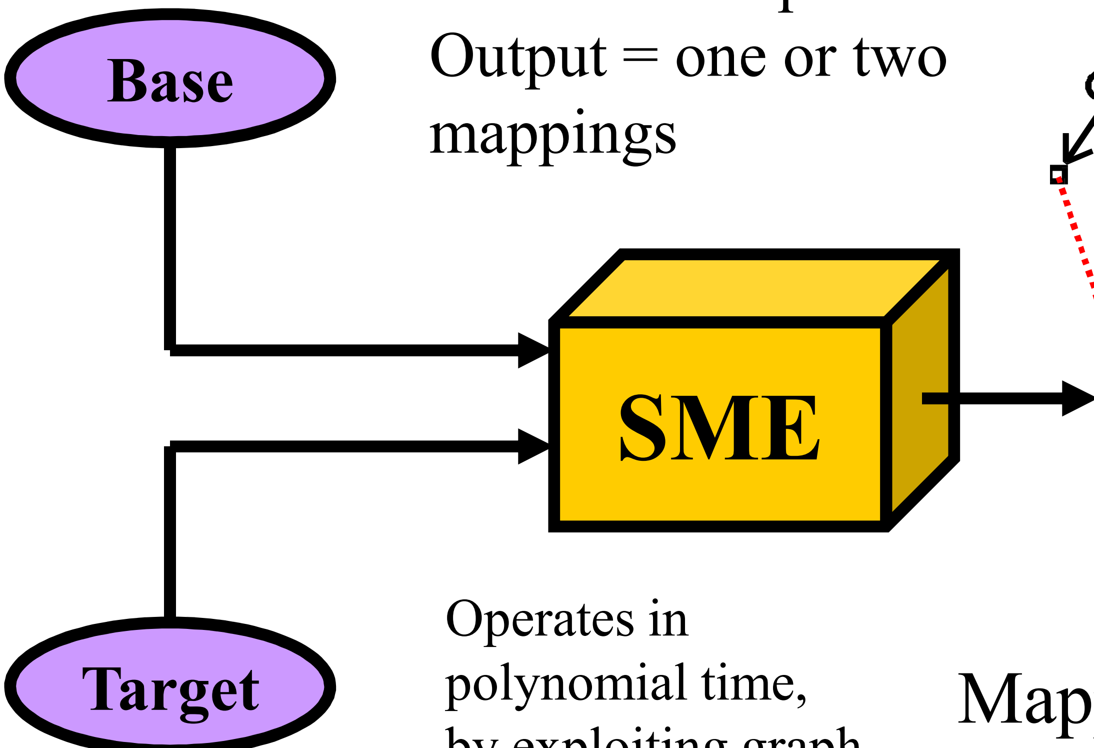
- Analogy and similarity involve
  - correspondences between *structured* descriptions
    - Feature vectors are inadequate to model human cognition
  - candidate inferences fill in missing structure in target

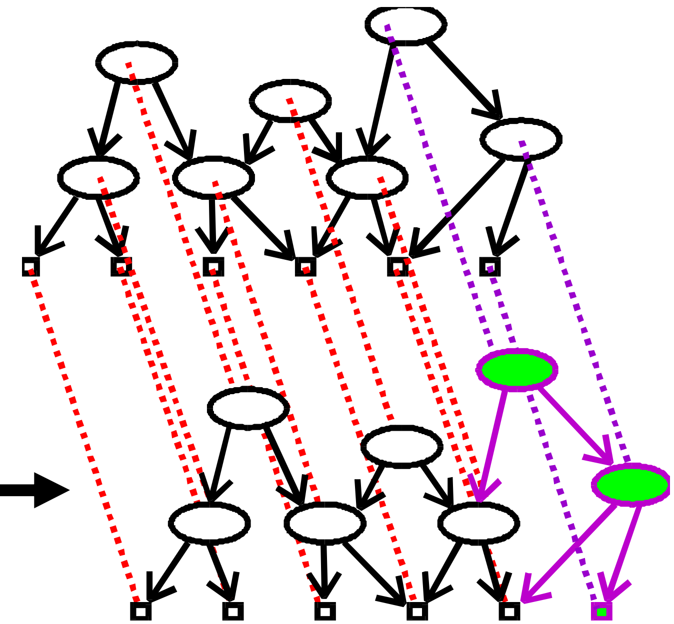


- Also provides account of similarity, metaphor
- Growing body of psychological evidence that same processes are used in perception, problem solving, and conceptual change

SME: Structure-Mapping Engine

Inputs = propositional descriptions, w/ incremental updates
Output = one or two mappings





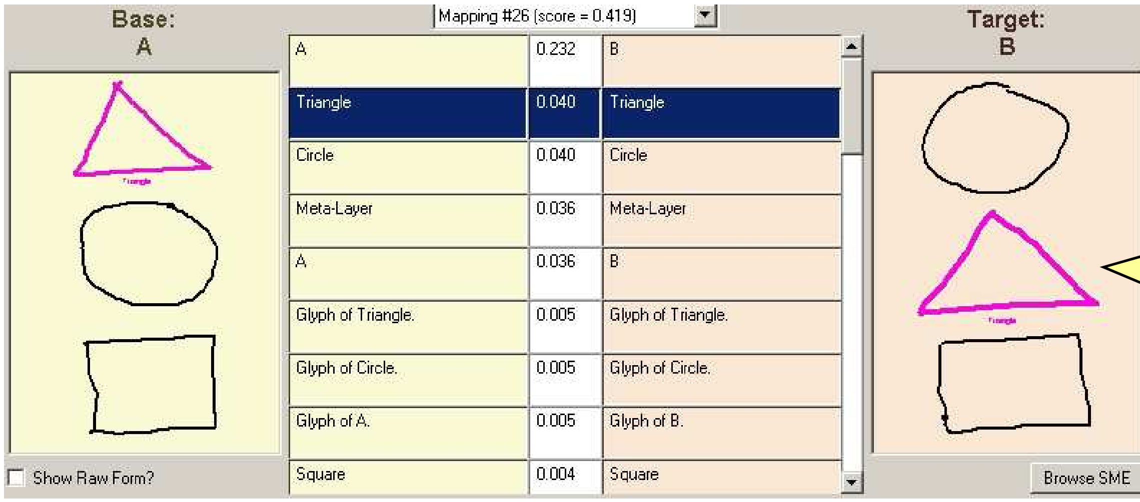
Operates in polynomial time, by exploiting graph labels & greedy algorithms

Mappings = correspondences
+ structural evaluation
+ candidate inferences

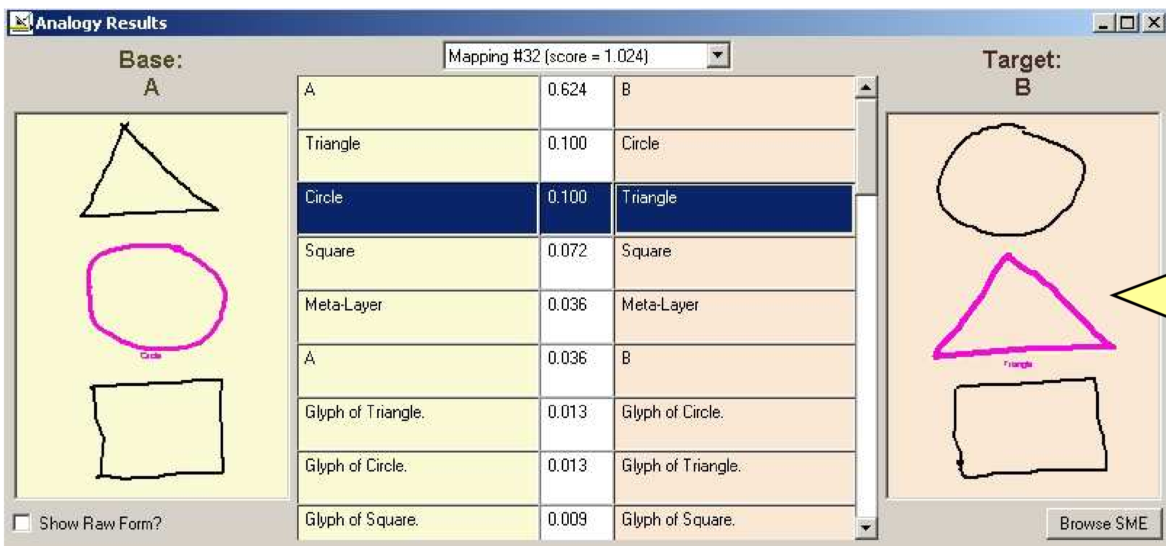
Computational Properties of SME

- Scalable
  - Cases can contain thousands of propositions
  - Cases can be dynamically constructed and expanded from knowledge base contents
- Flexible
  - Has been used with large knowledge bases developed by others (e.g., Cyc, KM)
- Supports Integration
  - Analogy ontology enables smooth integration with logical reasoners
- SME remains the only general-purpose cognitive simulation of analogical matching used with multiple knowledge systems and reasoners

SME can operate over visual structure

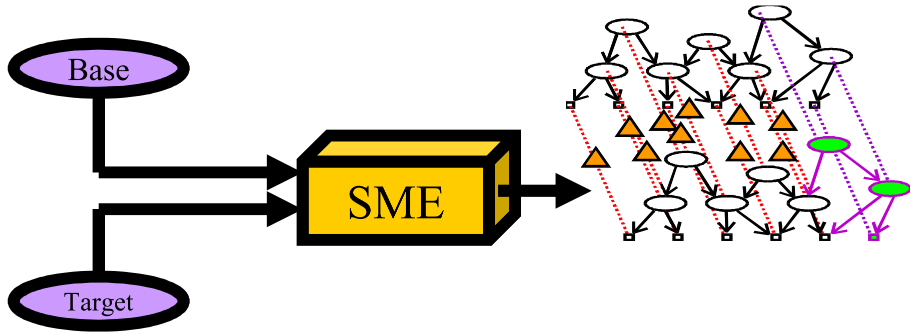


Corresponds to what people choose in fast response-time task

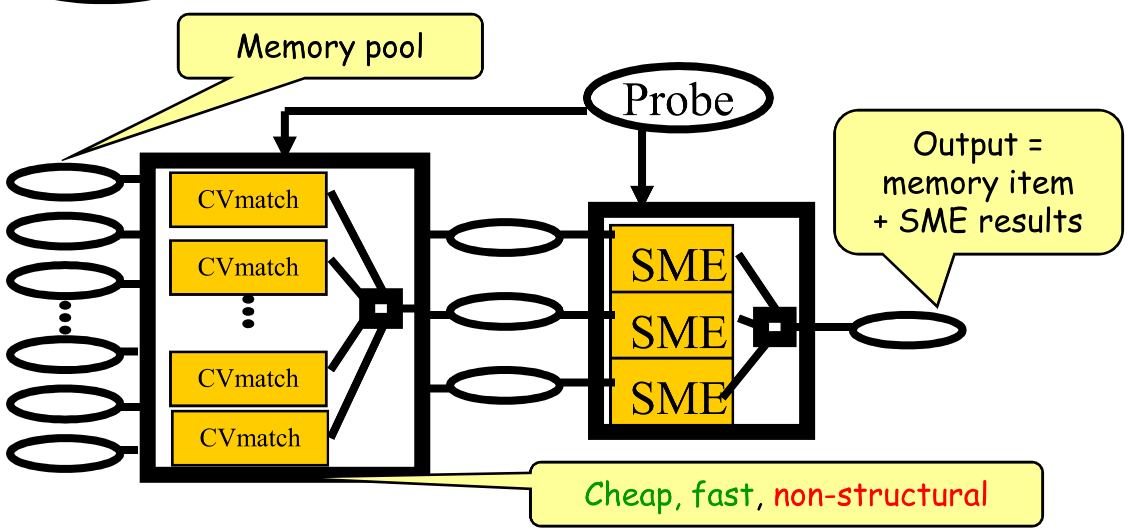


Corresponds to what people choose when given more time

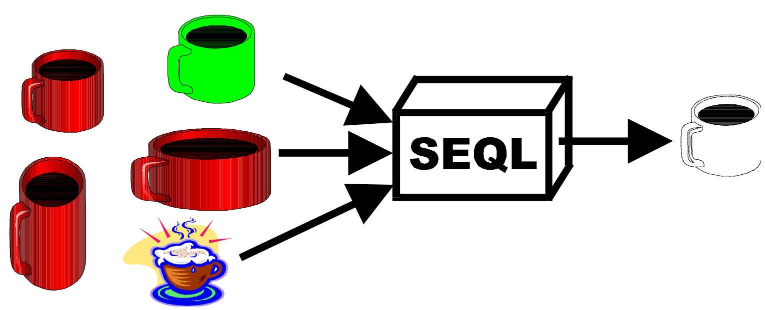
Building Blocks for Analogical Processing



SME = Matching



MAC/FAC = Retrieval



SEQL = Generalization w/probabilities

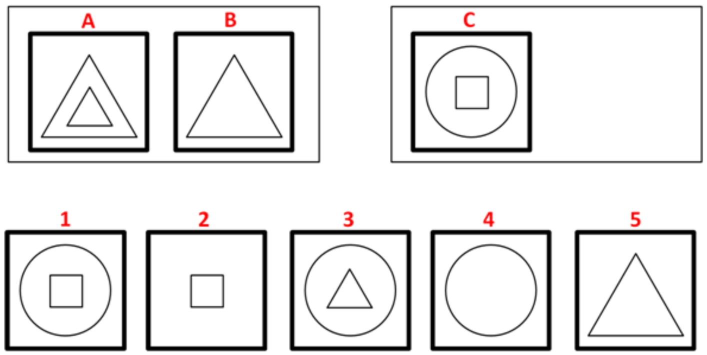
Psychological evidence (examples)

- Used to model existing findings
  - e.g., SME models effects of relational structure on similarity
  - e.g., MAC/FAC models dissociation between surface effects on reminding versus preference for deep structure in mapping
  - e.g., SEQL-based model of Marcus experiment still only one that learns in same span of stimuli as infants, and can handle noise
- Used to predict new findings
  - e.g., SME: Initial stage of metaphor processing is symmetric
  - e.g., SEQL: Can generate orders of presentation which can help/hurt concept learning
- A number of aspects not yet modeled
  - e.g., Working memory capacity limits

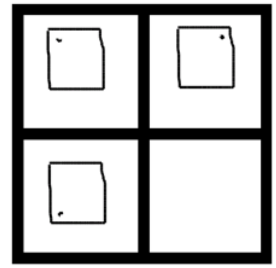
Some CogSketch Simulation Examples

Geometric Analogy

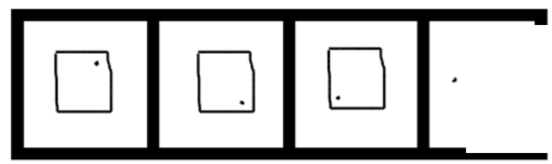
- Problems of the form "A is to B as C is to __?"



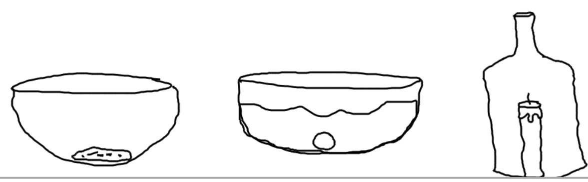
Raven's Progressive Matrices



- Used to measure intelligence
- Extensive data on human performance available

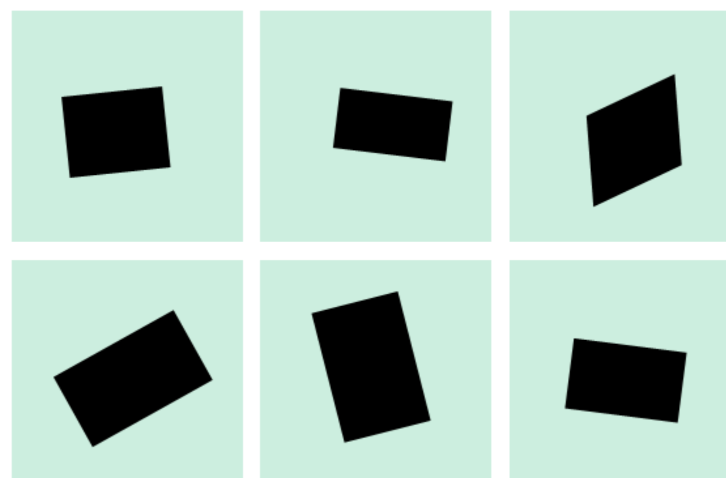


Learning spatial prepositions



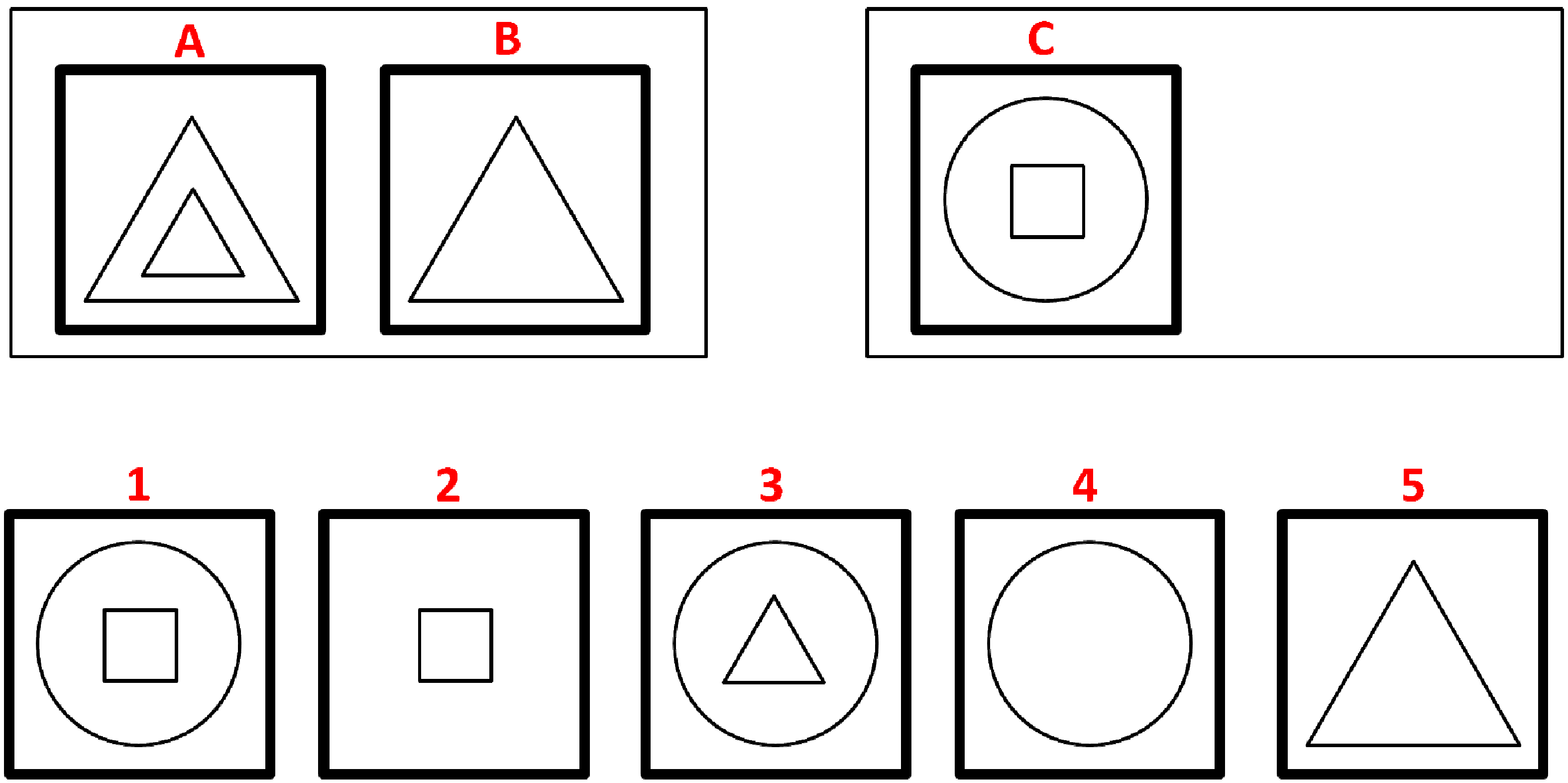
Best Generalization IN
Size: 3
(*candle in bottle, cookie in bowl, marble in water*)
--DEFINITE FACTS:
(rcc8-TPP figure ground)
--POSSIBLE FACTS:
33% (Basin ground)
33% (Bowl-Generic ground)

Visual Oddity Task



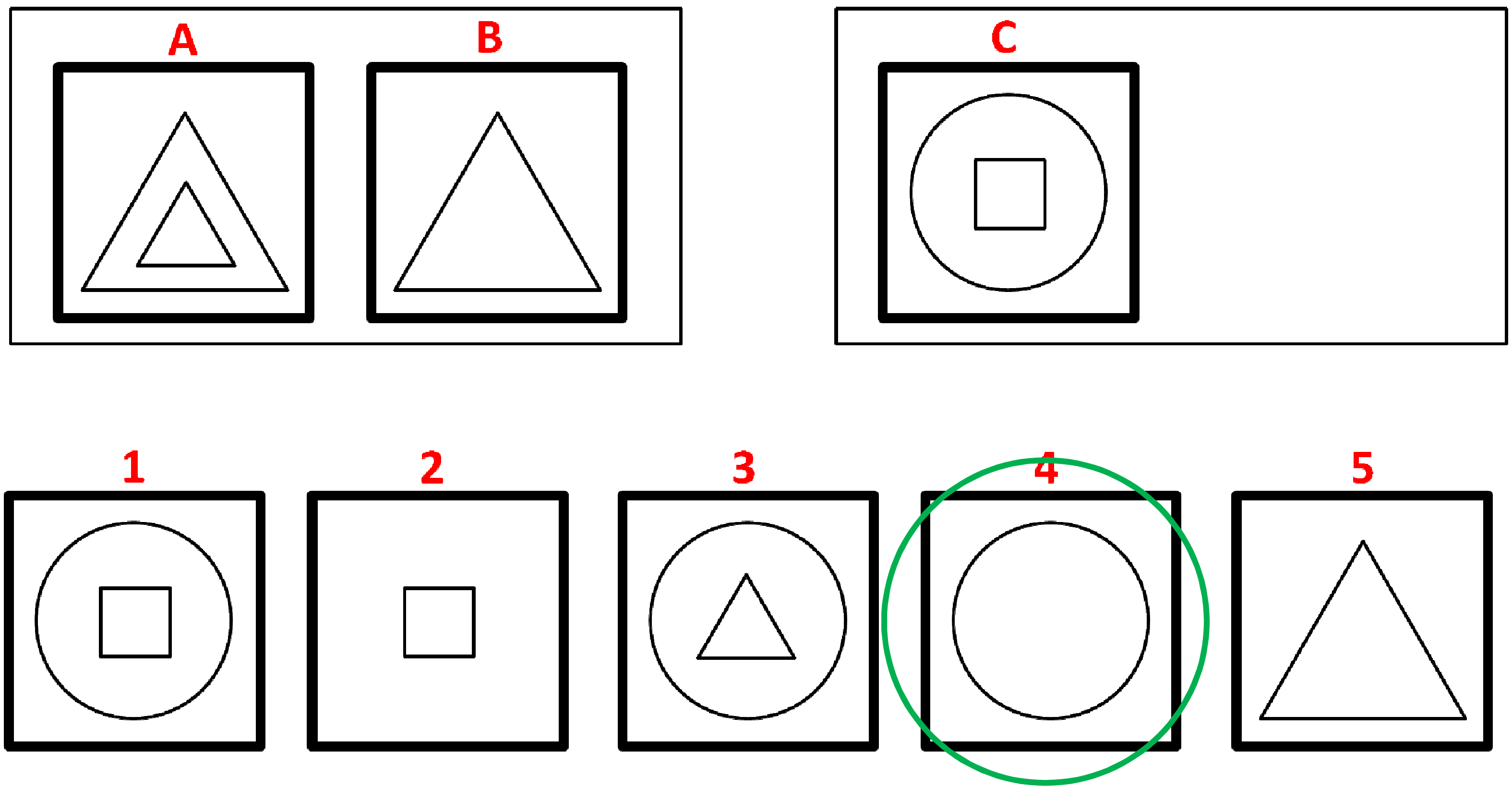
Geometric Analogy

- Problems of the form "A is to B as C is to __?"

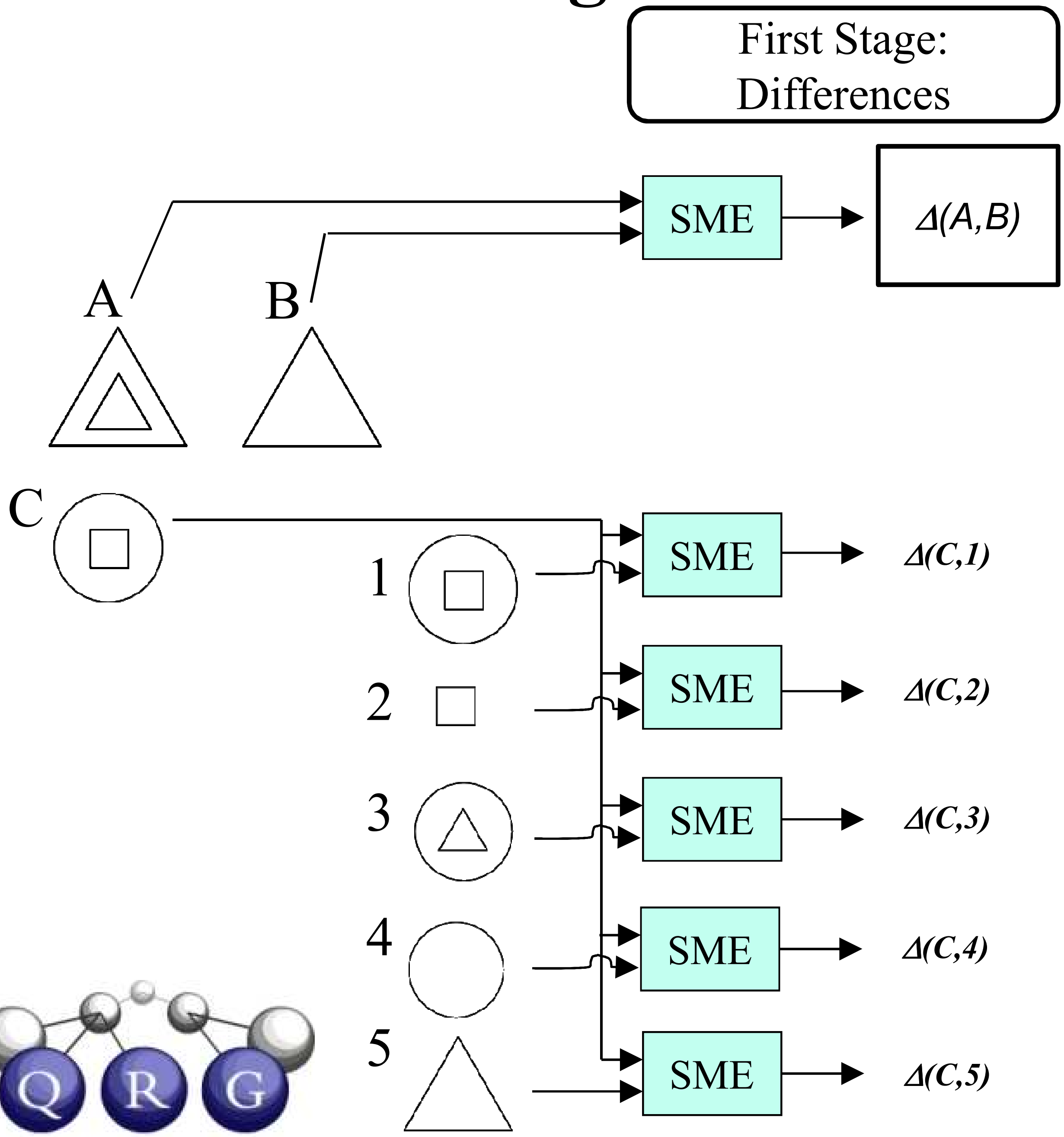


Geometric Analogy

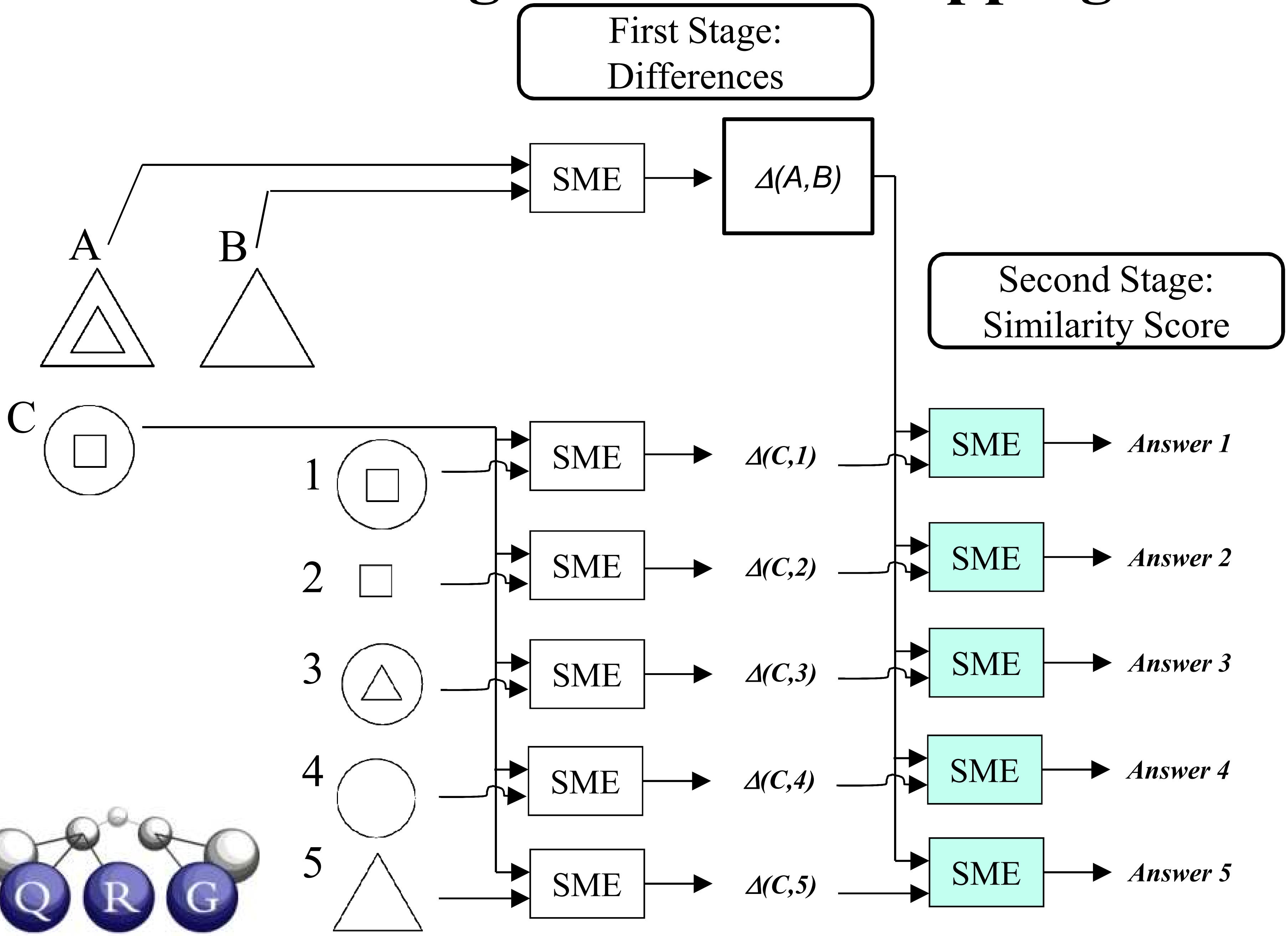
- Problems of the form "A is to B as C is to __?"



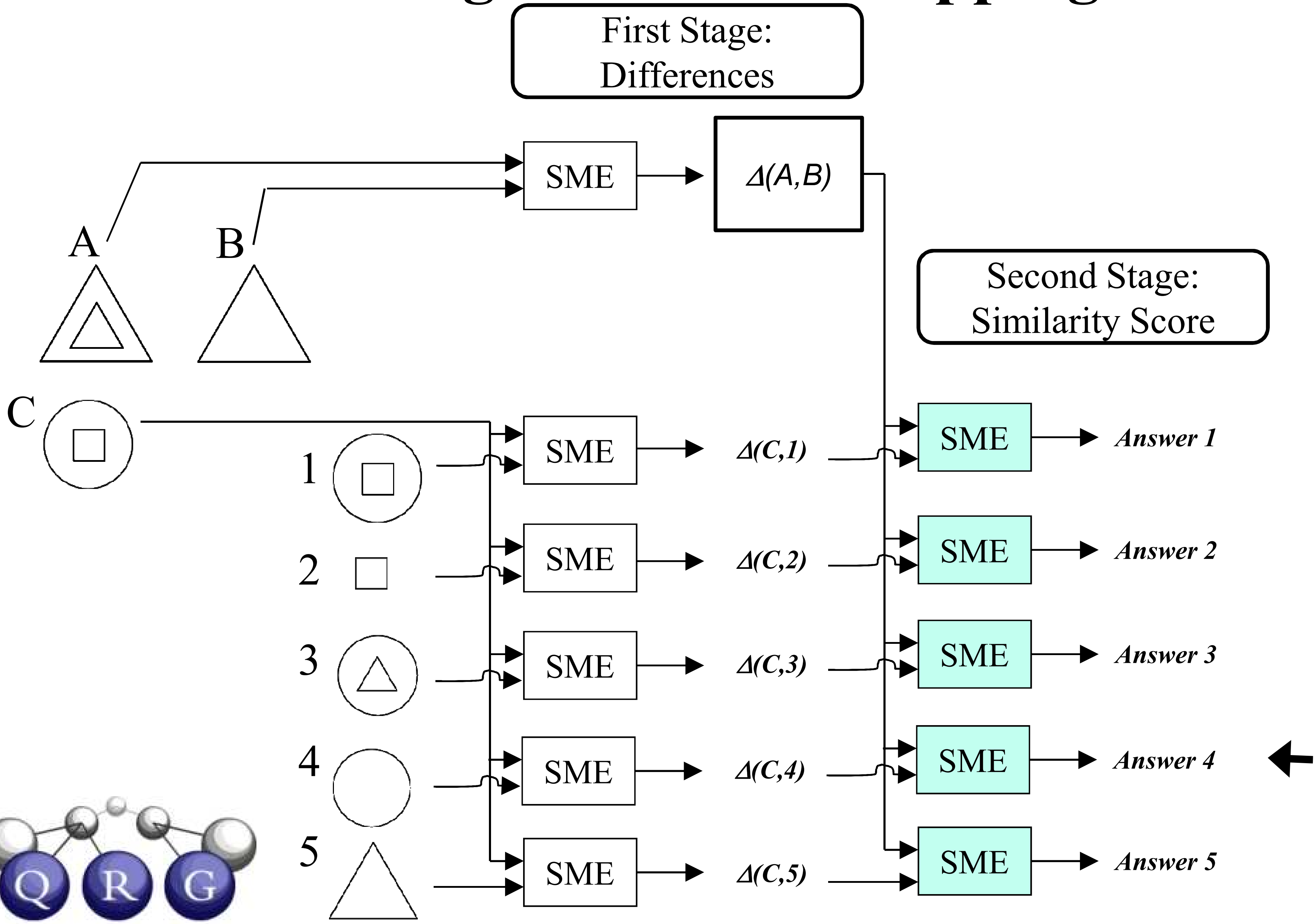
Two-Stage Structure Mapping



Two-Stage Structure Mapping



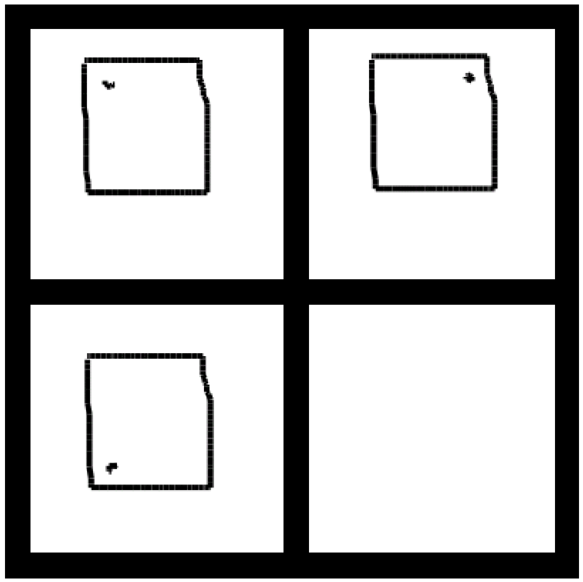
Two-Stage Structure Mapping



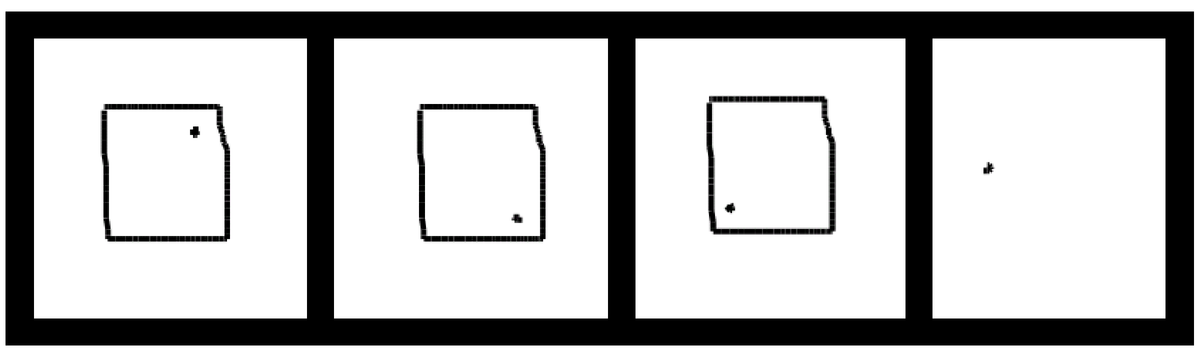
Evaluation

- Constructed 20 problems originally used by Evans (1968) with his landmark ANALOGY system
  - Gave problems to 34 participants
  - Ran problems on computational model
- Model chooses answer preferred by humans on all 20 problems
- Model shows a .75 correlation with human timing data

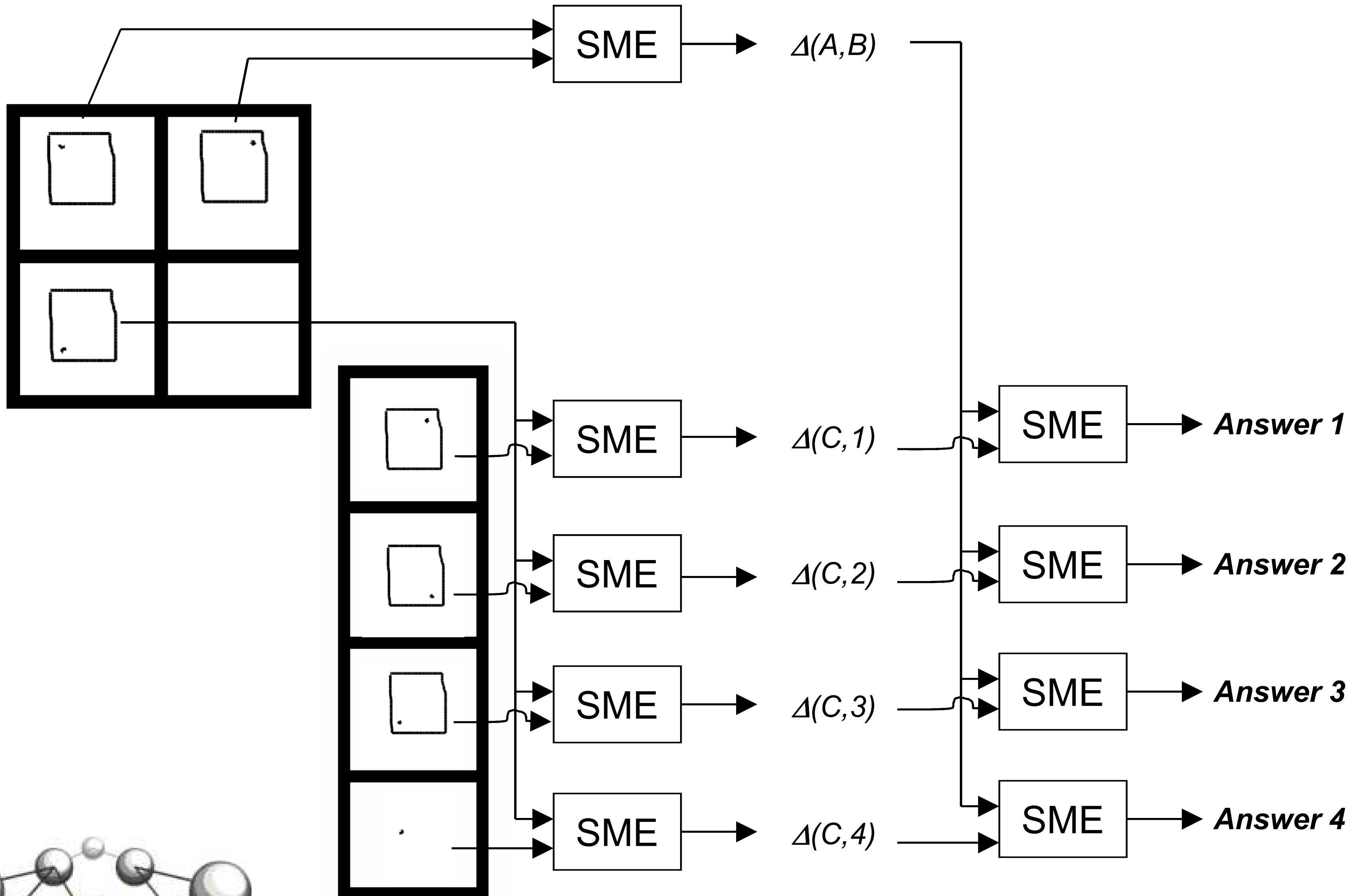
Raven's Progressive Matrices



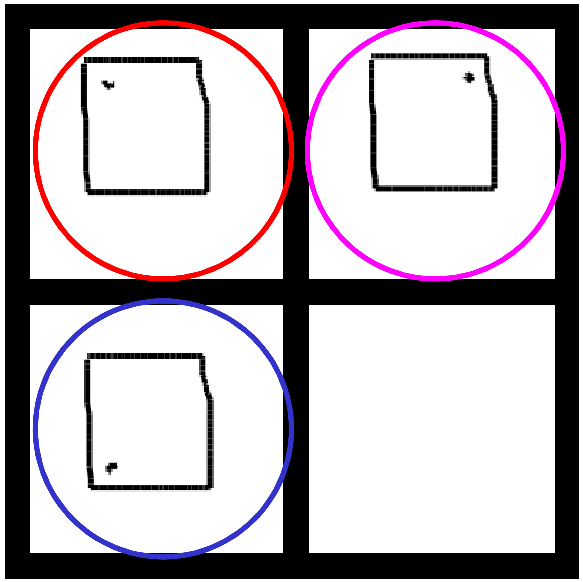
- Used to measure intelligence
- Extensive data on human performance available



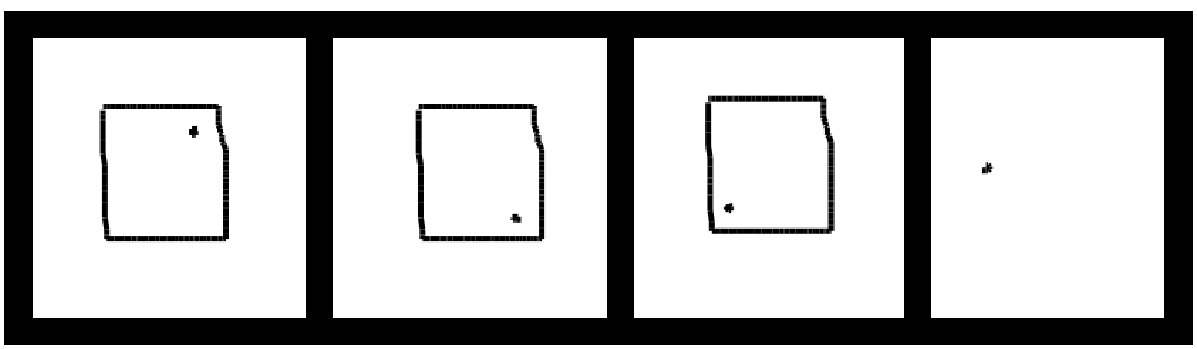
Two-Stage Structure Mapping



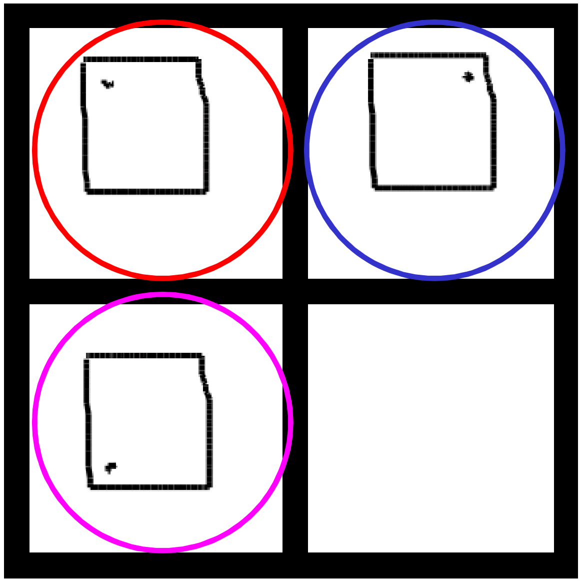
Solving the RPM



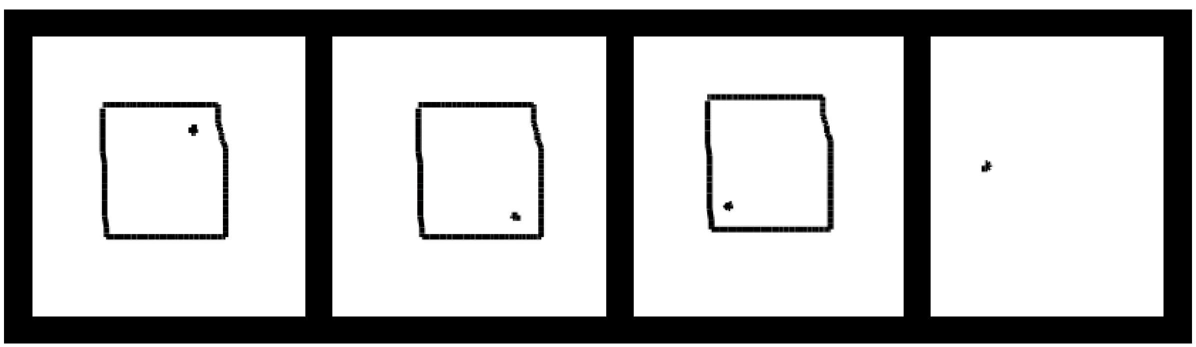
- Row solution
- A is to B, as C is to __?
- difference(A,B)
  - In A, the inner object is on the left side of the outer object
  - In B, the inner object is on the right side of the outer object



Solving the RPM



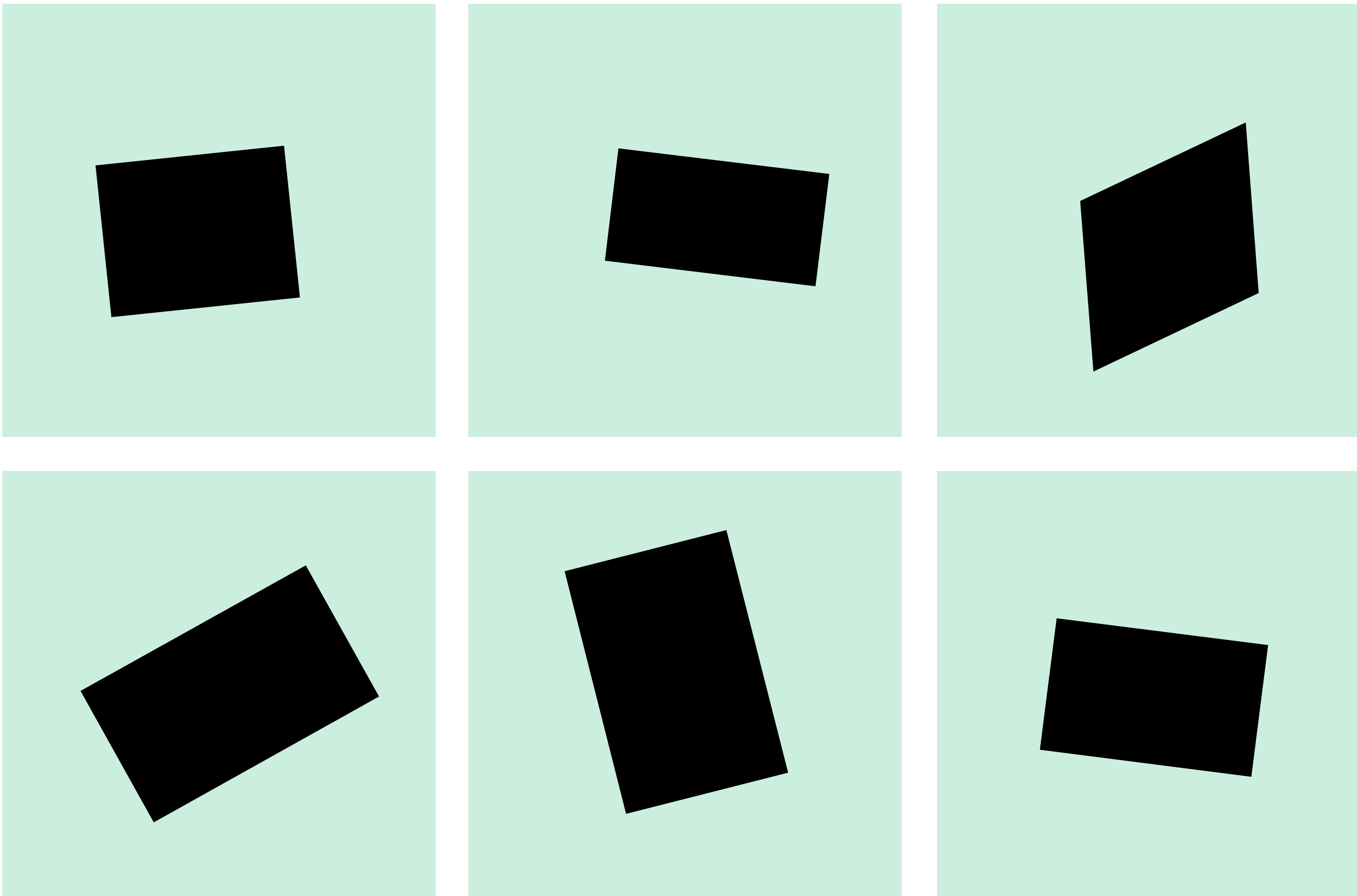
- Column solution
- A is to B, as C is to __?
- difference(A,B)
  - In A, the inner object is on the top half of the outer object
  - In B, the inner object is on the bottom half of the outer object



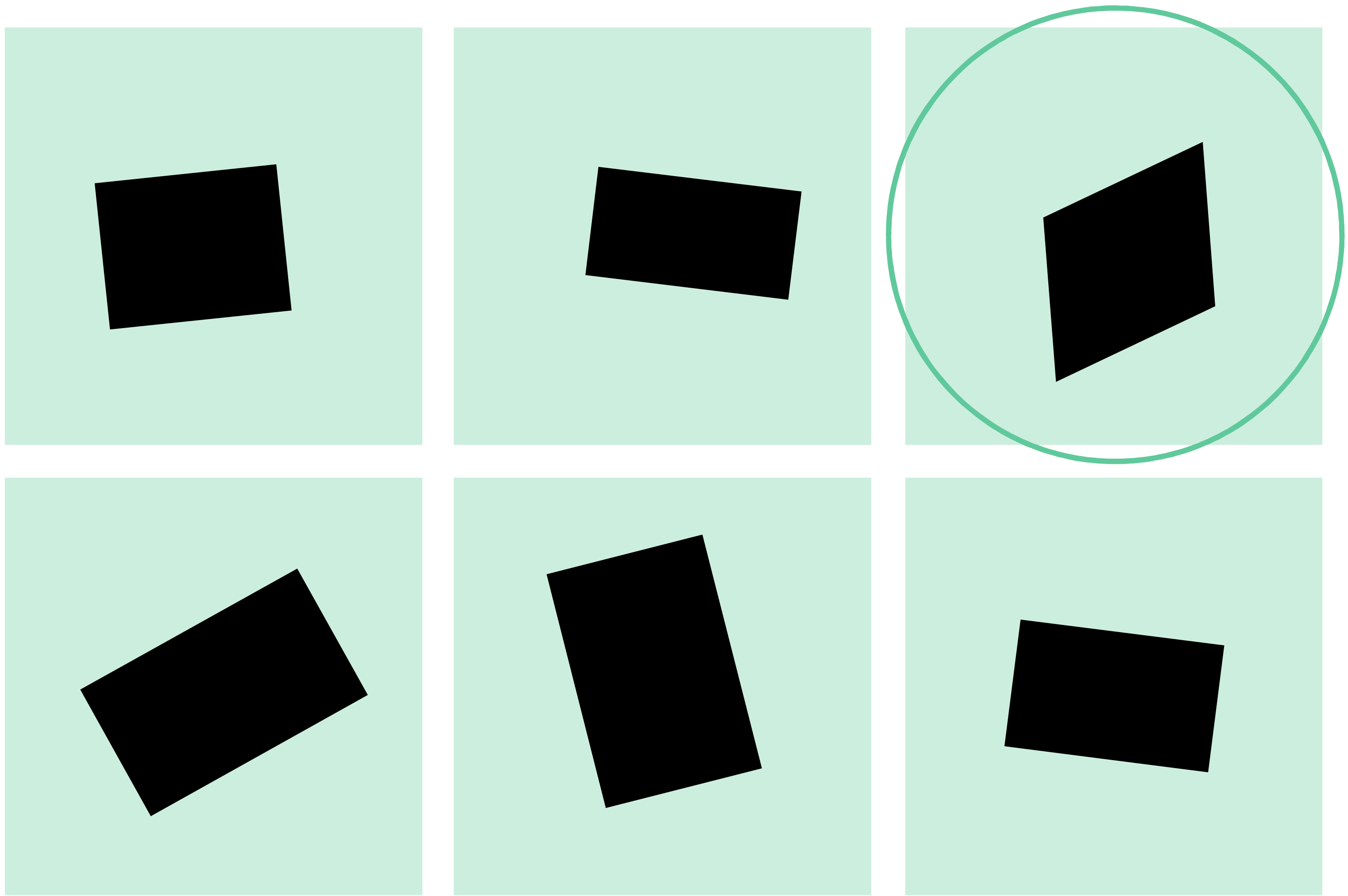
Evaluation

- Initially evaluated on two fairly easy sections of the standard RPM
  - B: 2x2 matrices, 6 possible answers
  - C: 3x3 matrices, 8 possible answers
- Performed at the level of the average American adult on those sections
  - B: 12/12
  - C: 10/12

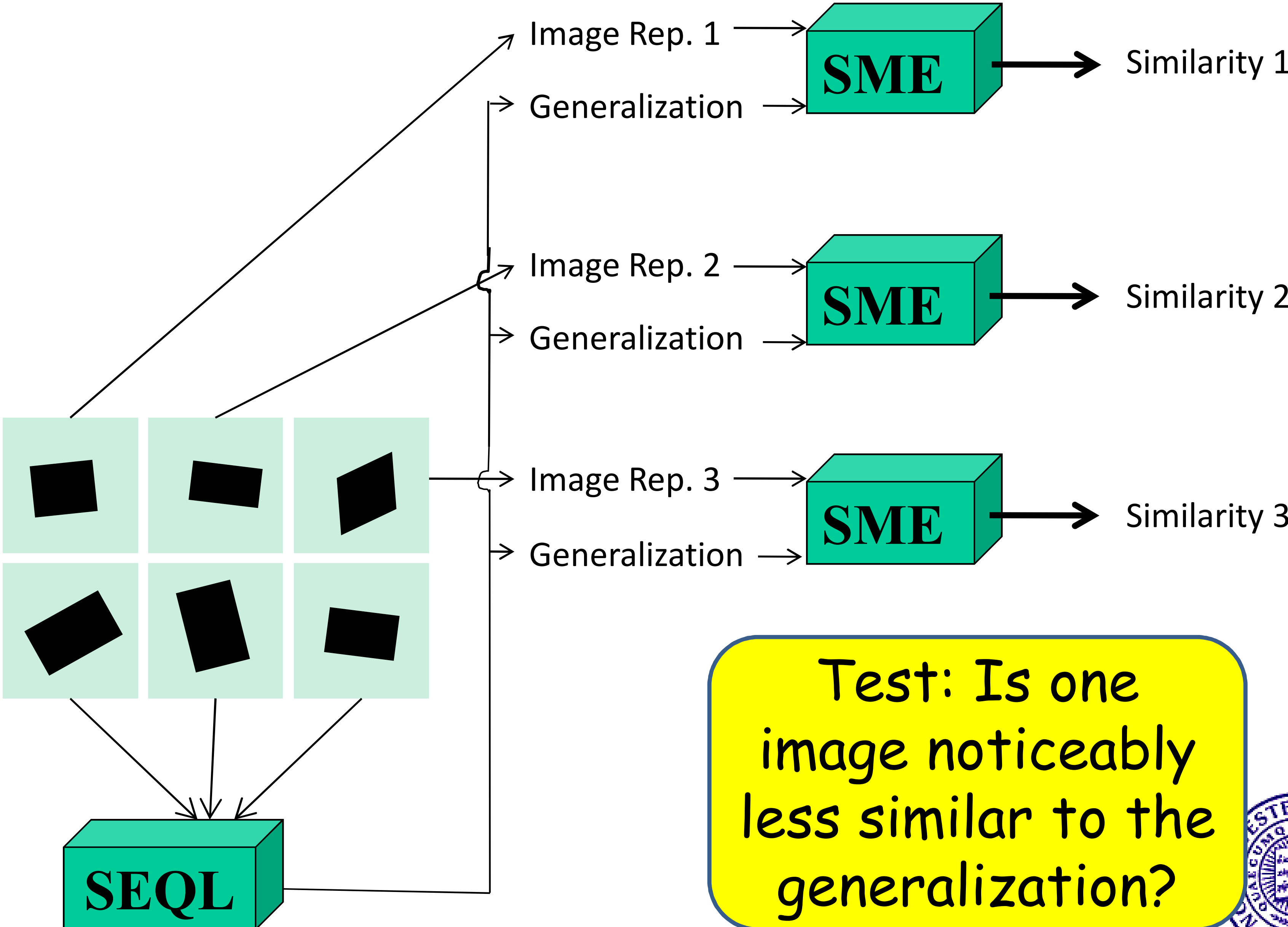
Visual Oddity Task



Visual Oddity Task

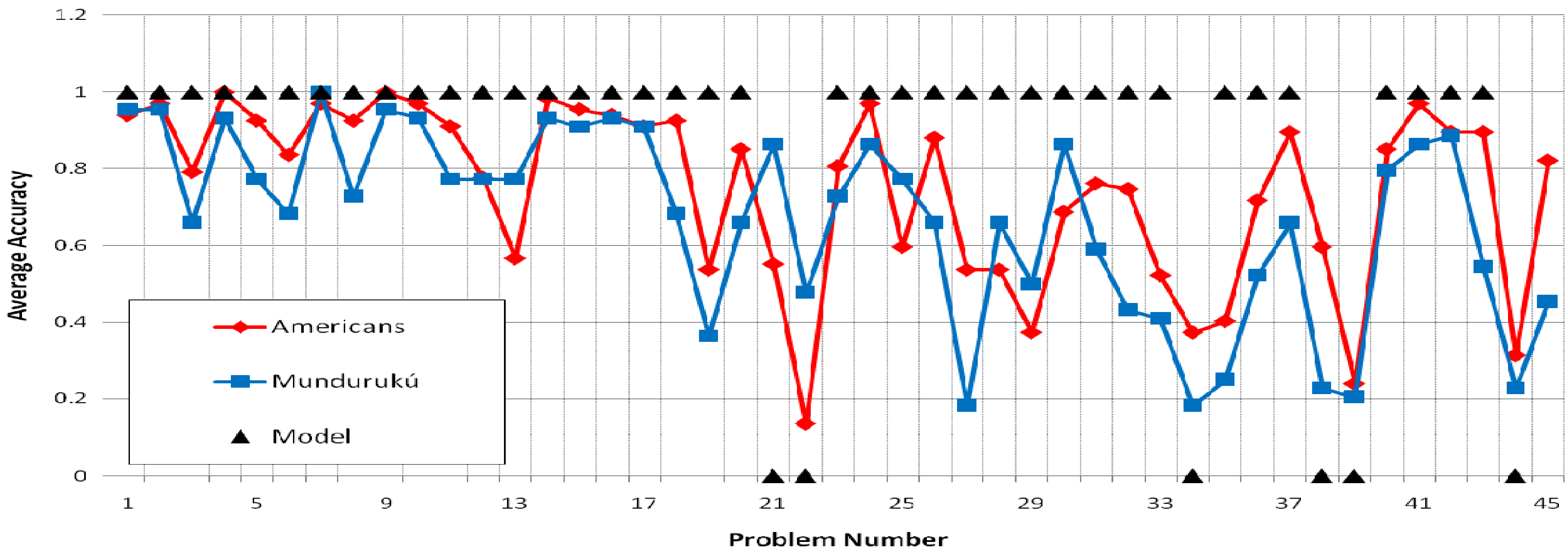


Our model for the Oddity Task

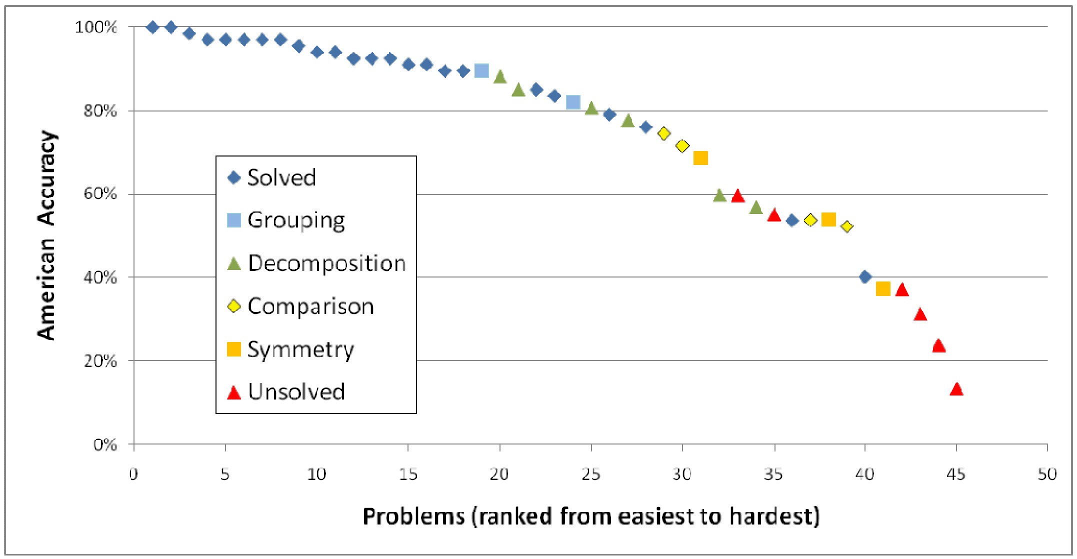


Test: Is one image noticeably less similar to the generalization?

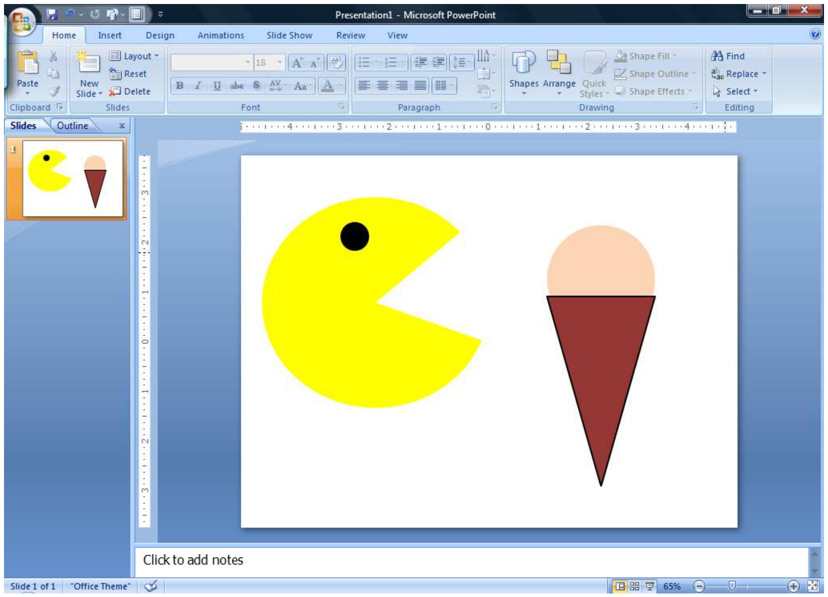
Results



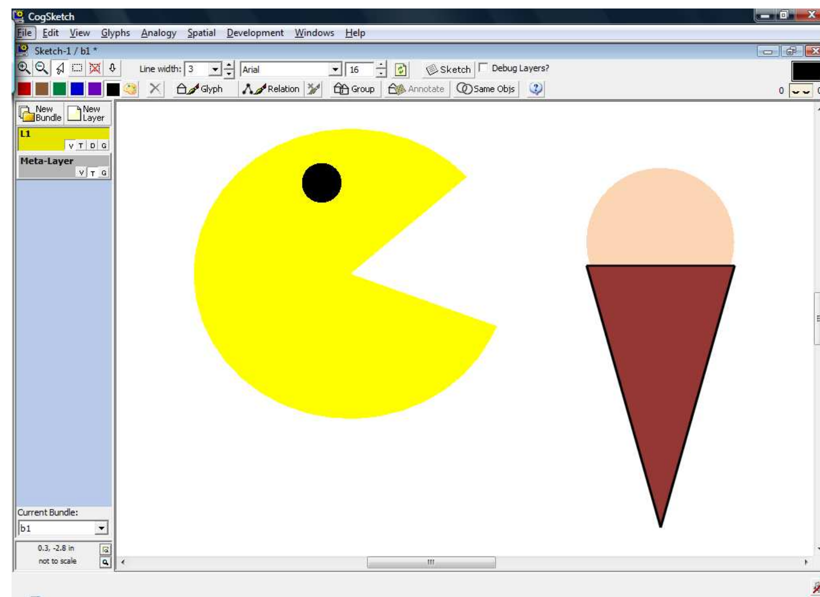
- Overall Performance: 39 / 45
- Correlation w/Americans: 0.656 (aged 8-13)
- Correlation w/Mundurukú: 0.493 (all ages)
- Model can be used to identify operations that contribute to problem difficulty



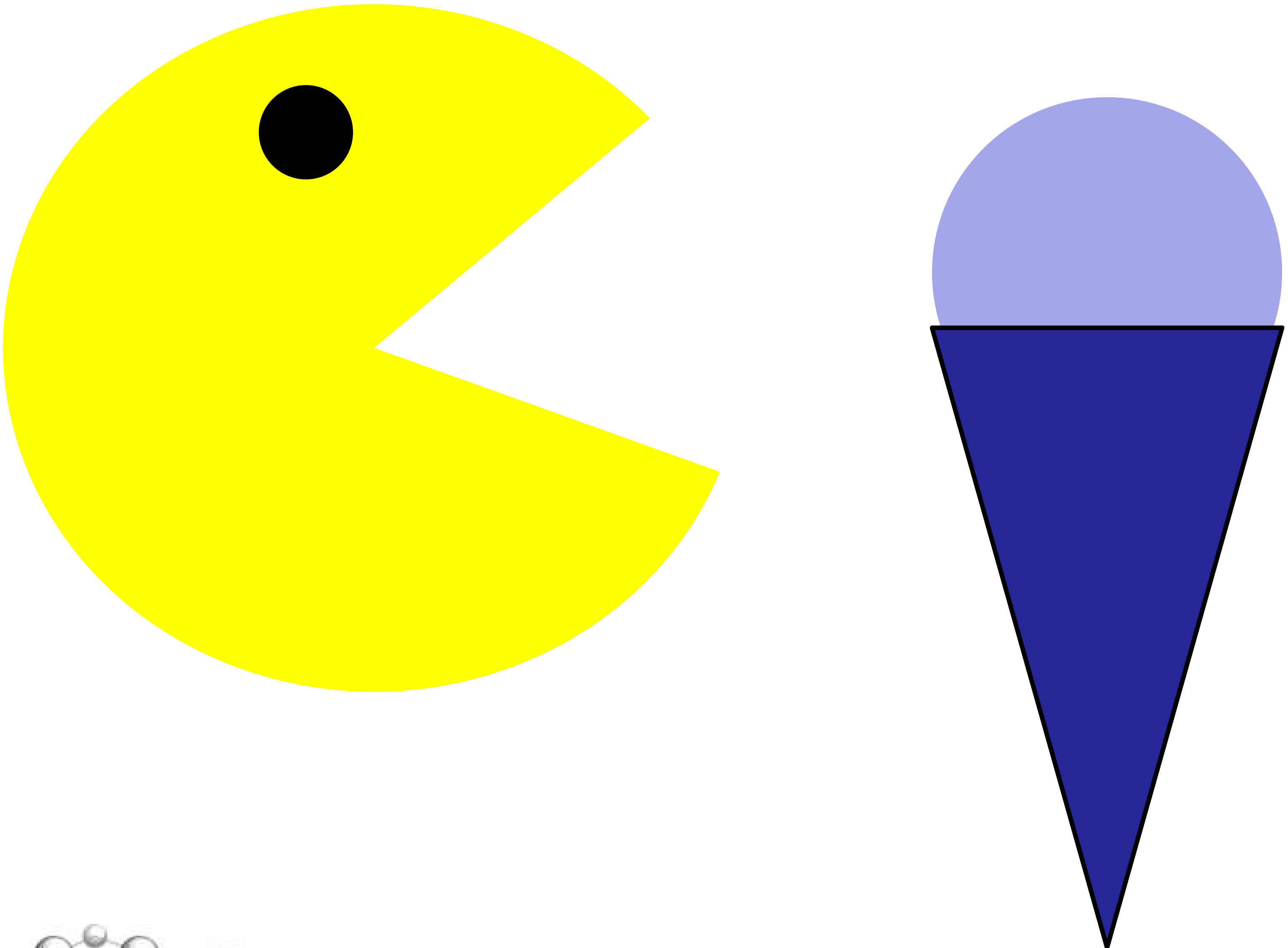
Copy/Paste from PowerPoint



N.B.: It can take a few moments for the ink processor to finish before the glyphs appear



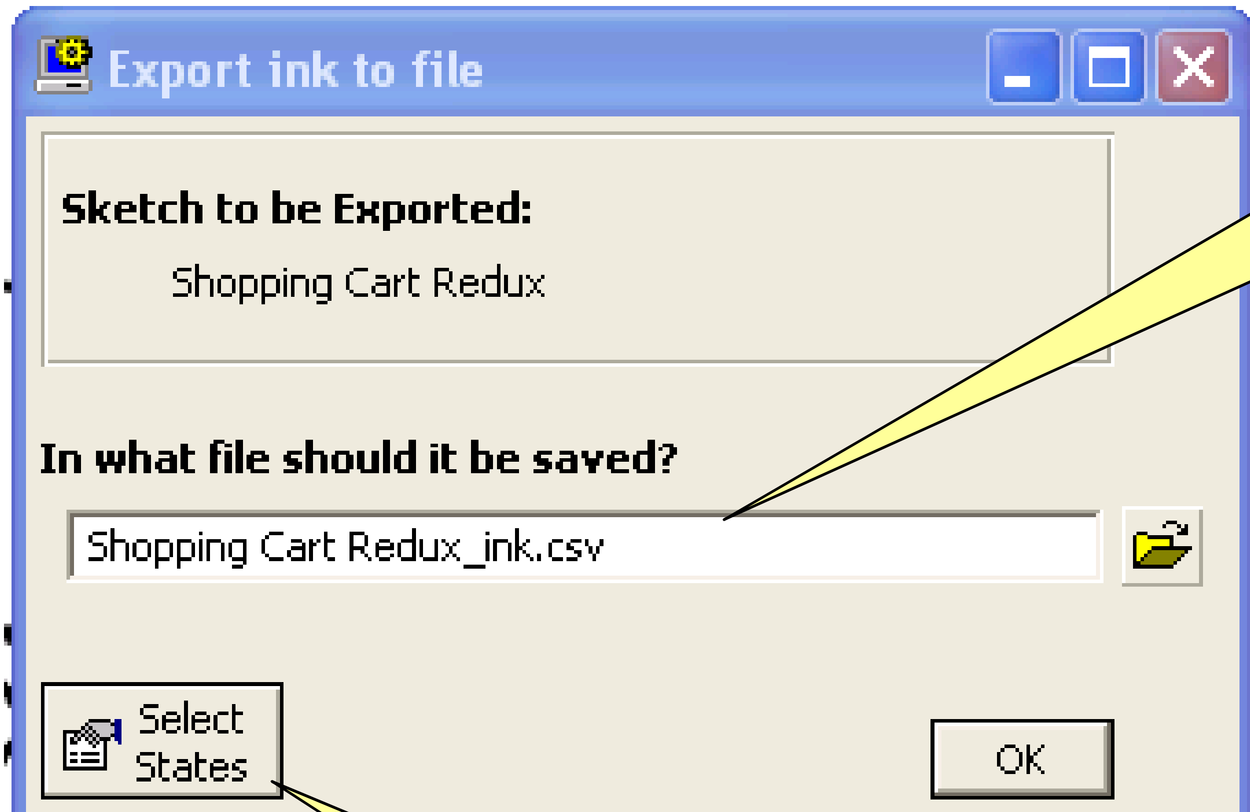
- Import shapes drawn in PowerPoint via copy/paste
- PowerPoint shape -> Glyph



Copy/Paste from PowerPoint

- What is supported
  - Most simple shapes
    - Straight/Curved, Open/Closed, Custom-drawn, etc
  - Line thickness, line color, fill color
  - Group PowerPoint shapes together to make them a single CogSketch glyph
- Not supported
  - More complex shapes
    - 3D shapes, shapes with multiple polygons, arrows
  - More complex attributes (shading, textures, etc)

Exporting Ink



Comma separated value files can be directly read by modern spreadsheets

Can pick which bundles are included

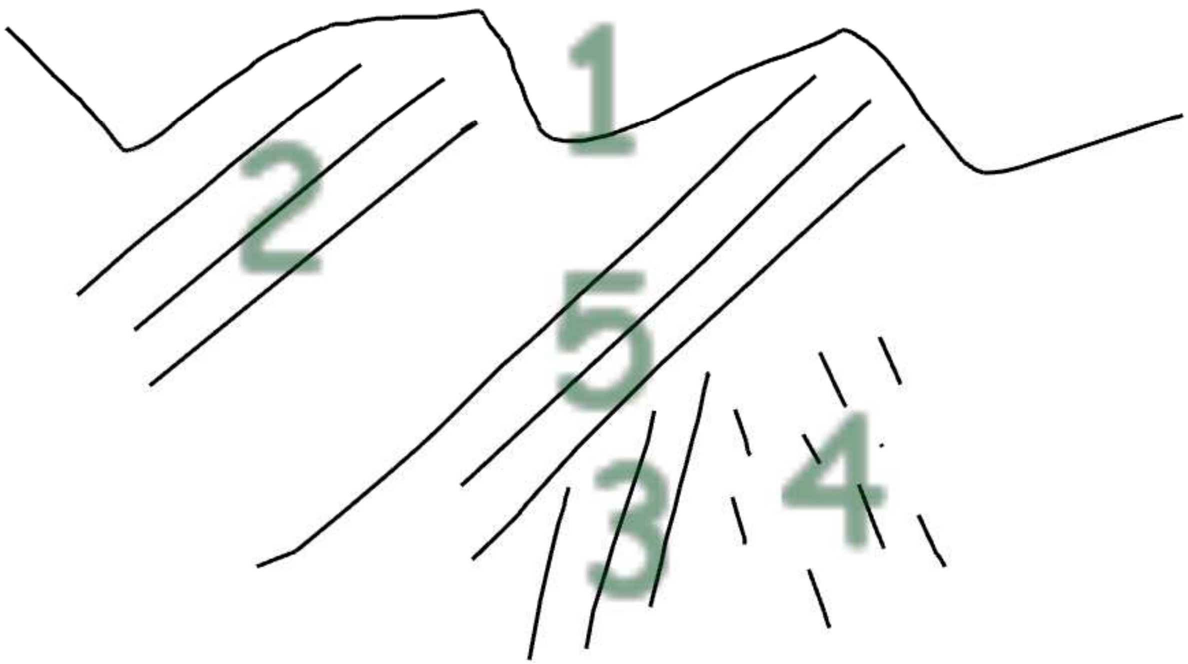
Result of Ink Export

Note: Timestamp information provided for every ink point in the sketch

Sketch Namestring	Sketch Case	Bundle Namestring	Bundle Case	Layer Namestring	Layer Objname	Glyph Namestring	Glyph Objname	Ink ID	Ink Type	Ink Point X	Ink Point Y	Ink Point Timestamp (s)
Shopping Cart Redux	Case-3429195339	"Shopping Cart Anatomy"	BCase-3429195452	"Physical"	ObjectL-225	"Handle"	Object-154	745	sketch-polyline	-2.24479	0.4583	252.35339d0
Shopping Cart Redux	Case-3429195339	"Shopping Cart Anatomy"	BCase-3429195452	"Physical"	ObjectL-225	"Handle"	Object-154	745	sketch-polyline	-2.24479	0.4583	252.64331d0
Shopping Cart Redux	Case-3429195339	"Shopping Cart Anatomy"	BCase-3429195452	"Physical"	ObjectL-225	"Handle"	Object-154	745	sketch-polyline	-2.24479	0.4687	252.6752d0
Shopping Cart Redux	Case-3429195339	"Shopping Cart Anatomy"	BCase-3429195452	"Physical"	ObjectL-225	"Handle"	Object-154	745	sketch-polyline	-2.26563	0.4791	252.70673d0
...	...	...	...	...	...	...	...	...	...	...	...	...

Viewing Timing Data

- Glyph order
  - Need to be in experimenter mode
- Ink replay

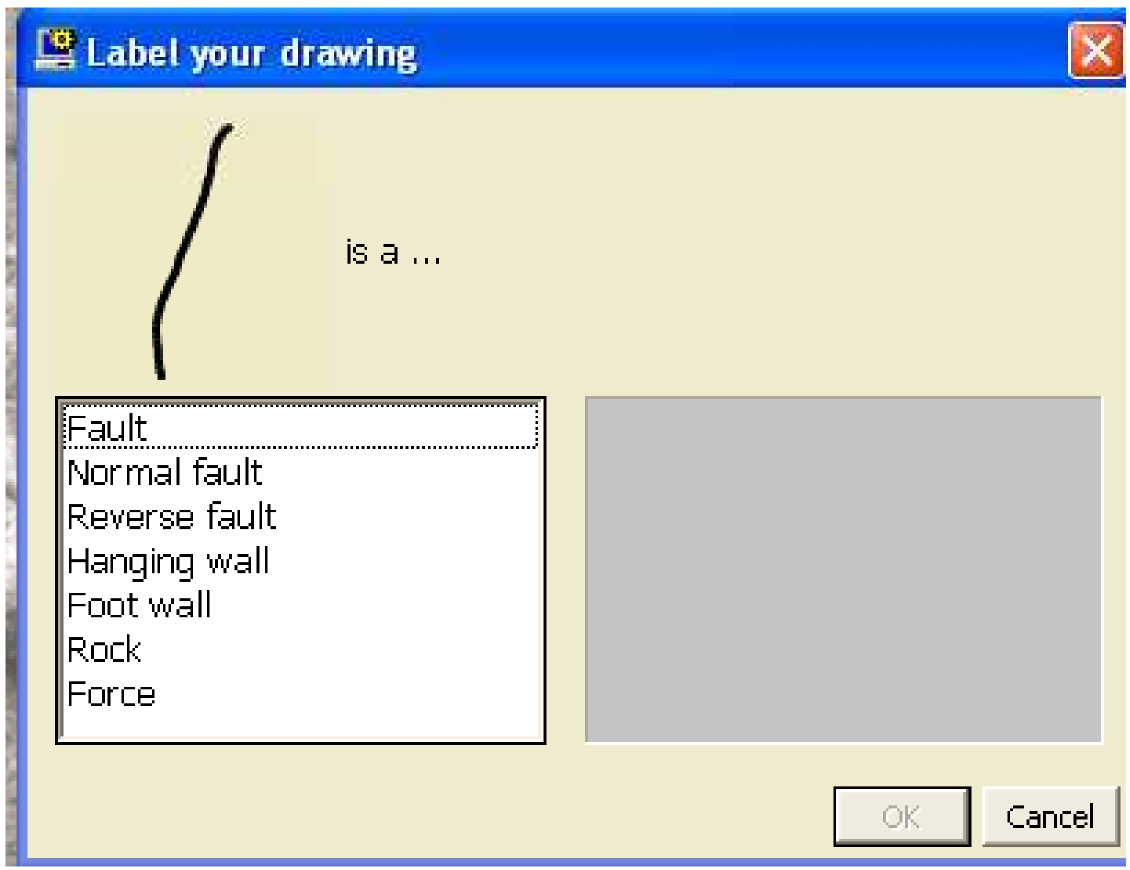


Gathering Sketch Data

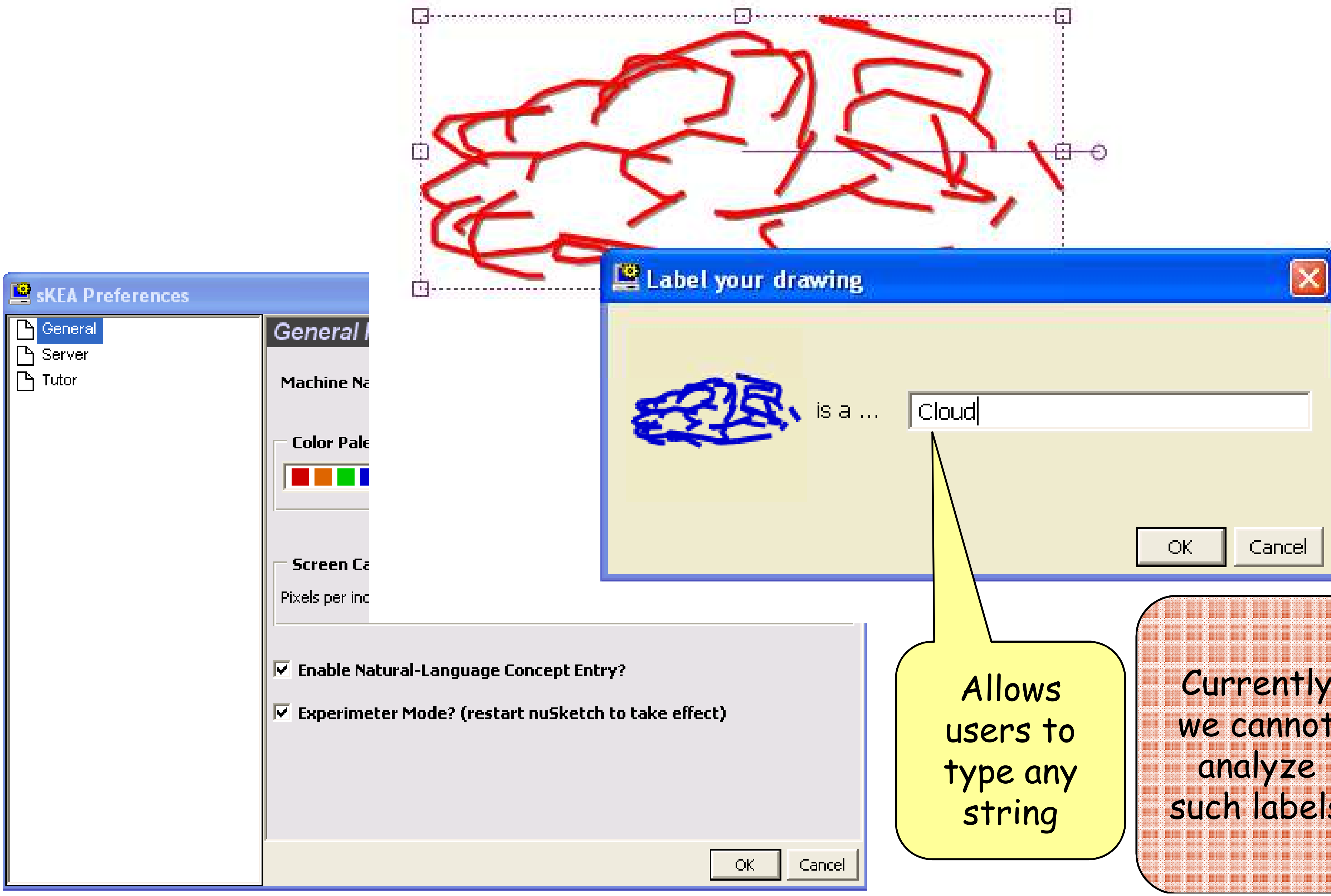
- Potential CogSketch advantages in data collection:
  - Easier to archive and transmit bits than dead trees
  - Captured digital ink is potentially easier to analyze than video or scanned bitmaps
    - Timing data automatically captured
  - Conceptual labeling could reduce subsequent hand-coding of data (“what’s that??”)
  - Visual/spatial processing could become calibrated enough with human judgments to automate some kinds of data analysis
- Potential disadvantages:
  - Teaching participants how to sketch with it
  - Overhead of conceptual labeling can be distracting

Simplifying Concept Labeling

- Worksheets use a simple list
- Can be done in any order



Experimental: Free-form NL input



Allows users to type any string

Currently we cannot analyze such labels

Coming up: Skins and Scripting

- Skins = ability to hide/expose capabilities in the interface
  - Often useful for participants to have fewer distractions
  - You can already choose skins when building worksheets
  - File format and documentation under development to let experimenters generate their own skins.
- Scripting
  - Want ability to run participants through a number of sketching exercises, with minimal or no experimenter intervention
  - Suggestions about what you need would be welcome

Questions?