Using Automatic Item Generation to Develop Practice Non-Verbal Reasoning Items for a High-Stakes Admissions Test

Ms Marita MacMahon Ball, ACER
Professor Mark Gierl, University of Alberta
The Story So Far...

- History
- The Challenge
- Trial and verdict
- Re-trial
- Verdict?
- What’s next?
The Medical Admissions Test

- What is the test? UMAT
- How high are the stakes? Very
- What needs to change? Paper to Computer
- What are the limitations? The need for differentiation
- Is there a solution? It is looking good
Middle of Sequence

The five figures can be rearranged to form a logical sequence.

Select the alternative that would most logically and simply be in the **middle** of the sequence.
In this set of frames, there is a total of eight different shapes occurring with differing frequencies - the arrow and the cube appear just once, the grey hexagon and arrowed cross appear twice, the heart and cross appear three times, while the circle is present in four frames. It is possible to arrange the set so that adjacent frames have three shapes in common. For example, A and B both have the cross, the arrowed cross and the arc. B and E also have in common the arc, the cross and the circle. E and D have the arc, the circle and the heart in common; D and C have the heart, the circle and the hexagon in common.

The simplest explanation of this pattern is that, between frames, one shape leaves the frame and is replaced by a different one which then remains. Frames A and C must start (or end) the sequence because both have only one other frame with three shapes in common, while B, E and D have two. Starting with A, between A and B the arrow is replaced by a circle; then between B and E the arrowed cross is replaced by a heart; then between E and D the straight cross is replaced by a hexagon. Finally, between D and C the arc is replaced by a cube. This gives the sequence ABEDC, making E the middle frame.
Select the alternative that most logically and simply continues the series.

Sample Item 2

Next in Sequence
In this sequence, note that each frame contains a large triangle and four small shapes - three at the corners, and one in the centre. Between frames, two of the shapes swap places. Between the 1st and 2nd frame, the swap is between the shape in the centre (a circle) and the bottom-left shape, a cross. Between the 2nd and 3rd frames it is between the centre shape (a cross) and the top shape (a square). Between the 3rd and 4th frames, it is between the square in the centre and the triangle in the bottom right corner.

The simplest explanation of this pattern is that, starting from bottom left and moving clockwise around the triangle, the central shape swaps places with one of the corner shapes. Once the swap has taken place, the corner shape remains in that position.

Therefore, between the 4th and 5th frames, the centre shape (the triangle), should swap with left-corner shape, the circle. So, in the answer frame the triangle should be at bottom left and the circle should be in the centre, with the positions of the cross and the square unchanged. The alternative that satisfies these conditions is B.

In C, the triangle and circle are in correct positions but the square and cross are reversed.
In A, only the circle and the cross are in correct positions.
In D, only the square is in correct position.
In E, none of the shapes is in the correct position.
The Rationale in Action
3. Select the alternative that most logically and simply continues the series.

4. Select the alternative that most logically and simply continues the series.
A New Approach

Writing the Rules

Example (5 Elements)

1. 1 shape + 4 moving objects
2. 3 objects outside and 1 object inside of shape
3. object 1 in the middle swaps with object 2 outside of shape
4. swapping continues clockwise so object 2, now in the middle, swaps with next object (object 3)
5. other objects do not move
IGOR’s New Challenge

• What are the benefits of partnerships? Matching stubborn problems to diverse problem solvers
• What are the pay-offs? Novel solutions
• What are the risks? Timelines and outcomes
• Problem with solution 1? Cloning
• Outcome of solution 2? New AIG method
What is AIG?

- Automatic item generation (AIG) is the process of using item models to generate test items with the aid of computer technology.

- Automatic item generation combines “art” (HUMANS: Judgment, Expertise, Experience) with “science” (MODERN COMPUTING) to create a new approach to item development.
George has a collection of 397 postage stamps. He got 61 of the stamps from Malaysia. Approximately what percentage of the stamps come from Malaysia?

- A 7%
- B 15%
- C 20%
- D 23%
<NAME> has a collection of <TOTAL> postage stamps. He has <SUBSET> of the stamps from Malaysia. Approximately what percentage of the stamps come from Malaysia?
Step 2: Identify Content

<table>
<thead>
<tr>
<th><strong>NAME</strong></th>
<th>George, Mike</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TOTAL</strong></td>
<td>11-201 (no multiples of 10)</td>
</tr>
<tr>
<td><strong>SUBSET</strong></td>
<td>5-75% of Total</td>
</tr>
</tbody>
</table>
George has a collection of 397 stamps. He got 61 stamps from Malaysia. Approximately what percentage of the stamps come from Malaysia?
A 7%
B 15%
C 20%
D 23%

Mike has a collection of 112 stamps. He got 45 stamps from Malaysia. Approximately what percentage of the stamps come from Malaysia?
A 40%
B 36%
C 52%
D 84%

Mike has a collection of 139 stamps. He got 70 stamps from Malaysia. Approximately what percentage of the stamps come from Malaysia?
A 50%
B 57%
C 87%
D 91%
A 1-layer item model permits the manipulations of a linear set of generative operations using a small number of elements.

<NAME> has a collection of <TOTAL> postage stamps. He has <SUBSET> of the stamps from Malaysia. Approximately what percentage of the stamps come from Malaysia?

<NAME>: George, Mike (1-Layer Element)
1-Layer Modeling (Cloning)

- Our nonverbal reasoning clones used with solution 1
• An n-layer item model permits the manipulations of a non-linear set of generative operations using a large number of elements

• The concept of n-layer item generation is adapted from the literature on syntactic structures of language where sentences are organized in a hierarchically—this organization, where elements are embedded within elements, is used as the principle to generate items
• A sample of our nonverbal reasoning n-layers used with solution 2
Select the alternative (A, B, C, D, E) that most logically and simply fits the missed value.
Reorder the five figures to form the simplest and most logical sequence possible. Then, select the alternative that is in the middle of the sequence.
Select the alternative (A, B, C, D, E) that most logically and simply fits the missed value.
Our current n-layer nonverbal reasoning model can generate 2786 unique items (1 model = 2786 items).

We can add more layers or more content within each layer to increase the generative capacity of this model—we can also create new models.

Big question: What do ACER content specialists think?
THANK YOU

Ms. Marita MacMahon Ball
(Marita.MacMahonBall@acer.edu.au)

Dr. Mark Gierl (mark.gierl@ualberta.ca)