Origin of Math Garden

- The Dutch version of Math Garden (Rekentuin.nl) was developed at the department of Psychological Methods of prof. dr. Han van der Maas of the University of Amsterdam (UvA).
- Development:
  - School year ’07–’08 and ’08–’09 Math Garden was part of a research project.
  - March 2009: Oefenweb.nl was founded to further develop Math Garden.
  - Today: >1700 schools in The Netherlands and Belgium.

Background ideas

The cognitive system in development is a complex system

Arithmetic and language learning are instances of cognitive expertise

- Lots of deliberate practice needed
  - Adjusted to individual (one-to-one guidance)
  - Direct feedback
Individual differences are huge

ICT allows new developments
- Laptops
- Tablets
- Fast internet
- Wifi

Web-based adaptive practice and monitoring systems
- Idea: Digital notebooks for daily work in classroom
- Choice practice items automatically adapted to child, differentiation
- Combining (playful) practicing and pupil monitoring
  - Less tests in classrooms
  - High frequency monitoring
- No checking, automatic progress reports
- Web-based (cloud)
- New type of adaptive testing (psychometrics)

CAT (Computer Adaptive Testing)
- Low (-)
- High (+)

Problems standard CAT
1. Requires pretesting of items > expensive
2. Participants and examiners do not know how to weigh speed and accuracy
3. Requires administration of items such that the chance of answering an item correctly is .5
  - Not motivating for children

How to find an equal opponent?
Elo ratings in chess

Low -- High

Elo ratings in chess

Low -- High

Elo ratings in Math Garden

Children (Ability)

Low -- High

Items (Difficulty)

Self-organizing system

Children (Ability)

Low -- High

Items (Difficulty)

Which item is more difficult?

8 + 9
2 + 2
1 + 70
9 + 66
19 + 58
3 + 6
80 + 10
44 + 3

High Speed High Stakes rule

- HSHS punishes quick guesses
- Easy to visualize
Advantages new CAT

- Self-organizing system
- No pretesting required (less time and money)
- All children (all ages) practice on their own level
- Includes response times with HSHS scoring rule
- Speed accuracy trade-off problem solved
- Guessing problem solved
- Every child answers 75% of the problems correctly

Demonstration Math Garden

Prizes

The games

Back-end

Aims:
- No marking and correcting of schoolwork needed.
- High frequency progress-monitoring
- Analysis of pupil’s typical errors and specific problems.
- Insights for individual tuition.

School class overview

Comparison with all other children of the same age that use Math Garden.
**Development**

![Graph showing playing frequency and scaling](image)

**Detailed individual reports**

<table>
<thead>
<tr>
<th>QUEST</th>
<th>QUESTION</th>
<th>CORRECT ANSWER</th>
<th>GIVEN ANSWER</th>
<th>RESPONSE TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>3 sec</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>100</td>
<td>100</td>
<td>3 sec</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>100</td>
<td>100</td>
<td>7 sec</td>
</tr>
<tr>
<td>20</td>
<td>4</td>
<td>100</td>
<td>100</td>
<td>3 sec</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>100</td>
<td>100</td>
<td>7 sec</td>
</tr>
</tbody>
</table>

**Usage**

- ± 300,000 active users
- ± 800 million item responses in 5 years
- ± 1,500,000 items per day
- ± 1800 schools
- ± 45,000 items

**User statistics**

![Graph showing playing frequency](image)

**Research and Development**

- Anonymous data
- Innovative psychometrics
- New insights
- Math development & learning processes

**Advantages**

- Ecological valid
  - Field study
  - Naturalistic learning environment
- Many domains (34 games in math, language, cognitive functions)
- Age range (4 to adult)
- High frequent measurements
Reliability

- Correlation m+n, n+m

![Graph showing correlation](image)

- Rating mirror item (n+m)

- Rating item (m+n)

Validity

- r = 0.88
- r = 0.98

![Validity graphs](image)

Difficulty level

- Self-adapted testing

- Easy: E(P) >= 9 (half coins)
- Normal: E(P) >= 7.5
- Difficult: E(P) >= 6 (double coins)

Preference for easy, normal, and difficult by grade and sex

![Preference graph](image)

Learning domains

- Math Garden (20 games)
- Language sea (15 games)
- Words&Birds (7 games)
- Typing Garden (8 games)
- Stas Plant (3 games)
- ?

Psychometrics

For large scale computer adaptive practice
Data representation I - educational measurement data

<table>
<thead>
<tr>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Data representation II - extensions in both directions

<table>
<thead>
<tr>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Data representation III - sparse data

<table>
<thead>
<tr>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>.</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>-</td>
<td>0</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>

Data representation IV - long data

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

38,564 1,032 1 223,879,987

Data representation IV - long data

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

38,564 1,032 1 223,879,987

Data representation IV - long data

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

38,564 1,032 1 223,879,987

Modeling

\[
(\theta_t, \tilde{\theta}_t) \quad \downarrow
\]  

\[
\hat{\theta}_t, \hat{\theta}_t \quad \rightarrow \quad \theta_{t+1}, \tilde{\theta}_{t+1}
\]

26-09-17
Elo ratings I

- Dynamic paired comparisons
- Roots in chess ability estimation
- Widely applied
- Connected to foundations of psychometrics
- Simple, fast and easy to apply
- Large scale adaptive testing useful in educational games

Elo ratings II

\[ R_n = R_o + K(W - W_e) \]

- \( R_n \) is the new rating after the event
- \( R_o \) is the pre-event rating
- \( W \) is the rating point value of a single game score
- \( W_e \) is the expected game score

Elo ratings III

Response times

- Speed-accuracy response model
- Prevents quick guessing
- Re-captures information

Summary

1. Data revolution
   1. Algorithms that we have don’t work anymore
   2. Conceptual and computational problem
2. Math garden quite small compared to national monitoring systems
3. Creative algorithms and data filters on educational ‘big data’

Applications

On user level
Project opportunities

Implement an Elo ratings system on math practice data to track student and item development.

Elo ratings are popular in many contexts, but also have found their application in learning contexts, as it allows for simple tracking of the development of both the ability of students, and the difficulty of items, over time.

This project will ask you to calculate and use Elo ratings for exploratory data analysis. How do development traces look like, can you visually identify different groups of students? Are students within schools more alike than students between schools? What patterns of study time can be charted? And more.
Detailed individual reports

<table>
<thead>
<tr>
<th>S/N</th>
<th>Test</th>
<th>Grade</th>
<th>Student</th>
<th>Response time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12 - 5</td>
<td>12</td>
<td>80</td>
<td>137</td>
</tr>
<tr>
<td>2</td>
<td>13 - 5</td>
<td>13</td>
<td>135</td>
<td>137</td>
</tr>
<tr>
<td>3</td>
<td>14 - 5</td>
<td>14</td>
<td>135</td>
<td>137</td>
</tr>
</tbody>
</table>

New window on development

Preference for easy, normal and difficult by grade and sex

More info?