Making Sense of Multimodal Learning Analytics (MMLA)

Marcelo Worsley
Assistant Professor, Learning Sciences & Computer Science

Northwestern
Agenda

- What is MMLA?
- Why use MMLA?
- How to use MMLA?
- Examples of MMLA research
- On-going challenges and future directions
- Resources
What is MMLA?

Learning Analytics—a set of multi-modal sensory inputs, that can be used to predict, understand and quantify student learning.  
(Worsley and Blikstein, 2011)
What is MMLA?

Multimodal learning analytics (MMLA) (Blikstein & Worsley, in press.; Blikstein, 2013; Worsley, 2012) sits at the intersection of three ideas: multimodal teaching and learning, multimodal data, and computer-supported analysis. At its essence, MMLA utilizes and triangulates among non-traditional as well as traditional forms of data in order to characterize or model student learning in complex learning environments. However, as we describe later, the ways that researchers utilize multimodal data vary widely.

(Worsley, Abrahamson, Blikstein, Grover, Schneider and Tissenbaum, 2016)
What is MMLA?

Most Popular MMLA Data Types

- Audio: 0.4
- Video: 0.3
- Biophysical data: 0.15
- Eye trackers: 0.2
- Digital interactions: 0.1
What is MMLA?
What is MMLA?

- Video
  - Affect
    - Frustration
    - Confusion
  - Body Language
    - Surprise
    - Gestures
  - Eye Gaze
    - Head Position
    - Arm Position
<table>
<thead>
<tr>
<th>VIDEO</th>
<th>AUDIO</th>
<th>TEXT</th>
<th>EYE TRACKING DATA</th>
<th>DEPTH CAMERA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pose</td>
<td>Speaker Diarization</td>
<td>Emotion/Affect</td>
<td>Fixations</td>
<td>Gestures</td>
</tr>
<tr>
<td>Gaze</td>
<td>Prosody</td>
<td>Complexity</td>
<td>Fixation Pattern</td>
<td>Body Position</td>
</tr>
<tr>
<td>Gestures</td>
<td>Spectrum</td>
<td>Cohesion</td>
<td>Pupil Dilation</td>
<td>Joint Tracking</td>
</tr>
<tr>
<td>Object Manipulation</td>
<td>Emotion/Affect</td>
<td>Semantics</td>
<td>Eye location</td>
<td></td>
</tr>
<tr>
<td>Facial Expressions</td>
<td>Intensity</td>
<td>Syntactics</td>
<td>Attention</td>
<td></td>
</tr>
<tr>
<td>Proxemics</td>
<td>Speaking Rate</td>
<td>Content</td>
<td>Head Pose</td>
<td></td>
</tr>
<tr>
<td>Scene Understanding</td>
<td>Collaboration/Turn-Taking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement</td>
<td>Voice Quality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heartate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object Tracking</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMARTPHONES</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sitting/Standing</td>
<td>Gestures</td>
<td>Arousal</td>
<td>Activity</td>
<td>fMRI</td>
</tr>
<tr>
<td>Location</td>
<td>Tool Usage</td>
<td>Cognitive Load</td>
<td></td>
<td>ELECTROCARDIOLOGY</td>
</tr>
<tr>
<td>Noise Level</td>
<td>Hand Movement</td>
<td>Body Temperature</td>
<td></td>
<td>ELECTROENCEPHALOGY</td>
</tr>
<tr>
<td>Smartphones</td>
<td>Finger Movement</td>
<td>Gestures</td>
<td></td>
<td>ELECTROMYOGRAPHY</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Heartate</td>
<td></td>
<td>ELECTROGASTROGRAPHY</td>
</tr>
<tr>
<td>Leap Motion</td>
<td></td>
<td>Stress</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDA Watch</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Logs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other modalities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fMRI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrocardiography</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrophysiology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electromyography</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrogastrography</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Why use MMLA?

- Teaching and learning are multimodal
- Study and support complex learning environments
  - See the hard to see
- Inform design of multimodal technologies
- Expand notions of learning to non-traditional modalities
  - Improve accessibility and inclusivity
- Triangulate across modalities
Why use MMLA?

- Visualizing/Representing information for human inference
- Prediction of indicators
- Data-driven interventions
- Evaluating conjecture-based learning designs
How to use MMLA?
How to use MMLA?

Collect ➔ Pre-process ➔ Extract ➔ Analyze ➔ Infer ➔ Apply
How to use MMLA?

DATA CAPTURE SOFTWARE
Open Social Signal Interpretation
Lab Streaming Layer
Open Pipe Kit
Open Broadcaster Software
Multisense (AV Recorder)
iMotions Attention Tool
How to use MMLA?

DATA CAPTURE SOFTWARE
Open Social Signal Interpretation
Lab Streaming Layer
Open Pipe Kit
Open Broadcaster Software
Multisense (AV Recorder)
iMotions Attention Tool
How to use MMLA?

DATA CAPTURE SOFTWARE
Open Social Signal Interpretation
Lab Streaming Layer
Open Pipe Kit
Open Broadcaster Software
Multisense (AV Recorder)
iMotions Attention Tool
How to use MMLA?

DATA CAPTURE SOFTWARE
Open Social Signal Interpretation
Lab Streaming Layer
Open Pipe Kit
Open Broadcaster Software
Multisense (AV Recorder)
iMotions Attention Tool
How to use MMLA?

DATA CAPTURE SOFTWARE
Open Social Signal Interpretation
Lab Streaming Layer
Open Pipe Kit
Open Broadcaster Software
Multisense (AV Recorder)
iMotions Attention Tool
How to use MMLA?
How to use MMLA?

Pre-processing is important for data synchronization, accounting for individual differences between participants, and getting data in the appropriate format for data extraction. It tends to vary by data type.
How to use MMLA?
<table>
<thead>
<tr>
<th>VIDEO</th>
<th>AUDIO</th>
<th>TEXT</th>
<th>EYE TRACKING DATA</th>
<th>DEPTH CAMERA</th>
</tr>
</thead>
<tbody>
<tr>
<td>OpenFace (gaze)</td>
<td>HTK</td>
<td>Natural Language Toolkit</td>
<td>PyGaze</td>
<td>OpenNUI</td>
</tr>
<tr>
<td>OpenFace (face recognition)</td>
<td>Praat</td>
<td>Lightside</td>
<td>Ogama</td>
<td>Kinect for Windows</td>
</tr>
<tr>
<td>Emotient FACET</td>
<td>OpenSmile</td>
<td>LIWC</td>
<td>EyeTrackingR</td>
<td>QSRLib</td>
</tr>
<tr>
<td>Noldus Face Reader</td>
<td>OpenEar</td>
<td>Cohmetrix</td>
<td>Tobii SDK</td>
<td>Libfreenect</td>
</tr>
<tr>
<td>OpenCV</td>
<td>Covarep</td>
<td>Tone Analyzer</td>
<td>SMI SDK</td>
<td></td>
</tr>
<tr>
<td>Affectiva</td>
<td>ELAN</td>
<td>Stanford Parser</td>
<td>iMotions Attention Tool</td>
<td></td>
</tr>
<tr>
<td>Microsoft Face API</td>
<td>ICES Diarizer</td>
<td>Wordnet</td>
<td>Pupil Dilation</td>
<td></td>
</tr>
<tr>
<td>Microsoft Emotion API</td>
<td>Matlab</td>
<td>Sentwordnet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intraface</td>
<td>LIUM Diarizer</td>
<td>Mallet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multisense</td>
<td>CMU Sphinx</td>
<td>Word2Vec</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eularian Magnification Code</td>
<td>Google ASR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>iMotions</td>
<td>AT&amp;T Watson ASR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDA WATCH</td>
<td>Bing Speech API</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ledalab</td>
<td>Audacity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DATA EXTRACTION SOFTWARE</td>
<td>Emovoice</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Open Social Signal Interpretation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
How to use MMLA?
How to use MMLA?

Human Coders

rapidminer

Weka GUI Chooser

Applications
Explorer
Experimenter
KnowledgeFlow
Simple CLI
Fusion

Let’s say that you want to study engagement and have audio, video, bio-physiological and gesture data available for analysis. How do you use these to compute a measure of engagement?
How to use MMLA?

1. Collect
2. Pre-process
3. Extract
4. Analyze
5. Infer
6. Apply
How to use MMLA?
Examples
Gather Multimodal Data → Identify States → Correlate state usage with dependent variable → Develop Inferences
Gather Multimodal Data

- **Xbox Kinect**
  - Audio (Talk)
  - Video (Pose/Gaze)
  - Gestures (Hand Movement)

Identify States

Correlate state usage with dependent variable

Develop Inferences

**RapidMiner**
- XMeans

**Researcher(s)**
Xbox Kinect
- Audio (Talk)
- Video (Pose/Gaze)
- Gestures (Hand Movement)

RapidMiner
- XMeans

Custom

Researcher(s)
Challenges & Future Directions
Simplifying Data Capture & Analysis
Better Visualization/Inference Tools
Best Practices Around Data Fusion and Data Analysis Pipelines
Applications

(especiallly as it relates to inclusive technology and providing feedback to learners and teachers)
Resources

Multimodal Learning Analytics Special Interest Group - [http://sigmla.org](http://sigmla.org)

CrossMMLA Workshops EC-TEL 2017, LAK 2018 (tentative)

MMLA Workshops @ ICMI, LAK, ICLS (2012 – Present)

SOLAR LASI – Multimodal Learning Analytics Tutorial Workshops

[Journal of Learning Analytics Special Section on Multimodal Learning Analytics](http://example.com)

[CIRCL Cyberlearning Report](http://example.com)


Marcelo Worsley
Assistant Professor
Electrical Engineering and Computer Science & Learning Sciences
marcelo.worsley@northwestern.edu
Marceloworsley.com
tiilt.northwestern.edu