



# Exploring the effectiveness of video viewing in an introductory x-MOOC of algebra

J. Triay<sup>1</sup>, J. Minguillón<sup>2</sup>, T. Sancho-Vinuesa<sup>2</sup>, V. Daza<sup>1</sup>

<sup>1</sup>Universitat Pompeu Fabra, Barcelona, Spain

<sup>2</sup>Universitat Oberta de Catalunya, Barcelona, Spain

# MOOCs

- **First MOOCs in 2008, very popular in 2012**
- **From OER to “truly” open education...**
- **...but very high dropout rates**
- **Huge MOOC platforms**
  - **edX**
  - **MiriadaX**
- **From individual experiences to institutional**
  - **Future Learn, Iversia, ...**
  - **UCATx (edX based)**

# “Decoding Algebra” course

- **Introductory course for university students**
- **5 weeks, one module / week, 3-5 hours / week**
- **98 videos ranging from 5' to 15'**
  - **Theory, exercises, challenges**
- **Quizzes with 8 or 10 questions**
  - **Three attempts**
  - **Best attempt**
  - **Feedback points back to videos**
- **Self-evaluation**

# Motivation

- **Exploratory analysis**
  - **Video consumption**
  - **Relationship with quiz attempts**
  - **Relationship with dropout**
- **Analyze the lack of instructional scaffolding**
- **Build a set of tools for analyzing UCATx courses**
- **Plenty of data available from UCATx**
  - **PhD, PostDoc opportunities at eLC / UOC**

# Data gathering and preprocessing

- 194 course participants
- Around 400000 events in six weeks
  - EdX tracking logs format
  - Video events: *play\_video*, *stop\_video*, *seek\_video*, *pause\_video*
  - Quiz events: *problem\_check*
- One binary variable for each video (23 in the first module, VM1\_1 ... VM1\_23)
- Target: predicting PASS / FAIL of first quiz (PFM)

# Data analysis (I)

- **Four groups according to topic (G1, ..., G4)**
  - **Natural numbers (4 videos)**
  - **Integer numbers (13)**
  - **Rational numbers (3)**
  - **Real / complex numbers (3)**
- **Two groups according to type (GT, GE)**
  - **Theory (12 videos)**
  - **Exercises (11)**

# Data analysis (II)

- Polychoric (tetrachoric) correlation between VM1\_i variables
  - Higher correlation by “smoothing”
  - Correlation matrix can be NPD
- Principal component analysis
  - Applied to each group (6)
  - Varimax rotation to reduce number of items in each component
- The first component is centered and scaled (MEAN=0, SD=1)

# Results (I)

- **Component explained variance is 40-45%**
- **Large kurtosis**
  - **Assymetrical distribution**
  - **Bimodality (d1)**
- **All component items have significant weights (d2)**
  - **Weights follow an increasing pattern in most cases (d3)**
  - **No difference between theory and exercise weights within topic groups (d4)**



# Results (II)

- 41.8% of students pass the first test
- Two generalized linear models (logit) built for each category
  - According to topic
    - Only G2 and G4 are significant ( $p < 0.05$ ) (d5)
  - According to type
    - Both GT and GE are significant ( $p < 0.05$ ) (d6)
- No real predictive power

# Discussion

- (d1) Most students do nothing or watch all videos, showing no planning or strategy
- (d2) For each group, all videos are considered to be equally important
- (d3) Nevertheless, in some cases last videos seem to be more important (accumulative?)
- (d4) Theory and exercise videos within a group seem to be similar (exercises slightly higher)
- (d5) Nat + Int / Rat + Real + Com “groups”
- (d6) Both theory and exercise are needed

# Conclusions and future research

- **Students face the MOOC with no instructional scaffolding, they take a look and decide whether following the course or not**
- **Quizzes may be biased towards some topics**
- **Measure MOOC impact on the Algebra course**
- **Include more data from video consumption**
  - **Use real time spent on each video**
  - **Analyze video usage (pause, seek)**
  - **Take into account feedback from quizzes**
- **This was just a first step!**

# Thank you!

jminguillona[at]uoc[dot]edu  
@jminguillona

<http://oer.uoc.edu/LAIKA/>

