Student Retention and Complexity Theory

Jonas Forsman (jonas.forsman@physics.uu.se)
Physics Education Research, Department of Physics and Astronomy
Outline

• Research Rationale + Research Questions

• Short introduction to Student Retention Research

• Introduction to Complexity Theory

• Empirical examples of the use of Complexity Theory in education
  – Institutional action in a Complex System of HE
  – Students’ networked interaction
Research Rationale

Societal Need – Engineers and Scientists: A Swedish example

Percentage of Students That Have Finished Their Degree Within Five and Seven Years

Degree completion (%)

Starting Year


years

5

7
Research Rationale

The "right kind" of student – Recruitment?

*It is "very unlikely that there is another hidden pool of students that we might magically discover if we change or further improve our selection procedures" (Allie et al., 2009, p. 3).*

The UK example (Smith, 2010; European Commission, 2004)
- No significant increase in choosing
- Completion numbers are still declining.

A student who stays is equivalent to 2-3 recruited students.
Research Rationale

What can we do when students are at the university?

Modelling efforts

acknowledge "the complexity of student retention".

"regards the decision to leave a particular social system [i.e. studies in higher education] as the result of a complex social process" (Spady 1971, p. 38)

call for the use of social network analysis

Student retention research needs to employ “network analysis and/or social mapping of student interaction... [To]...better illuminate the complexity of student involvement” (Tinto, 1997, p. 619)
Research Questions

**General Aim:**

*How to conceptualize and carry out analysis of student retention for university physics students using a Complexity Theory perspective?*

Research Question 1: *In order to explore viable options for real world practice to enhance student retention, how can informative modelling of action within the complex system be established?*

Research Question 2: *In terms of university physics being a complex system, what roles of student interaction patterns emerge vis-à-vis (1) the core concepts of student retention and (2) students’ grade achievement?*
Introduction to Student Retention – Historical Overview

- University structure

- Social Integration perspective (~1960)

- Psychological perspective (~1980)

- Learning communities (~2000)
Student Integration Model

Outline
- Research
- Rationale

Student Retention
- Research

Complexity
- Theory

Empirical Example
- RQ1: Institutional Action
- RQ2: Students’ Networked Interaction

Expanded Student Integration Model (Reconstructed from Tinto, 1997, p. 615)
Student Integration Model

Social systems (Durkheim, 2004 [1961])
- Individual ↔ Social system

Social system (Tinto, 1997; Spady 1970;1971)
- Composed of social norms, expectations, discourses.

Academic system (Tinto, 1997; Spady 1970;1971)
- Composed of Academic norms, expectations, discourses.
Why focus on Complexity?

Idea of a complexity component has been ubiquitous in the field of student retention

Researchers have identified a multitude of critical factors

Empirical inconsistencies found in factors affecting student retention can be seen as problematic
What is Complexity Theory?

Complexity Theory

... aims to describe and understand complex systems and their capacity to show order, pattern, and behaviour.

... has taken root and emerged in a wide range of disciplines, (see Waldrop, 1992)

what is central in describing or understanding a complex system is identifying the components, their interactions, and the higher order behaviours and properties that emerge from the complex system
The Role of Components

Outline
Research Rationale
Student Retention Research
Complexity Theory
Empirical Example – RQ1: Institutional Action
Empirical Example – RQ2: Students’ Networked Interaction
The Role of Components and Their Interaction

Contact network (Stehlé et al., 2011) containing classes of pupils (coloured clusters) and teachers (white nodes) and their interactions during one day of primary school. (Raw data available from SocioPatterns Collaboration at http://www.sociopatterns.org)
The Role of Multilevel Systems

Stratification (nestedness) of Higher Education

Outline
- Research Rationale
- Student Retention Research
- Complexity Theory
- Empirical Example – RQ1: Institutional Action
- Empirical Example – RQ2: Students’ Networked Interaction

System effects

Students’ networked interaction

- University
- Schools
- Classrooms
- Social and Academic Groups
- Students

Stratification (nestedness) of Higher Education
There are a number of different 'types' of complexity.

...the problem of measuring complexity is the problem of describing electromagnetism before Maxwell's equations. In the case of electromagnetism, quantities such as electric and magnetic forces that arose in different experimental contexts were originally regarded as fundamentally different. Eventually it became clear that electricity and magnetism were in fact closely related aspects of the same fundamental quantity, the electromagnetic field. Similarly, contemporary researchers in architecture, biology, computer science, dynamical systems, engineering, finance, game theory, etc., have defined different measures of complexity for each field.

Lloyd (2001, p.7)
Description of an educational system as a complex system,

Educational systems are made up of a number of agents (students, teachers, etc.) and components (study behaviour, social factors, financial factors, etc.) that interact.

Interaction networks emerge that have different functions throughout the system.

Educational systems adapt to both internal and external influences and evolve over time though, for example, policy, curriculum reforms, and economy.
Empirical Example: Research Question 1

In order to explore viable options for real world practice to enhance student retention, how can informative modelling of action within the complex system be established
Rationale

Interest in Higher Education as a “complex” system is increasing (cf. Morrison 2005; Davis & Sumara 2006; Lemke & Sabelli 2008; Mason 2008; Maroulis et al. 2010).

Previous educational research lacks a methodology to pose “what if” questions. (cf. Davis & Sumara 2006; Sabelli et al. 2013)

Two ways of producing a ”skeleton” – Theoretically or Empirically – in order to enable such methodology
# Data

**Data** - 573 Questionnaire respondents (First year students)

<table>
<thead>
<tr>
<th>Questionnaire Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students’ age</td>
</tr>
<tr>
<td>Students’ Parents’ education</td>
</tr>
<tr>
<td>Students’ biological gender</td>
</tr>
<tr>
<td>Students’ Housing situation</td>
</tr>
<tr>
<td>Students’ impairments</td>
</tr>
<tr>
<td>Stem profile combination (Beta-Mentality Model)</td>
</tr>
<tr>
<td>Students’ prior exposure to university PR</td>
</tr>
<tr>
<td>Students’ previous achievement in mathematics</td>
</tr>
<tr>
<td>Students’ previous achievement in physics</td>
</tr>
<tr>
<td>Students’ prior education Teacher expectations (2 [Expec])</td>
</tr>
<tr>
<td>University facilities (5 [Fc])</td>
</tr>
<tr>
<td>Scheduling (6 [N])</td>
</tr>
<tr>
<td>Course materials (4 [Oo])</td>
</tr>
<tr>
<td>Teacher behaviours (7 [Tc])</td>
</tr>
<tr>
<td>Travel time to campus</td>
</tr>
<tr>
<td>Assessment and feedback (9 [Ts])</td>
</tr>
<tr>
<td>Students’ reenrolment expectations</td>
</tr>
<tr>
<td>Student experiences of university facilities (2 [Fcs])</td>
</tr>
<tr>
<td>Degree importance (2 [Important])</td>
</tr>
<tr>
<td>Language skills (2 [Language])</td>
</tr>
<tr>
<td>Fraternity membership</td>
</tr>
<tr>
<td>Student experience of course materials (2 [Oos])</td>
</tr>
<tr>
<td>Students’ study behaviour (20 [Sb])</td>
</tr>
<tr>
<td>Students’ self-evaluated skills (3 [Skill])</td>
</tr>
<tr>
<td>Student’s credits achieved [1 item]</td>
</tr>
</tbody>
</table>
Method Overview

Commonly used estimation methods in Bioinformatics:

• Bayesian Networks (cf. Friedman, et al. 2000)

• Correlation Network (cf. Langfelder & Horvath, 2008)

• Partial Correlation Networks (cf. Peng et al., 2009)

• Multidimensional Scaling (proposed as a way to estimate network structure in an educational system [Forsman, et al. 2014a])

Method Overview

Simulation tests
• Random network
  – 40-150 node size
  – 200-1200 simulated Likert-type samples
  – Noise ratios (1:100, 1:50, 1:10)

\[
\text{Accuracy} = \frac{\sum \text{True positive} + \sum \text{True negative}}{\sum \text{All outcomes and predicted outcomes}}
\]

<table>
<thead>
<tr>
<th>Estimation Method</th>
<th>Mean Accuracy</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMST</td>
<td>83% (70% / 59%)</td>
<td>2% (1% / 1%)</td>
</tr>
<tr>
<td>Bayesian</td>
<td>80% (69% / 58%)</td>
<td>3% (2% / 1%)</td>
</tr>
<tr>
<td>Correlation</td>
<td>80% (69% / 58%)</td>
<td>1% (1% / 1%)</td>
</tr>
<tr>
<td>Multidimensional Scaling</td>
<td>78% (68% / 58%)</td>
<td>1% (1% / 2%)</td>
</tr>
<tr>
<td>Partial Correlations</td>
<td>27% (36% / 44%)</td>
<td>2% (12% / 5%)</td>
</tr>
</tbody>
</table>
Multilayer Minimum Spanning Tree Analysis

(Grönlund, Bhalerao, & Karlsson, 2008, p. 317)
Method

Multilayer minimum spanning tree analysis (Grönlund et al., 2008)

+ significance test

+ positive / negative ties
Outline

Research Rationale

Student Retention Research

Complexity Theory

Empirical Example – RQ1: Institutional Action

Empirical Example – RQ2: Students’ Networked Interaction

Visualization
“General Targets”?

Topological and Cluster Diversity (cf. Eagle, Macy and Claxton, 2010)

Exploring System Effects – Estimation

Rationale

Estimation should take into account the nonlinear feedback and interaction effects (cf. Davis & Sumara 2006; Sabelli et al. 2013)

Multiple parts of the system adapt to the suggested implementation (cf., Stephens & Richey 2011).

We propose to use Gibbs sampling in a networked system

However, not everything in a Higher Education system can be changed
## Categorization

<table>
<thead>
<tr>
<th>Constant</th>
<th>First Order</th>
<th>Second Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students’ age</td>
<td>Teacher expectations (2 [Expec])</td>
<td>Students’ reenrolment expectations</td>
</tr>
<tr>
<td>Students’ Parents’ education</td>
<td>University facilities (5 [Fc])</td>
<td>Student experiences of university facilities (2 [Fcs])</td>
</tr>
<tr>
<td>Students’ biological gender</td>
<td>Scheduling (6 [N])</td>
<td>Degree importance (2 [Important])</td>
</tr>
<tr>
<td>Students’ Housing situation</td>
<td>Course materials (4 [Oo])</td>
<td>Language skills (2 [Language])</td>
</tr>
<tr>
<td>Students’ impairments</td>
<td>Teacher behaviours (7 [Tc])</td>
<td>Fraternity membership</td>
</tr>
<tr>
<td>Stem profile combination (Beta-Mentality Model)</td>
<td>Travel time to campus</td>
<td>Student experience of course materials (2 [Oos])</td>
</tr>
<tr>
<td>Students’ prior exposure to university PR</td>
<td>Assessment and feedback (9 [Ts])</td>
<td>Students’ study behaviour (20 [Sb])</td>
</tr>
<tr>
<td>Students’ prior education</td>
<td></td>
<td>Students’ self-evaluated skills (3 [Skill])</td>
</tr>
<tr>
<td>Students’ previous achievement in mathematics</td>
<td></td>
<td>Student’s cedits achieved [1 item]</td>
</tr>
<tr>
<td>Students’ previous achievement in physics</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Methodology

Estimation through Gibbs sampling from a normal distribution

\[ n_i \sim N(\mu_i^*, \sigma_i^2) \]

\[ \mu_i^* = \frac{\sum_j w_{ij} n_j}{\sum_j w_{ij}} \]

\[ \sigma_i^2 = \frac{\sum_j w_{ij}^2}{\left(\sum_j w_{ij}\right)^2 - \left(\sum_j w_{ij}^2\right)} \sum_j w_{ij} (n_j - \mu_i^*)^2 \]

Where \( w_{ij} \) is equal to the edge weight between aspect \( i \) and \( j \), and \( n_j \) is the value of aspect \( j \). \( \mu_i^* \) is the weighted mean, and \( \sigma_i^2 \) is the unbiased mean square weighted estimator of sample variance.
## Estimation Results

<table>
<thead>
<tr>
<th>First Order Aspects</th>
<th>Change (%)</th>
<th>St.Dev (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher expectations - Expec_difficulties</td>
<td>11</td>
<td>30</td>
</tr>
<tr>
<td>Course materials - Cm_material</td>
<td>9</td>
<td>32</td>
</tr>
<tr>
<td>Teacher behaviours - Tb_empathize</td>
<td>8</td>
<td>30</td>
</tr>
<tr>
<td>Teacher behaviours - Tb_content</td>
<td>8</td>
<td>30</td>
</tr>
<tr>
<td>Course materials - Cm_feedback</td>
<td>8</td>
<td>30</td>
</tr>
<tr>
<td>Course materials - Cm_late</td>
<td>7</td>
<td>30</td>
</tr>
<tr>
<td>Teacher behaviours - Tb_enthusiasm</td>
<td>6</td>
<td>29</td>
</tr>
<tr>
<td>Teacher behaviours - Tb_explain</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>Assessment &amp; feedback - Af_level</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>Assessment &amp; feedback - Af_constr</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>Teacher behaviours - Tb_available</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>Teacher expectations - Expec_interest</td>
<td>5</td>
<td>28</td>
</tr>
<tr>
<td>Scheduling - N_lectures*</td>
<td>5</td>
<td>80</td>
</tr>
</tbody>
</table>

Empirical Example: Research Question 2

*In terms of university physics being a complex system, what roles of student interaction patterns emerge vis-à-vis (1) the core concepts of student retention and (2) students’ grade achievement?*
Social and Academic Interaction Networks – “rules of interaction”

Data

2 first year courses (~ 60 + 60) Interaction network survey together with grades and course completion

Rationale

Student retention research (eg. Tinto, 1975; Spady, 1972)

Social interaction ↔ Grades (cf. Thomas, 2000)

Social system ↔ Social Integration ↔ Complexity

Rationale cont.

Divide interaction into social and academic interaction

Social systems (eg. Durkheim, 2004 [1961])

Social system ↔ Social interaction (cf. Sawyer, 2005)

Social system ↔ Academic system (eg. Tinto, 1975; Spady, 1972)

Research Aims

Can these "systems" in effect "govern" interaction?

i.e., are these networks random?

Are the social and academic systems two distinct systems?

i.e., are these networks significantly different?
## Course Networks

<table>
<thead>
<tr>
<th>Course Networks</th>
<th>Simulation of random network - mean clustering coefficient</th>
<th>Clustering Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 S&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.027 ± 0.005&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.364</td>
</tr>
<tr>
<td>C1 A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.026 ± 0.005&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.247</td>
</tr>
<tr>
<td>C2 S&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.041 ± 0.004&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.424</td>
</tr>
<tr>
<td>C2 A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.034 ± 0.005&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.374</td>
</tr>
</tbody>
</table>
Two systems of "rules of interaction"?

Social and Academic Interaction Networks – grade achievement

Data

2 first year courses (~ 60 + 60) Interaction network survey together with grades and course completion

Research Aim

Can different types of networks predict grade achievement in a better way than a non-divided network
Methodology

- Testing 9 types of network measurements
  - Two different networks

- Multiple logistic regression’s → testing FDR
  \[ P(k) \geq \frac{k}{m \cdot c(m)} \alpha \]
  - keeping only models passing BHY-procedure
    (Benjamini & Yekutieli, 2001)

- AICc criterion minimization of remaining models
Results

\[ GA \propto EC.s - BC.s + ID.a + EV.a \]

- Academic achievement is a function of both social and academic processes.
- Students need to be socially integrated into the social networked interactions but not at the centre.
- Students need to be structurally integrated in the academic networked interaction.

Summary

Complexity Theory is a viable option to investigate possible effects of institutional action to enhance grade achievement, and ultimately student retention.

Complexity Theory poses a powerful theoretical framework to frame social network analysis.

Trade-off between system effects and stable effects when identifying targets for institutional action.

A handful of possible targets for institutional action – be careful of system effects.

Students’ social network and academic networks can be framed as originating from two *social systems*.

Students’ structural position in the social network as well as in the academic network are of importance for their grade achievement.


References


References


References
