Student Retention and Complexity Theory

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

UPPSALA UNIVERSITET

Outline

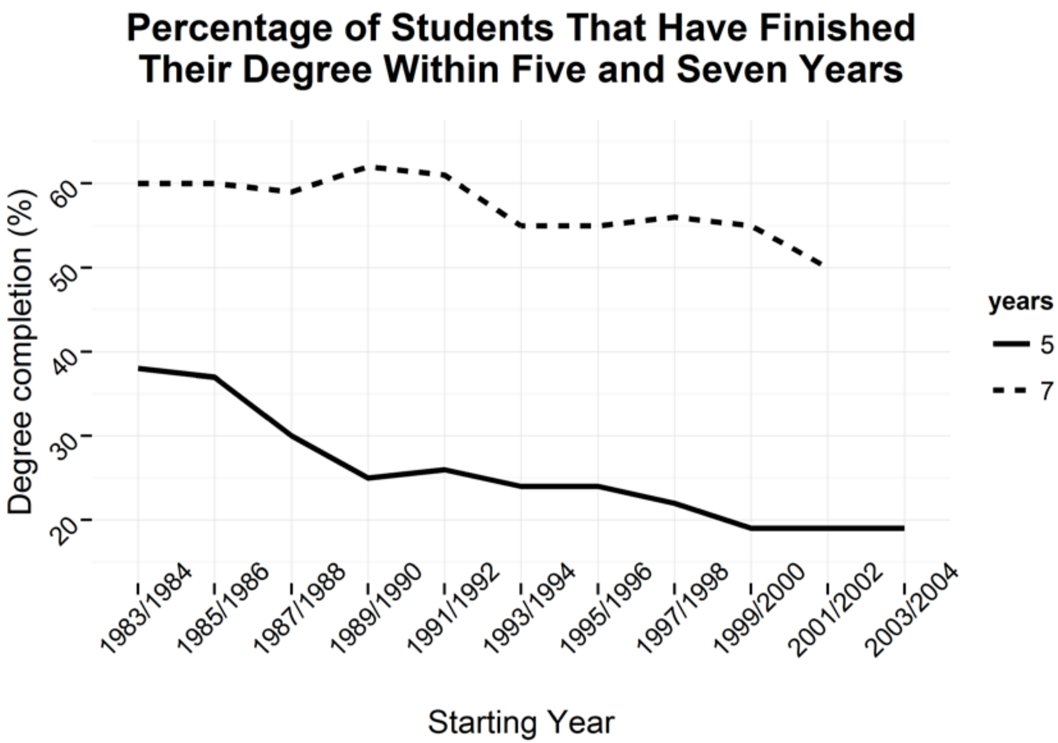
- Research Rationale + Research Questions
- Short introduction to Student Retention Research
- Introduction to Complexity Theory
- Empirical examples of the use of Complexity Theory in education
  - Institutional action in a Complex System of HE
  - Students' networked interaction

UPPSALA UNIVERSITET

Research Rationale

Societal Need – Engineers and Scientists:

A Swedish example



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.

UPPSALA UNIVERSITET

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?*

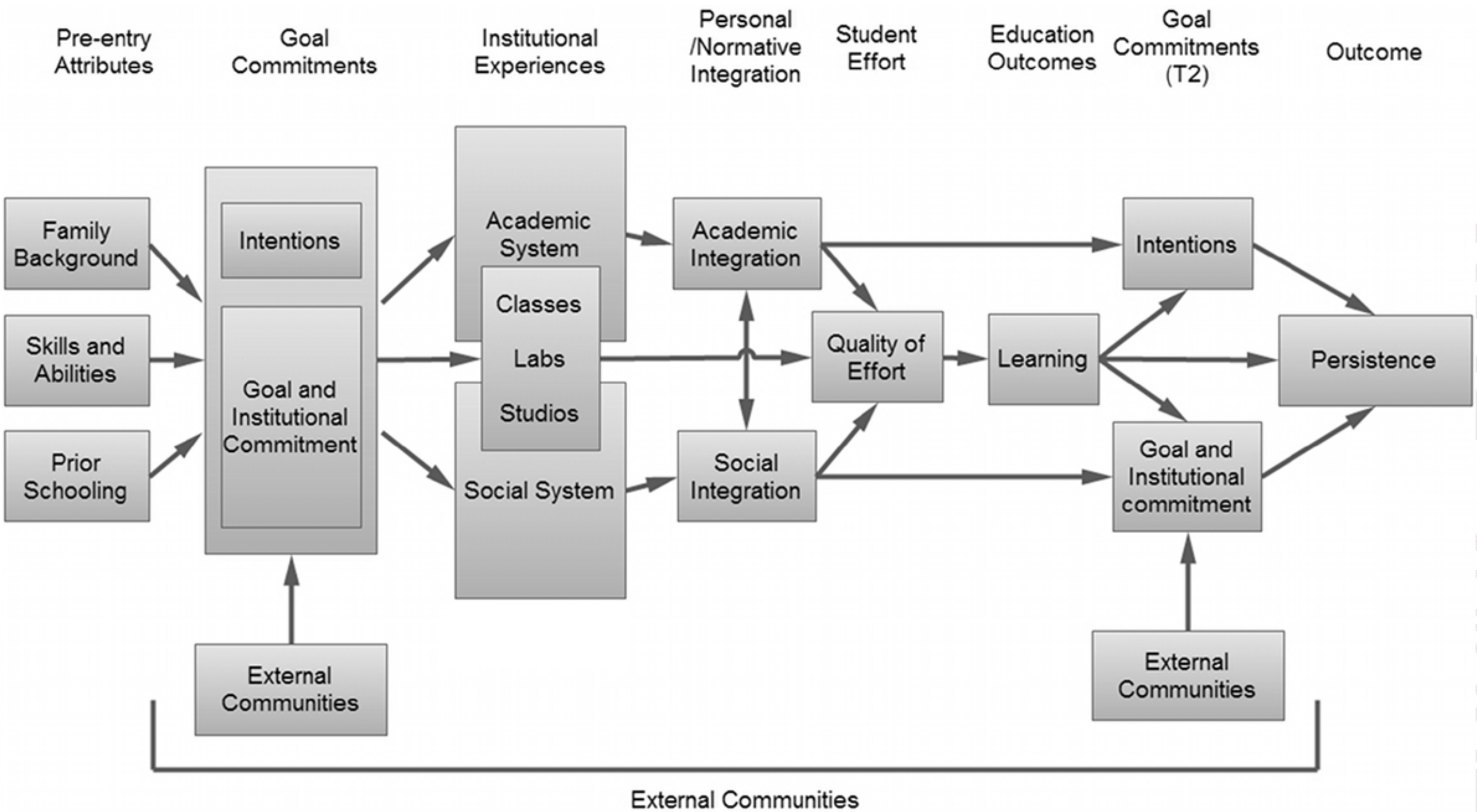
Research Question 2: *In terms of university physics being a complex system, what roles of student interaction patterns emerge vis-à-vis (1) the core concepts of student retention and (2) students' grade achievement?*

UPPSALA UNIVERSITET

Introduction to Student Retention – Historical Overview

- University structure
- Social Integration perspective (~1960)
- Psychological perspective (~1980)
- Learning communities (~2000)

Student Integration Model



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

UPPSALA UNIVERSITET

Outline

Research Rationale

Student Retention Research

Complexity Theory

Empirical Example – RQ1: Institutional Action

Empirical Example – RQ2: Students' Networked Interaction

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.

UPPSALA UNIVERSITET

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

UPPSALA UNIVERSITET

What is Complexity Theory?

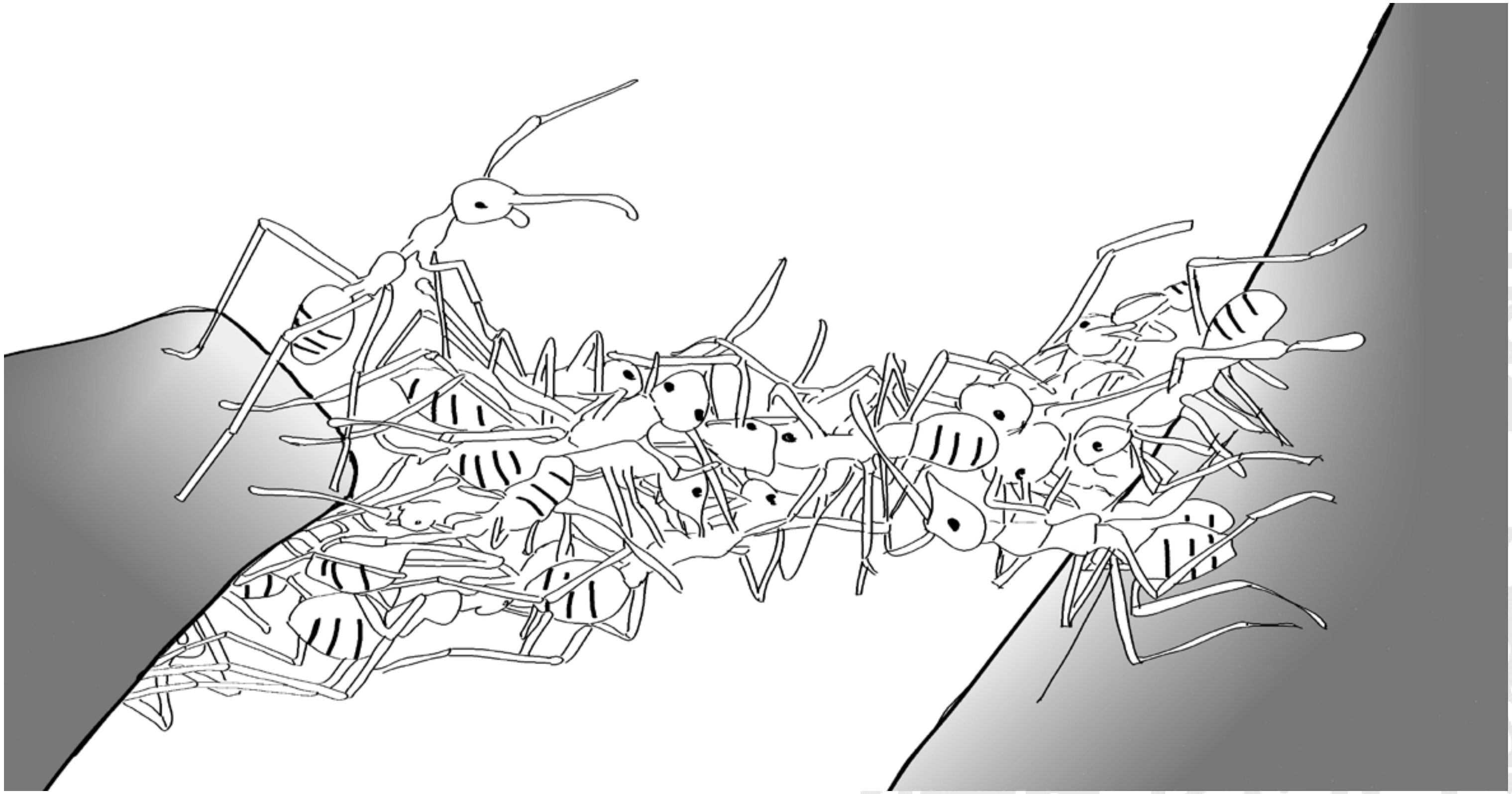
Complexity Theory

… 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



UPPSALA UNIVERSITET

Outline

Research Rationale

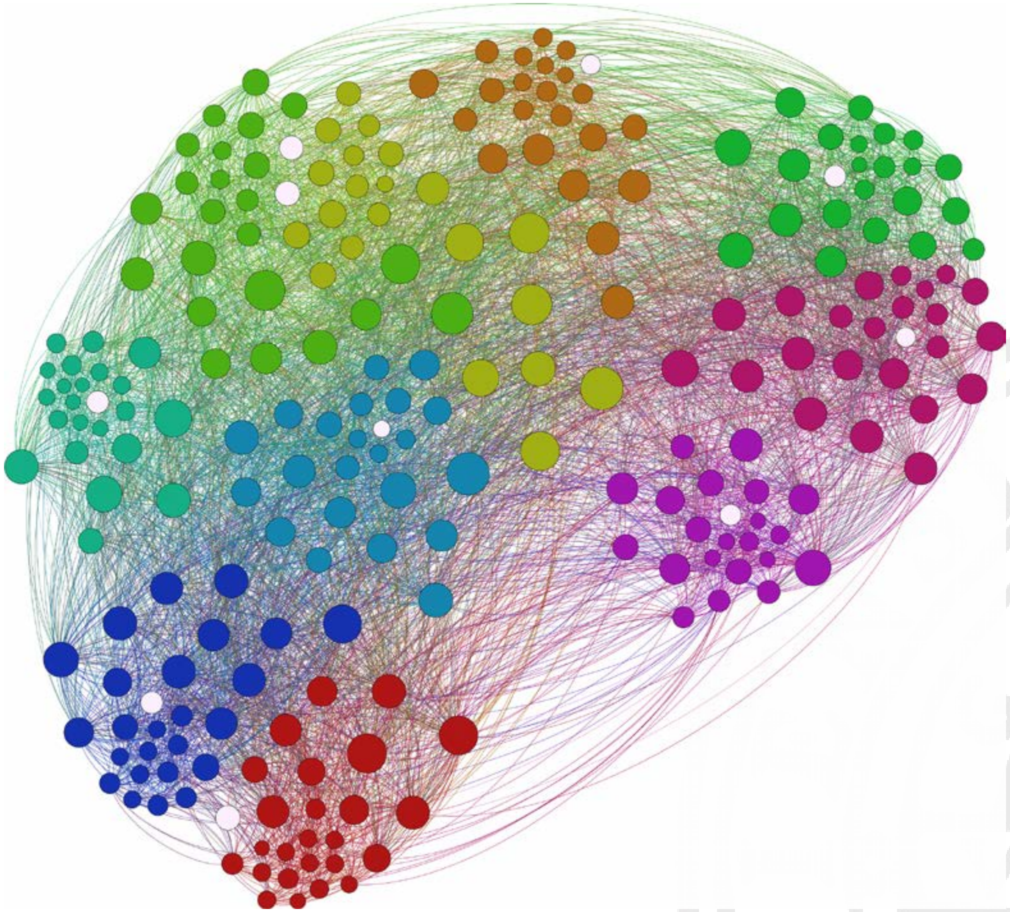
Student Retention Research

Complexity Theory

Empirical Example – RQ1: Institutional Action

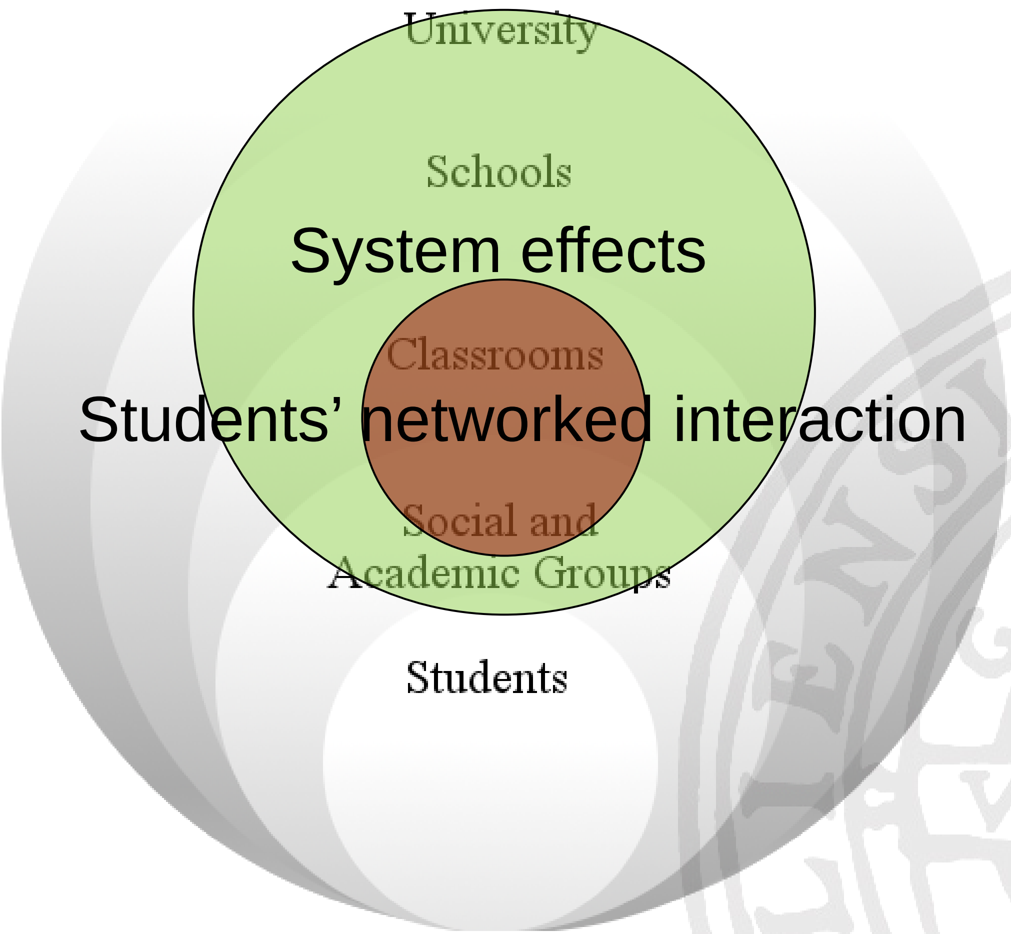
Empirical Example – RQ2: Students' Networked Interaction

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

UPPSALA UNIVERSITET

Characterization of the Term Complexity

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)

UPPSALA UNIVERSITET

Characterization of Complexity in Educational Systems

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.

UPPSALA UNIVERSITET

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

UPPSALA UNIVERSITET

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

UPPSALA UNIVERSITET

Outline

Research Rationale

Student Retention Research

Complexity Theory

Empirical Example – RQ1: Institutional Action

Empirical Example – RQ2: Students' Networked Interaction

Data

Data - 573 Questionnaire respondents (First year students)

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

UPPSALA UNIVERSITET

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])
- Multilayer Minimum Spanning Tree Analysis (cf. Grönlund, et al. 2009; Forsman, et al. 2014b; 2014c)

UPPSALA UNIVERSITET

Method Overview

Simulation tests

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

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

Estimation Method	Mean Accuracy	Standard deviation
MMST	83% (70% / 59%)	2% (1% / 1%)
Bayesian	80% (69% / 58%)	3% (2% / 1%)
Correlation	80% (69% / 58%)	1% (1% / 1%)
Multidimensional Scaling	78% (68% / 58%)	1% (1% / 2%)
Partial Correlations	27% (36% / 44%)	2% (12%/ 5%)

UPPSALA UNIVERSITET

Outline

Research Rationale

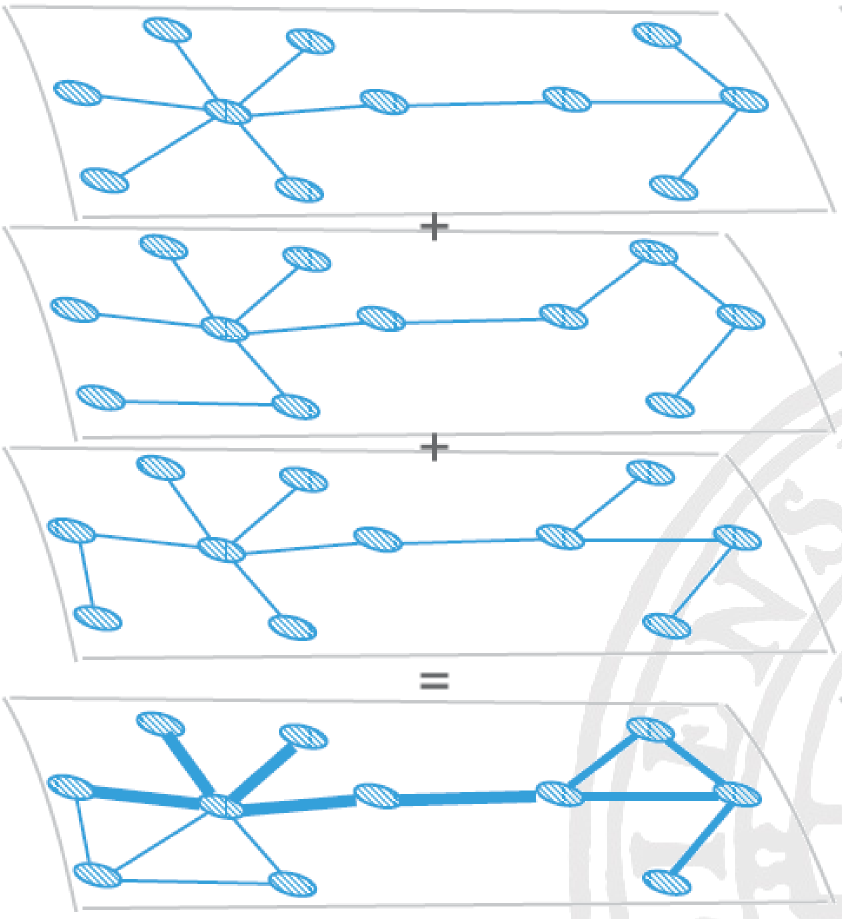
Student Retention Research

Complexity Theory

Empirical Example – RQ1: Institutional Action

Empirical Example – RQ2: Students' Networked Interaction

Multilayer Minimum Spanning Tree Analysis



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

UPPSALA UNIVERSITET

Method

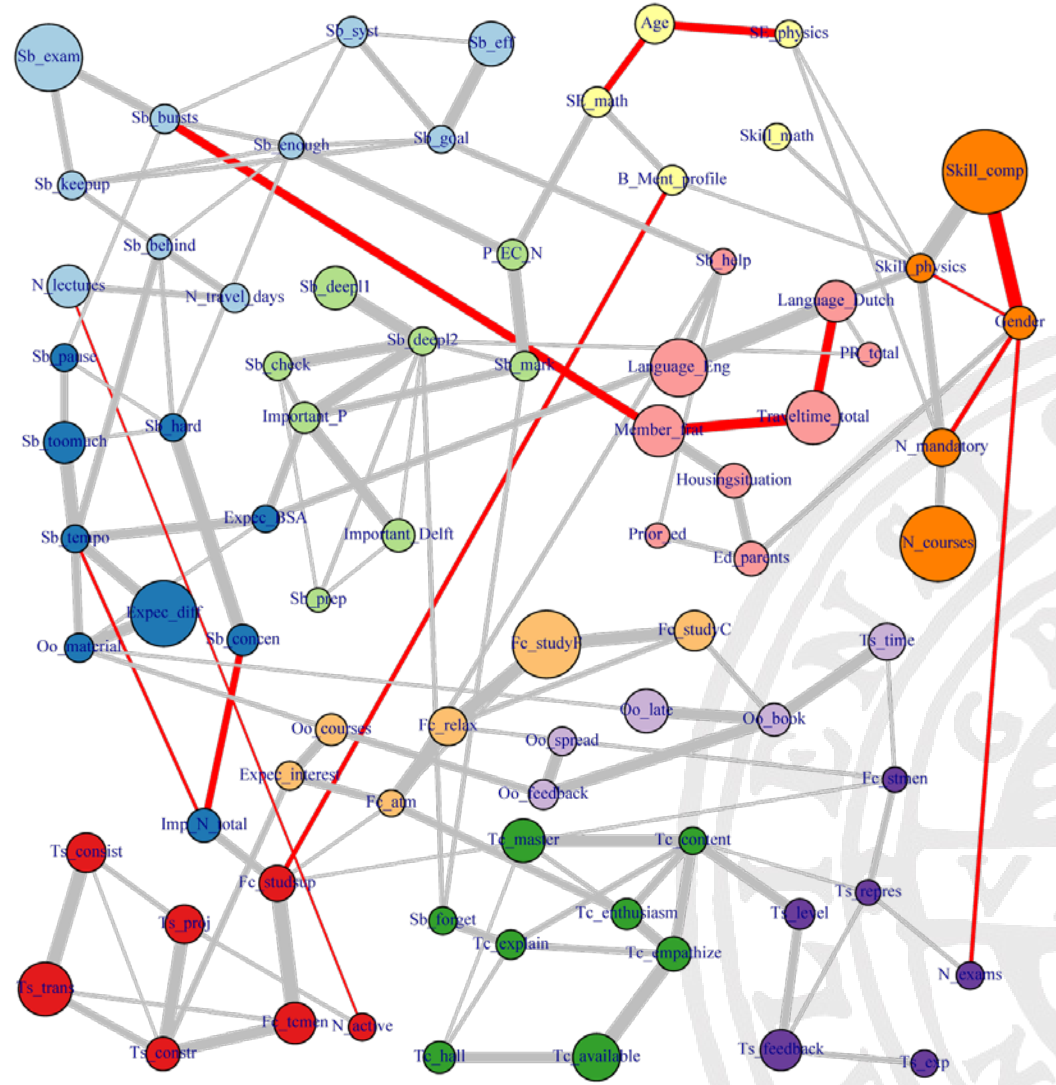
Multilayer minimum spanning tree analysis (Grönlund et al., 2008)

+ significance test

+ positive / negative ties

UPPSALA UNIVERSITET

Visualization



UPPSALA UNIVERSITET

Outline

Research Rationale

Student Retention Research

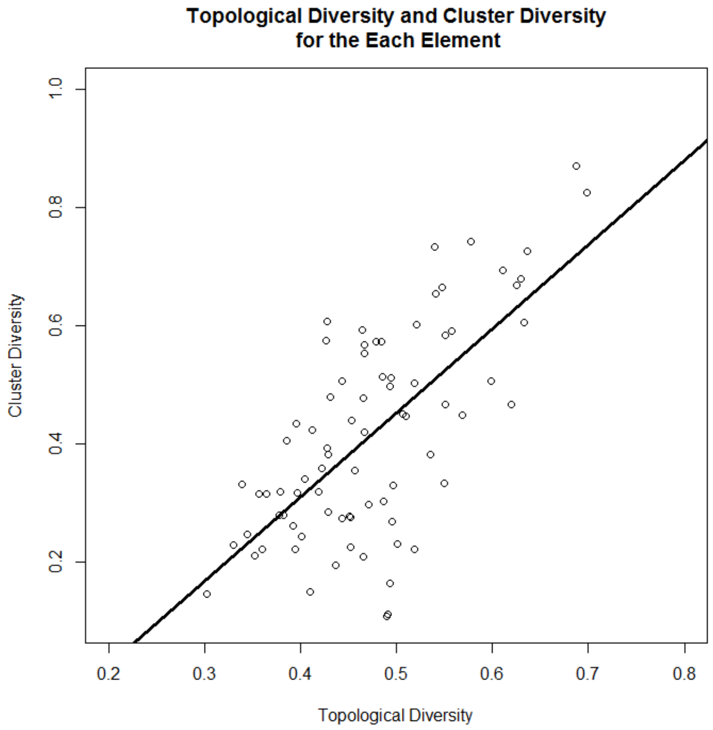
Complexity Theory

Empirical Example – RQ1: Institutional Action

Empirical Example – RQ2: Students' Networked Interaction

"General Targets"?

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



Forsman, J., Van den Bogaard, M., Linder, C., & Fraser, D. (2014). Considering Student Retention as a complex system: A possible way forward for enhancing Student Retention. *European Journal of Engineering Education*. doi: 10.1080/03043797 .2014.941340.

UPPSALA UNIVERSITET

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

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

UPPSALA UNIVERSITET

Outline

Research Rationale

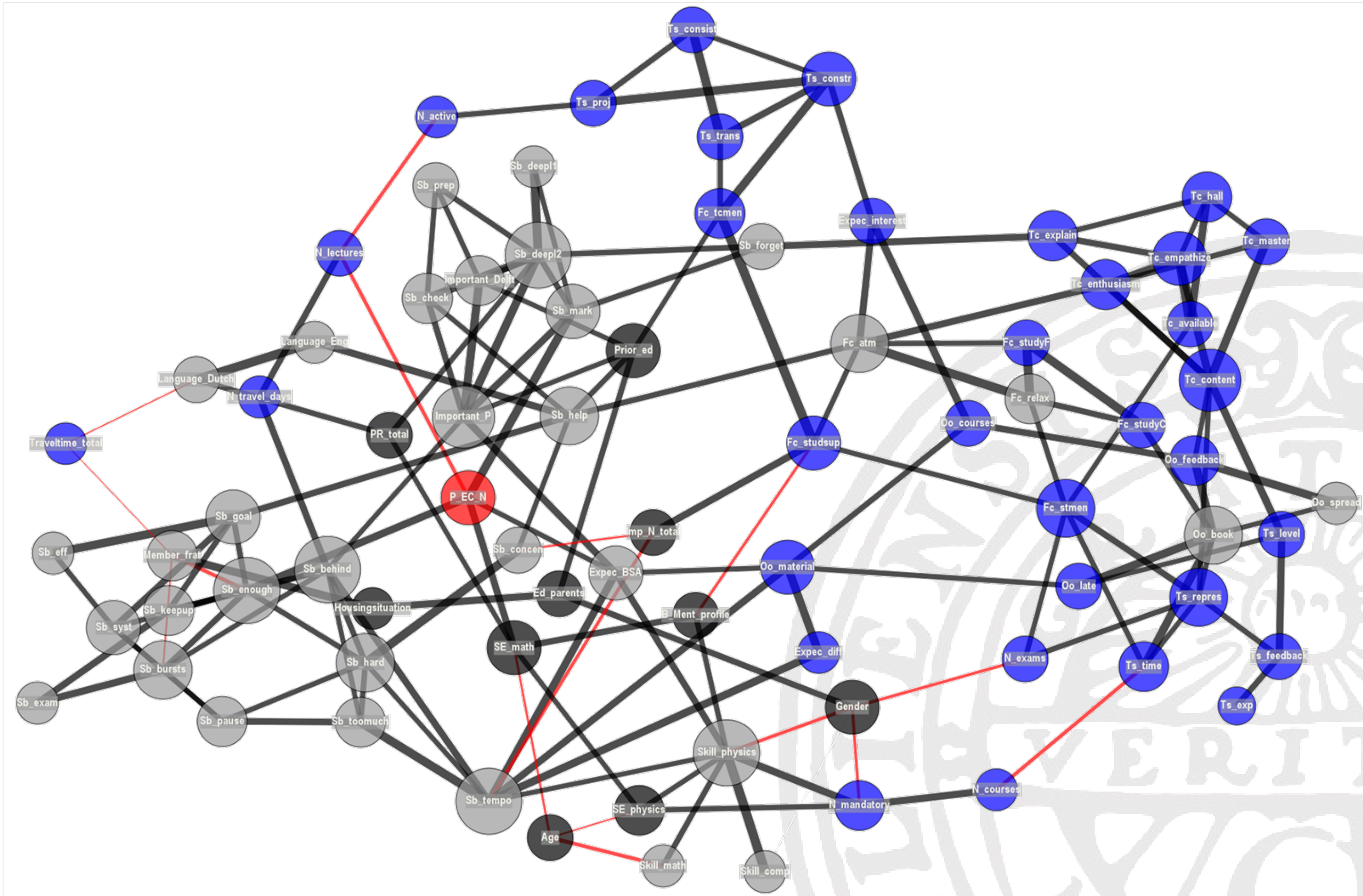
Student Retention Research

Complexity Theory

Empirical Example – RQ1: Institutional Action

Empirical Example – RQ2: Students' Networked Interaction

Visualization

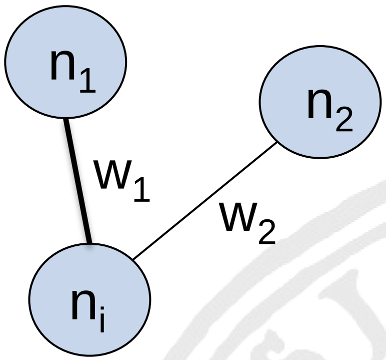


UPPSALA
UNIVERSITET

Methodology

Estimation through Gibbs sampling from a normal distribution

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



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

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

Where w_{ij} is equal to the edge weight between aspect *i* and *j*, and n_j is the value of aspect *j*. μ_i^* is the weighted mean, and σ_i^2 is the unbiased mean square weighted estimator of sample variance

UPPSALA UNIVERSITET

Outline

Research Rationale

Student Retention Research

Complexity Theory

Empirical Example – RQ1: Institutional Action

Empirical Example – RQ2: Students' Networked Interaction

Estimation Results

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

Forsman, J., Mann, R. P., Linder, C., & Van den Boogard, M. (2014). Sandbox University: Estimating influence of institutional action. *PLoS ONE,* 9(7), doi: 10.1371/journal.pone. 0103261.

UPPSALA UNIVERSITET

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
(Forsman, J., Moll, R., & Linder, C. (2014). Extending the theoretical framing for PER: An illustrative application of complexity science, *Physical Review Special Topics-Physics Education Research*, 10(2), 020122.)

UPPSALA UNIVERSITET

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?

UPPSALA UNIVERSITET

Outline

Research Rationale

Student Retention Research

Complexity Theory

Empirical Example – RQ1: Institutional Action

Empirical Example – RQ2: Students' Networked Interaction

Random?

Course Networks	Simulation of random network - mean clustering coefficient	Clustering Coefficient
C1 S[b]	$0.027 \pm 0.005^{*}$	0.364
C1 A[c]	$0.026 \pm 0.005^{*}$	0.247
C2 S[b]	$0.041 \pm 0.004^{*}$	0.424
C2 A[c]	$0.034 \pm 0.005^{*}$	0.374

UPPSALA UNIVERSITET

Outline

Research Rationale

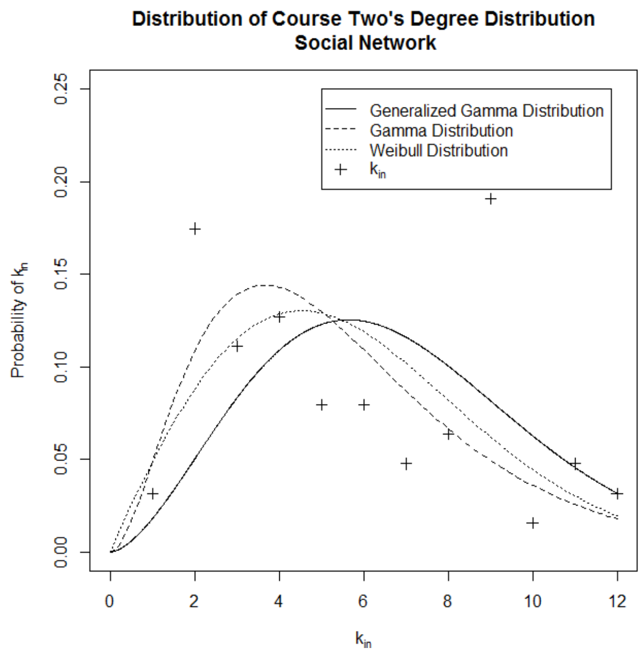
Student Retention Research

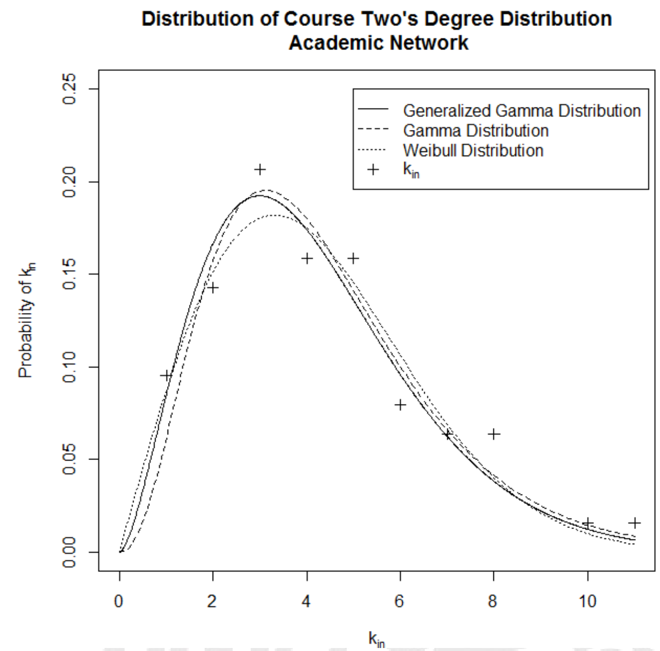
Complexity Theory

Empirical Example – RQ1: Institutional Action

Empirical Example – RQ2: Students' Networked Interaction

Two systems of "rules of interaction"?





Forsman, J., Moll, R., & Linder, C. (2014). Extending the theoretical framing for PER: An illustrative application of complexity science, *Physical Review Special Topics-Physics Education Research*, 10(2), 020122.

UPPSALA UNIVERSITET

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

UPPSALA UNIVERSITET

Methodology

- Testing 9 types of network measurements
  - Two different networks

- Multiple logistic regression's → testing FDR

$$P_{(k)} \geq \frac{k}{m \cdot c(m)} \alpha \qquad \mathrm{c(m)} = \sum_{i=1}^{m} \frac{1}{i}$$

  - keeping only models passing BHY-procedure (Benjamini & Yekutieli, 2001)

- AICc criterion minimization of remaining models

UPPSALA UNIVERSITET

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.

Forsman, J., Linder, C., & Moll, R. (Submitted). Exploring Indicators for Academic Success Using Complexity Thinking and Network Analysis to Investigate Students' Social and Academic Network Structures. *Social Networks*

UPPSALA UNIVERSITET

Summary

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

Complexity Theory pose 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

UPPSALA UNIVERSITET

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

UPPSALA UNIVERSITET

References

Spady, W. 1971. Dropouts from higher education: Toward an empirical model. *Interchange*, 2(3), 38 – 62.

Tinto, V. (1997). Classrooms as communities: Exploring the educational character of student persistence. *Journal of Higher Education*, 68(6), 599 – 623.

Bean, J. P. (1980) Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education,* 12(2), 155 – 187.

Bean, J. P. (1982). Student attrition, intentions, and confidence: Interaction effects in a path model. *Research in Higher Education,* 17(4), 291 – 320.

Waldrop, M. M. (1992). *Complexity: The emerging science at the edge of order and chaos*. New York: Simon & Schuster.

Schäfer, J., Opgen-Rhein, R., & Strimmer, K. (2006). Reverse engineering genetic networks using the GeneNet package. *R News*, 6(5), 50 – 53.

Csardi, G., & Nepusz, T. (2006) The igraph software package for complex network research, *InterJournal, Complex Systems*, 1695-1704.

R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Schmidt-Heck, W., Guthke, R., Toepfer, S., Reischer, H., Duerrschmid, K., & Bayer, K. (2004). Reverse engineering of the stress response during expression of a recombinant protein. In: *Proceedings of the EUNITE 2004 European Symposium on Intelligent Technologies, Hybrid Systems and their Implementation on Smart Adaptive Systems*, June 10-12, 2004, Aachen, Germany, Verlag Mainz, Wissenschaftsverlag, Aachen, 2004, 407-441 (ISBN 3-86130-368-X).

UPPSALA UNIVERSITET

References

Mouritsen, K. N., Poulin, R., McLaughin, J. P., & Thieltges, D. W. (2011) Food web including metazoan parasites for an intertidal ecosystem in New Zealand, *Ecology*, 92(10), 2006.

Stehlé, J., Voirin, N., Barrat, A., Cattuto, C., Isella, L., Pinton, J. F., ... & Vanhems, P. (2011). High-resolution measurements of face-to-face contact patterns in a primary school. *PloS one*, 6(8), e23176.

Lloyd, S. (2001). Measurements of complexity: A nonexhaustive list. *IEEE Control Systems Magazine.* 21(4), 7 – 8.

Morrison, K. (2005). Structuration Theory, Habitus and Complexity Theory: Elective Affinities or Old Wine in New Bottles? *British Journal of Sociology of Education*, 26(3), 331 – 326.

Davis, B., & Sumara, D. (2006). *Complexity and education: Inquiries into learning, teaching and research*, New Jersey & London: Lawrence Erlbaum Associates.

Lemke, J. L., & Sabelli, N. H. (2008). Complex systems and educational change: Towards a new research agenda. *Educational Philosophy and Theory*, 40(1), 118-129.

Mason, M. (2008). What is complexity theory and what are its implications for educational change?. *Educational Philosophy and Theory,* 40(1), 35-49.

Maroulis, S., Guimera, R., Petry, H., Stringer, M. J., Gomez, L. M., Amaral, L. A. N., & Wilensky, U. (2010). Complex systems view of educational policy research*. Science(Washington)*, 330(6000), 38-39.

Sabelli, N., Lemke, J., Cheng, B., & Richey, M. (2013). *Education as a complex adaptive system: Report of progress*. Available: http://msce.sri.com/publications/Education_As%20A_Complex_Adaptive%20System.pdf. Accessed 30 September 2013.

Friedman, N., Linial, M., Nachman, I., & Pe'er, D. (2000). Using Bayesian networks to analyse expression data. *Journal of computational biology*, 7(3-4), 601-620.

UPPSALA UNIVERSITET

References

Langfelder, P., & Horvath, S. (2008). WGCNA: an R package for weighted correlation analysis, *BMC Bioinformatics*, 9(559).

Peng, J., Wang, P, Zhou, N., & Zhu, J. (2009). Partial correlation estimation by joint sparse regression models. *Journal of American Statistical Association*, 104(486), 735 – 746.

Forsman, J., Linder, C., Moll, R., Fraser, D., & Andersson, S. (2014a). A new approach to modelling student retention through an application of complexity thinking. *Studies in Higher Education*, 39(1), 68–86.

Grönlund, A., Bhalerao, R. P., and Karlsson, J. (2009). Modular gene expression in poplar: a multilayer network approach. *New Phytolologist,* 181, 315 – 322.

Forsman, J., Van den Bogaard, M., Linder, C., & Fraser, D. (2014b). Considering Student Retention as a complex system: A possible way forward for enhancing Student Retention. *European Journal of Engineering Education. doi*: 10.1080/03043797 .2014.941340.

Forsman, J., Mann, R. P., Linder, C., & Van den Boogard, M. (2014c). Sandbox University: Estimating influence of institutional action. *PLoS ONE*, 9(7), doi: 10.1371/journal.pone. 0103261.

Eagle, N., Macy, M., & Claxton, R. (2010). Network Diversity and Economic Development. *Science*, 328, 1029 - 1031.

Stephens, R., & Richey, M. (2013). Accelerating STEM Capacity: A Complex Adaptive System Perspective. *Journal of Engineering Education*, 100(3), 417-423.

Thomas, L. S. (2000). Ties that bind: A social network approach to understanding student integration and persistence. *The Journal of Higher Education*, 71(5), 591 - 615.

Durkheim, E. (2004). *Suicide: a Study in sociology*, (M. Johansson, Trans). Lund: Argos (Original work published in 1961).

UPPSALA UNIVERSITET

References

Sawyer, R. K. (2005). *Social emergence societies as complex systems*. New York: Cambridge University Press.

Benjamini, Y., & Yekutieli, D. (2001). The Control of the False Discovery Rate in Multiple Testing under Dependency, *The Annals of Statistics*, 29(4), 1165-1188, Stable url: http://www.jstor.org/stable/2674075