Using complexity thinking to develop a new model of student retention

Jonas Forsman[a]*, Cedric Linder[a], Rachel Moll[b], Duncan Fraser [c]

[a]*Division of Physics Education Research, Department of Physics and Astronomy, Uppsala University, Uppsala, Sweden*
[b]*Faculty of Education, Vancouver Island University, Vancouver, British Columbia, Canada*
[c]*Centre for Research in Engineering Education, University of Cape Town, Rondebosch, 7701 South Africa*

Abstract

This paper proposes a new approach to the modeling of student retention in higher education, namely the use of Complexity Thinking, in conjunction with Exploratory Factor Analysis and Multidimensional Scaling. To illustrate our proposal we analyse a small data sample collected from undergraduate engineering students at a highly regarded traditional Swedish university. This analysis shows that issues affecting student retention should be viewed as nested, interconnected systems, in which certain components are more influential than others, rather than in linear terms.

Keywords: Student retention, modeling systems, complexity thinking

*Duncan Fraser. Tel.: +27-21-650-2515
  E-mail address: Duncan.Fraser@uct.ac.za

1. Motivation

Student retention has long been an area of research in higher education. This research has led to models of student retention that are widely used (Bean, 1982; Tinto, 1997). Although "complex" aspects of the relationships between the variables in these models have been acknowledged by many workers in the field, they have not been explicitly incorporated into these models. As a result the existing models are easily interpreted in linear ways. To address this critique we are proposing a different approach to modeling student retention by drawing on complexity thinking as a conceptual framework to develop a new structure for the modeling of student retention. Complexity thinking, when employed in educational research as a conceptual framework (Davis &Sumara, 2006), can identify the structure and dynamics that characterize complex systems, such as the adaptive and decentralized systems of variables that influence student retention.

To illustrate what we are proposing we have used data consisting of questionnaire responses made by Swedish students most of whom were in Masters of Science in Engineering (Physics, and Materials Physics) programmes. We chose this setting context because of international concern over the critical increases in demand for new engineers and scientists,coupled with a decline in interest in careers in science and technology in many countries, particularly in the developed world (CSEPP, 2007; OECD, 2009).

Moreover, in many countries the percentage of students who manage to successfully complete their degree requirements in science and engineering programmes in minimum time, or who graduate at all, is decreasing (CSEPP, 2007; OECD, 2009). In Sweden, for example, the percentage of university students who complete the Master of Science in Engineering degree in minimum time has decreased from 30% to 19% from 1987/8 to 2003/4, and the overall completion rate is also declining (see Figure 1).
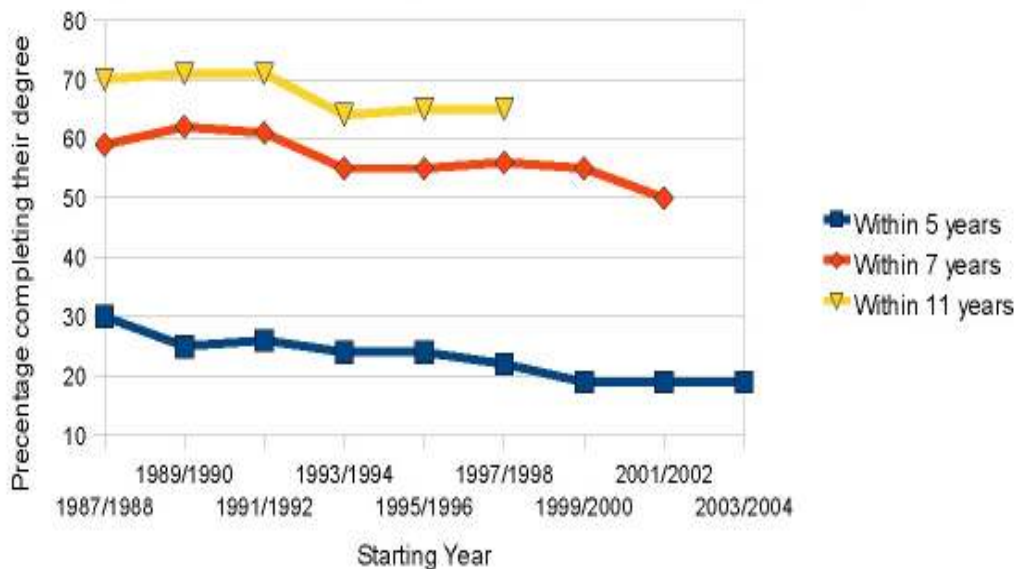

Figure 1. Percentage of students completing their degree within a set time frame [37, 38]

The current initiatives and the decline of graduation rates in both the EU and the US, especially in science and engineering, has created a renewed challenge for higher education institutions to create conditions that are more likely to enhance student progression. It is our aim in this study to develop a model of student retention that will be able to inform such institutions of the most critical features affecting student retention, together with how they are related to one another, so as to better enable them to make holistic decisions with regard to improving retention.

2. Models of student retention

For the purpose of this paper, we start with an overview of the models of student retention and the issues that have driven their development, in order to highlight the importance of introducing complexity thinking into such models. More extensive historical overviews of the field in general are available (Metz, 2004; Summerskill, 1962; Tinto, 1975).

The notion of "complexity"[†] has been apparent in the development of meaningful models of student retention, but it has not been brought to the fore up till now. For example, Spady (1971, p. 38) argues that the formulation of a truly comprehensive model of student retention needs a perspective that "regards the decision to leave a particular social system [studies in higher education] as the result of a complex social process". More recently Bean (2005, p. 238) has argued that "students' experiences are complex, and their reasons for departure are complex".

This brings us to our starting point for the theoretical development for this study. Currently two models of student retention are widely used: The Student Integration Model of Tinto (1975, 1987 and 1997) and The Student Attrition Model of Bean (1980, 1982 and 2005). While these models are seen by some researchers as being two separate systems, we agree with Cabrera et al. (1992a, p. 145) that both "regard persistence [retention] as the result of a complex set of interactions over time."

Yorke and Longden (2004) describe how the focus of early studies of student retention in higher education were on university structures, such as libraries, schedules, courses and exam timetables. Subsequently there was a shift in focus towards incorporating a social integration perspective, influenced largely by the work of Spady (1970 and 1971).

The social integration perspective posits that becoming integrated within a social system requires learning the norms, value-systems, and beliefs through interactions within the system. This played a major role in the development of Spady's theoretical model (1970 and 1971), which sees social integration as a process that encompasses many aspects of students' everyday lives (such as friendships, family support, the student's feeling of satisfaction, and the student's intellectual development). Spady's model also includes student characteristics such as grade performance, family background, and academic potential. Spady argued that students needed to become a part of the social world of the university if departure rates are to decline.

Tinto (1975) published an expanded version of Spady's model, which made a distinction between the social system of the university and the academic system. Tinto argued that students also need to become academically integrated to persist in their studies. He posited that some interactions that lead to social integration do not necessarily lead towards integration into the academic system of the university. In Tinto's conceptual framework the academic system contains the rules, norms and expectations that direct academic life within an institution.

――――――

[†]   Complexity should not be seen as a synonym for "complicated", but rather as a characteristic of a system that responds adaptively to stimuli in a way that changes the system itself.

Bean (1980) critiqued Tinto's model for its lack of external factors, such as economy and housing. The point of departure for Bean's model was that student attrition should be seen as analogous to work turn-over in an employment setting. From this, students' attitudes and behaviour are shaped by factors such as social experiences, the experience of the quality of the institution, and family approval.

Cabrera et al. (1992a, 1992b and 1993) evaluated Bean's and Tinto's models by surveying 2453 full-time American freshman students. They found that the two student retention models have common ground and they support each other in explanatory value. The questionnaire they designed was made up of 79 questions, selected from well validated instruments previously used in the field of student retention (see Bean, 1982; Pascarella&Terenzini, 1979).

Eaton and Bean (1995) expanded Bean's earlier model by adding approach and avoidance behavioural theory, on the basis that students' experiences shape their individual behavioural approaches towards university life. Some students' experiences lead towards avoidance behaviour, and some towards an approach behaviour, both of which affect academic integration and thus the students' intention to leave or stay.

Tinto (1997) then expanded his model by introducing the notion of "internal" and "external" communities that affect student integration into university life. He asserted that within classrooms there are "internal" learning communities where both social and academic systems coexist. The concept of learning communities, together with the presence of "external" communities, opened up new constructs that could help to improve student retention.

Braxton (2000, p. 258) argued that due to the wide variations within the empirical findings associated with Tinto's model, it should be "seriously revised". Braxton and Lien (2000) compiled empirical results on academic integration and concluded that Tinto's claim (1975 and 1997) that it is a central construct has not been demonstrated empirically. Braxton and Hirshy(2004) provided empirical data to support their proposal to incorporate three additional factors that may influence social integration: commitment of the institution to student welfare, institutional integrity, and communal potential. Braxton (2000) suggested that a new foundation for modelingstudent retention needs to be developed and Tinto himself (2010) argued for the need to develop models that aim towards informing the institutional action of universities.

The next step is then to put forward a modeling system that includes the constructs of the earlier models, can adapt to variations within empirical findings, and empowers universities in their actions toward enhancing student retention. Complexity thinking is a conceptual framework that can help achieve these aims and has the potential to suggest changes to educational practice.

3. Conceptual framework

In this section, we will present the concepts that we draw upon from complexity thinking to produce a more powerful and holistic modeling system of student retention. As indicated above, the complex nature of student retention has been recognised by the inclusion of a wide set of constructs from different fields of study (such as sociology and psychology) in recent models. What is lacking in dealing with the complexity of student retention is the explicit incorporation of complexity thinking, with its potential as a trans-disciplinary theory that can embrace constructs from a wide array of perspectives and can provide insight into how these constructs interact and influence each other. We will use network theory, exploratory factor analysis and multidimensional scaling to develop a new characterization of the structure and dynamics of the complex system of student retention in higher education.

3.1. Complexity thinking

Complexity thinking aims to describe and understand complex systems and their capacity to show order, patterns, and structure. Especially important is how these orders, patterns and structures seem to emerge spontaneously from interactions between components of systems. Complexity thinking has emerged and taken root in a wide range of disciplines, generating a theory that essentially "transcends disciplines" (Waldrop, 1992). For more details on the historical development of complexity thinking, see Waldrop (1992), and for an overview of current applications of complexity thinking in a wide array of fields, see Mitchell (2009).

Complexity thinking is often pitted against "classical science", which is, in turn, portrayed in terms of efforts to condense phenomena into their simplest components. However, to obtain a reasonable portrayal of a complex phenomenon, an understanding of the properties of the components alone is not sufficient. Thus, what is central in describing or understanding a complex system is identifying the components, their interactions, and what *emerges* from the complex system: system behaviours, properties and structures or the "structuring structures" (Bourdieu, 1984) of the complex system (for example, see Davis &Sumara, 2006).

One can conceptualize the essential aspects of complex systems' structure, dynamics, and predictability through metaphors (for example, see Gilstrap, 2005), computer simulations (for example, see Brown &Eisenhardt, 1997) and systems of modeling (for example, see Mowat& Davis 2010). From this perspective, the essential aspects of complex systems, and what have given rise to complexity thinking's ubiquitous emergence, are that all complex systems share similar structure and dynamics. Although the behaviour of complex systems such as society, organisms, or the internet can only be conceptually discussed as somewhere in-between complete order and complete disorder, any attempt to measure or distinguish one system as more complex than another often breaks down (Mitchell, 2009). If a system is to be identified as being a complex system what needs to be investigated is the presence of structures and dynamics that are common among complex systems, not the complexity itself(Davis &Sumara, 2006).

3.1.1. The structure of complex systems

Complex systems have *decentralized networked structure*, which means that there are a few components or nodes that are much more connected than others. This kind of structure can be contrasted to two other types of networks: (1) centralized networks with only one central node with every other node only connected to that central node; and, (2) distributed networks where all nodes have the same connectivity in the network. Information is spread effectively in centralized networks, but they are vulnerable to break down due to the dependency on the central node. On the other hand, distributed networks are robust to break downs but inefficient in spreading information. In the case of decentralized networks, when a highly connected component is removed or breaks down, then the whole system will suffer considerable damage. The system will remain stable, however, with the removal of any of the many less important or less connected nodes.

Due to their decentralized structure, all complex systems are networked with other complex systems. Moreover, components within a complex system can be considered to be complex systems themselves, thus complex systems are *nested*. Nested systems have similar structure and dynamics but operate on different scales (time, size and so forth). For example, mathematics learning-for-teaching has been modelled as several nested systems: subjective understanding, classroom collectivity, curriculum structure, and mathematical objects (Davis &Sumara, 2006). Each level of such nested complex systems exhibits similar structures and dynamics but operates within different time-scales (for example, subjective understanding has a faster rate of change than mathematical objects) and/or at different levels of analysis (such as the level of an individual, or the level of a group of individuals, or the level of a particular culture, or the level of all human beings).

3.1.2. Dynamics of complex systems

One key aspect of complex systems is that they are continually changing as the components in the system interact with the external environment and with one another. This means that complex systems are *adaptive* and *self-*

organize; properties, behaviour and structure all emerge without an external system or an internal "leader system" that controls the complex system.

Components of complex systems interact mainly locally via *neighbour-interactions*, which can fuel processes that lead to emergence such as positive feedback (brings the system to a non-equilibrium state) or equilibrium through negative feedback. Positive feedback tends to amplify, and negative feedback tends to dampen properties, behaviours and structures. Depending on how "connected" each component is with other components within the system, the positive or negative feedback can be greatly amplified or dampened, which gives rise to the possibility of emergence. Complexity thinking has established that decentralized network structure is a key element in facilitating emergence in complex systems. Through the concept of neighbour-interactions and the decentralized network structure we can argue that nested systems that are highly connected can be seen as close to each other (Davis &Sumara, 2006).

Complexity thinking is not characterized by a particular method but by a methodological perspective that employs a range of methods to study complex phenomena (Davis &Sumara, 2006). Complex systems are networked constellations of components, which in our example are the students' viewpoints of their experience of higher education in the first year. Each component such as students' attitudes towards their program and their financial stability, is considered to emerge from and be situated within multiple complex systems. Analysis of the structure and dynamics of the complex system of students' retention is possible through, but not constrained to, the following tools used in complexity studies: exploratory factor analysis, multidimensional scaling and network theory.

3.2. Exploratory factor analysis

Exploratory factor analysis is used to study patterns and order within complex data by comparing angles between points in a multidimensional space. A useful way to view exploratory factor analysis is to see it as essentially what Hoftstede et al. (1990, p. 299) has called "ecological factor analysis"; an analysis where the stability of the analysis does "…not depend on the number of aggregate cases but on the number of independent individuals who contributed to each case".

The components used in the analysis are the retention questionnaire responses plus other student-specific information. Exploratory factor analysis identifies those components that have "commonalities" (Kim & Mueller, 1978) by using the covariance between the components. Components with higher covariance are grouped into a number of factors, with the number being determined by the groupings that arise. Using a complexity thinking perspective, these factors were interpreted as a self-organized pattern of different nested systems, in our case of the complex system of student retention in higher education.

Exploratory factor analysis will normally reveal that some of the components are present in more than one of the factors. From a complexity thinking perspective, this was interpreted as evidence of neighbourinteractions between the nested systems through their shared components.

3.3. Multidimensional scaling

As denoted by the conceptual framework, components of a complex system interact locally (Davis &Sumara, 2006) and thus components that have a high relative closeness to other components in the multidimensional scaling analysis can be regarded as being connected[2] and within each other's "zone of influence". In the multidimensional

―――――

[2]   "Connected" is used as a broad term that encompasses the interaction, communication, and dependence between the different components of the system.

scaling analysis of the questionnaire data, the answers and their proximities are used to create a representation of the emergent network structure of the complex system. The components may be seen as vertices connected by edges, which form a basis for visualization and allow for measurements of component interaction through the use of network theory.

A good way to determine the relative proximity of components to one another is to use multidimensional scaling because it offers a way to calculate the distances between points of data in multidimensional space. The relative closeness ("multidimensional proximity") of components to one another is the "likeness" or "similarities" (Schiffman et al., 1981) of those components.

3.4. Network theory

The orienting emphasis in network theory is "structural relations" (Knocke& Yang, 2008). From such a framing the essential elements of a network are the *nodes* (vertices) and the links (connections) between nodes. In the current study, nodes are components examined in the retention questionnaire and data collected from students. Network theory is thus a powerful analytic tool to explore and illustrate structure connectivity that was produced using multidimensional scaling.

3.4.1. Network theory concepts

The nodes represent the components of a network (i.e. items on the retention questionnaire and student data), and the edges represent the relationships between the nodes. When two nodes are directly connected the two nodes are adjacent. A *path* is a way through a sequence of nodes that begins with a starting node, follows adjacent nodes through the network and ends at an end. When every node in the network is reachable (i.e., a path exists between every node) the network is connected. If there are many paths between two nodes, the *shortest path* between them is the one with the fewest connections made through other nodes (Freeman, 1978). Visualization and analysis of networks, and therefore complex systems, is made possible by using these constructs of network theory.

3.4.2. Network measurements and interpretation

The network to represent the system of student retention was formed using multidimensional scaling. We assumed that we had an undirected network where the connections between the nodes did not have a specific direction of influence. Analysis of the created network was done by using Statnet(Handcock et al., 2003), a free package designed for analysing networks which works with the "R" statistical computing and graphics program. Identification of "important" nodes was done by calculating each node's *centrality*.

In this study we distinguish between *closeness centrality* and *betweenness centrality* (Bernardsson 2009). *Closeness centrality* is an ordinal measure of how "close" every other node is, and it is calculated through finding the shortest path between nodes. Information can be spread to the whole network more effectively via nodes with high *closeness centrality*(Freeman, 1978). *Betweenness centrality* is the frequency that one particular node is a part of the shortest path between every other node. Nodes that are more frequently a part of the shortest path between nodes may be interpreted as having a high degree of *"control* of communication" (Freeman, 1978, p. 224) in the network.

4. Method

Data was collected from two sources: firstly student records were used to obtain student demographic information such as age, gender, higher education credits achieved within and outside the programme, and student retention, and secondly, a questionnaire with 29 questions was developed to explore influences on student retention (see Appendix). The questionnaire was largely based on the work of Cabrera et al. (1992a, 1992b and 1993). Students answering the questionnaire were asked to mark their level of agreement (or disagreement) with 29

statements on a five-point Likert scale. Each separate piece of information (record data and question responses) was an item in the analyses, giving 34 items altogether.

As a preliminary study to verify our approach, the questionnaire was administered to 51 students (39 of whom were in two Engineering Programmes, and the remaining 12 in a Physics Programme) participating in a second semester Physics course at a traditional Swedish university. Re-enrolment in the second year (third semester) was used as a measurement of student retention and was found to be 82.4%.

5. Results

Firstly, exploratory factor analysis was used to identify subsidiary complex systems within the broader complex system, and to demonstrate their nested structure. Secondly, multidimensional scaling was used to show the connectedness of the components and to visualize how the components of the complex system interact with one another.

Exploratory factor analysis showed the subsidiary systems that are nested within the larger complex system of student retention. Four sub-systems were identified, with some items that overlapped between them, which demonstrated the nestedness of the sub-systems. Other items dropped out altogether.

This pilot study was too small to draw any conclusions about the nature of these four sub-systems, which appear to correspond to different levels of analysis of the whole complex system such as the level of the individual student, the level of groups of students, the level of the institution, etc. It is tempting to try to match these sub-systems with the systems identified earlier, such as the internal social and academic systems and the external system proposed by Tinto (1997), but we believe that this would tend to limit what could emerge from an analysis such as this one.

5.1. Exploratory Factor Analysis

Having satisfied ourselves that we had data items grounded in the literature we started with an exploratory factor analysis. This was to identify the complex systems that make up the greater system of student retention through the identification of the factors within the overall system. Our analytical tool was the Statistical Package for the Social Sciences, SPSS (Predictive Analytic SoftWare, PASW, version 18.0). Our starting point was the normalized matrix of the questionnaire data together with the students' higher education credits achieved within and outside their programme, retention (re-enrolment in the second year), age, and gender.

The following three measures were used together to achieve an appropriate correlation matrix of items to be used for exploratory factor analysis (Dziuban&Shirkey, 1974):

1. Kaiser-Meyer-Olkin's (KMO) measure of sampling adequacy. Items were removed recursively from the data until a value of 0.68 was obtained, close to the guideline 0.7 recommended by Kaiser & Rice (1974).

2. Bartlett's (1950) test of sphericity. This had a significance of less than 0.001 when guideline 1 had been achieved.

3. The anti-image correlation measure of sampling adequacy (MSA). Items with an MSA less than 0.5 were removed (Kaiser, 1970).

As a result, eleven items (of a total of 34) were removed and are listed in Table 1. These items were interpreted as having little effect on the system of student retention, at least as far as this illustrative study is concerned, given the limited data set and the high level of retention from first year to second year.

In deciding the number of factors in the model we used a scree test (see scree plot in Figure 2). "The scree test involves examining the graph of the eigenvalues … and looking for the natural bend or break point in the data where the curve flattens out. The number of data points *above* the 'break' ... is usually the number of factors to retain" (Costello, 2005, p. 3). Every item is treated as a vector that has an eigenvalue of 1.0 before the iterative rotations and calculation of vector projection on an axis. An eigenvalue of 7 (One Factor) provides us with the information that all significant loadings in One Factor can be grouped,providing us with 7 times as much information as a single variable and also that the items in the factor share traits.

Table 1. Items removed from the Exploratory Factor Analysis

Question	Item
-	Age
-	Gender
-	Credits passed that are not a part of the programme of study
6	My possibility to continue with my studies is dependent on me working while I study.
9	It is important for me to graduate at my University.
13	I have achieved the study-results I expected during the first year.
18	It is important for me to get a university degree.
20	I have developed a good relationship with my teachers in the courses I have studied.
26	First year physics courses have been inspiring.
27	University physics courses are much different from my previous physics. Courses
28	First year physics courses have had a clear connection to everyday life.

This led us to choose a Four Factor solution for the model (Hofstede, 2001). The cut-off at Four Factors, and not Five (although they have nearly the same eigenvalue) was guided by seeing that a Five Factor solution provided one factor with only two variables that had significant loading (more than 0.32) which according to the analysis-method is not appropriate for such a factor solution. The "extra" factor wouldn't give much more information than adding one or two other variables to the analysis or the questionnaire.

Significant item loadings for each factor were identified by using a minimum loading of 0.32 on each item, which corresponds to a 10% shared variance between items (Tabachnick&Fidell, 2001). Question 12 was retained at a loading of 0.313 (which is very close to 0.32).
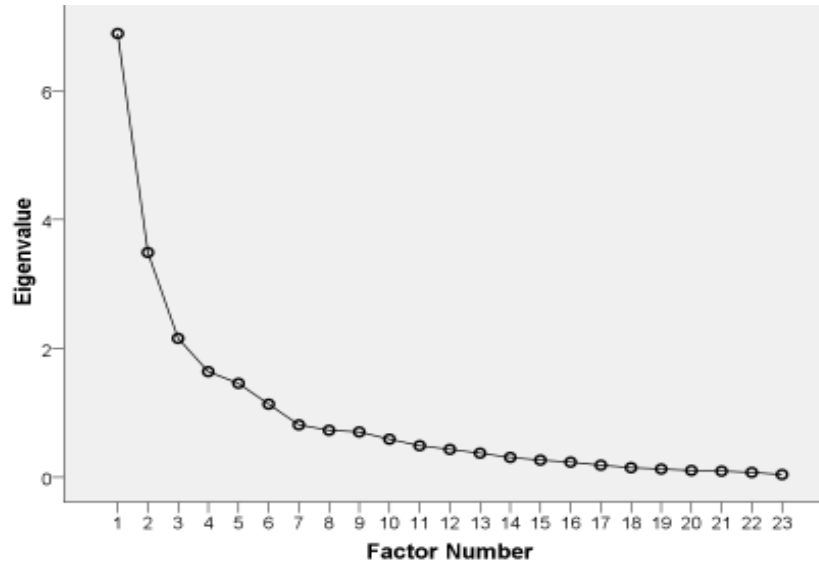

Figure 2. Scree plot showing eigenvalues after rotation

The results of the exploratory factor analysis are shown in Table 2. Note that these results differ from the normal result in exploratory factor analysis where unique variables are sought for each factor. It is tempting to try to characterize the four factors in terms of the systems identified by others (such as university academic systems and social systems (Tinto, 1975) and support systems (Bean, 1980)) but this cannot be done because of the small sample size. What Table 2 does show is that there is overlap of items between the four factors, each of which is a complex system in itself. This illustrates both the complexity and the nestedness of the system of student retention as a whole. It also highlights the existence of neighbour interactions between the four nested systems, as well as the fact that they have fuzzy boundaries.

We can only provide a very tentative characterization of these four factors, due to the small sample used in the analysis. The items with the highest loadings in Factor 1 have to do with the status of the programme the students are studying. Factor 2 seems to be characterized by a sense of belonging. It is hard to find any strong theme in Factor 3 which has particularly fuzzy boundaries with Factor 2. Financial issues clearly dominate Factor 4. One might therefore identify Factor 2 as pertaining to the individual student, Factor 1 with the institution and Factor 4 with the external system, but note the comments in this regard at the start of the Results section.

Table 2. Loading from the exploratory factor analysis giving the factors sorted adjacently by the number of shared items.

Item	Factor 1	Factor 2	Factor 3	Factor 4

H.E. credits within programme (HECwP)			0.542	
Retention			0.934	
Q1. Best university programme	0.788			
Q2. Family approval		0.472		
Q3. Satisfied with finances				0.836
Q4. Finances - focus on studies				0.833
Q5. Finances - teacher demands				0.796
Q7. Satisfied with curriculum	0.328	0.458		
Q8. Close friends encouragement		0.580		
Q10. I belong at my university		0.637	0.447	
Q11. Future employment	0.464	0.390		
Q12. My close friends rate this institution as high quality			0.313	
Q14. Satisfied with experience of higher education		0.687	0.411	
Q15. Easy to make new friends.		0.842		
Q16. Right choice - university		0.683	0.399	
Q17. Right choice - programme	0.758			
Q19. It is important to get a degree from this programme	0.708			
Q21. Initiation weeks		0.855		
Q22. First year courses fit together				0.459
Q23. Previous knowledge		0.385		
Q24. Clear educational trajectory		0.447	0.396	
Q25. Faculty support		0.345	0.322	0.461
Q29. I intend to re-enroll	0.835			

Note: Light grey shading denotes the items that have a loading above 0.32 in more than one factor.

5.2. Multidimensional scaling

Multidimensional scaling was used to visualize the network of items that influence student retention (using the same 34 items that were used for the Exploratory Factor Analysis). Network theory data analysis tools and complexity thinking were used to interpret the results.

5.2.1. Network Creation

The multidimensional scaling analysis was used on the data to determine the distances between items arising from this data. A dimensional solution ranging from two to four dimensions was explored, because multidimensional scaling usually provides a solution that has fewer dimensions than exploratory factor analysis on the same data (Schiffman et al., 1981). The multidimensional scaling analysis converged to a solution after 15 iterations. Figure 3 shows the resulting visualization of the network.
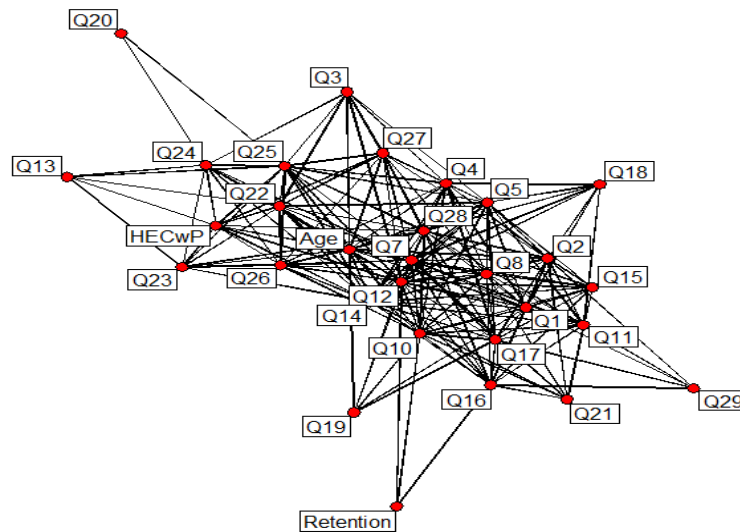

Figure 3.Network visualization from Multidimensional Scaling analysis. Note that this visualisation shows just the connections, and not the actual proximities.

Note how in Figure 3 the items that are less connected (and hence less important) are on the periphery, whereas the more central items clearly lie on the paths between many other items and are thus more significant to the operation of the network.

We used the multidimensional proximities between the items to identify items with relative closeness or proximity. Using the neighbour-interactions concept and an understanding of the structure of decentralized networks (Davis &Sumara, 2006) from complexity thinking, we recursively lowered the cut-off for proximities and network visualizations were produced. The statistical computing and graphics "R" program, together with Statnet package (Handcock et al., 2003), were used for visualization and measurements. Iterations were run as long as the network continued to resemble a decentralized network, but were ended before the network broke down and ceased to be connected (Freeman, 1978).

Two items were considered to be within each others' "zone of influence" when their proximity was below 0.25. The analysis was complete when the majority of the items had proximities less than 0.25. To retain the connectedness of the system Retention needed to have a higher cut-off of 0.5. HEPoP, Gender Q6 (studies dependent on working) and Q9 (importance of achieving a degree from this university) all dropped out at this cut-off level. These items are four of the eleven items that were dropped from the Exploratory Factor Analysis.

Note that as the cut-off level is lowered further the system becomes less and less connected. At a cut-off proximity of 0.1 less than half the items remain connected to one another.

Multidimensional scaling was used to calculate the proximities between items pertaining to student retention and these were visualized in a 2-D network. The network was found to be decentralized in structure. Three items that

were clearly outliers from multidimensional scaling were also items that dropped out of the exploratory factor analysis. However, not all items that were dropped in the exploratory factor analysis were loosely connected in the network. Moreover, three particularly influential items (nodes) were identified. These three items were each present in two of four factors in the exploratory factor analysis.

5.2.2. Influential items

In this study we distinguish between *closeness centrality* and *betweenness centrality* (Bernhardsson 2009). Closeness centrality is an ordinal measure of how 'close' every other node is, and it is calculated through finding the shortest path between nodes. Information can be spread to the whole network more effectively from nodes with high closeness centrality (Freeman 1978). Betweenness centrality is the frequency that one particular node is a part of the shortest path between every other node. Nodes that are more frequently a part of the shortest path between nodes may be interpreted as having a high degree of *'control* of communication' (Freeman 1978, p. 224) in the network.

Consideration of Figure 4 shows that there are seven items with relatively high *betweenness centrality* as well as relatively high *closeness centrality*: Q12 (friends' opinion of institutional quality), Q7 (satisfaction with one's course curriculum), Q25 (faculty support), Age of the students, Q14 (students' satisfaction of being at the university), Q10 (the feeling of belonging at the university), and Q28 (physics is connected to everyday life). Item Q25 (faculty support) is interesting in that it seems to lie outside the broad band of points showing higher Betweenness centrality vs higher Closeness centrality: It has a much higher Betweenness centrality than the rest of the items in the band. The same is also true for Item Q24 (clear educational trajectory). Connections between exploratory factor analysis results and multidimensional scaling results will be discussed below.
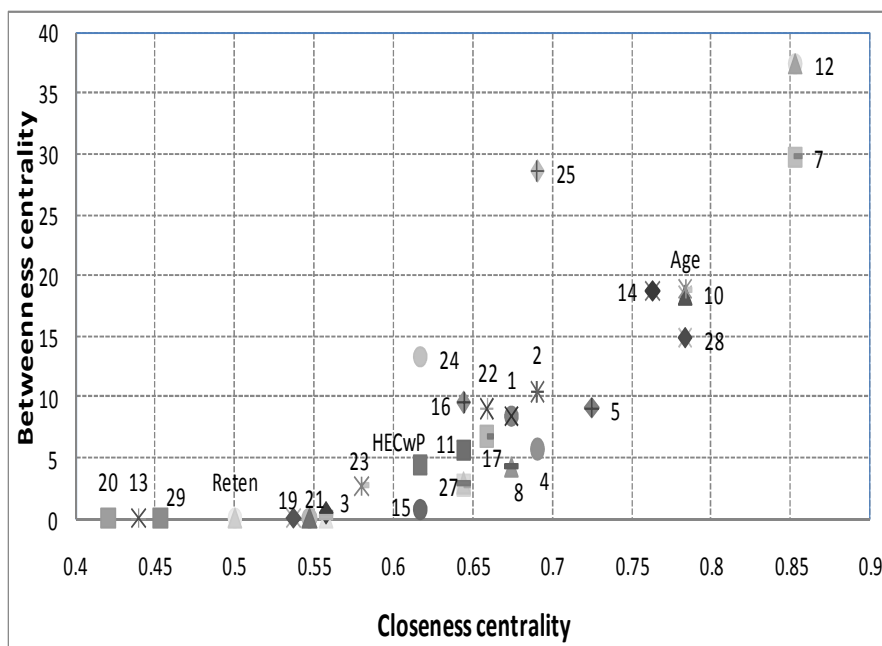

Figure 4.*Closeness centrality* and*Betweenness centrality* scatter plot of the network created by the Multidimensional Scaling analysis proximities of items.

13

6. Discussion

This pilot study was too small to draw any conclusions about the nature of the four sub-systems identified through exploratory factor analysis. It is tempting to try to match these sub-systems with systems previously identified, such as the internal university academic, social, and support systems and the external system (Bean, 1980; Tinto, 1975 and 1987), but we believe that this would tend to limit what could emerge from an analysis such as this one.

From multidimensional scaling and the visualisation of the network shown in Figure 3, it is clear that this is a decentralized network. The three most influential components were each present in two of four factors in the exploratory factor analysis. The four components that dropped out of the multidimensional scaling and three of the outliers were among the eleven components that were dropped from the exploratory factor analysis. Thus both sets of analyses produce congruent results.

7. Conclusions

What is new in our example is that we were able to identify certain items as influencing *the complex system as a whole*. This means that they should not be seen as direct linear influences, but as influences *mainly through other items*. This implies that things are more interconnected than previously acknowledged by the existing models of student retention. In this way, these items, their emergent patterns, and their interactions, combine to form a dynamic model of student retention with nested sub-systems.

While some previous researchers have acknowledged the complex nature of interactions of elements relevant to student retention, our analytic example shows how the *structure* and *dynamics* of the complex systems that influence retention can be brought to the fore empirically. We would suggest that it is unlikely that either faculty or students are aware of the extent of the complexity that underpins how student retentionemerges as a result of the complex interaction between the components of the nested systems. In our modeling, the higher education experience is shown to have its greatest influence manifested through the complex dynamics of these different nested systems. This is in stark contrast to linear thinking about the experience of higher education and it can provide insight into how to better manage it to improve student retention.

For this work to be more meaningful it will clearly need to be extended to a much larger data sample in which retention is determined over a longer time frame. It will also need to be triangulated with more in-depth qualitative studies. This preliminary study has shown the potential of this approach to modeling student retention to uncover a deeper insight into student retention in programmes such as engineering and thus inform institutional actions aimed at improving retention.

Acknowledgements

References

Bartlett, M.S. (1950). Tests of significance in factor analysis.*British Journal of Psychology3* 77 – 85.
Bean, J.P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education12*(2) 155 – 187.
Bean, J.P. (1982). Student attrition, intentions, and confidence: Interaction effects in a path model. *Research in Higher Education17*(4) 291 – 320.

Bean, J.P. (2005). Nine themes of college student retention. In Seidman, A., (Ed.) (2005). *College student retention: Formula for student success.* pp. 215 – 244. Westport: Praeger Publishers.

Bernhardsson, S. (2009). *Structures in complex systems: Playing dice with networks and books.* PhD diss., Umeå University.

Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste.* Cambridge: Harvard University Press.

Braxton, J.M. (2000). *Reworking the student departure puzzle.* Nashville: Vanderbilt University Press.

Braxton, J.M. & Hirshey,A.S. (2004). Reconceptualizing antecedents of social interactions in student departure.In Yorke,M. & Longden,B. (Eds.) (2004). *Retention student success in higher education* (p. 89-102). Maidenhead: England: Open University Press.

Braxton, J.M. & Lien,A. L. (2000). The viability of academic integration as a central construct in Tinto's interactionalist theory of college student departure. In J. M Braxton (Ed.), *Reworking the student departure puzzle* (p. 11-28). Nashville: Vanderbilt University Press.

Brown, S.L. & Eisenhardt,K.M. (1997). The art of continuous change: Linking complexity theory and time-paced evolution in relentlessly shifting organizations.*Administrative Science Quarterly42*(1) 1 – 34.

Cabrera, A.F., Castañeda, M.B., Nora, A. & Hengstler,D. (1992a). The convergence between two theories of college persistence. *Journal of Higher Education63*(2) 143 – 164.

Cabrera, A.F., Nora, A. & Castañeda,M.B. (1992b). The role of finances in the persistence process: A structural model. *Research in Higher Education33*(5) 571 – 593.

Cabrera, A.F., Nora, A. & Castañeda,M.B. (1993). College persistence: Structural equations modeling test of an integrated model of student retention. *Journal of Higher Education64*(2) 123 – 39.

Committee on Science, Engineering, and Public Policy. (2007). *Rising above the gathering storm: energizing and employing America for a brighter economic future.* Washington, D.C.: The National Academies Press.

Costello, A.B. (2005).Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation10*(7) 1 – 9.

Davis, B. & Simmt,E. (2006). Mathematics-for-teaching: An ongoing investigation of the mathematics that teachers (need to) know. *Educational Studies in Mathematics61* 293 – 319.

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

Dziuban, C.D. & Shirkey,E.C. (1974). When is a correlation matrix appropriate for factor analysis? Some decision rules.*Psychological Bulletin81*(6) 358 – 361.

Eaton, S. & Bean, J. (1995). An approach/avoidance behavioral model of college student attrition.*Research in Higher Education36*(6) 617 – 645.

Freeman, L.C. (1978). Centrality in social networks conceptual clarification.*Social Networks1*(3) 215 – 239.

Gilstrap, D.L. (2005). Strange attractors and human interaction: Leading complex organizations through the use of metaphors.*Complicity: An International Journal of Complexity & Education2*(1) 55-69.

Handcock, M.S., Hunter, D.R., Butts, C.T., Goodreau, S.M. & Morris,M. (2003). *Software tools for the statistical modeling of network data. Version 2.1-1.* Project home page at http://statnet.org, URL http://CRAN.R-project.org/package=statnet.

Hofstede, G.B. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Hofstede, G.,B. Neuijen, D.D. Ohayv, & G. Sanders. (1990). Measuring organizational cultures: A qualitative and quantitative study across twenty cases. *Administrative Science Quarterly35*(2) 286 – 316.

Kaiser, H.F. (1970). A second generation Little jiffy. *Pscyhometrika 35* 401 – 416.

Kaiser, H.F., & J. Rice. 1974. Little jiffy, Mark IV. *Journal of Educational and Psychological Measurement34*(1) 111 – 117.

Kim, J.-O., & C. W. Mueller. (1978). *Factor analysis: Statistical methods and practical issues.* London: Sage Publications.

Knocke, D., & So.Yang. (2008). *Social network analysis quantitative applications in the social sciences.* (2nd ed.). California: Sage Publications.

Metz, G.W. (2004). Challenge and changes to Tinto's persistence theory: A historical review. *Journal of College Student Retention Research Theory and Practice 6*(2) 191 – 207.

Mitchell, M. (2009). *Complexity: A guided tour.* New York: Oxford University Press.

Mowat, E., & B. Davis. (2010). Interpreting embodied mathematics using network theory: Implications for mathematics education. *Complicity: An International Journal of Complexity and Education 7*(1) 1 – 31.

Organisation for Economic Cooperation and Development (OECD). (2009). *Education at a glance 2009: OECD Indicators.* Paris: OECD Publishing.

Pascarella, E.T., & P.T. Terenzini. (1979). Interaction effects in Spady's and Tinto's conceptual models of college dropout. *Sociology of Education 52* 197-210.

Schiffman, S.S., L.M. Reynolds, & F.W. Young. (1981). *Introduction to multidimensional scaling theory, methods, and applications.* Orlando: Academic Press, Inc.

Spady, W. 1970. Dropouts from higher education: An interdisciplinary review and synthesis. *Interchange1*(1) 64 – 85.

Spady, W. (1971). Dropouts from higher education: Toward an empirical model.*Interchange2*(3) 38 – 62.

Statistics Sweden & National Agency for Higher Education. (2001). Higher education. Undergraduate education. Throughput. Results up to 1999/2000 inclusive. Series: Utbildning och forskning [education and research] UF20 SM0102. SCB: Svante Öberg. Retrieved from: http://www.scb.se/statistik/UF/UF0205/2000I02/ UF20SM0102.pdf

Statistics Sweden & National Agency for Higher Education Higher Education. (2009). Students and graduated students at first and second cycle studies 2007/08. Series: Utbildning och forskning [education and research] UF20 SM0902. SCB: Mats Vadman. Retrieved from: http://www.hsv.se/download/18.1db453f212664e532507ffe302/UF20SM0902.pdf (1 Dec, 2010)

Summerskill, J. (1962). Dropouts from college. In *The American college: a psychological and social interpretation of higher learning,* ed. N. Sanford, 627-657. New York: Wiley.

Tabachnick, B.G., & L.S. Fidell. (2001). *Using multivariate statistics.* Boston: Allyn and Bacon.

Tinto, V. 1975. Dropout from higher education: A theoretical synthesis of recent research.*Review of Educational Research45*(1) 89 – 125.

Tinto, V. (1987). *Leaving college: Rethinking the causes and cures of student attrition.* Chicago: University of Chicago Press.

Tinto, V. (1997). Classrooms as communities: Exploring the educational character of student persistence.*Journal of Higher Education68*(6) 599 – 623.

Tinto, V. (2010). From theory to action: Exploring the institutional conditions for student retention. *Higher Education: Handbook of Theory and Research25* 51 – 89.

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

Yorke, M., & B. Longden. (2004). *Retention and student success in higher education.* Maidenhead: Open University Press.

Appendix: Questionnaire

Q1. I am studying one of the best programmes at the university.

Q2. My family approves of my attending my University.

Q3. I am satisfied with my financial situation.

Q4. My financial situation allows me to focus on my studies as much as I want.

Q5. My financial situation allows me to focus on my studies as much as the teachers demand.

Q6.My possibility to continue with my studies is dependent on me working while I study.

Q7. I am satisfied with my course curriculum.

Q8. My close friends encourage me to continue attending my University.

Q9. It is very important for me to graduate at my University.

Q10. I feel I belong at my University.

Q11. My degree at this university will help me secure future employment.

Q12. My close friends rate this university as a high quality institution.

Q13. I have achieved the study-results I expected during the first year.

Q14. I am satisfied with my experience of H.E.

Q15. It has been easy for me to meet and make friends with other students at this university.

Q16. I am confident I made the right decision in choosing to attend at my university.

Q17. I was right when choosing to study this programme.

Q18. It is important for me to get a university degree.

Q19. It is important for me to get a degree from this particular programme.

Q20. I have developed a good relationship with my teachers in the courses I have studied.

Q21. The initiation weeks were a good start for my program studies.

Q22. It is clear to me how the courses during the first year fit together.

Q23. The teaching has corresponded well with my previous knowledge.

Q24. My educational trajectory is clear for me.

Q25. Faculty staff have provided me with the support I needed to succeed in my studies.

Q26. First year physics courses have been inspiring.

Q27. University physics courses are much different from my previous physics courses.

Q28. First year physics courses have had a clear connection to everyday life.

Q29. I will re-enrol at this programme of study next autumn.