

Engagement and undergraduate retention: social network analysis and student social ecologies

Geoff Woolcott
Southern Cross University, Australia

Robyn Keast
Southern Cross University, Australia

Daniel Chamberlain
La Trobe University, Australia

Ben Farr-Wharton
University of Technology Sydney, Australia

ABSTRACT

Undergraduate student retention is known to be complex and subject to multiple factors acting together on individual students in different contexts. This study examines an innovative multifactorial model that provides a student-centered and place-based identification of the duality of risk/opportunity in undergraduate education. The model, developed using a novel alignment of social network theory and ecological systems theory, is here applied to analysis of a large archive of behavioral data sets related to 4065 undergraduate students at a regional university in Australia. The model provided illustrations of how previously identified risk factors are connected within social ecologies, both for individuals and groups. The analysis also provided additional risk factors, calculated from the social ecology data sets. All risk factors were examined against the background of the student social ecology network with significant results: the identified risk factors can be seen to be linked in some individuals and sub-cohorts; the newly-identified risk factors can also be linked; and, all risk factors can be connected to other proximal factors. Implications for intervention and support are discussed.

Keywords: engagement, retention, undergraduate, social network analysis, social ecologies

Copyright statement: Authors retain the copyright to the manuscripts published in AABRI journals. Please see the AABRI Copyright Policy at <http://www.aabri.com/copyright.html>

INTRODUCTION

The broad diversity of the university student population, either arriving from secondary schools or increasingly through less direct and alternative pathways, has given rise to an equally diverse range of challenges related to retention and attrition (Hayden, 2010; Maher & Macallister, 2013; Price Waterhouse Coopers, 2007; Museus & Jayakumar, 2012; Nelson, Clarke, Kift & Creagh, 2011; Quaye & Harper, 2014; Gilardi & Guglielmetti, 2011). There is a considerable body of research supporting the view that attrition is disadvantageous to both students and universities, as well as to society at large, and retention and course completion arguably advantageous.

In response to these challenges, universities generally design intervention strategies and programs for retention on models derived from analysis of sets of factors considered collectively across university environments. Such design has had mixed results, with some successes but also some poorly targeted, ineffective and inefficient interventions and programs (Clarke Nelson & Stoodley, 2011; Eckles & Stradley, 2012; Nichols, 2010; Thomas, 2011).

There is a considerable body of research, however, supporting the view that multiple and varied factors affect the duality of retention/attrition on university engagement over multiple timeframes (Braxton, Hirschy & McClendon 2011; Coates & McCormick, 2014; Dumbrigue, Moxley & Najor-Durack, 2013; Kezar, 2014). This has not, unfortunately, resulted in the development of models with which to examine engagement or risk factors, much less development of any longitudinal predictive multifactorial models, such that individual student needs are examined in particular contexts with a view to retention support and associated structures.

This study examines multiple factors identified as related to undergraduate retention in data sets obtained from a commencing student cohort at a university in regional Australia. The study examines the potential of an interrelationship model, derived from social network theory and social ecology theory, that explores the connections between the multiple factors related to student connections to their community—the student community ecology. The model applies newly developed conjunctions between theoretical research on social network analysis (SNA) (Borgatti, Everett & Johnson, 2013; Newman, 2010) and developmental social ecologies (DSEs) (Bronfenbrenner, 1979; Bronfenbrenner & Morris, 2006). The interrelationship model developed here is centered on the social ecology network of the individual, such as seen in statistical (Lipina, Insúa & Echeverría, 2015) and theoretical exploration using social network analytic methods (Neal & Neal, 2013). The SNA model, however, also has application to groups of individuals, based on selection of connected group features.

Studies of undergraduate retention based in SNA are already appearing in the literature (Dawson, 2008, 2010; Rizzuto, LeDoux & Hatala, 2009; Thomas, 2000), but, a unique feature of this study is the source of the social network data and the size of the group analyzed—archived behavioral data sets from an entire year of an undergraduate cohort, a total of 4065 students. To date, other studies have used SNA of selected archival data or they have used smaller student cohorts, for example, archived data from a student information system for 460 students at a small college (Eckles & Stradley, 2012).

BACKGROUND

This study, therefore, builds on the acknowledgement that a key issue in undergraduate engagement and retention resides in the knowledge that academic risk is actually very complex, but that interventions to improve retention are based on single factors

or markers, sometimes considered in groups (Coates, 2009; Radloff & Coates, 2013; Nelson et al., 2011). Use of such factor analysis presents a number of issues, not the least of which is that data are slowly collected and collated, in Australia producing a two- or three-year lag in national data production. Additionally, it is sometimes assumed that factors can be applied to any university or any student. The most significant issue, however, is that these factors are known to interact with each other and with other factors in students' lives.

The response to retention/attrition taken in this study addresses the need to acknowledge the complexity of student interactions and adopt methodologies and models that are multifactorial and dynamic, where each factor informs and influences others. As well the study examines a large number of the factors that inform and influence student life at a particular university, using methodologies and models that take into account particular attributes of personality and place. Current approaches to student retention and attrition do not effectively address the needs of individual students at their place of study and do not sufficiently acknowledge the agency of students in their choice of engagement with university life.

We look here primarily at engagement as readiness for interaction, being 'in gear' and ready to interact, with an emotional connection that influences attitudes, behaviour and interactions directed towards learning (Chapman, 2003). Coates and colleagues (e.g., Coates, 2009; Coates & McCormick, 2014), in conducting the Australasian Survey of Student Engagement, use categories such as academic challenge, active learning, student and staff interactions, enriching educational experiences, supportive learning environments and work integrated learning, within which to examine engagement of first-year university students.

The model developed here is underpinned by SNA, where a system is reduced to a set of actors (or actants) called 'nodes' and a set of relationships called 'edges' that link the nodes together (Newman, 2010). Network analysis takes system elements and social network structures and their relationships as the fundamental unit of analysis, rather than individuals (Wasserman & Faust, 1994). The exploration of networks, and the connectivity of nodes within them (using empirical data), has developed rapidly in recent years, largely because the rules governing the relationships within such networks remain independent of the nature of the subjects being linked (Hanneman & Riddle, 2005) and because of rapid advances in software for analyzing large data sets (Borgatti, 2012).

Studies such as Eckles and Stradley (2012) argue that social networks should matter in studies of student retention. This view is based on sociological and higher education theories of such luminaries as Tinto (1975, 1997) and Bean (Bean 2005; Bean & Eaton, 2002), and their exploration of student connections and networks, including social interactions, in relation to social integration at university. Eckles and Stradley (2012), in fact, examined social integration in terms of relationships between first and second year retention and social network variables in a cohort of first-year students in order to determine how friends' retention and attrition behaviors impacted on retention.

The current study takes a broader approach, derived from recent adaptations of social ecology theory (Bronfenbrenner & Morris, 2006; Lipina et al., 2013) and in so doing suggests that social integration may sit as a complex component within social interaction. This study, therefore, targets the interactions or behaviors of students more generally, with a focus on proximal interactions (interactions that are directly connected to the student) as recorded in archival data sets. In order to do this, the study adapts a recent theoretical study that applied SNA in combining conceptualizations of social circles from Simmel (1955[1922]) with Bronfenbrenner's social ecologies (1979) to produce nested networks. This nested network adaptation of DSEs allows a structural picture to be developed of the interrelationships of individuals and elements/factors. In this network adaptation, the micro, meso and other systems of DSEs are redefined based on social interactions and student behaviors, and the

authors argue that the DSEs may be more useful if thought of in terms of overlapping configurations of interconnected ecological systems.

The current study takes the extra step of using both SNA and social ecologies to develop a model that can be applied to studies of academic risk, based on examination of proximal factors in a social ecology. This strategy was used partly to avoid a computer-intensive algorithm approach (Lipina et al., 2015), which requires considerable statistical analysis as well as dedicated computer programming. Instead, this study has adapted SNA methodologies used in other studies by Author and colleagues (Author). As a result, the nested network adaptation does not strictly adhere to the psychology-based construction of Neal and Neal (2013), but rather conforms to standard SNA methodologies in relating factors in a 1-modal analysis, where like is related to like on a one-to-one basis. For example, one proximal factor used here is student age, which can be related on a one-to-one basis to the proximal factor of student current socioeconomic status.

Social ecology networks and consensus interrelationship maps

The study introduces two novel conceptualizations related to SNA: 'social ecology networks' (SENs); and, related consensus interrelationship maps (CIMs). The consensus interrelationship maps (CIMs), effectively show weighted SENs that represent large numbers of students, hence the term 'consensus'. These conceptualizations are examined for their use in identifying groups of multiple factors that may contribute to academic risk in undergraduate university education. The SENs and CIMs, along with associated risk factor networks, are examined also for their use in enabling comparison of students with equivalent profiles, that is, where one student is at risk and another was not.

The SENs/CIMs were used to examine the relationship of previously identified risk factors, for either selected groups or for individuals, with other proximal factors. In particular, interrelated risk factors, or risk factor networks, were examined in a context of retention and attrition. Attention was given to the relatedness of risk factor networks to withdrawal and failure grades, or to course continuity. An additional dimension of this study, however, was the identification of novel risk factors from the archived data, using analysis of relative risk (Osborne, 2006; Szulimas, 2010). These risk factors were compared with those previously identified.

The CIMs arise from consideration of all of the factors that are common to groups of students, allowing a structural picture to be developed of the interrelationships of individuals and factors. Both the SENs and CIMs are, in fact, nested networks that can be used to look at the intersection of student experiences and events/locations of interactions as part of the social ecology mapping and risk assessment. Another advantage of SNA in constructing SENs/CIMs is the ability to link together the various layers of the ecology and focus this to the student level—whereas previously the executive/administrative level and the student level remained separate and distant elements.

Validation of the SNA approach, in co-operation with university equity practitioners, was achieved using a sample of students from equity groups considered to be at risk of failure or withdrawal. This included Aboriginal and Torres Strait Islander (ATSI) students, low socioeconomic status (low-SES) students, students with disability and students from regional and remote locations. Although a four-system adaptation of Bronfenbrenner's (1994) social ecology mapping was utilized (Table 1, Appendix), this study focused on the proximal factors or micro system elements, although the analysis also contained some meso, exo and macro system elements.

Research questions

The study was developed to answer the following research questions, related primarily to student interrelationships.

Question 1

How can SNA be used to construct SENs and related CIMs to examine academic risk against the background of a student social ecology?

Question 2

How are proximal factors or elements of university students related to established risk factors for retention, considered either singly or in combination?

Question 3

How are risk factors for retention related to each other, based on comparison across a student social ecology? In other words, are there a risk factor networks for a given ecology?

METHOD

Data collection

The archived student data sources comprised a de-identified, interdisciplinary sample comprising 4065 students, the undergraduate population of commencing domestic students in a single calendar year at a regional university. Commencing students in a calendar year were considered to be those who had been offered a university placement in that calendar year and who had enrolled at the university for the first time in at least one subject in that year, but did not exclude students who had transferred from one course to another. The cohort included both mature-aged students and students who left high school at the end of the previous calendar year and comprises largely first-year students. A summary of the diversity of the cohort is given in Table 2 (appendix).

The student-centered sampling of behavioral factors was conducted across existing archived data sets for these largely first-year domestic undergraduates. There are other data sets that could potentially be accessed, but these were not available in the time frame of the study. Archives (generally data bases) from which data sets were obtained included: Academic Skills appointments, Advocacy case records, Learning Management System – Blackboard, Data Warehouses, Management Information System, Student Management System, Indigenous Australian Student Services, events participation, Library Services patron usage, Equity and Diversity Databases, Mentoring program participation, careers service appointments, off-campus accommodation service usage, student loan service usage and cultural and sporting participation lists.

Data analysis

SNA was employed to examine the complex relationships that appeared to exist between proximal factors and students.

Data matrix construction

In order to construct a suitable network from a data matrix, nodes were considered as behavioral or demographic factors (including their categorizations) and edges as connections of relationships between factors for each individual student, for example, an individual connected mid socioeconomic status to a node of age 22–26. By considering these nodes and edges the data sets were used as the basis for constructing a data matrix, subsequently used to

determine SENs and related CIMs. The SENs and CIMs were, therefore, based on sociocultural factors in the environment of the sampled students.

As a first step in the matrix construction the data were categorized using a list of factors (elements) based primarily around the proximal level (micro level) of a social ecology (Table 3, Appendix).

These proximal factors/elements were then elaborated into the data matrix from the collected archival data. Data sets for each factor category were sub-categorized, for example 'student age' was subdivided into sub-categories, such as 'age 22-26'. The matrix was constructed as an adjacency matrix, with individual students as rows and factor subcategories as columns. The data matrix, therefore, was extremely large, with over 4000 rows and about 1500 columns after categorization and data cleaning was completed. The matrix was coded as presence=1 and absence=0 for each student intersection with a subcategory. Design of the initial data matrix allowed data to be added to the matrix as the project progressed. Although this process was labor intensive, the data matrix proved suitable for analysis and for generation of nested SENs, CIMs and associated risk factor networks as part of the SNA.

Social network analysis of the matrix

The network analysis software UCINET (Borgatti, 2012) was used to quantitatively analyze the complex layering of the ecologies and to produce nested network maps that represented the various SENs of interest, as well as the CIMs. These maps provided also visual overviews and illustrations of the system's structure. In order to visualize connections as network maps in the form of SENs or CIMs, pairwise connections were examined between students and categories (or subcategories) of factors. A network map was then generated that showed connections for the relevant student/factor connections.

FINDINGS

Social ecology networks and consensus interrelationship maps

SNA of the data matrix facilitated the elaboration of a number of aspects of the relatedness of proximal factors, including the identified risk factors. Networks of proximal factors, the SENs, obtained for the cohort and sub-cohorts, were weighted so that CIMs were obtained. These CIMs effectively show the SENs for large numbers of students. The following illustrations exemplify this elaboration in order to answer the research questions of this examination of multiple factors related to undergraduate retention. The identified risk factors considered in this analysis are set out in Table 4 (appendix) and comprise demographic, academic and engagement indicators. The demographic category, for example, comprises such indicators as SES, disability, international, non-English speaking background (NESB), and ATSI. The engagement and academic factors listed here correspond in a general sense to engagement factors, such as those listed in Coates (2009).

Proximal factor relationships

Proximal factor relationships are student-centered and connected for each student at the university, and are hence place-based. This interrelated focus represents a significant feature of this approach that few past studies have embraced with respect to undergraduate retention, except in a few cases based in small sub-populations of a university (Dawson, 2008, 2010; Eckles & Stradley, 2012; Rizzuto, LeDoux & Hatala, 2009; Thomas, 2000). Studies have otherwise been based on forming universal generalizations from case studies or

from studies that have merged data and de-emphasized student-centered or place-based factors.

Figure 1 (Appendix) illustrates the proximal connections for a particular student from this cohort, one who withdrew. In this cases the proximal factors are shown connected to a particular factor category, current low-SES (current SES rather than that at the parent/guardian home), rather than to the student. Such ‘star diagrams’ offer real benefits in relation to diagnosing student issues related to retention and determining intervention strategies. Here, such diagrams provided a visual confirmation of the differences in proximal factors, including identified risk factors, and their differential relationships between at-risk students and students who are not considered at risk. In Figure 1 (Appendix), for example, there are a number of identified risk indicators (included with other proximal factors) present in the star diagram and listed in Table 4 (Appendix), including remoteness, disability, ATSI and female gender. Identifying these factors grouped for particular students may be a powerful tool in determining funding for risk interventions.

Star maps for continuing students, however, also contains several such risk indicators and there may be a number of other factors that may be mitigating against failure or withdrawal. The analysis that follows demonstrates how this mitigation may be occurring and also how there are currently hidden ‘covariate risk factor networks’, based in relative risk estimations, that may be contributing to risk overall for both of these students and other students in the cohort.

The cohort matrix allows a star diagram to be reproduced for any student in the cohort and such diagrams may be useful in examining particular interventions related to particular factors, such as current low-SES students. These diagrams have the potential to be used also as consensus diagrams, with weightings of nodes and edges to show the most common proximal connections for all students. For current low-SES students, for example, these may show the connections that may appear to be most important across the cohort.

Social ecologies and risk factor networks

This project was designed to examine a number of issues related to the compounding of risk for students in undergraduate retention studies, issues that are currently addressed inefficiently using percentage contributions of risk. The SNA here provides a set of correlations of any selected risk factor or group of factors, but set against the background of the social ecology of the student cohort. This latter point is critical since generalizations of risk and particular generalizations across cohorts (even in the same institution in different years) may provide an inaccurate view of the risk and an inadequate and inefficient allocation of resources to risk alleviation and/or prevention. Thus, the SNA provides an important breakthrough in both visualizing and understanding the set of risk factors involved in retention/attrition and how they may be differentially connected for each student or sub-cohort, or in the entire cohort.

Risk factor networks in the study cohort

SNA of the data matrix has provided sets of multiple correlations of risk factors from within a meaningful context of proximal factors of students in the study cohort. Figure 2 (Appendix) shows an example of a SEN for a continuing student. In the top diagram, identified risk factors (the connected red squares at nodes) are set against a SEN, while in the lower diagram, the risk factor network is shown without the connections to other proximal factors of that SEN. This figure shows that risk factors can be identified against the background of proximal factors. These diagrams quickly become quite dense and,

consequently, the focus here is on the risk factor networks themselves, given that it is only in this study that risk factors have been shown to be entwined in a complex way with other proximal factors.

Figure 3 (Appendix) shows an example of a risk factor network, derived from a SEN of proximal factors for a student who has withdrawn. This figure highlights the difficulty in using previously identified factors to identify students at risk of attrition. Other examples of risk factor networks in this study, for example, show that some students with similar profiles (e.g., limited use of Blackboard, low home and current SES, engaged in distance education courses) withdrew while the others continued.

While it's not possible to determine the reasons underlying one student's decision to withdraw, it is possible to build a profile of factors based on proximal relationships to give a sense of student agency. An important finding in this social ecology SNA, therefore, is that a given risk factor may not act on its own, since it is always associated with a set of other factors that would not be considered necessarily in a simple statistical analysis, such as a calculation of percentage risk.

For example, SENs, such as seen in in Figure 3 (Appendix, connect low-SES as a factor with a number of other factors in both withdrawing and continuing networks, implying that low-SES by itself needs to be considered against this background of other risk factors, as well as other proximal factors. By way of contrast, low-SES is considered a factor, on statistical grounds, that is implicated as important in student retention, even when considered in isolation (Richardson, Bennett & Roberts, 2016). This kind of detailed and deeper retention/attrition analysis can identify the important factors in this background for the cohort, or a given sub-cohort, through use of a consensus risk factor networks, as well as identifying all of the factors for a single student—for example risks A and B may act in combination with other factors, whereas for another student risk may also involve risk C.

Relative risk and risk factor networks

The undergraduate student dataset included information on whether the student had withdrawn from the system. This information was used to find the subgroup of students ($n=1062$) towards whom interventions may be targeted. The characteristics of the subgroup were compared to the overall network and to the subgroup of students that completed or continued their studies ($n=3003$). Initial analysis here allowed compilation of a list of factors more likely to be associated with someone who had withdrawn to be compiled, which served as a preliminary starting point for the analysis of risk. These steps/processes were the basis for considering connections between proximal factors captured in an analysis of the network matrix data, where the likelihood (through correlation) of these connections occurring for the entire cohort was compared with the likelihood of these connections occurring for those who have withdrawn.

This analysis was conducted using relative risk (probability ratio) estimates for the cohort matrix, where it was treated as an empirical correlation matrix (Chu & Davis, 2011). The aim was to provide a greater nuance to the initial set of previously identified risk factors and allow a configuration of covariate risk factors, or 'covariate risk factor networks' that may assist in better directing resources towards intervention, including early intervention. The estimation of relative risk and odds ratios allowed use of the outcome 'withdrawing from the course' as a variable of constant effect in the regression analysis—this outcome is commonly used when analyzing undergraduate cohorts for retention/attrition.

The relative risk for each factor was calculated from the cohort matrix and the results, combined with results from SNA analysis, allowed the construction of 'covariant risk factor networks' derived from the matrix as a whole and which, in turn, may be cohort specific. The

diagram in Figure 4 (Appendix), for example, shows covariant risk factor networks for an individual and for a selected sub-cohort, illustrating the connections across several of the highest covariant risk factors, as calculated using relative risk.

The figure shows how these are connected to the previously identified risk factors and illustrates how the two sets of risk factors may be considered together. An important implication is that the more nuanced set of covariant risk factors can be connected to the broader risk markers currently used at a particular university. This may allow for more specific targeting of sub-cohorts that are currently withdrawing, but who may be poorly identified.

From risk factor networks to consensus risk factor networks

As was the case with the star diagrams, the risk factor networks observed within the SENs can be mapped and weighted as consensus risk factor networks (a type of CIM) in order to illustrate the most common associations of factors for these networks, as well as the relative importance of each identified risk factor. Figure 5 (Appendix) shows such a consensus risk factor network, illustrating risk factors seen in Figure 3 (Appendix).

Not all of the identified risk factors are, in fact, present in the consensus risk factor networks if connections of less than 100 students are included (as a form of weighting), as indicated in Figure 5 (Appendix). This figure shows the identified risk factors for the continuing cohort as only comprising connections between the demographic factors related to low-SES (red squares at node), the attendance factors related to online and distance education (yellow squares), the engagement factors related to being a mentee (black) and low numbers of Blackboard logons (blue). The other proximal factors in this diagram are drawn from the data matrix and were not identified as risk factors.

Significantly, the demographic factors related to ATSI, Disability, NESB and International are not connected in this diagram (see the list on left side of the figure that shows factor exclusions), meaning that if they were present as connected factors for withdrawing students then this was for less than 100 students. A diagram with a different weighting, say less than 50 students connected, may offer a more nuanced view. Such consensus risk factor networks may be used to examine the view that risk factors present in over half the cohort (e.g., low- to mid-SES, distance mode or female gender) may not be considered as adequate determinants of risk since they do not sufficiently discriminate across the cohort or are not present exclusively in withdrawing student risk factor networks when CIMs are examined.

The consensus risk factor networks of themselves, or used in conjunction with the background networks, may go some way towards finding a common ground against which to direct and test intervention strategies, as well as to test the effectiveness of current strategies. This is because there is no longer a requirement to generalize the strategy such that it is only related to statistical data derived from averaging over large numbers of students on multiple campus locations. Instead, a strategy can be devised for a particular risk factor network and a particular sub-cohort, given the appropriate data entered into the matrix from which the maps are generated.

DISCUSSION

Student social ecologies

Students at the study university, like those at many regional universities, are drawn largely from rural or peri-urban populations spread across a wide geographic area within the

university footprint. Such regional centres offer flexible course delivery, often with a strong online presence, and deliver inclusive learning to diverse learners, as indicated in the summary in Table 2. The student demographic at the study university is characterised by a large proportion of students from low-SES backgrounds, with 20% of the undergraduate student population from low-SES backgrounds and 53% from mid-SES backgrounds. There are 67% of students who are female and 3% identify as Indigenous. Additionally, 42% of SCU students are the first in their family to attend to university and this is thought to present practical challenges, particularly to students from low- or mid-SES backgrounds as they transit into university studies.

The findings show the pronounced benefit that may be obtained using the new and innovative interrelationship models piloted in this study to examine academic risk of individuals in this university location. More generally, the findings successfully demonstrate the potential of using this model in representing, in a dynamic and connected way, sets of behavioral factors related to student life at university, where each factor informs and influences others. This is, in fact, the first reported analysis that shows how identified risk factors, which have previously been examined singly, can be shown to be connected and examined in combination in the context of behavioral relationships identified from the proximal factors related to university student lived experiences.

The current study is significant in its novel application of SNA in a social ecology context, such that it allows connections to be viewed dynamically—that is, with high dimensionality—so that a number of different student-based factors can be examined on a large-scale.

Risk factor networks

A highlight of the study is the finding that SENS/CIMS that include risk factors can be established from archived data of proximal factors that connect undergraduates directly to the world that includes the university. These networks can be used to establish risk factor networks for a particular sub-cohort. These SENS/CIMS can be established as networks in combination from dyads upward in order to determine subgroups of risk and inadequate determinants of risk. Such analysis may enable a more accurate evaluation of the interconnection of factors related to previously identified risks of withdrawal or failure, or attrition, as well as a corresponding evaluation of the interconnection of factors related to continuance or retention.

Risk factors clearly need to be considered as part of the collective of proximal factors to which a student is connected and which is situated in their world and subject to their interactions or individual agency. While these risk factors are connected to the student, they may also be connected to each other. The SENS, as well as the related CIMS, help determine the connection of factors, including risk factors and can be constructed not only for the entire cohort but any part of the cohort or a selected sub-cohort (e.g., sub-cohorts of distance, ATSI or disability). Relative risk based on the SNA matrix may give a more nuanced evaluation of risk factor networks by considering covariate risk factors derived from calculations of the relative risk from the student cohort matrix.

Risk factor networks and academic risk hierarchies in a social ecology

This report has illustrated some of the ways that risk can be examined where risk factors are connected as risk factor networks across a number of differing proximal factors in a particular student community ecology, such as the commencing student cohort examined here. The study indicates that, although previously identified risk factors may act in risk

factor networks, relative analysis appears to be a novel, yet highly effective, way to determine the risk factors that may be acting across particular community ecologies. Future student-centred and place-based studies such as the current study may do well to collect a broader and more complete set of data across demographics, academic and engagement categories, including survey data as well as archived data, so that risk factor networks may be more closely tied to particular social ecologies.

This type of analysis could be extended in future studies to enable the construction of a risk hierarchy, where factors within risk factor networks could be weighted to acknowledge their usefulness for targeted interventions, such as those in use currently. Female gender, for example, may not be a useful risk factor at the study university since it does not serve to discriminate continuing and withdrawing students across the cohort. Being female, however, may be a risk factor when combined with other risk factors, such as ATSI and single study location or mode of study, and a more appropriately targeted intervention may best serve if directed at the students with such risk factor networks. The study analysis suggests that construction of risk factor hierarchies could accommodate the relevance of particular risk factors to particular sub-cohorts, and could also be used in interventions that target students who have large risk factor networks and poor support networks. These students may currently not be engaged in establishing success networks.

Data collection is related to risk factor networks

The use of SENs or CIMs requires a construction process based on data collection around the factors that have an impact on a student's life, including their life at university. There are few retention studies that have collected or processed such data (Eckles & Stradley, 2012). The data collection in this study was derived from archives, but may need to be supplemented by data collection that is more immediately relevant and which may need to be collected specifically with social ecologies in mind. Data collection also needs to be consistent across cohorts for cohort comparison in terms of related social ecology networks. This may mean that removal of data from the matrix is needed where data is considered to be redundant.

Future work on SENs and CIMs, including exploring risk factor networks based on known risk factors of more nuanced covariate risk factors derived from calculation of relative risk, would benefit from automation of data sorting into the matrix in order to avoid the time-consuming data cleaning and categorization used in the current process. This would benefit also from use of dedicated data collection and entry into dedicated computer software, allowing for more targeted searches and more efficient servicing of student academic need. There are studies exploring real-time data collection and analysis, but in relation to how students emotional connections influence their studies (Wang et al., 2015).

Novel and innovative retention futures based on social ecology networks

The nested network adaptation of the student social ecologies, based on the theoretical modeling of Neal and Neal (2013), allowed a structural picture to be developed of the interrelationships of individuals and factors. The nested networks, visualized as SENs/CIM, will allow practitioners to look at the intersection of student experiences and events/locations of interactions, highlighting points of vulnerability and need, which may require appropriate interventions. The network maps provide a dynamic snapshot of interactions across factor categories and these, and associated network metrics, offer considerable potential in enabling diagnosis and evaluation for use in planning of intervention and support.

The dynamic and integrated approaches in this study, and the SENs/CIMs produced,

are expected to be extremely useful in discussing intervention and support practices with equity practitioners and change managers, and potential policy changes with both equity practitioners and a university executive. In particular, the approaches may be extremely useful in connection with university programs concerned with resilience and wellbeing, in addition to the development of early indicators of risk. The findings show that this approach lends itself to strategies that are based on the personal needs of individual students and which may be used to support students in following their own self-determined life choices, a growing area of research around agency (Jääskelä et al., 2016).

Such SENs/CIMs, if constructed with larger data sets of non-proximal factors, may be used also for comparative assessment of academic risk as well as planning and policy development related to integrated intervention and support for undergraduate students at universities, as well as in the higher education sector more broadly.

CONCLUSION

This study explores the notion that set of interconnected factors than single, isolated factors, need to be examined in order to examine student interrelationships in relation to retention/attrition. The approach used here is novel, bringing together major theoretical and empirical studies on both social ecology and SNA in order to examine the complex system that is student interrelationships. Accordingly, the study has examined previously identified risk factors and has constructed risk factor networks (i.e., sets of connected risk factors) set against proximal factors to more accurately depict a student-centered framework for retention studies—one that is also place-based as social ecology studies suggest (Lipina et al., 2013, Lipina et al., 2015).

Large archived data sets were used to construct what is called a ‘social ecology network’ of a student cohort—the first time that this approach has been used to examine student retention. The analysis has indicated that, while the process is innovative and informative, to more fully reach its potential, data collection may need to be focused on extending the social ecology in relevant areas. The study suggests that additional studies are needed in order to establish mechanisms and research to enhance retention based around the findings presented here.

The study, however, has successfully developed and completed an evaluation of academic risk that involved a multifactorial and dynamic model developed using SNA. This approach provided a comprehensive examination of how factors in each student’s life at university are interrelated, both for individuals and groups. Some of these factors are, in fact, those identified as academic risk factors (risk of withdrawal) from national analyses and are currently used by Australian universities to implement support and intervention. The SNA, however, also provided additional covariate risk factors, calculated as student-centered, place-based factors exclusive to the study university.

All risk factors were examined against the background of the student social ecology network with significant results: the identified risk factors can be seen to be linked in some individuals and sub-cohorts; the newly-identified risk factors can also be linked; and, all risk factors can be connected to other proximal factors.

This study has demonstrated that a way forward may be to more closely examine support/engagement as a way of optimising opportunity through disseminating the positive aspects of the social ecologies of successful or continuing students. This, in turn, may reduce risk. Arguably, such opportunities may benefit students through feedback—feeding results back into the student ecology may allow us to examine whether such knowledge would enable change in the community, including positive changes to retention profiles. Students may be empowered by this additional information/insight, enabling their own agency in

making life choices. Identifying risk factors and related two-way interactions may also be particularly useful in allowing students to self-monitor, or to receive support, for example through the development of markers for early indicators of risk.

What this means for the future

The multifactorial methods used here should allow equity practitioners to base retention interventions on a more nuanced system that takes into account localized data collection and analysis. Risk factor estimates, based on this type of SNA, can be place-based and student-centered, and related to other factors in a student's university life, with risk factors calculated from social ecology data sets. This means that methods can now be developed to compare the social ecology networks of students who may be at risk of failure, based on comparison of their network with consensus networks of students who are not at risk, with these being idealized as success networks based on both survey and archival data sets. Additionally, these same networks could potentially be viewed at different temporal intervals, giving a continuing representation of risk across a cohort.

This adaptation of SNA is being used by the authors to develop dedicated computer software in order to readily generate both individual and group networks of factors from real-time data entry and to provide dedicated outputs of risk networks for use in planning both interventions and future data collection. This planning would take into account the other non-proximal elements/factors of the student social ecology not considered in the current study, but which would enable the efficiency of support services provided at a university to be examined and assessed.

One goal of any future study looking at commencing student cohorts would be to use a similar style of SNA to establish the effectiveness of current support networks. This expansion of the current project would require a dedicated software program that would use SNA to map the links of proximal factors to meso, exo and macro-systems in an expanded social ecology. Such an expansion has been an integral component of the work done by Lipina and others (Lipina et al., 2013), but currently requires dedicated computer programming and does not have the comprehensive capability that a software program adapted to use with SNA may offer.

Such a dedicated response would offer equity professionals and the executive a much-needed tool that would enable measurements required for determining efficiency strategies and the related removal of redundancies. This development would potentially lead to more effective use of resources, including workloads and budgets. Such a response has a strong theoretical base in complexity modelling (Forsman et al., 2014) and builds on past models of retention and attrition, such as those of Bean (2005) and Tinto (1975, 1997), where these are related to retention as a result of complex system interactions (Cabrera, Nora & Castañeda, 1992).

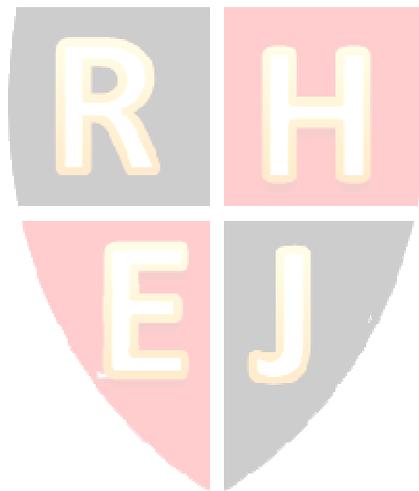
REFERENCES

- Bean, J. P. 2005. Nine themes of college student retention. In A. Seidman (Ed.), *College student retention: Formula for student success*, pp. 215–44. Westport, CT: Praeger.
- Bean, J., & Eaton, S. (2002). The psychological underlying successful retention practices. *Journal of College Student Retention*, 3(1), 73-89.
- Borgatti, S. (2012). *UCINET: Graph Visualization Software, v6.509*. Harvard: Analytic Technologies.
- Borgatti, S., Everett, M. G., & Johnson, J. C. (2013). *Analyzing social networks*. Los Angeles, CA: Sage.
- Borgatti, S., Everett, M. G., & Johnson, J. C. (2013). *Analyzing social networks*. Los

- Angeles, CA: Sage.
- Braxton, J. M., Hirschy, A. S., & McClendon, S. A. (2011). *Understanding and Reducing College Student Departure: ASHE-ERIC Higher Education Report, Volume 30, Number 3* (Vol. 16). John Wiley & Sons.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Bronfenbrenner, U. (1994). Ecological models of human development. In *International Encyclopedia of Education (2nd ed.)*, pp. 793-828. Oxford: Elsevier.
- Bronfenbrenner, U. & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner (Ed.), *Theoretical models of human development* Volume 1 of *Handbook of Child Psychology (6th ed.)*, pp., 793-828. Wiley Online Library.
- Cabrera, A.F., Nora, A., & Castañeda, M. B. (1992). The role of finances in the persistence process: A structural model. *Research in Higher Education, 33*(5), 571–93.
- Chapman, E. (2003). Alternative approaches to assessing student engagement rates. *Practical Assessment, Research and Evaluation, 8*(13), 1-10.
- Chu, J. S. G., & Davis, G. F. (2011). Who killed the inner circle? The breakdown of the American corporate elite network, 1999-2009.
http://opensiuc.lib.siu.edu/pnconfs_2011/1/?utm_source=opensiuc.lib.siu.edu%2Fpnconfs_2011%2F1&utm_medium=PDF&utm_campaign=PDFCoverPages
- Clarke, J., Nelson, K., & Stoodley, I. (2011). *Capability maturity models: A discussion paper and a work in progress*. SESR-MM Report 001 prepared for the ALTC Project, Establishing a framework for transforming student engagement, success and retention in higher education institutions. Brisbane: Queensland University of Technology.
- Coates, H. B. (2009). *Engaging students for success: Australasian student engagement report*. Camberwell, Australia: Australian Council for Educational Research Ltd..
- Coates, H., & McCormick, A. C. (Eds.). (2014). *Engaging university students: International insights from system-wide studies*. New York, NY: Springer.
- Dawson, S. (2008). A study of the relationship between student social networks and sense of community. *Educational Technology and Society, 11*(3), 224–238.
- Dawson, S. (2010). “Seeing” the learning community: An exploration of the development of a resource for monitoring online student networking. *British Journal of Educational Technology, 41*(5), 736–752.
- Divjak, B., & Peharda, P. (2010). Social network analysis of study environment. *Journal of Information and Organizational Sciences, 34*(1), 67-80.
- Dumbrigue, C., Moxley, D., & Najor-Durack, A. (2013). *Keeping students in higher education: Successful practices and strategies for retention*. New York: Routledge.
- Eckles, J. E., & Stradley, E. G. (2012). A social network analysis of student retention using archival data. *Social Psychology of Education, 15*(2), 165-180.
- Forsman, J., Linder, C., Moll, R., Fraser, D., & Andersson, S. (2014). A new approach to modelling student retention through an application of complexity thinking. *Studies in Higher Education, 39*(1), 68-86.
- Gilardi, S., & Guglielmetti, C. (2011). University life of non-traditional students: Engagement styles and impact on attrition. *The Journal of Higher Education, 82*(1), 33-53.
- Hanneman, R. A., & Riddle, M. (2005). *Introduction to social network methods*. Riverside, CA: University of California, Riverside. Retrieved from <http://faculty.ucr.edu/~hanneman/>.
- Hayden, M. (2010). Student attrition from higher education institutions. In E. Baker, P. Peterson & B. McGaw (Eds.), *The International Encyclopedia of Education, Vol. 4* (3rd Ed.), pp. 467-472. Oxford: Elsevier.

- Jääskelä, P., Poikkeus, A. M., Vasalampi, K., Valleala, U. M., & Rasku-Puttonen, H. (2016). Assessing agency of university students: validation of the AUS Scale. *Studies in Higher Education*, 1-19 (online).
- Kezar, A. (2014). Higher education change and social networks: A review of research. *The Journal of Higher Education*, 85(1), 91-125.
- Lipina, S.J., Insúa, I., & Echeverría, H. translated by V.P. Melia (2015). Manual and computerised analysis and application of the 'Diversified Derivation Algorithm (DDA)'. ACUMAR: Division of Environmental Health, Section of Cognitive Development, Validation and Assessment, Buenos Aires, Argentina.
- Lipina, S.J., Segretin, M.S., Hermida, J., Prats, L., Fracchia, C., Camelo, J.L., & Colombo, J. (2013). Linking childhood poverty and cognition: Environmental mediators of non-verbal executive control in an Argentine sample. *Developmental Science*, 16(5), 697-707.
- Maher, M., & Macallister, H. (2013). Retention and attrition of students in higher education: Challenges in modern times to what works. *Higher Education Studies*, 3(2), p62.
- Museus, S. D., & Jayakumar, U. M. (Eds.) (2012). *Creating campus cultures: Fostering success among racially diverse student populations*. New York: Routledge.
- Neal, J. W., & Neal, Z. P. (2013). Nested or networked: Future directions for ecological systems theory. *Social Development*, 22(4), 722-737.
- Nelson, K., Clarke, J., Kift, S., & Creagh, T. (2011). *Trends in policies, programs and practices in the Australasian first year experience literature 2000-2010*. Brisbane: QUT.
- Newman M. E. J. (2010). *Networks*. New York, NY: Oxford University Press.
- Nichols, M. (2010). Student perceptions of support services and the influence of targeted interventions on retention in distance education. *Distance Education*, 31(1), 93-113.
- Osborne, J.W. (2006). Bringing balance and technical accuracy to reporting odds ratios and the results of logistic regression analyses. *Practical Assessment Research & Evaluation*, 11(7). Available online at <http://pareonline.net/getvn.asp?v=11&n=7>.
- Price Waterhouse Coopers. (2007). *Research Report: The economic benefits of a degree*. London: Universities UK. Retrieved September 18, 2007, from <http://www.bournemouth.ac.uk/conservation/newsandevents/UUKfulltext>
- Quaye, S. J., & Harper, S. R. (Eds.) (2014). *Student engagement in higher education: Theoretical perspectives and practical approaches for diverse populations*. New York: Routledge.
- Radloff, A., & Coates, H. (2013) Monitoring and improving student engagement. In S. Marginson (Ed.), *Tertiary education policy in Australia* (pp. 21-30). Melbourne, Australia: Centre for the Study of Higher Education.
- Richardson, S., Bennett, D., & Roberts, L. (2016). *Investigating the Relationship between Equity and Graduate Outcomes in Australia*. Curtin University, Australia: National Centre for Student Equity in Higher Education.
- Rizzuto, T. E., LeDoux, J., & Hatala, J. P. (2009). It's not just what you know, it's who you know: Testing a model of the relative importance of social networks to academic performance. *Social Psychology of Education*, 12(2), 175-189.
- Simmel, G. (1955 [1922]). The web of group affiliations. In R. Bendix (Trans.), *Conflict and the web of group affiliations* (pp. 126-195). Glencoe, IL: Free Press.
- Szumilas, M. (2010). Explaining odds ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 19(3), 227-229.
- Thomas, L. (2011). Do pre-entry interventions such as 'aim higher' impact on student retention and success? A review of the literature. *Higher Education Quarterly*, 65(3), 230-250.

- Thomas, S. L. (2000). Ties that bind: A social network approach to understanding student integration and persistence. *Journal of Higher Education*, 71(5), 591-615.
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 5(1), 9-125.
- Tinto, V. (1997). Classrooms as communities: Exploring the educational character of student persistence. *Journal of Higher Education*, 68(6), 599–623.
- Wang, R., Harari, G., Hao, P., Zhou, X., & Campbell, A. T. (2015). SmartGPA: How smartphones can assess and predict academic performance of college students. In *UbiComp'15, Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 295-306). New York, NY: ACM.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. New York: Cambridge University Press.



Appendix

Table 1:
System levels used for academic risk assessment

The four system levels	Examples of system factors (elements)
Micro system – proximal institutions and groups directly impacting the student, e.g., family, university, clubs	Educational background, economic circumstances, social circles, including clubs and societies (teams)
Meso system - relational aspects of proximal institutions and groups involved in the micro-system	family cultural background, family financial stress, parent mental health
Macro system - social setting (immediate context) e.g., commerce, industry, government	community resources, access to social support, social mobility
Exo system – societal beliefs, norms and values	discrimination, stigmatization, inclusion/exclusion

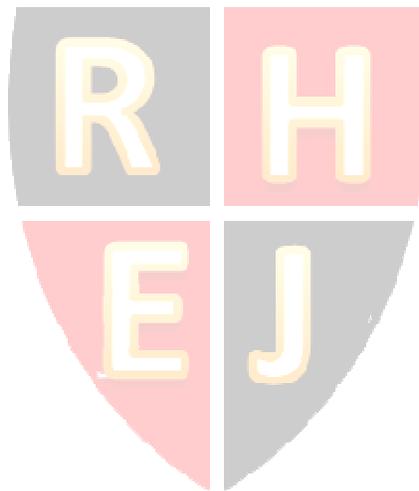


Table 2:
Summary of student diversity in the study cohort

Sub-cohort description	Total in commencing cohort
Female	2772
Male	1293
Age 21 or under	880
Age 22 to 35	2200
Origin in regional Australia	2284
Current socio-economic status low to mid	3450
First in family	2003
Secondary education	769
Less than 3 years since completion of Year 12	1321
Study mode is online only	1594
Study mode is mixed	1059
ATSI	173
Disability	345
Non-English speaking	218

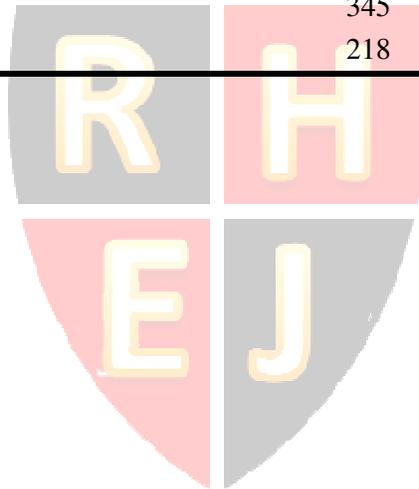


Table 3:
Examples of proximal system factors (elements) and categories in this study

The system level	Examples of groupings of proximal system factors (elements)	Previously identified proximal factor categories in this study
Micro system – proximal, e.g., family, university, clubs	Educational background prior to university entry, family economic circumstances, accommodation and travel, social circles, including clubs and societies (teams)	<p>Demographic (ATSI code, international indicator, SES indicator, NESB indicator, disability, gender, age, remote location, first in family)</p> <hr/> <p>Academic (attendance mode, high attrition unit, course preference, withdrawn unit, GPA, application category, high load, failed units, scholarship, access LMS system)</p> <hr/> <p>Engagement (Blackboard logons, Customer Relationship Management (CRM) cases, mentorship participation)</p>

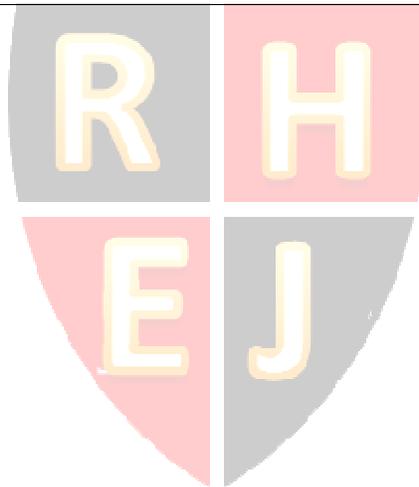
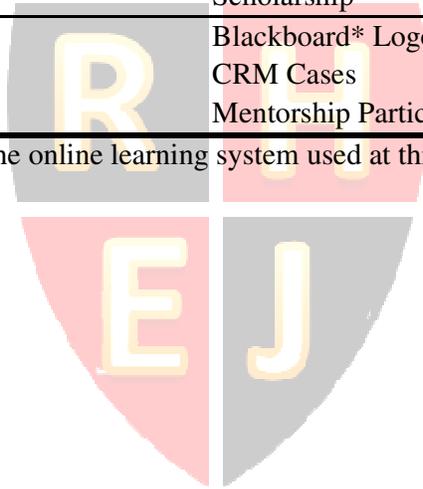


Table 4:
Previously identified risk indicators for the undergraduate cohort

Indicator Category	Indicator
Demographic	Disability NESB Indicator ATSI Code International Indicator SES Indicator
Academic	Course Preference Attendance Mode High Attrition Unit Withdrawn Units Grade Point Average (GPA) Application Type Category High Load Failed Units Scholarship
Engagement	Blackboard* Logons CRM Cases Mentorship Participation

*Blackboard Collaborate is the online learning system used at this university.



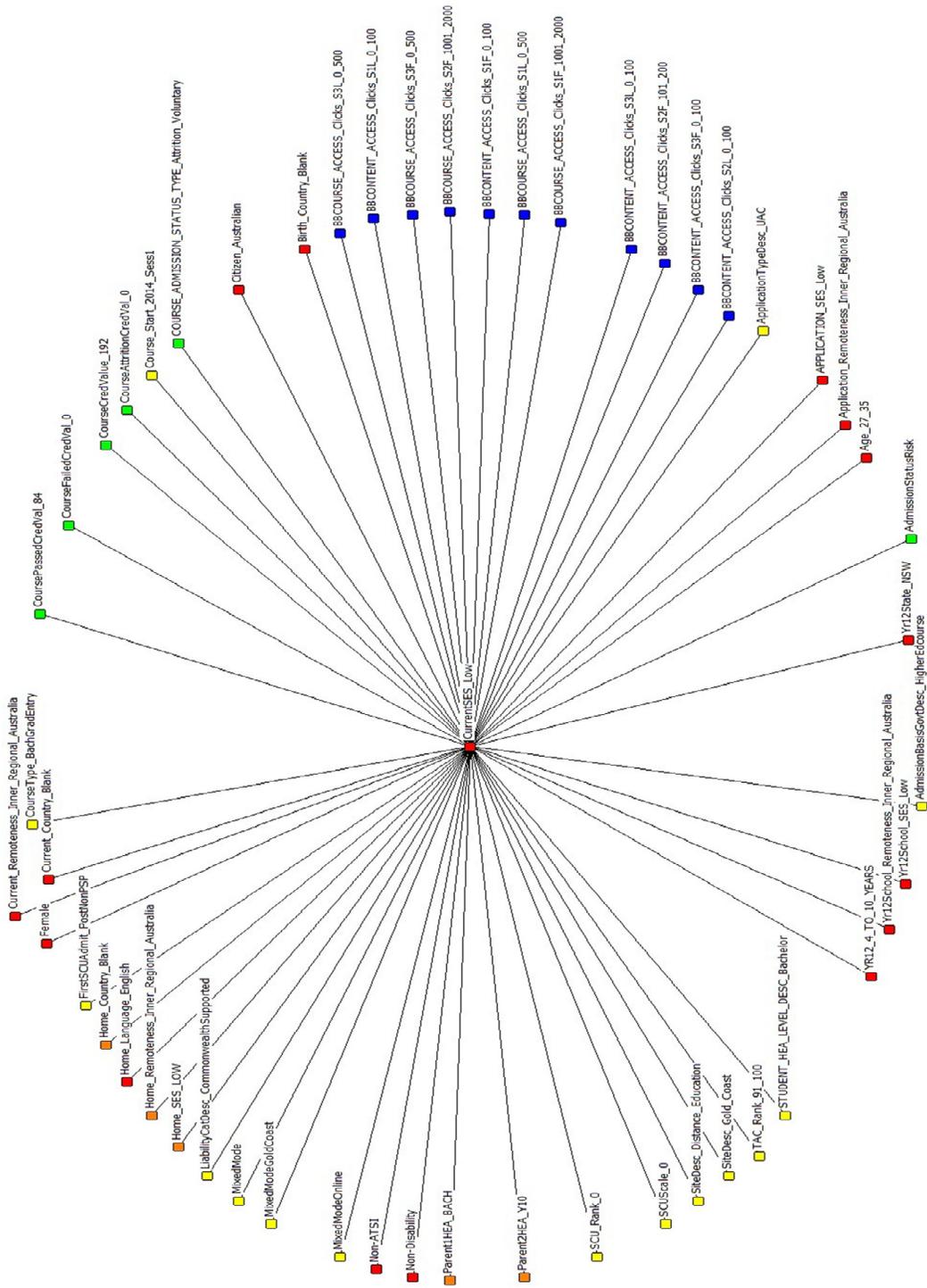


Figure 1: Star diagram for a withdrawn student linking current low-SES to other factors

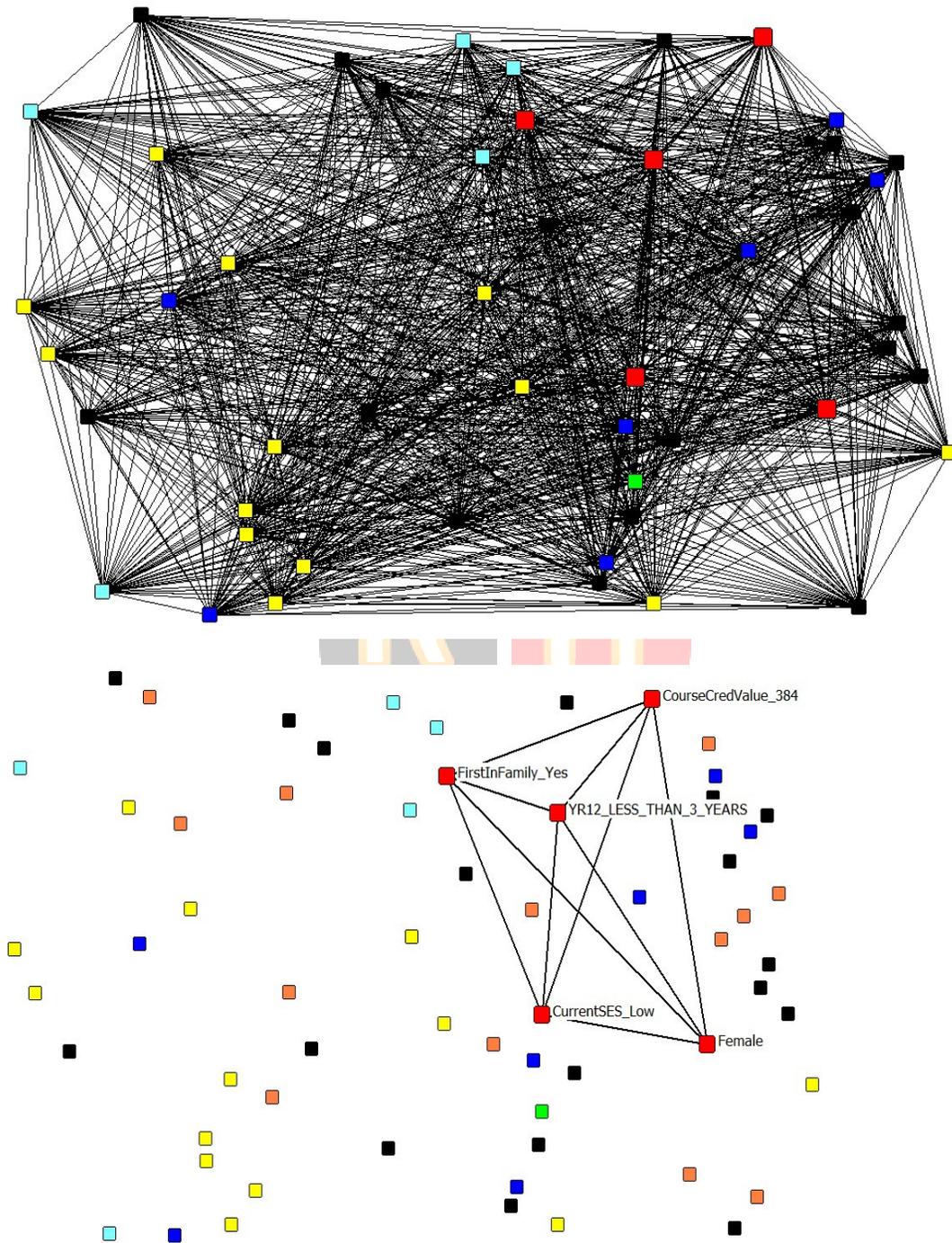


Figure 2: Social ecology network (SEN) of a continuing student showing connections between identified risk factors, including current low-SES (red squares at nodes) and other proximal factors for this individual (top diagram). In the lower diagram only the risk factors are connected.

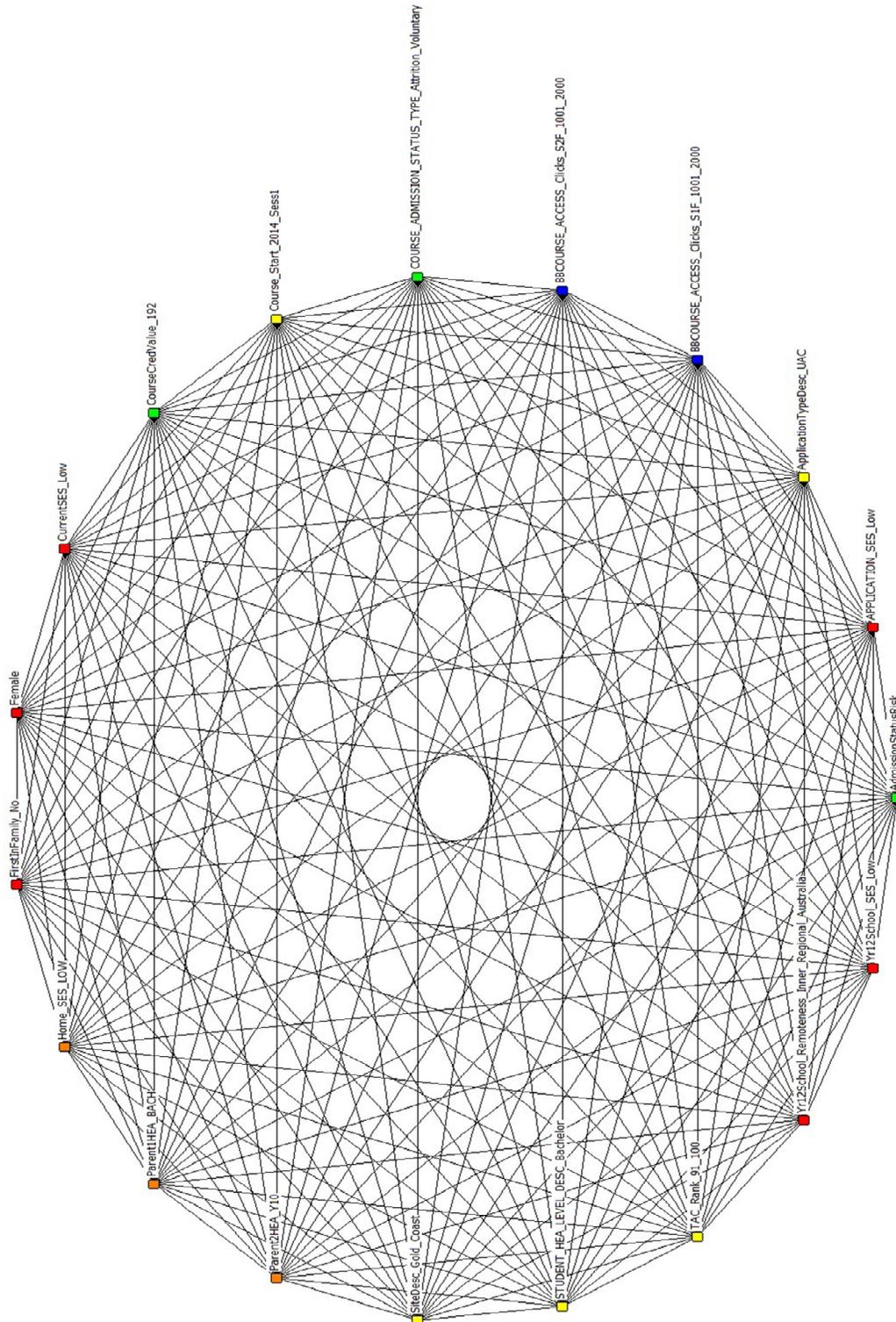


Figure 4: Risk factor network for a withdrawn student based on factors calculated from relative risk. Covariant risk factors include internal mode of study and tertiary entrance ranking. Previously identified risk factors include low home SES and blackboard logons.

