JOSSAE (Journal of Sport Science and Education)

OPEN ACCESS
E-ISSN Online: 2548-4699
http://journal.unesa.ac.id/index.php/jossae/index
http://dx.doi.org/10.26740/jossae.v7n1.p7-16

The Determinants of Student Attrition in an Undergraduate Sport and Exercise Science Degree

Lyn Kee[a]*, Minh Huynh[b], Paul Xanthos[c], Charlie Davids[d], Lachlan James[e]
[a]Swinburne University of Technology, Australia
[b,c,d,e]La Trobe University, Australia

Correspondence: ykee@swin.edu.au
Received: 19 Oct 2021 **Accepted**: 18 Apr 2022 **Published:** 30 Apr 2022

Abstract
The purpose of this study was to present a novel approach to identify factors that influence student attrition in an undergraduate Sport and Exercise Science (SpES) degree. A non-linear statistical technique was used to analyze demographic indicators of students enrolled in the first year of the SpES bachelor's degree at an Australian university (*n* = 312). Student dropout was successfully predicted using Chi-Squared Automatic Interaction Detection (CHAID). Using only four variables (i.e., Region, ATAR, Admission, & VCEBus), the model achieved an 87.50% classification accuracy. The article concludes with an endorsement of the proposed analysis to predict student dropout. Administrative and teaching-focused university staff may adopt the novel approach to identify the relevant demographic indicators for their respective institutions or programs of study in the management of student attrition. The limitations of this study are also discussed.

Keywords: attrition; dropout; higher education; sport and exercise science; chi-squared automatic interaction detection

1. Introduction

The attrition rate of higher education in Australia has been fluctuating at around 15% since 2005 (Department of Education, 2020). This lack of improvement should raise concerns, as a higher number of enrolments in recent years implies that the number of students leaving university without qualification is also increasing each year (Cherastidtham et al., 2018). The costs of attrition not only affect universities but also the students in terms of wasted time, foregone employment and earning opportunities (Commonwealth Productivity Commission, 2019). The first year of university study is especially critical to students' subsequent persistence, as it serves as a transition period, during which students adapt to a new learning environment and learn to engage with the university and peers (Baik et al., 2015).

Research in higher education has reported mixed results on the first-year experience. On the positive side, Baik et al. (2015) found that students' first-year experience has improved over the years after analysing data spanning two decades. It was reported that levels of engagement with study, satisfaction with university courses and teaching quality, and relationships with teaching staff have all improved. However, there was a significant proportion of students who reported low levels of motivation and that coping with university study is challenging (Baik et al., 2015). This is particularly true for students entering higher education with low Australian Tertiary Admission Ranks (ATAR; Baik et al., 2015). This finding is in line with other studies that reported the attrition risk rises as ATAR falls (e.g., Cherastidtham et al., 2018; Kemp & Norton, 2014). Despite there has been some improvement in the first-year experience, student dropout remains an issue.

Apart from the ATAR, research in student attrition also investigates the influence of demographic factors. A recent study in the United States by (Aulck et al., 2016) for example, found that age is a strong predictor of eventual attrition. Aulck et al., (2016) argued that dropout can be predicted as early as students' first semester at university, as the study's prediction model obtained a 67% accuracy rate. In the Australian context, much focus is placed on students' socioeconomic status (SES), residential location, and indigeneity. An analysis of the national data that tracks university students from the commencement to completion of their study, for example, found that remote and indigenous students have the lowest completion rate nine years after commencing their degree (Edwards & Mcmillan, 2015).

The main objective of this study is to present a novel approach to identify factors that influence student attrition in an undergraduate Sport and Exercise Science (SpES) degree and identify the various factors that contribute to attrition risk. Although various statistical analysis methods have been adopted previously for dropout prediction, the most commonly used methods are often linear statistical techniques such as the logistic regression models (for a review, see Lykourentzou et al., 2009). This paper presents an alternative analysis method, using Chi-squared Automatic Interaction Detection (CHAID) to predict dropout. CHAID is a decision tree technique that works similarly to regression analysis in terms of producing prediction models. An advantage of decision tree induction is that they are not bound by linear assumptions and provide an intuitive method of communicating statistical information to those unfamiliar with data mining approaches. Furthermore, the output can provide greater insight than linear models. As such, it can be a desirable method for communicating dropout information to teaching or administration focused academics.

It is acknowledged that many personal factors such as demographic characteristics are not malleable and are largely out of universities' control. However, gaining a better understanding of the various risk factors will enable the establishment of early warning systems and allow appropriate assistance to be provided to students before problems become overwhelming. Further, as a result of the Australian government's effort to promote allied health and medical research, courses relating to science and broader areas of health have received increased interest in recent years (Anderton, 2017). Given the growing interest in these areas, it is important to identify what factors contribute to attrition in related fields of study.

Demographic variables, the basis of admission, and prior achievement (ATAR) will be included in the analysis. It is expected that ATAR predicts dropout because it reflects students' academic ability and preparedness for further study. This is in line with Edwards and Mcmillan (2015) cohort analysis exploring equity group outcomes at university. Along with remote and indigeneity status, ATAR is also shown to influence the completion rate. Further, Anderton et al., (2017) suggested that proficiency in secondary mathematics and science is a good indicator of academic performance at university. However, it is not clear whether this applies to all areas of study in science, technology, engineering and mathematics (STEM). As such, along with demographic variables, the basis of admission, and ATAR, subjects studied in the final year of high school will also be included in the analysis.

2. Method

A cross-sectional study design applying a non-linear statistical approach was used to classify dropouts from commonly collected demographic indicators. Deidentified data were requested via the university's business analytics department and the study was approved by the university's human research ethics committee. The dataset contained demographic information for each student enrolled in the first year of the bachelor of sport and exercise science degree at La Trobe (HBSES) in 2019 (n = 312). The indicators included in the dataset are presented in Table 1. The dataset also described course status at the commencement of 2020. Status categories were as follows: Admitted (n = 252), AWOL (n = 5), Leave of Absence (n = 17), Withdrawn (n = 55). Status data were missing from two students and these were removed from the analysis. Dropout was defined as any student withdrawn from the course who was also not enrolled in another La Trobe course in 2020. AWOL students were also classified as Dropout. Continuing were all admitted students and 'Withdrawn' students who were enrolled in another La Trobe degree. Students on 'Leave of Absence' were removed from subsequent analysis.

Descriptive statistics for each of the demographic indicators were acquired from the 2019 cohort. A chi-squared automatic interaction detection (CHAID) was employed to model relationships between the demographic indicators and dropout. A minimum number of cases for a parent node was set at 10, while 5 cases or greater was allowed for a child node. 10-fold cross-validation. This statistical approach is a form of decision tree analysis that helps to determine how variables best merge to explain the outcome in the given dependent variable. The technique creates all possible cross-tabulations for each categorical predictor until the best outcome is achieved and no further splitting can be performed (see Figure 1). CHAID has the advantage of not being constrained by linear assumptions, being capable of handling missing data and providing an output that is interpretable to the end-user (Han et al., 2012).

3. Result

Descriptive statistics for independent variables are presented in Table 1. The results from the CHAID model are presented in Figure 1. Dropout (Dropout = YES) occurrence was 13.5% (42/312 – Node 0). Of the eight independent variables specified, the model successfully classified Dropout using only four variables (Region, ATAR, Admission and VCEBus). Region has the largest impact on classification. Students who were from Melbourne - South and South East; Melbourne - North West; Bendigo and Surrounding Districts; Else, dropped out 24.2% of occasions (16/66 – Node 2). The model achieved an 87.5% classification accuracy (See Table 2).

Table 1. Descriptive Statistics for Independent Variables

Indicator	Dropout (%)	Non-Dropout (%)
	n = 42 (13.5%)	*n* = 270 (86.5%)
Gender		
Male	29 (14.5%)	171 (85.5%)
Female	13 (11.6%)	99 (88.4%)
Region		
Melbourne - North West	8 (23.5%)	26 (76.5%)
Interstate other than NSW/ACT	1 (12.5%)	7 (87.5%)
Melbourne - South and South East	6 (20%)	24 (80%)
Melbourne - West	3 (8.6%)	32 (91.4%)
Melbourne - North East	12 (12.6%)	83 (87.4%
Melbourne - East	6 (11.8%)	45 (88.2%)
Melbourne - Inner	4 (12.9%)	27 (87.1%)

Indicator	Dropout (%)	Non-Dropout (%)
Other Rural Victoria	0 (0%)	12 (100%)
Albury-Wodonga & Surrounding Districts	0 (0%)	2 (100%)
Greater Sydney	0 (0%)	1 (100%)
Bendigo and Surrounding Districts	1 (100%)	0 (0%)
International	0 (0%)	7 (100%)
Other Rural NSW	0 (0%)	2 (100%)
Shepparton & Surrounding Districts	0 (0%)	1 (100%)
Geelong and Surrounding Districts	0 (0%)	1 (100%)
Else	1 (100%)	0 (0%)
SES		
Low	11 (23.9%)	35 (76.1%)
Medium	19 (11.3%)	149 (88.7%)
High	11 (12.2%)	79 (87.8%)
Else	1 (12.5%)	7 (87.5%)
Basis of Admission		
School Leaver (ATAR)	27 (12.4%)	190 (87.6%)
School Leaver (No ATAR)	3 (20%)	12 (80%)
Previous Higher Education	9 (16.7%)	45 (83.3%)
TAFE	2 (13.3%)	13 (86.7%)
Other	1 (9.1%)	20 (90.9%)
VCE Business Subjects		
Yes	11 (8.3%)	121 (91.7%)
No	31 (17.2%)	149 (82.8%)
VCE Health and Physical Education Subjects		
Yes	31 (12.6%)	215 (87.4%)
No	11 (16.7%)	55 (83.3%)
VCE Maths Subjects		
Yes	34 (12.2%)	244 (87.8%)
No	8 (23.5%)	26 (76.5%)
VCE Science Subjects		
Yes	33 (13.6%)	209 (86.4%)
No	9 (12.9%)	61 (87.1%)
VCE Technology Subjects		
Yes	3 (12.5%)	21 (87.5%)
No	39 (13.5%)	249 (86.5%)

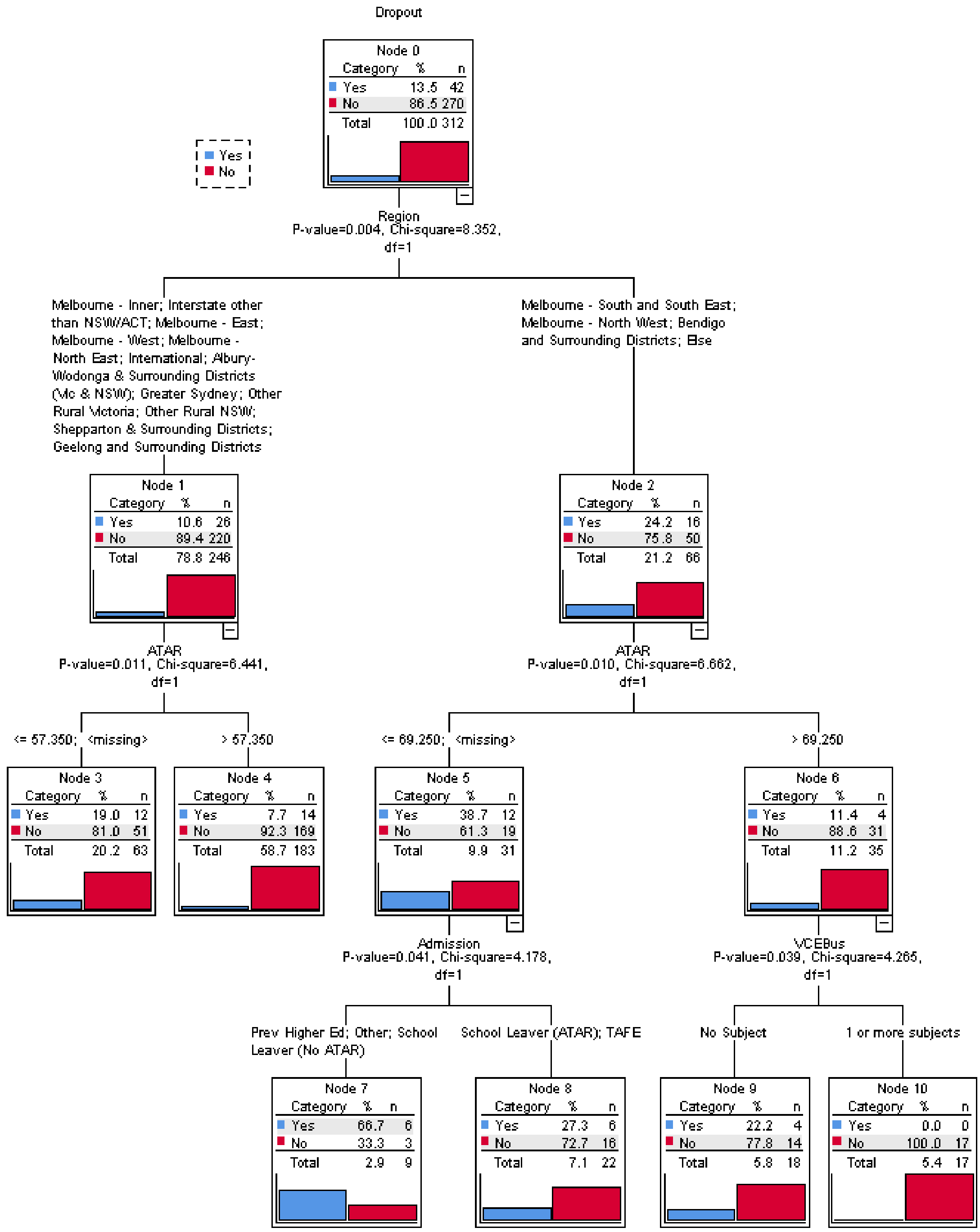

Figure 1. CHAID Output: Node 0 presents the dropout frequency and count for the data set (13.5% dropout). Reigion has the biggest impact on model classification (with 24.2% of students from the locations listed above Node 2 dropping out), with ATAR also highly influencial. Admission and VCEBus further contributed to classification.

Table 2. Confusion Matrix for Classification Demonstrating The Relationship Between Model Predictions and Actual Outcomes.

Observed	Predicted		
	Yes	No	Percent Correct
Yes	6	36	14.29%
No	3	267	98.89%
Overall Percentage	2.88%	97.12%	87.50%

4. Discussion

The present study sought to present a novel approach to identify factors that influence student attrition in an Australian Sport & Exercise Science (SpES) degree. Chi-squared automatic interaction detection (CHAID) was employed to model the relationships between key demographic variables and dropout. Student dropout was successfully predicted using the model, which achieved a classification accuracy of 87.50%. Region was featured at the first level of the tree and, as such, the hypothesis that ATAR would be the primary predictor cannot be accepted. However, ATAR was represented in the second level and was one of only four (of the nine input variables) featured in the model. Basis of Admission and VCEBus were the remaining two variables forming the model. The analysis also revealed multiple combinations of variables that are characteristic of dropout. These findings will permit academic administrators to focus their attention on a few key early-stage indicators to identify those students most at risk of not progressing to the second year of university study.

Exploring student attrition and/or academic performance at tertiary institutions is an important area of study and one which has received much attention in the literature. Previous studies have predominantly used common statistical approaches, such as logistic regression or correlation analysis, to model relationships between demographic variables and relevant outcomes; including student dropout (Aulck et al., 2016) or first-year grade point average (Anderton, 2017). For example, Aulck et al., (2016) employed a regularized logistic regression model to predict student attrition and yielded a prediction accuracy of 66.59%. In the present study, we utilised an alternative approach to predicting attrition involving CHAID, which is not bound by linear assumptions. Importantly, our CHAID predictive model yielded a superior classification accuracy of 87.50%. Thus, the findings of the present study endorse the use of this novel predictive model (alongside or in place of logistic regression) to identify at-risk students in their first year of study with a high degree of accuracy.

The model identified region as the primary predictor of dropout. In the present study, students were predominantly located throughout Victoria, Australia; however, there were also a small number of interstate and international students. Students who reported living in South and South East Melbourne; North West Melbourne; Bendigo and surrounding districts were more likely to drop out than students living in Inner Melbourne; East Melbourne; West Melbourne; North East Melbourne and other areas. There have been previous suggestions that region may serve as a contributing factor to dropout (Baik et al., 2015; Li & Carroll, 2020). However, these posit that those in rural or remote settings are at greater risk of attrition than those in metropolitan areas. Thus, the discrepancies across regions of metropolitan Melbourne observed in the present study are somewhat intriguing. It would be reasonable to speculate that regions associated with higher dropout risk are also associated with lower SES (Bradley et al., 2008). However, the CHAID model used in the present study found that SES did not contribute meaningfully to explain student dropout.

As anticipated, a lower ATAR was also identified by the model as a significant indicator of dropout for students entering a SpES degree. In regions associated with a higher dropout rate, an ATAR of 69.25 or lower increased dropout risk. From these regions, 11.4% of students with an ATAR greater than 69.25 dropped out while 39.7% of students with an ATAR of 69.25 or lower dropped out. A similar result was noted in regions associated with a lower dropout rate. 7.7% of students with an ATAR greater than 57.35 dropped out from these regions while 19.0% of students with an ATAR of 57.35 or lower dropped out from these regions. This is not surprising as previous research has identified that students with low ATAR scores feel less prepared for university, experience less enjoyment of their courses, have lower levels of academic engagement, are more likely to have difficulties with their studies, and are more likely to consider deferring or withdrawing (Baik et al., 2015). These findings support the consensus that a lower ATAR increases dropout risk (Cherastidtham et al., 2018; Edwards & Mcmillan, 2015; Li & Dockery, 2015). Cherastidtham et al., (2018) reported on data from public Australian universities, highlighting that a student with an ATAR between 70 and 79 is almost twice as likely to drop out compared to a student with an ATAR of 90 or above and that ATAR has a high impact on students completing bachelor-level university courses. Similarly, Edwards and Mcmillan (2015) examined Australia-wide data, identifying that an ATAR of less than 60 contributed to a lower likelihood of students completing their course within nine years. Importantly, the current study highlights that this phenomenon is also evident specifically in a SpES cohort and can be detected using CHAID.

It is interesting to note the large discrepancy identified by the model in the ATAR score predicting dropout between regions associated with higher dropout and regions associated with lower dropout (69.25 & 57.35 respectively). It has been suggested that a lower ATAR reflects lower levels of academic preparedness (Baik et al., 2015). Similarly, university entry scores are a predictor of student's GPAs at the end of the first semester in their course of study (McKenzie & Schweitzer, 2001). If these premises are accepted, then it might be that academic preparedness is even more important for students from regions associated with higher dropout (i.e., "less advantaged" regions). For example, it appears that a student with an ATAR of 65.00 is less likely to drop out if they are from a "more advantaged" region compared to a "less advantaged" region. As such, there might be benefits of residing in "more advantaged" regions that potentially accommodate for lower academic preparedness. Again, this points to the possibility that the reasons for dropout for students across different regions may be tied to a confounding variable that impacts academic preparedness, and thus warrants further investigation. In the United States, Johnson (2008) reported that high school characteristics such as socio-demographic characteristics influence odds of enrolment, persistence, and graduation from university. In an Australian setting, Anderton (2017) reported that academic success in first-year allied health students was superior in those that attended government (i.e., public) schools when compared to non-government schools (i.e., catholic or independent). The distribution of government and non-government schools across metropolitan Melbourne may provide some explanation for those regions associated with a higher risk of dropout. Perhaps secondary schools that provide a higher quality of teaching and academic preparedness are clustered in the "more advantaged" regions of Melbourne.

Basis of admission further contributed to dropout risk for SpES students from "less advantaged" regions combined with a relatively low ATAR score of 69.25 or less (i.e., higher risk students). 66.7% of higher-risk students admitted to the course with a previous higher education admission or admitted as a school leaver not on the basis of their ATAR dropped out. In contrast, only 27.3% of those students admitted to the course as a school leaver on the basis of their ATAR or with a previous TAFE admission dropped out. These results suggest the probability of higher-risk students dropping out is compounded by having a previous admission in a higher education course or via another non-ATAR pathway. The influence of the highest obtained qualification on student withdrawal from their current course has previously been investigated. Cherastidtham et al., (2018) highlighted that students completing a bachelor-level course that has already obtained a bachelor qualification (or higher) are at a slightly higher risk of non-completion, citing that the cost of completing a subsequent bachelor course might outweigh the benefits (e.g., improved job prospects) of obtaining a subsequent degree. In addition, the authors also describe that students that have completed a lower than bachelor-level post-school qualification (such as courses offered by TAFE institutes) are more likely to complete a bachelor course, citing that these students have displayed the academic ability and effort required to complete a diploma-level course prior to admission to the course (Cherastidtham et al., 2018). This commentary appears to support the results of our study on SpES students. It should also be noted that students commencing a bachelor course that was previously admitted in another course might have transferred into their current course, rather than having previously completed a bachelor course. Cherastidtham et al., (2018) explain that relatively poor previous university performance, rather than course transfer (i.e., withdrawal) from another bachelor course, is more likely to increase the risk of dropout and course transfer as an isolated factor does not appear to influence dropout risk. Our admission data based on enrolment in a previous higher education course does not distinguish between students that previously completed a bachelor-level course and students that transferred from their previous course. As such, the influence of transfer from another bachelor course and previous completion of a bachelor course on subsequent dropout from a current course warrants further investigation.

For the higher risk population, a non-ATAR admission [School Leaver (No ATAR)] was more likely to result in dropout according to the CHAID model. These students are more likely to be older given that relatively young students are more likely to enter the course on the basis of their ATAR following high school. This finding reflects previously published Australian-wide data. The latest available national Australian data exploring completion and dropout from 2005- to 2016-commencing cohorts over a four-year period indicates that commencers over the age of 25 consistently complete their course at a lower rate than commencers under the age of 25. These data also show that commencers over the age of 25 consistently drop out at a higher rate compared to commencers under the age of 25. As previously postulated, older students are more likely to have work and family commitments than younger students. Work and family commitments have previously been identified as prevalent factors contributing to student dropout (Beer & Lawson, 2017), requiring students to make decisions regarding how much time to invest in study vs work or family commitments (Burke et al., 2017). In contrast, qualitative research has identified that students with families to care for feel encouraged to remain in their course to provide a good example for their children (Kirk, 2018). However, the same work highlighted that having a family to care for increases responsibility and thus pressure placed on the student (Kirk, 2018). The added pressure might also contribute to an increased risk of dropout among this population.

5. Conclusion and Recommendation

There were some limitations to this study. First, data from this study included students enrolled in a single course from one university. There is a possibility that the results of this study are not transferable to students enrolled at other universities or in other courses. However, given that data specific to SpES students is lacking in the literature, this could also be viewed as a strength of the study. Regardless, the results of the study do appear to reflect previously published findings. Second, it was not possible to identify students that transferred from La Trobe University to a bachelor's course at another institution. These students were classified as dropouts in this study although these students had not dropped out of bachelor-level study per se.

To conclude, we employed and presented a novel predictive model that classified region and ATAR as significant predictors of dropout for SpES students, suggesting that residing in "less advantaged" areas and a lower ATAR increases the risk of students dropping out of university. In addition, admission to a university course with a previous higher education course admission or via a non-ATAR pathway also increased the risk of dropout among higher-risk students. The present study presented a novel predictive model, and we endorse the use of CHAID to successfully predict student dropout with a high degree of accuracy. Administrative and teaching-focused university staff may adopt this novel approach to identify the relevant demographic indicators for their respective institutions or programs of study in the management of student attrition.

Acknowledgement

All authors listed have made a direct and intellectual contribution to this research.

References:

Anderton, R., Hine, G., & Joyce, C. (2017). Secondary school mathematics and science matters: Academic performance for secondary students transitioning into university allied health and science courses. *International Journal of Innovation in Science and Mathematics Education*, *25*(1).

Anderton, R. S. (2017). Identifying factors that contribute to academic success in first year allied health and science degrees at an Australian University. *Australian Journal of Education*, *61*(2). https://doi.org/10.1177/0004944117713321

Aulck, L. S., Velagapudi, N., Blumenstock, J. E., & West, J. D. (2016). Predicting Student Dropout in Higher Education. *ArXiv*, *abs/1606.0*.

Baik, C., Naylor, R., & Arkoudis, S. (2015). The first year experience in Australian universities: Findings from two decades, 1994-2014. In *Melbourne Centre for the Study of Higher Education*.

Baik, C., Naylor, R., Arkoudis, S., & Dabrowski, A. (2019). Examining the experiences of first-year students with low tertiary admission scores in Australian universities. *Studies in Higher Education, 44*(3), 526-538. https://doi.org/10.1080/03075079.2017.1383376

Beer, C., & Lawson, C. (2017). The problem of student attrition in higher education: An alternative perspective. *Journal of Further and Higher Education*, *41*(6). https://doi.org/10.1080/0309877X.2016.1177171

Bradley, D., Noonan, P., Nugent, H., & Scales, B. (2008). Review of Australian Higher Education. In *Final Report* (Issue December). www.deewr.gov.au/he_review_finalreport

Burke, P. J., Bennett, A., Bunn, M., Stevenson, J., & Clegg, S. (2017). *It's about time: Working towards more equitable understandings of the impact of time for students in higher education.* https://www.ncsehe.edu.au/publications/its-about-time-workingtowards-more-equitable-understandings-of-the-impact-of-time-for-students-in-higher-education/

Cherastidtham, I., Norton, A., & Mackey, W. (2018). *University attrition: what helps and what hinders university completion?* https://grattan.edu.au/wp-content/uploads/2018/04/University-attrition-background.pdf

Commonwealth Productivity Commission. (2019). *The Demand Driven University System: A Mixed Report Card* (Issue June). https://www.pc.gov.au/research/completed/university-report-card/university-report-card.pdf

Department of Education, S. and E. (2020). *Completion Rates of Higher Education Students - Cohort Analysis, 2005-2019*.

Edwards, D., & Mcmillan, J. (2015). Completing University in Australia: A Cohort Analysis exploring equity group outcomes. *Joining the Doots - Research Briefing Paper*, *3*(3).

Johnson (2008). Enrolment, persistence and graduation of in-state students at a public research university: does high school matter? *Research in Higher Education, 49*(8), 776-793. doi:10.1007/s11162-008-9105-8

Han, J., Kamber, M., & Pei. J. (2012). Data Mining: Concepts and techniques (3rd ed.). San Francisco, CA: Morgan Kaufmann.

Kemp, D., & Norton, A. (2014). *Review of the Demand Driven Funding System.*

Kirk, G. (2018). Retention in a Bachelor of Education (Early childhood studies) course: students say why they stay and others leave. *Higher Education Research and Development*, *37*(4). https://doi.org/10.1080/07294360.2018.1455645

Li, I. W., & Carroll, D. R. (2020). Factors influencing dropout and academic performance: an Australian higher education equity perspective. *Journal of Higher Education Policy and Management, 42*(1), 14-30. https://doi.org/10.1080/1360080X.2019.1649993

Li, I. W., & Dockery, A. M. (2015). Does School Socio-economic Status Influence University Outcomes ? *Australian Journal of Labour Economics*, *18*(1), 75–94.

Lykourentzou, I., Giannoukos, I., Nikolopoulos, V., Mpardis, G., & Loumos, V. (2009). Dropout prediction in e-learning courses through the combination of machine learning techniques. *Computers and Education*, *53*(3). https://doi.org/10.1016/j.compedu.2009.05.010

McKenzie, K. & Schweitzer, R. (2001) Who Succeeds at University? Factors predicting academic performance in first year Australian university students. *Higher Education Research & Development*, 20:1, 21-33, doi: 10.1080/07924360120043621

Niewoudt, J. E., & Pedler, M. L. (2021). Student retention in higher education: why students choose to remain at university. *Journal of College Student Retention; Research, Theory & Practice, 0*(0), 1-24. http://dx.doi.org/10.1177/1521025120985228