Exploring Retention Themes in Student Evaluation of Teaching Survey Comments of Engineering Subjects

Sam Cunningham, Rick Somers, James Chalmers and Sarah Dart
Queensland University of Technology
Corresponding Author Email: sam.cunningham@qut.edu.au

ABSTRACT

CONTEXT

Improving student retention is a key theme of the recent Australian Universities Accord. However, developing strategies to enhance retention has proven challenging due to numerous contributing and interacting factors.

PURPOSE

Student Evaluation of Teaching (SET) surveys are a common mechanism for gathering student feedback on their learning experiences. Open-ended responses in these surveys can offer valuable insight into aspects of student experience with potential links to retention. This research aims to answer the research question: "How are key themes of attrition and retention discussed in SET survey comments from engineering subjects?" This forms an initial step in enabling development of evidence-based actions for improving retention at the individual subject, course, or faculty level of engineering programs. To identify and extract insights at scale, large numbers of comments need to be considered.

METHODS

Several frameworks related to attrition and retention were explored for this study. Acknowledging that factors contributing to retention can be categorised as external, internal, and pre-university, this study focussed on the internal factors that institutions maintain control over. Dividing these internal factors down further, this research focussed on sub-factors of student support, student engagement, and sense of belonging. Natural language processing techniques (both dictionary and zero-shot models) were then used to identify the presence of each sub-factor within SET survey comments from engineering subjects at a large Australian university.

OUTCOMES

Models were applied to identify whether survey comments referenced the selected sub-factors. Student support was the most prevalent factor, followed by student engagement and sense of belonging. This research demonstrates the capability of applying a combined dictionary and zero shot machine learning approach to analysing SET survey comments at scale.

CONCLUSIONS

By identifying key elements of attrition and retention in SETs, key themes from student feedback can be identified and can inform faculty or university strategies. For further research, additional factors beyond SETs can be explored to link to students' demographics, and academic success.

KEYWORDS

Student Evaluation of Teaching Surveys, Attrition, Retention, Text Analysis

Introduction

Improving retention rates to ensure that students are given the best chance of completing their university degrees is a key theme of the recent Australian Universities Accord (Australian Government Department of Education, 2024). Attrition rates in Australian universities have tracked at approximately 15% for the last two decades (Australian Government Department of Education and Training, 2017). Developing effective strategies to enhance this rate has proven challenging given the complexity of numerous contributing and interacting factors that impact students' decisions about whether to continue in their university studies (Kerby, 2015; Tinto, 1975).

Student Evaluation of Teaching (SET) surveys are a common mechanism for gathering student feedback on their learning experiences (Spooren et al., 2017). Open-ended responses in these surveys can offer valuable and actionable insights from a large group of students into aspects of their learning experience (Dart & Cunningham, 2023; Laundon et al., 2023), with a potential link to retention. However, most research on SET survey comments has focused on identifying and visualising broad themes within datasets (Cunningham-Nelson et al., 2019; Cunningham-Nelson et al., 2020; Cunningham & Dart, 2021) and assessing bias (Heffernan, 2021; Tucker, 2014).

This study aims to answer the research question: "How are key themes of attrition and retention discussed in SET survey comments from engineering subjects?" This forms a preliminary step in enabling the development of evidence-based actions for improving retention at the individual subject, program, or faculty level of engineering programs.

Background

Student Attrition and Retention

A large body of research has worked to develop theoretical models that explain and predict student attrition. Foundational work by Tinto (1975) characterised attrition as being dependent on students' background attributes (such as skills and demographics) and their level of academic and social integration developed through university experiences. These variables in turn influenced students' goal and institutional commitment, ultimately determining their attrition outcome. Subsequent work has often built upon this model by introducing additional variables, such as those relating to the broader environment (e.g. financial support to study and transfer opportunities) (Bean & Metzner, 1985; Kerby, 2015) and the institution (e.g. policies, academic standards, and learning and teaching culture) (Kerby, 2015; Pascarella, 1980).

Research has sought to validate theoretical models by applying statistical analysis to real world data sets. An example of this performed by the Australian Government Department of Education and Training (2017) used statistical regression techniques to identify which selected variables contributed most significantly to student attrition. The study concluded that the "institution is a more important factor in explaining attrition than [student background factors including] the basis of admission, the student's ATAR [Australian Tertiary Admission Rank], type of attendance, mode of attendance or age" (Australian Government Department of Education and Training, 2018, p. 5). However, only 22.5% of variation in the attrition outcome was able to be explained by the selected variables (Australian Government Department of Education and Training, 2017), implying that other neglected factors played a substantial role in students' attrition decisions.

Kerby (2015) was selected as the key framework guiding this study, given its holistic, detailed, and contemporary theoretical model of student attrition. Kerby (2015) categorised attrition factors as external, pre-college, and internal. External factors are those beyond the scope of individual students and institutions. Pre-college factors relate to students' backgrounds and are principally beyond the control of universities. Institutional factors are those that universities maintain control over. Thus, they present the obvious area for universities to focus on when developing retention strategies, further supported by the conclusions of the statistical analysis noted above (Australian Government Department of Education and Training, 2017). A range of specific institutional factors have been identified, with those most frequently referenced in Kerby (2015) relating to *student* *engagement*, *student support*, and *sense of belonging*. These will be the focus of the present study. It is worth noting that student experiences at the 'educational interface' play a dominant role in students' development in relation to these factors (Kahu & Nelson, 2018), which suggests that the themes should be evident in students' SET survey comments.

Text Analysis of SET Survey Comments

The open-ended comment component of SET surveys provides an opportunity for students to provide feedback on their subject experience. These comments often contain useful feedback, however, analysis on a large scale can be challenging (Cunningham-Nelson et al., 2019). Natural language processing methods and machine learning approaches can support in the systematic and large-scale analysis of these comments. Various processes exist for analysing text comments. This includes traditional models trained on large amounts of data to complete a particular task, versus zero-shot models trained on a large amount of data, but where predictions can be made without seeing any examples, only types of categories (Pourpanah et al., 2022). Another common approach used when identifying themes in textual data is a dictionary match style approach (Albaugh et al., 2014). This involves a more rudimentary search of terms or words within responses.

Method

This study took place in a large metropolitan university in Australia. The SET survey for the university contained five quantitative Likert questions, and two open-ended questions asking for a free-text response. The first open-ended question asked students to comment on elements of the subject done well, and the second question asked about aspects of the subject that could be improved. Responses to these questions is the focus of the present research. Three years of data (2021 to 2023) from Semester 1 cohorts studying engineering subjects was analysed. This amounted to a total of 17,746 comments, with an average comment length of 128 and 288 characters, for the best aspects and need improvement questions respectively. Ethics was granted by Queensland University of Technology's Human Research Ethics Committee (approval 8266).

The text analysis approach utilised in this study combined both dictionary and machine learning (zero-shot model) approaches. Using an ensemble approach allowed identification of exact word matches through a dictionary and more complex relationships through the zero-shot model (Cunningham et al., 2022). Using a zero-shot model also provided the opportunity to investigate the sentiment of the comments. The following two pre-trained zero-shot models were used for this study. Note that both these models can be tested using the Application Programming Interface (API) via the following links:

- Sentiment zero shot model - https://huggingface.co/CouchCat/ma_sa_v7_distil
- Category / class zero shot model - https://huggingface.co/facebook/bart-large-mnli

Table 1 shows the list of terms used in the dictionary component of the approach. The terms were split into the three selected themes of support, engagement, and belonging as introduced above. The words were developed by the authors though collaborative brainstorming and discussion. As the focus of this research was on exploring applications of machine learning methods to SET comments, this process was deemed sufficient to generate initial word lists that would enable this testing. It should be noted that some terms represent the titles of specific services from the institution (e.g. 'HiQ' manages general enquiries, and the 'Oodgeroo Unit' provides support for Indigenous students). These institutional terms have been labelled with a hash ([#]). The dictionary matching meant checking if exact matches to these terms were identified in student responses. This was performed through an automated Python script.

Table 2 shows the chosen zero-shot terms and their respective thresholds used in the machine learning component of the approach. Like the dictionary, these terms were collaboratively developed by the authors and through consultation of the literature. This included consideration of the SEQuery tool, developed by the Social Research Centre who run the national Student Experience Survey, which was designed to classify words through a dictionary approach (Social

Research Centre, 2019). Thresholds were determined through preliminary testing with small numbers of comments and expected matches for each of the themes. The zero-shot model outputs both potential classifications and confidences, so the threshold can be used to refine the output. Testing terms for the zero-shot models was important, as zero-shot models are trained to have a broad understanding of English language. Consequently, some terms may be used in different ways for higher education or student survey contexts.

Table 1 – Dictionary of terms used; note “*” represents wildcard character

Student Support	Student Engagement	Sense of Belonging
hiq[#]	professional	respect
wellbeing	relevant	respectful
counselling	constructive	respected
retention	respect	respective
phone call	structure	motivat*
stimulate	pace	participat*
studiosity	work integrated learning	anxi*
library	wil	
dropin	organis*	
drop-in	frustra*	
oodgeroo[#]		

Table 2 – Zero-Shot Terms and Thresholds

Student Support		Student Engagement		Sense of Belonging	
Label	**Thres-hold**	**Label**	**Thres-hold**	**Label**	**Thres-hold**
services	0.950	active learning	0.900	peer relations	0.910
facilities	0.900	academic challenge	0.940	friendships	0.975
mentor	0.890	enriching experience	0.990	group tasks	0.925
support	0.890	connection	0.900	socialisation	0.940
manageable workload	0.970	collaborative class	0.890	learning communities	0.965
quality teacher	0.650	peer engagement	0.900	human connection	0.985
enrolment	0.910	teacher communication	0.880	community feeling	0.990
induction	0.950	enthusiasm	0.905	peer connection	0.970
discrimination	0.850	interactions	0.915	inclusion	0.940
		compulsory	0.940	peer support	0.975
		expectations	0.880	feeling like students matter	0.910
		interest	0.905	relevant to future	0.990
		boredom	0.800	interest	0.905
		happiness	0.600	relevant to career	0.930
		elation	0.800	career	0.860
		anxiety	0.795	challenge	0.980
		satisfaction	0.940	confidence	0.905
				satisfaction	0.940
				dissatisfaction	0.940
				sense of belonging	0.960

As an example of the classification process, one student comment stated “This unit was structured really well. Very applicable to industry. The drafts were very effective and helped with learning. Lots of work in the unit but I think it was all relevant.” This comment was flagged as reflecting the themes of: (1) Sense of Belonging, from zero shot model’s ‘career’ label; (2) Student Engagement, from zero shot model’s ‘enriching experience’ label and use of the dictionary terms ‘relevant’ and ‘structure*’); (3) and Student Support, from zero shot model’s ‘quality teacher’ label). The automated categorisation process was repeated for every comment.

Results

Results are first presented for overall sentiment of comments. This is followed by a summary of how the three institutional factors are represented, with further exploration of the zero-shot terms in each of the themes. The zero shot terms were chosen for detailed expansion in the results section, as they offer more abstract ideas, instead of the more defined dictionary terms. It was also clear that most identified themes came from the zero shot models, attributed to the relatively limited set of terms considered in the dictionary approach.

Figure 1 shows sentiment for responses to the open-ended questions in the survey as determined by the zero-shot model. As expected, when responding to “What was done well?”, negative sentiment was least represented in students’ responses, with neutral sentiment slightly more frequent, and positive sentiment accounting for the largest proportion. In contrast, for the question “What could be done better?”, positive sentiment was least represented, with neutral and negative sentiment accounting for similar, but larger proportions. Although this broadly reflects sentiment expectations, the specific proportions vary substantially in comparison between questions.

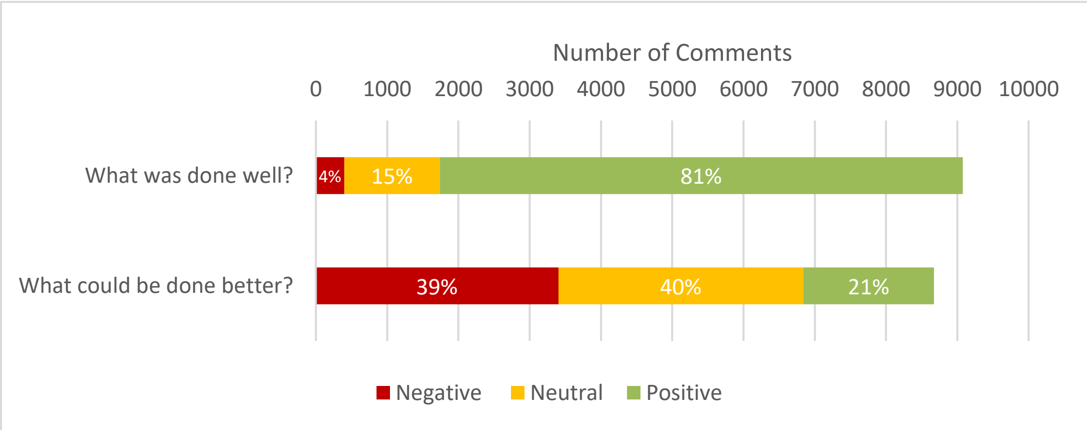

Figure 1 – Sentiment for Responses to Open-Ended Questions

Figure 2 shows how frequently the three institutional themes were mentioned when using the combined dictionary and machine learning zero-shot approach. Overall, *Student Support* was the most prevalent theme across both survey questions, followed by *Student Engagement*, and finally *Sense of Belonging*. Also worth noting is that all themes are more prevalent in response to the question “What was done well?”, despite the often longer average comment length in response to the question “What could be done better?”.

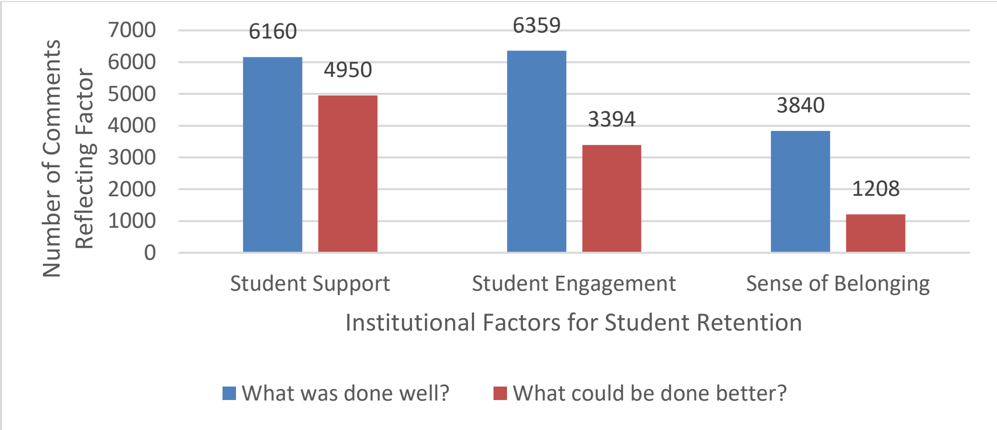

Figure 2 – Institutional Factors for Student Retention by Question using the Combined Dictionary and Machine Learning Approach

Table 3 shows the counts for each label under the theme of *Student Support*. Labels are ordered from most to least for total counts, and results are also shaded to highlight frequency. This theme is the most different of the three in terms of the size of label counts between the two questions students were responding to. The order of labels across both questions is similar, with '*quality teacher'* and '*support'* the most frequently mentioned across both questions. In terms of relative importance, *'quality teacher'* accounts for 72% of "What was done well" label counts, 36% of "What could be done better", and 66% of total counts. The next most prevalent label '*support'* accounts for only 15% of "What was done well" label counts, 27% of "What could be done better", and 17% of total counts.

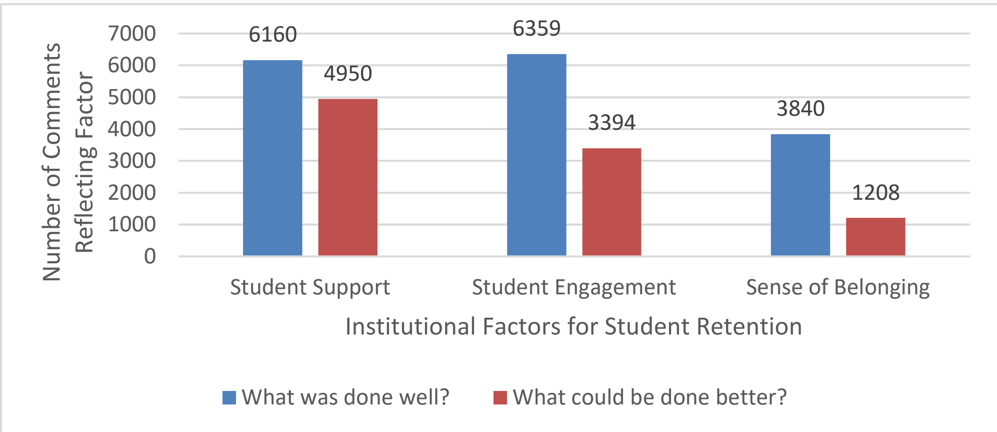

Figure 2 – Institutional Factors for Student Retention by Question using the Combined Dictionary and Machine Learning Approach

Table 3 – Number of comments mentioning Student Support labels by survey question

Label	"What was done well" Count	"What could be done better" Count	Total
quality teacher	5587	520	6107
support	1142	387	1529
manageable workload	702	126	828
facilities	208	137	345
discrimination	8	153	161
services	79	37	116
enrolment	16	58	74
mentor	36	13	49
induction	16	9	25

Table 4 shows the counts for each zero-shot label under the theme *Student Engagement*, again ordered from most to least total counts of labels. '*Enriching experience'* was the most prevalent label overall, with 96% of total counts for this label coming from the question "What was done well?". Like with *Sense of Belonging*, the importance of these labels varied with the context of the questions, with '*enriching experience'* accounting for 37% of "What was done well" label counts but only 28% of total counts. Similarly, '*academic challenge'* accounted for 29% of "What could be done better" label counts, but only 11% of total counts. While the order of labels shows more differences for *Student Engagement* (as compared to *Sense of Belonging*), again only key labels were dominant for each question: '*enriching experience'* for the positive aspects and '*academic challenge'* for the negative aspects of *Student Engagement*.

Table 5 shows the counts for each zero-shot label under the theme *Sense of Belonging*. The pattern of counts for *Sense of Belonging* appears similar to *Student Engagement*. Overall, '*dissatisfaction'* was the most prevalent label for the theme, with 95% of comments with this label coming in response to the question "What could be done better?". The most prevalent label in response to the question "What was done well?" was '*feeling like students matter'*, which has less than half the count of '*dissatisfaction'*. '*Dissatisfaction'* represents 47% of the responses to the question "What could be done better", but it accounts for only 26% of the total responses. Similarly, '*feeling like students matter'* constitutes 24% of the responses to "What was done well", but only 16% of the total responses. Therefore, the significance of these labels changes depending on the framing of the question. Additionally, for the least prevalent labels, the order of counts across the questions is very similar, suggesting that '*feeling like students matter*' and '*interest'* were the two labels most important to the positive aspects of *Sense of Belonging* and that '*dissatisfaction'* and '*challenge'* were most important to the negative aspects of the theme.

Table 4 – Number of comments mentioning Student Engagement labels by survey question

Label	"What was done well" Count	"What could be done better" Count	Total
enriching experience	4499	199	4698
academic challenge	515	1293	1808
active learning	1447	219	1666
interest	1051	284	1335
collaborative class	947	239	1186
connection	716	330	1046
teacher communication	501	316	817
interactions	534	261	795
peer engagement	530	152	682
satisfaction	474	171	645
elation	373	128	501
expectations	95	404	499
enthusiasm	231	36	267
happiness	108	125	233
boredom	30	169	199
anxiety	4	107	111
compulsory	24	32	56

Table 5 – Number of comments mentioning Sense of Belonging labels by survey question

Label	"What was done well" Count	"What could be done better" Count	Total
dissatisfaction	165	3399	3564

feeling like students matter	1526	672	2198
challenge	313	1258	1571
interest	1051	284	1335
relevant to future	402	361	763
relevant to career	504	238	742
satisfaction	474	171	645
group tasks	312	278	590
sense of belonging	308	83	391
socialisation	233	83	316
peer relations	173	75	248
peer connection	184	51	235
community feeling	192	10	202
inclusion	119	49	168
career	114	48	162
peer support	123	32	155
human connection	111	26	137
confidence	65	44	109
learning communities	19	8	27
friendships	13	6	19

Discussion and Concluding Remarks

This study investigated how three key themes related to retention were discussed in SET survey comments from engineering subjects. *Student Support* was most prevalent, followed by *Student Engagement*, and then *Sense of Belonging*. The study shows how natural language processing techniques can be used to identify these themes at scale. The breakdown of terms within each theme provides an insight into the most frequent concepts that are mentioned by students in the SET survey context.

Several limitations exist for this study. The zero-shot models have their own set of limitations, especially since they are trained on general English language instead of a higher education-specific context. Some concepts, such as sense of belonging, may be underrepresented due to the nature of the survey, whereby students were not directly asked about this theme as part of the survey. The development of the dictionary terms could also be completed in a more systematic way through robust consultation, for example by consulting several individuals and experts. Additionally, context clearly plays a very important role in the presentation of these results too. The survey context, subject context, and context in which the model is trained all influence the results.

These exploratory findings demonstrate the potential usefulness of applying these types of models to identify themes of retention in SET comments at scale, as well as the general capabilities of a combined dictionary and zero-shot model classification. Future work should work to improve classification accuracy through further development of term lists, categories, and model refinement. Additionally, future work should consider linking the SET survey textual data with identified themes to other institutional data, such as student engagement or achievement. This further analysis could provide more opportunities to explore how these complex measures interact with each other.

References

Albaugh, Q., Soroka, S., Joly, J., Loewen, P., Sevenans, J., & Walgrave, S. (2014). Comparing and combining machine learning and dictionary-based approaches to topic coding. *the Annual Comparative Agendas Project (CAP) Conference*, Konstanz, Germany, Australian Government Department of Education. (2024). *Australian Universities Accord Final Report*. Retrieved from: https://education.gov.au/australian-universities-accord/resources/final-report

Australian Government Department of Education and Training. (2017). Discussion Paper on Improving Completion, Retention and Success in Higher Education. Retrieved from: https://docs.education.gov.au/system/files/doc/other/final_discussion_paper.pdf

Australian Government Department of Education and Training. (2018). *Higher Education Standards Panel Final Report* - Improving Retention, Completion and Success in Higher Education. Retrieved from: https://docs.education.gov.au/system/files/doc/other/final_report_for_publishing.pdf

Bean, J. P., & Metzner, B. S. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research*, *55*(4), 485-540.

Cunningham-Nelson, S., Baktashmotlagh, M., & Boles, W. (2019). Visualizing Student Opinion Through Text Analysis. *IEEE Transactions on Education*, *62*(4), 305-311.

Cunningham-Nelson, S., Laundon, M., & Cathcart, A. (2020). Beyond satisfaction scores: visualising student comments for whole-of-course evaluation. *Assessment & Evaluation in Higher Education*, 1-16.

Cunningham, S., & Dart, S. (2021). What do students care about?: An analysis of topics impacting student evaluation survey results in engineering. Proceedings of the Research in Engineering Education Symposium and Australasian Association for Engineering Education Annual Conference (REES AAEE 2021),

Cunningham, S., Laundon, M., Cathcart, A., Bashar, M. A., & Nayak, R. (2022). First, do no harm: automated detection of abusive comments in student evaluation of teaching surveys. *Assessment & Evaluation in Higher Education*, *1*, 1-13.

Dart, S., & Cunningham, S. (2023). Using institutional data to drive quality, improvement and innovation. In M. D. Sankey, H. Huijser, & R. Fitzgerald (Eds.), *Technology Enhanced Learning and the Virtual University* (pp. 1-24). Springer

Heffernan, T. (2021). Sexism, racism, prejudice, and bias: a literature review and synthesis of research surrounding student evaluations of courses and teaching. *Assessment & Evaluation in Higher Education*, 1-11.

Kahu, E. R., & Nelson, K. (2018). Student engagement in the educational interface: understanding the mechanisms of student success. *Higher Education Research & Development*, *37*(1), 58-71.

Kerby, M. B. (2015). Toward a New Predictive Model of Student Retention in Higher Education: An Application of Classical Sociological Theory. *Journal of College Student Retention: Research, Theory & Practice*, *17*(2), 138-161.

Laundon, M., Cunningham, S., & Cathcart, A. (2023). Institutional approaches to evaluation of learning and teaching: A sector scan of Australasian universities. *Journal of Higher Education Policy and Management*, *45*(5), 511-528.

Pascarella, E. T. (1980). Student-faculty informal contact and college outcomes. *Review of Educational Research*, *50*(4), 545-595.

Pourpanah, F., Abdar, M., Luo, Y., Zhou, X., Wang, R., Lim, C. P., Wang, X.-Z., & Wu, Q. M. J. (2022). A Review of Generalized Zero-Shot Learning Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1-20.

Social Research Centre. (2019). *2018 Student Experience Survey Methodological Report*.Retrieved from: https://www.qilt.edu.au/docs/default-source/resources/ses/2018/2018-ses-methodological-report.pdf?sfvrsn=1f5a55a7_1

Spooren, P., Vandermoere, F., Vanderstraeten, R., & Pepermans, K. (2017). Exploring high impact scholarship in research on student's evaluation of teaching (SET). *Educational Research Review*, *22*, 129-141.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, *45*(1), 89-125.

Tucker, B. (2014). Student evaluation surveys: anonymous comments that offend or are unprofessional. *Higher Education*, *68*(3), 347-358.

Acknowledgements

Funding which supported this research project was provided by a Faculty of Engineering learning and teaching grant at the Queensland University of Technology. The authors would also like to thank the students who spend time providing feedback in their SET survey responses.

Copyright statement