Relations Between Graduates' Learning Experiences and Employment Outcomes: A Cautionary Note for Institutional Performance Indicators

This is a post-peer-review, pre-copyedit version of an article published in the International Journal of Educational and Vocational Guidance. The final authenticated version is available online at: http://doi.org/10.1007/s10775-021-09477-0

The recommended citation is:

Brown, J. L., Hammer, S. J., Perera, H. N., & McIlveen, P. (2021). Relations between graduates' learning experiences and employment outcomes: a cautionary note for institutional performance indicators. *International Journal for Educational and Vocational Guidance*. https://doi.org/10.1007/s10775-021-09477-0

Relations Between Graduates' Learning Experiences and Employment Outcomes: A Cautionary Note for Institutional Performance Indicators

Student learning experiences and graduate outcomes are, in part, evaluated using institutional performance indicators. Publicly available data from Australian government national surveys of university graduates was accessed. We explored whether the subscales of the Course Experience Questionnaire (CEQ): Good Teaching Scale (GTS); Graduate Skills Scale (GSS); and Graduate Qualities Scale (GQS) predicted graduates' decisions to take further studies, their employment status, and their overall satisfaction. For the specific subscales, GTS was found to predict graduates' overall satisfaction with their course experience, and the GSS and GQS subscales combined were found to predict negligible increases in employment outcomes. Our findings highlight the imperative for higher education leaders to critically examine discourse about the link between graduates' skills and qualities with their employment outcomes. We suggest future research focus instead on investigating alternative conceptions of employability that are concerned with the relations between psycho-social capital and employment outcomes.

Keywords: graduate attributes; skills; qualities; employment; employability; outcomes; course satisfaction; performance-based funding

Introduction

Recently the Australian Government released a report that recommends the introduction of performance based funding of universities linked to achievement of four performance indicators, namely: student success, equity group participation, graduate outcomes, and student experience (Commonwealth of Australia 2019). The purpose of this new funding arrangement is to "to ensure universities focus sufficient attention on the quality of their teaching and student support to ensure the best possible graduate outcomes" (p. 3). Allocation of funding would be provided for institutions improving performance across the four performance indicators (Commonwealth of Australia 2019). Two of the performance indicators of interest in this present investigation include the graduate outcome indicator and the student experience indicator. These two performance indicators are collected via the Graduate Outcome Survey (GOS) which includes items on destination outcomes (e.g., further study and employment), and the Course Experience Questionnaire (CEQ) scales which measures graduates' perception of teaching quality and development of generic skills and graduate qualities (QILT 2020). Thus, universities' scores on these indicators represent a "high stake" outcome with respect to performance-based funding.

This paper examines the implicit assumption in the Australian Government plan that quality of teaching and student experiences is linked to graduate outcomes as measured by achievement of employment. We begin by reviewing relevant literature on employability and employability skills followed by a brief history of the CEQ (Ramsden 1991), which provided the foundation for existing national surveys.

Employability

While performance funding mechanisms and the salience of the CEQ are issues specific to Australia, graduate employment outcomes and enhanced graduate employability are

expected of universities globally. In higher education, the teaching and assessing of graduate attributes, or generic skills and qualities, is explicitly connected with employability (Hammer et al. 2020; Barrie 2006). Despite continued disagreement and confusion in higher education about the concept of employability, the focus on the development of human capital via teaching and assessing graduate attributes remains dominant in higher education (Hammer et al. 2020; Oliver and Jorre de St Jorre 2018). However, many scholars take a broader view of employability that includes human, social, psychological, and cultural capital (Tomlinson 2017; Clarke 2018), and the importance of the process of career self-management to support the achievement of employment (Okay-Somerville and Scholarios 2017; Bridgstock 2009; Jackson and Bridgstock 2020). We take the definition of dispositional employability (Fugate et al. 2004): that employability is a psychosocial process that facilitates proactive behaviours of individuals to adapt or respond to opportunities or challenges in the labour market. The dimensions of dispositional employability include human and social capital, personal and career adaptability, and career identity (Fugate et al. 2004). Dispositional employability has been found to predict positive emotions and affective commitment to organisational change (Fugate and Kinicki 2008), job search intensity (Tomas and Maslić Seršić 2017; McArdle et al. 2007), and self-esteem and re-employment of unemployed workers (McArdle et al. 2007).

Learning Experiences and Perceived Skill Development

The Australian Government uses its Tertiary Education Quality and Standards Agency (TEQSA) to regulate and monitor the quality of higher education. The student experience and graduate outcomes are, in part, evaluated using data collected from annual national surveys of graduates' reports of satisfaction and outcomes. Established in the early 1990s, the CEQ (Ramsden 1991; Wilson et al. 1997) has provided an

enduring foundation for the Australian national surveys. The CEQ has been evaluated in other countries and appears to be a valid measure in different cultures (Byrne and Flood 2003; Stergiou and Airey 2012; Law and Meyer 2011; Richardson 1994). Since the original development of the CEQ, it has expanded beyond the original subscales measuring good teaching, clear goals, appropriate workload, and assessment, to include the *generic skills scale* (Wilson et al. 1997; Ramsden 1992). The new subscale would address a growing interest in graduates' employability and capacity for lifelong learning, and the ability to identify desirable generic skills for the workplace (Wilson et al., 1997, p. 36). Further subscales were added to the CEQ battery for *graduate qualities*, and a range of others related to student support and motivation (McInnis et al. 2001; Griffin et al. 2003). These progressive amendments to the CEQ, particularly the generic skills scale, affirmed an important shift in higher education toward an era of greater accountability and, moreover, graduate employability.

A key response from universities has been to embed graduate attributes, including generic or employability skills into university curricula. Recent research affirms the continued importance of this strategy (Hammer et al. 2020), as does Australia's 2015 Higher Education Standards Framework (Commonwealth of Australia, 2015). Continued sector consensus about the importance of graduate attributes, including employability skills, for student employability has influenced the development and use of government-funded quality assurance instruments, such as the CEQ. These instruments measure, amongst other things, graduates' perceptions of their generic skills development. The significance of this quality assurance strategy for the Australian Higher Education sector is amplified by Australian government plans to allocate additional funding on the basis of university performance on these indicators. Students' perceptions of their skill development and the quality of their university and

degree have been found to be important contributors to confidence in achieving employment outcomes (Rothwell et al. 2008; Rothwell et al. 2009; Álvarez-González et al. 2017).

The Present Research

The research presented here focuses specifically on the subscales of the CEQ, the Good Teaching Scale (GTS), Generic Skills Scale (GSS), and the Graduate Qualities Scale (GQS). Research analysing the GOS dataset in the past decade has found increased probability of obtaining an employment outcome for students with higher scores on the GSS and GSQ scales, albeit with small effect sizes (Jackson 2014), as well as significant differences between a range of demographic variables and scores on the GSS and GQS scales (Jackson 2016). However, Jackson's research did not conduct confirmatory factor analysis on the CEQ before testing the relations between the CEQ scales and employment outcomes, and since those studies were conducted the government contract to administer the CEQ and GOS has been awarded to another research team who made some changes to the way employment outcomes were measured. Therefore, it is timely to re-examine data from the GOS and CEQ to independently validate the psychometric properties of the CEQ and to test the relations between the CEQ and employment outcomes.

The present research had two main aims with respect to the validity of the CEQ. First, we aimed to determine the factor structure of the CEQ subscales by using recent GOS datasets (i.e., 2015 to 2017). Second, we aimed to test whether the CEQ subscales predicted graduates' overall satisfaction with studies, and graduate outcomes as measured by overall employment and further education. The presence of positive predictive relations would be additional evidence of their validity and, moreover, their utility as indicators of graduate outcomes.

Method

Participants

Graduates of Australian higher education institutions voluntarily provided their responses to the national GOS approximately four months after completing their degrees. The present study focused on the data for Australian citizens who were graduates of undergraduate bachelor's degrees completed at Australian institutions, in line with the proposed performance indicators which intends to use data from domestic undergraduates (Commonwealth of Australia 2019). The combined dataset included responses from N = 110685, participants. The median age was 24 years for the three surveys separately and combined; the means for age were M = 26.00 (SD = 8.2) in 2015, M = 24.84 (SD = 87.8) in 2016, M = 24.35 (SD = 7.5) in 2017, and M = 24.72 (SD = 7.7) for the three years combined. The participant characteristics appear in Table 1. In reading the participant characteristic table, it is important to note that the GOS commenced in 2015 as a limited pilot, thus, there is a relatively lower count of participants for that year. The similarity of proportions of these demographic variables is an indication of their equivalence across the survey years. The count and proportions for fields of study are summarised in Table 2.

Measures

GOS datasets are available for public access, analysis, and reporting from higher education institutions and the Social Research Centre (www.qilt.edu.au) which manages administration of the GOS on behalf of the government and institutions. Complete descriptive reports are available online at the public portal www.qilt.edu.au. We extracted the following data because they are the most relevant from the CEQ applicable to the aims of this research.

Course Experience Questionnaire Scales

Three scales from the CEQ: GTS; GSS; and GQS were included in this research. For each scale, the scores were calculated as an average of all items. Participants rated their level of agreement with each item by using a five-category Likert-type scale ranging from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). We used Cronbach's α coefficient to assess the internal consistency of the scales and found them to be high for each survey year.

Good Teaching Scale (GTS). The GTS measures graduates' perceptions of the quality of aspects of good teaching practice such as provision of feedback, making learning interesting, and motivating students to learn (Byrne and Flood 2003). A sample item is "the teaching staff normally gave me helpful feedback on how I was going". The GTS was found to have good internal consistency (α = .90 to α = .91) across the three years of data collected in this current study.

Graduate Skills Scale (GSS). The GSS measures graduates' perceptions of the extent to which their courses aided the development of generic skills including "problem-solving, analytic skills, teamwork, confidence in tackling unfamiliar situations, ability to plan work and written communication skills" (Wilson et al. 1997). A sample item is "the course improved my skills in written communications". The GSS was found to have good internal consistency (α = .85 to α = .86) across the three years of data collected in this current study.

Graduate Qualities Scale (GQS). The GQS measures graduates' perceptions of "whether the course generated higher-order outcomes and perspectives related to lifelong learning" (Graduate Careers Australia 2016). A sample item is "higher

education stimulated an enthusiasm for learning". The GQS was found to have good internal consistency ($\alpha = .87$) for each of the three years of data collected in this current study.

Outcome Variables

Three outcome variables included in the analysis were overall satisfaction, employment status, and enrolment in further study. These outcome variables are important indicators of higher education quality.

Overall Satisfaction. A single item measuring graduates' overall satisfaction with the course, using a Likert-type scale ranging from 1 = (*Strongly Disagree*) to 5 (*Strongly Agree*). The data analysis required this item to be converted to a binary categorical item, accordingly responses were recoded (0 = not in agreement, 1 = in agreement).

Employment status. The survey coded respondents as (1 = employed, 2 = unemployed, 3 = not in labour market). For the purposes of data analysis, a binary categorical item for labour force status was required. Responses were recoded to (0 = unemployed or not in labour force, 1 = employed).

Further study. A single categorical item requesting graduates to specify whether they had commenced further study (1 = full-time study, 2 = part-time study, 3 = not undertaking further study). The responses were recoded to a binary categorical item for data analysis (0 = no, 1 = yes).

Plan for Data Analysis

Our first aim was to test the factor structure of the CEQ subscales using recent data sets. We explored two potential measurement models: distinct correlated factors (CF) and

bifactor (BF) models. CF modelling assumes that the GSS, GQS, and GTS subscales are distinct but nonetheless correlate due to their items' similarities as indicators of graduate outcomes. On the other hand, BF modelling assumes that the items of the three subscales may be organised as one general factor (G-factor) and, alternatively, as three distinct subscales. In addition, we produced another set of models to determine if the GSS and GQS are measures of distinct constructs, as named (i.e., graduate skills, graduate qualities), or measures of the same construct. *Following this initial modelling,* we addressed the second aim of the research to explore relations of the CEQ subscales, employment status, further study, and overall satisfaction.

Analyses were conducted in a general latent variable modeling (GLVM) framework. Analyses were conducted using Mplus 8.0 (Muthén and Muthén 1998 - 2017). Solutions were estimated using the Weighted Least Squares Mean-and-Variance adjusted (WLSMV) estimator. All models were estimated while accounting for students' nesting within universities using the design-based correction of standard errors, operationalised via the complex design option in Mplus (Muthén and Muthén 1998 - 2017). Fit assessment was inclusive and involved an evaluation of fit indices, parameter estimates, and alternative models. As the χ^2 can be oversensitive to minor model misspecifications given even moderate-sized samples and contains a restrictive hypothesis test (i.e., exact fit), three approximate fit indices were considered: RMSEA, < .050 and .080 for close and reasonable fit; Comparative fit index (CFI) and Tucker-Lewis Index (TLI), > .900 and .950 for acceptable and excellent fit, respectively (Marsh et al. 2004). For nested model comparisons, because the adjusted χ^2 difference (MD $\Delta\chi^2$) test appropriate for the WLSMV estimator also tends to be sensitive to even trivial differences, changes in the CFI (ΔCFI) and RMSEA (ΔRMSEA) were primarily used. A decrease in the CFI and increase in the RMSEA of less than .010 and .015,

respectively, are indicative of support for a more restrictive model (Cheung and Rensvold 2002; Chen 2007). It should be noted that, under WLSMV in Mplus, the default link function is the probit function, which has an inherently unintuitive interpretation. Accordingly, to enhance interpretability of the probit regressions of the binary outcomes on the CEQ dimensions, we computed predicted probabilities from the probit regression coefficients using standard formulas (Muthén and Muthén 1998 - 2017).

Results

The results are organised in three sections. First, there are descriptive statistics and tests to exclude questions about potential confound effects associated with year of collection. Second, the measurement models are tested to confirm the best factor structure. Third, the models are tested to determine their utility to predict to the graduate outcome variables.

Descriptive Statistics

The means, standard deviations and correlations between variables is displayed in Table 3. It is possible that variations in the year of data collection (e.g., job market) would influence the modelling. Inspection of the descriptive statistics for the GTS, GSS, and GQS in Table 4 reveal some minor differences in mean levels. Thus, we had to exclude the possibility of bias due to survey year and employment status. Accordingly, we used a 3x3 analysis of variance (ANOVA) with the three levels of year of collection (2015, 2016, 2017) and three levels of employment status (employed, unemployed, not in the labour force) as the independent variables. The presence of a statistically significant interaction effect is an alert to potential bias. The large sample sizes increase the chance of finding a statistically significant effect; therefore, we included a measure of effect

size to ensure that any significant differences were practically meaningful, rather than merely due to the large sample size.

For the GTS, there was an interaction effect, $F(4) = 3.23$, $MS = 2.15$, $p = .012$. There were higher mean levels of GTS for those not in the labour market, followed by the unemployed, and then the employed respondents. The interaction effect arose from mean levels for the employed respondents being relatively higher in 2015. Nonetheless, the effect size partial $eta^2 = .000$ renders the interaction effect practically meaningless. There was a significant interaction for GSS, $F(4) = 2.88$, $MS = 1.28$, $p = .021$. The interaction emerged from differences between the means of those not in the job market, whose means were higher than those were employed or unemployed. Again, the effect size partial $eta^2 = .000$ rendered the interaction practically meaningless. Similar results were found for the GQS. There was a significant interaction effect, $F(4) = 4.60$, $MS = 2.12$, $p = .00$, due to equivalent means between those employed and unemployed, but relatively higher for those who were not in the job market. Again, the effect size partial $eta^2 = .000$ indicates nothing practically meaningful. The findings of statistically significant interactions effects are due to the large sample size. What matters most is that the zero effect sizes suggest that the differences are so marginally small as to be trivial and practically irrelevant. Thus, on the whole, there is no reason to suspect that the survey data were biased by labour market differences across the three years.

Measurement Models

First, we tested the correlated factor models which assume moderate correlations among three distinct factors; then we tested the bifactor models which assume a general factor mirrored by three separate factors. Table 5 shows the test statistics and fit indices for the measurement structures across both sets of models.

The three-factor CF-CFA provided an acceptable fit to the data; however, in this model, the correlation between the GSS and GQS factors was .989. Such a high correlation suggests the factors are dimensionally redundant (i.e., measuring the same construct). This finding tentatively supports specification of a revised models in which these two factors are combined into one. The BF-CFA of general factor and three specific factors did not produce to an admissible solution. Given the very high correlation between the GQS and GSS, we combined these factors into one. The test of the CF-CFA model resulted in an acceptable fit to the data, and, notably, no appreciable degradation in fit relative to the three-factor CF-CFA, Model Set 1 (ΔCFI = +.002, ΔTLI = +.004, ΔRMSEA = -.001).

In the two-factor CF-CFA, the correlation between the GTS and new GSQS factors was strong (r = .772), which indicated the possibility of a general factor underlying responses to all items. Accordingly, we used bifactor modelling because a BF-CFA model can accommodate construct-relevant multidimensionality due to the presence of both general and specific factors underlying response data. The test of the Model Set 2 BF-CFA model resulted in an excellent fit to the data, and an appreciable improvement in fit relative to Model Set 2 CF-CFA model (ΔCFI = +.028, ΔTLI = +.026, ΔRMSEA = -.009).

In the BF-CFA model, the general factor was very well-defined with uniformly strong standardised loadings (see Table 6). The strength of the general factor is remarkable given that the items of the CEQ were intended to index three distinct factors. Beyond the general factor, the standardised loadings for GSQS factor were also mostly moderate to strong. For the GTS factor, three of the six standardised loadings exceeded a value of .15, with two exceeding .40, suggesting adequate specificity beyond the general factor. Given the superior fit of the Model Set 2 BF-CFA model, the

well-defined general factor, and reasonably-well-defined specific factors, this model was retained for further analysis.

Prediction of Graduate Outcome Variables

To address the third aim of the research, we used the final BF-CFA model to test whether the CEQ subscales GTS and GSQS predicted graduates' employment status, further study, and course satisfaction outcomes. The overall model provided an excellent fit to the data, χ^2 (162) = 10077.496, $p < .001$, CFI = .976, TLI = .968, RMSEA = .024, 95% CI [.024, .025]. Figures 1, 2, and 3 show predicted probabilities of being employed, further study, and being satisfied overall with the course with respect to the G-Factor, GTS, and GSQS. Specific probit regression coefficients are reported in the narrative hereafter.

Relations Between CEQ Scales and Employment Outcomes

The G-Factor was negatively related to employment status (γ = -0.073, SE = .008, $p < .001$). Although statistically significant, this relation is practically negligible. Similarly, the relation between the combined GSQS factor and employment status was statistically significant (γ = 0.086, SE = .006, $p < .001$) but practically negligible. As shown in Figure 1, for increases in the G-factor from the mean to one standard deviation above the mean (+1SD), the predicted probability of employment (the closer the number is to 1, the more likely a graduate is to be employed) decreased from .871 to .855. For increases in the specific GSQS factor from the mean to +1SD, the predicted probability of employment increased from .871 to .888. These results indicate that gains in graduates' perceptions of their skills and qualities is likely to only marginally increase the probability of employment. The GTS factor was not significantly related to employment status (γ = 0.003, SE = .008, $p > .05$).

Relations Between CEQ Scales and Further Study

For further study, the G-factor was a significant and positive predictor ($\gamma = 0.131$, SE = .010, $p < .001$). For an increase in this factor from the mean to +1SD, the predicted probability of further study increased from .271 to .316 (see Figure 2). The specific GTS ($\gamma = -0.024$, SE = .013, $p > 05$) and GSQS ($\gamma = 0.013$, SE = .011, $p > 05$) factors did not significantly predict further study.

Relations Between CEQ Scales and Overall Satisfaction

The G-factor ($\gamma = 0.775$, SE = .004, $p < .001$) and specific GTS ($\gamma = 0.054$, SE = .007, $p < .001$) and GSQS ($\gamma = 0.325$, SE = .007, $p < .001$) factors were significantly predictive of course satisfaction. For increases in the G-factor from the mean to + 1SD, the predicted probability of overall satisfaction increased from .795 to .945. For increases in the specific GTS and GSQS factors from the mean to +1SD, the predicted probabilities of overall satisfaction increases from .795 to .810 and from .795 to .874, respectively (see Figure 3). Ramsden (1991) found significant correlations between the CEQ scales and overall satisfaction, but insufficient statistical support for the CEQ scales to be used as a proxy for overall satisfaction. Our findings suggest that with this most recent collection of responses to the CEQ, the G-factor, in particular, and the GTS and GSQS scales to a lesser extent, are significant predictors of overall satisfaction.

Discussion

The present research offers new findings about high-profile metrics currently used to appraise universities' performance, based on three consecutive years of national data. First, the CEQ (as measured by the general factor model) predicts graduates' overall satisfaction with their educational experience and enrolment in further study, which affirms the CEQ's utility as an indicator of graduates' satisfaction with teaching and

skill development. Second, we found that the measures of graduate skills and qualities—GSS and GQS—are collinear and are better modelled as a combination measure of the same factor (i.e., GSQS). This is an important finding as other research has treated the skills and graduate qualities scales as measuring different variables that have distinct effects on employment outcomes (Jackson 2016, 2014). Third, and most importantly, the GSQS marginally predicts employment outcomes but at such a low level of likelihood that its practical utility effect is trivial. In other words, these measures of graduate skills and qualities are measuring something, but that something has little direct effect on graduates' chances of being employed.

Our findings challenge the commonly accepted link between students' perceived qualities and skills development, and employment outcomes. The negligible relations between graduate skills and qualities and employment outcomes highlights the need to further investigate other predictors of employment outcomes. Self-perceived employability (Rothwell et al. 2008; Rothwell et al. 2009), which includes perceptions of knowledge, skills, and confidence in job search processes, has been found to predict employment status and job quality (Okay-Somerville and Scholarios 2017), however, the authors found that career self-management was a more important predictor of employment status and quality. This points to other opportunities for higher education institutions to develop students' employability through embedding careers and employability learning in the curriculum (Brown et al. 2019; Bridgstock et al. 2019).

Our findings also highlight the imperative for higher education leaders to critically examine calls from industry and government that universities should focus on the development of graduate skills for employability purposes. As an academic community, we should critically scrutinise claims about direct relations between the development of relevant graduate attributes and employability skills, and the

employment outcomes of graduates. Scholars have found that employers are more likely to hire graduates who have the required knowledge and technical skills foremost (Humburg and van der Velden 2015), whilst others view interpersonal qualities as more important (Dicker et al. 2018). In addition, Moore and Morton (2017) argue that it is not possible to systematically prepare students for the specific requirements they find themselves in as employers' requirements for the deployment of generic skills is highly contextualised.

The CEQ scales provide a valid measure of graduates' perceptions of the quality of teaching and development of skills and qualities. It is within the control of universities to implement strategies to improve teaching quality and make curricula changes to support students to develop generic skills. Indeed, universities over the past decade have invested in strategies to enhance student employability via embedding the teaching and assessing of graduate skills and attributes in the curriculum. Based on our findings, such efforts are likely to translate into higher graduate course satisfaction. However, attainment of learning outcomes is not the same thing as the achievement of employment outcomes. Yet current usage of these CEQ measures risks conflating graduate work-readiness, or the possession of skills and attributes, with graduate employability.

When it comes to improving the employment outcomes of graduates, the present findings are clear: there is a negligible relation between GSS and GQS scores and employment status (i.e., employed, unemployed or not in the job market). A key implication of the failure to find a predictive relation between the GSS, GQS and employment outcomes is that these measures should not be touted as proxies for graduate employability. Although "higher education provides no guarantee of actual employment" (Nilsson 2017), other strategies within the agency of universities can be

utilised to support graduate employment outcomes. For example, the inclusion of career development learning in the curriculum to increase students' career self-management skills (King 2004), initiatives that help students to develop social networks (Bridgstock 2019; Bridgstock and Jackson 2019), and through balancing the supply of enrolments in degree programs in anticipation of demand for qualifications in the labour market (Nilsson 2017).

Finally, our findings highlight significant methodological and ethical concerns with the planned use by the Australian Government of CEQ measures as key performance indicators for the tertiary sector. More specifically, the use of current measures of course experience to evaluate and reward high levels of institutional performance, or to provide this data to prospective students as a measure of quality employment outcomes for universities, in the absence of evidence, raises doubt about the use, perhaps misuse, of the GTS, GSS and GQS.

Limitations

A limitation to this study is the relatively short period of time after graduation which the survey is taken (i.e., four months), particularly given that 11.5% were unemployed and 11.4% were not in the labour market at the time of the survey. These graduates may have been searching for work that is substantively relevant to their qualification or dropped out of the labour market. Furthermore, a combined total of 38.5% reported that their qualification was "not at all important" or "not important" for their employment at the time of the survey. Thus, it is important to read the present findings with the caveat that the respondents may not have been in a personally suitable position. Moreover, the cross-sectional design means that it is not possible to identify causal pathways between variables. For example, it could be that employed graduates are more satisfied with their course experience than unemployed graduates because they achieved a desirable

outcome from their course.

Future Research

If GTS, GSS and GQS do not predict employment outcomes then what does? There is an emerging body of research into the psychosocial predictors of employment outcomes, based in the vocational psychology and organizational psychology literatures (Healy et al. 2020). Dispositional employability (Fugate et al. 2004; Fugate and Kinicki 2008) explores psychosocial capital (Koen et al. 2013; McArdle et al. 2007) that predicts employment-related behaviours and outcomes in graduates (González-Romá et al. 2018; Augustsson 2016; Lim et al. 2016). Research into the relations between psychosocial capital and employment outcomes would enhance stakeholders' understanding of what factors do and do not lead to employment outcomes.

GSS and GQS are self-assessment measures of graduates' self-perceptions of their skills and qualities, rather than measures of knowledge, skills, or other attributes sought by employers. Although the present findings indicate no substantive relation between GQS and GSS and graduates' actual employment outcomes, we note that the GSS and GQS have not been explored in relation to extant measures of graduates' self-perceived employability (Dacre Pool et al. 2014; Rothwell 2015). Based on Okay-Somerville and Scholarios (2017), a potential line of research would assess whether career adaptive behaviours mediate the relations between GSS, GQS and employment outcomes.

Conclusion

In many countries, employment outcomes for university graduates will no doubt persist as an important indicator of university education quality. There is, therefore, a pressing need for effective, transparent quality assurance measures that fairly assess related

performance and progress of higher education institutions. The present findings provide an appraisal of the validity of the CEQ and its application within QILT. Our findings are an informative contribution to debates about the conceptualisation and measurement of graduate employability—the CEQ subscales effectively do not predict graduates' employment status. These findings challenge university stakeholder discourse that conflates institutional performance against CEQ subscales with evidence of graduate employment outcomes and provide a platform for innovative thinking about employability by practitioners and policy leaders.

Acknowledgements

This work was supported in part through an Australian Government Research Training Program Scholarship.

References

Álvarez-González, P., López-Miguens, M. J., & Caballero, G. (2017). Perceived employability in university students: developing an integrated model. *Career Development International, 22*(3), 280-299, doi:10.1108/cdi-08-2016-0135.

Augustsson, G. (2016). The client company marginally utilises the knowledge of highly skilled temporary staff. *Journal of Education and Work, 29*(5), 519-539, doi:10.1080/13639080.2014.997681.

Barrie, S. C. (2006). Understanding what we mean by the generic attributes of graduates. *Higher Education, 51*, 215-241, doi:10.1007/s10734-004-6384-7.

Bridgstock, R. (2009). The graduate attributes we've overlooked: Enhancing graduate employability through career management skills. *Higher Education Research & Development, 28*(1), 31 - 44.

Bridgstock, R. (2019). Graduate Employability 2.0: Learning for life and eork in a docially networked world. In J. Higgs, G. Crisp, & W. Letts (Eds.), *Education for employability (Volume 1): The employability agenda*. Leiden, The Netherlands: Brill.

Bridgstock, R., Grant-Iramu, M., & McAlpine, A. (2019). Integrating career development learning into the curriculum: Collaboration with the careers service for employability. *The Journal of Teaching and Learning for Graduate Employability, 10*(1), 56-72, doi:10.21153/jtlge2019vol10no1art785.

Bridgstock, R., & Jackson, D. (2019). Strategic institutional approaches to graduate employability: Navigating meanings, measurements and what really matters. *Journal of Higher Education Policy and Management, 41*(5), 468-484, doi:10.1080/1360080x.2019.1646378.

Brown, J. L., Healy, M., McCredie, T., & McIlveen, P. (2019). Career services in Australian higher education: Aligning the training of practitioners to contemporary practice. *Journal of Higher Education Policy and Management, 41*(5), 511-533, doi:10.1080/1360080x.2019.1646380.

Byrne, M., & Flood, B. (2003). Assessing the teaching quality of accounting programmes: An evaluation of the course experience questionnaire. *Assessment & Evaluation in Higher Education, 28*(2), 135 - 145, doi:10.1080/0260293032000044491.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural equation modeling, 14*(3), 464-504, doi:10.1080/10705510701301834.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural equation modeling, 9*(2), 233-255, doi:10.1207/S15328007SEM0902_5.

Clarke, M. (2018). Rethinking graduate employability: the role of capital, individual attributes and context. *Studies in Higher Education, 43*(11), 1923-1937, doi:10.1080/03075079.2017.1294152.

Commonwealth of Australia (2019). Performance-based funding for the Commonwealth grant scheme: Report for the Minister for Education - June 2019. Canberra, Australia.

Dacre Pool, L., Qualter, P., & Sewell, P. J. (2014). Exploring the factor structure of the CareerEDGE employability development profile. *Education+ Training, 56*(4), 303-313, doi:10.1108/ET-01-2013-0009.

Dicker, R., Garcia, M., Kelly, A., & Mulrooney, H. (2018). What does 'quality' in higher education mean? Perceptions of staff, students and employers. *Studies in Higher Education, 44*(8), 1425-1441, doi:10.1080/03075079.2018.1445987.

Fugate, M., & Kinicki, A. J. (2008). A dispositional approach to employability: Development of a measure and test of implications for employee reactions to organizational change. *Journal of Occupational and Organizational Psychology, 81*(3), 503-527, doi:10.1348/096317907x241579.

Fugate, M., Kinicki, A. J., & Ashforth, B. E. (2004). Employability: A psycho-social construct, its dimensions, and applications. *Journal of Vocational Behavior, 65*(1), 14-38, doi:10.1016/j.jvb.2003.10.005.

González-Romá, V., Gamboa, J. P., & Peiró, J. M. (2018). University graduates' employability, employment status, and job quality. *Journal of Career Development, 45*(2), 132-149, doi:10.1177/0894845316671607.

Graduate Careers Australia (2016). Graduate course experience 2015: A report on the course experience perceptions of recent graduates.

Griffin, P., Coates, H., McInnis, C., & James, R. (2003). The development of an extended course experience questionnaire. *Quality in Higher Education, 9*(3), 259-266, doi:10.1080/135383203200015111.

Hammer, S., Ayriss, P., & McCubbin, A. (2020). Style or substance: How Australian universities contextualise their graduate attributes for the curriculum quality space. *Higher Education Research & Development*, 1-16, doi:10.1080/07294360.2020.1761304.

Healy, M., Hammer, S., & McIlveen, P. (2020). Mapping graduate employability and career development in higher education research: A citation network analysis. *Studies in Higher Education*, 1-13, doi:10.1080/03075079.2020.1804851.

Humburg, M., & van der Velden, R. (2015). Skills and the graduate recruitment process: Evidence from two discrete choice experiments. *Economics of Education Review, 49*, 24-41, doi:10.1016/j.econedurev.2015.07.001.

Jackson, D. (2014). Factors influencing job attainment in recent bachelor graduates: evidence from Australia. *Higher Education, 68*(1), 135-153, doi:10.1007/s10734-013-9696-7.

Jackson, D. (2016). Skill mastery and the formation of graduate identity in Bachelor graduates: evidence from Australia. *Studies in Higher Education, 41*(7), 1313-1332, doi:10.1080/03075079.2014.981515.

Jackson, D., & Bridgstock, R. (2020). What actually works to enhance graduate employability? The relative value of curricular, co-curricular, and extra-curricular learning and paid work. *Higher Education*, doi:10.1007/s10734-020-00570-x.

King, Z. (2004). Career self-management: Its nature, causes and consequences. *Journal of Vocational Behavior, 65*(1), 112-133, doi:10.1016/s0001-8791(03)00052-6.

Koen, J., Klehe, U.-C., & Van Vianen, A. E. M. (2013). Employability among the long-term unemployed: A futile quest or worth the effort? *Journal of Vocational Behavior, 82*(1), 37-48, doi:10.1016/j.jvb.2012.11.001.

Law, D. C. S., & Meyer, J. H. F. (2011). Adaptation and validation of the Course Experience Questionnaire in the context of post-secondary education in Hong Kong. *Quality Assurance in Education, 19*(1), 50-66, doi:10.1108/09684881111107753.

Lim, R. H., Lent, R. W., & Penn, L. T. (2016). Prediction of job search intentions and behaviors: Testing the social cognitive model of career self-management. *Journal of Counseling Psychology, 63*(5), 594-603, doi:10.1037/cou0000154.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers

in overgeneralizing Hu and Bentler's (1999) findings. *Structural equation modeling, 11*(3), 320-341.

McArdle, S., Waters, L., Briscoe, J. P., & Hall, D. T. (2007). Employability during unemployment: Adaptability, career identity and human and social capital. *Journal of Vocational Behavior, 71*(2), 247-264, doi:10.1016/j.jvb.2007.06.003.

McInnis, C., Griffin, P., James, R., & Coates, H. (2001). *Development of the course experience questionnaire (CEQ)*. Canberra, Australia: Department of Education, Training and Youth Affairs.

Moore, T., & Morton, J. (2017). The myth of job readiness? Written communication, employability, and the 'skills gap' in higher education. *Studies in Higher Education, 42*(3), 591-609, doi:10.1080/03075079.2015.1067602.

Muthén, L. K., & Muthén, B. O. (1998 - 2017). *MPlus Users Guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Nilsson, S. (2017). Employability, employment and the establishment of higher education graduates in the labour market. In M. Tomlinson, & L. Holmes (Eds.), *Graduate employability in context: Theory, research and debate* (pp. 65-86). London, United Kingdom: Palgrave Macmillan.

Okay-Somerville, B., & Scholarios, D. (2017). Position, possession or process? Understanding objective and subjective employability during university-to-work transitions. *Studies in Higher Education, 42*(7), 1275-1291, doi:10.1080/03075079.2015.1091813.

Oliver, B., & Jorre de St Jorre, T. (2018). Graduate attributes for 2020 and beyond: Recommendations for Australian higher education providers. *Higher Education Research & Development, 37*(4), 821-836, doi:10.1080/07294360.2018.1446415.

QILT (2020). Quality Indicators for Learning and Teaching. https://www.qilt.edu.au/.

Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The Course Experience Questionnaire. *Studies in Higher Education, 16*(2), 129-150, doi:10.1080/03075079112331382944.

Ramsden, P. (1992). *Learning to teach in higher education*. London, UK: Routledge.

Richardson, J. T. E. (1994). A British evaluation of the Course Experience Questionnaire. *Studies in Higher Education, 19*(1), 59-68, doi:10.1080/03075079412331382143.

Rothwell, A. (2015). Employability. In *APA handbook of career intervention, Volume 2: Applications.* (pp. 337-350): American Psychological Association.

Rothwell, A., Herbert, I., & Rothwell, F. (2008). Self-perceived employability: Construction and initial validation of a scale for university students. *Journal of Vocational Behavior, 73*(1), 1-12, doi:10.1016/j.jvb.2007.12.001.

Rothwell, A., Jewell, S., & Hardie, M. (2009). Self-perceived employability: Investigating the responses of post-graduate students. *Journal of Vocational Behavior, 75*(2), 152-161, doi:10.1016/j.jvb.2009.05.002.

Stergiou, D. P., & Airey, D. (2012). Using the Course Experience Questionnaire for evaluating undergraduate tourism management courses in Greece. *Journal of Hospitality, Leisure, Sport & Tourism Education, 11*(1), 41-49, doi:10.1016/j.jhlste.2012.02.002.

Tomas, J., & Maslić Seršić, D. (2017). Searching for a job on the contemporary labour market: The role of dispositional employability. *Scandinavian Journal of Work and Organizational Psychology, 2*(1), doi:10.16993/sjwop.9.

Tomlinson, M. (2017). Forms of graduate capital and their relationship to graduate employability. *Education + Training, 59*(4), 338-352, doi:10.1108/et-05-2016-0090.

Wilson, K. L., Lizzio, A., & Ramsden, P. (1997). The development, validation and application of the course experience questionnaire. *Studies in Higher Education, 22*(1), 33-53, doi:10.1080/03075079712331381121.

Table 1. Participant characteristics by year of collection and total years combined.

	Year of Collection							
	2015		2016		2017		Years Combined	
Participant characteristics	*n*	%	*n*	%	*n*	%	*n*	%
All participants	9385	100	52798	100	48502	100	110685	100
Male	3683	39.2	18238	34.5	16849	34.7	38770	35.0
Female	5702	60.8	34554	65.4	31651	65.3	71907	65.0
Indigenous	110	1.2	610	1.2	558	1.2	1278	1.2
Non-English speaking background	191	2.0	954	1.8	915	1.9	2060	1.9
Person with a disability	651	6.9	3292	6.2	3009	6.2	6952	6.3

Table 2. Count and Percentage of Participants' Degrees by Fields of Education

	Year of Collection							
	2015		2016		2017		Years Combined	
Fields	*n*	%	*n*	%	*n*	%	*n*	%
Natural and Physical Sciences	1046	11.1	6252	11.8	6219	12.8	13517	12.2
Information Technology	348	3.7	1448	2.7	1201	2.5	2997	2.7
Engineering and Related Technologies	556	5.9	1920	3.6	1402	2.9	3878	3.5
Architecture and Building	160	1.7	880	1.7	784	1.6	1824	1.6
Agriculture, Environmental and Related Studies	173	1.8	842	1.6	745	1.5	1760	1.6
Health	929	9.9	9571	18.1	10351	21.3	20851	18.8
Education	573	6.1	4566	8.6	4145	8.5	9284	8.4
Management and Commerce	2166	23.1	9260	17.5	7811	16.1	19237	17.4
Society and Culture	2782	29.6	13801	26.1	11978	24.7	28561	25.8
Creative Arts	652	6.9	4258	8.1	3866	8.0	8776	7.9

Table 3. Mean, Standard Deviation (SD), and Correlations for each Measure.

Measure	Mean	SD	GTS	GSS	GQS
GTS	3.55	.82			
GSS	3.89	.67	.668		
GQS	3.96	.68	.700	.859	
OSI	3.94	.67	.696	.710	.754

Note. GTS = Good Teaching Scale, GSS = Graduate Skills Scale, GQS = Graduate Qualities Scale, OSI = Overall Satisfaction.

Table 4. Mean and Standard Deviation (SD) for the Each Measure by Year of Collection and Labour Force Status.

Year	Labour Force Status	GTS		GSS		GQS	
		Mean	SD	Mean	SD	Mean	SD
2015	Employed	3.49	.83	3.87	.67	3.94	.68
	Unemployed	3.61	.83	3.92	.72	3.99	.75
	Not in labour force	3.74	.77	3.94	.72	4.08	.71
2016	Employed	3.53	.81	3.90	.65	3.97	.66
	Unemployed	3.59	.81	3.88	.71	3.94	.73
	Not in labour force	3.73	.78	3.97	.68	4.08	.68
2017	Employed	3.52	.82	3.87	.66	3.93	.68
	Unemployed	3.63	.82	3.89	.72	3.95	.74
	Not in labour force	3.73	.80	3.96	.67	4.08	.67

Note. GTS = Good Teaching Scale, GSS = Graduate Skills Scale, GQS = Graduate Qualities Scale.

Table 5. Model Fit Statistics and Indices for the CFA Models of the CEQ Data

Model	χ^2	df	CFI	TLI	RMSEA	RMSEA 90% CI	MD χ^2 (df)
			Model Set 1a				
CF-CFA	22487.915***	132	.937	.927	.042	.042, .043	-
BF-CFA[b]	-	-	-	-	-	-	-
			Model Set 2a				
CF-CFA	21694.653***	134	.939	.931	.041	.041, .042	[c]828.983 (2)***
BF-CFA	11686.651***	117	.967	.957	.032	.032, .033	[d]10019.621 (17)***

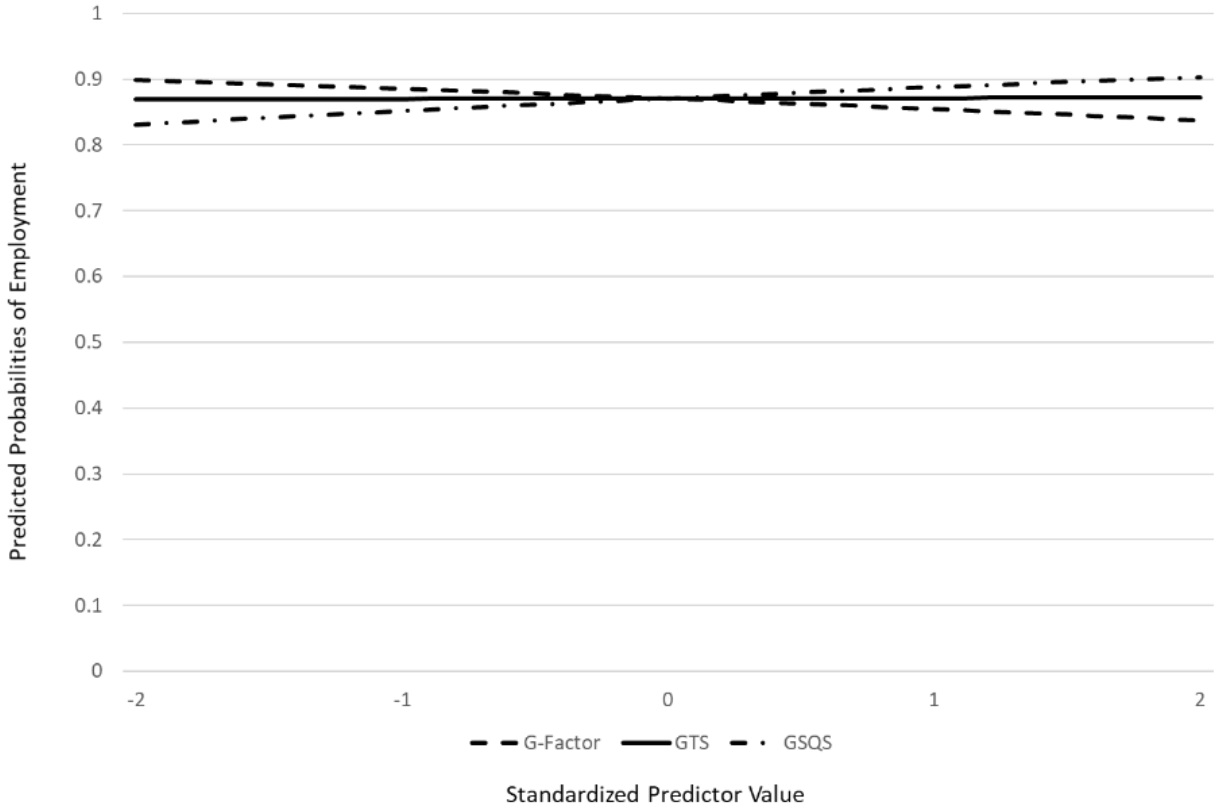
Note. [a]Model Set 1 refers to models in which GQS and GSS are distinct factors whereas Model Set 2 refers to models in which GQS and GSS have been collapsed. [b] The Model 1 BF-CFA did not converge. [c] This comparison is between the Model Set 1 and Model Set 2 CF-CFA models. *** $p < .001$. [d] This comparison is between the Model Set 2 BF-CFA and CF-CFA models.

Table 6. Completely Standardised Factor Loadings from the Retained BF-CFA Model

Item	G-factor	GTS S-Factor	GQS S-factor	h^2
GTS-01	.765	.454	–	.209
GTS-03	.795	.507	–	.112
GTS-10	.866	.101	–	.580
GTS-15	.854	-.034	–	.240
GTS-16	.858	-.027	–	.418
GTS-27	.778	.159	–	.378
GSS-06	.610	–	.219	.269
GSS-14	.651	–	.445	.263
GSS-23	.648	–	.515	.319
GSS-32	.569	–	.447	.315
GSS-42	.596	–	.615	.370
GSS-43	.567	–	.582	.440
GQS-11	.676	–	.355	.476
GQS-17	.733	–	.380	.306
GQS-30	.653	–	.365	.328
GQS-36	.627	–	.549	.267
GQS-40	.587	–	.573	.340
GQS-48	.589	–	.504	.399

Note. h^2 = model-based communality estimates. All loadings are statistically significant at $p < .001$, which is not unexpected given the very large sample.

Figure 1. Plot of predicted probabilities for employment status across a range of values for the (standardised) CEQ latent variables



.

Figure 2. Plot of predicted probabilities for further study status across a range of values for the (standardised) CEQ latent variables.

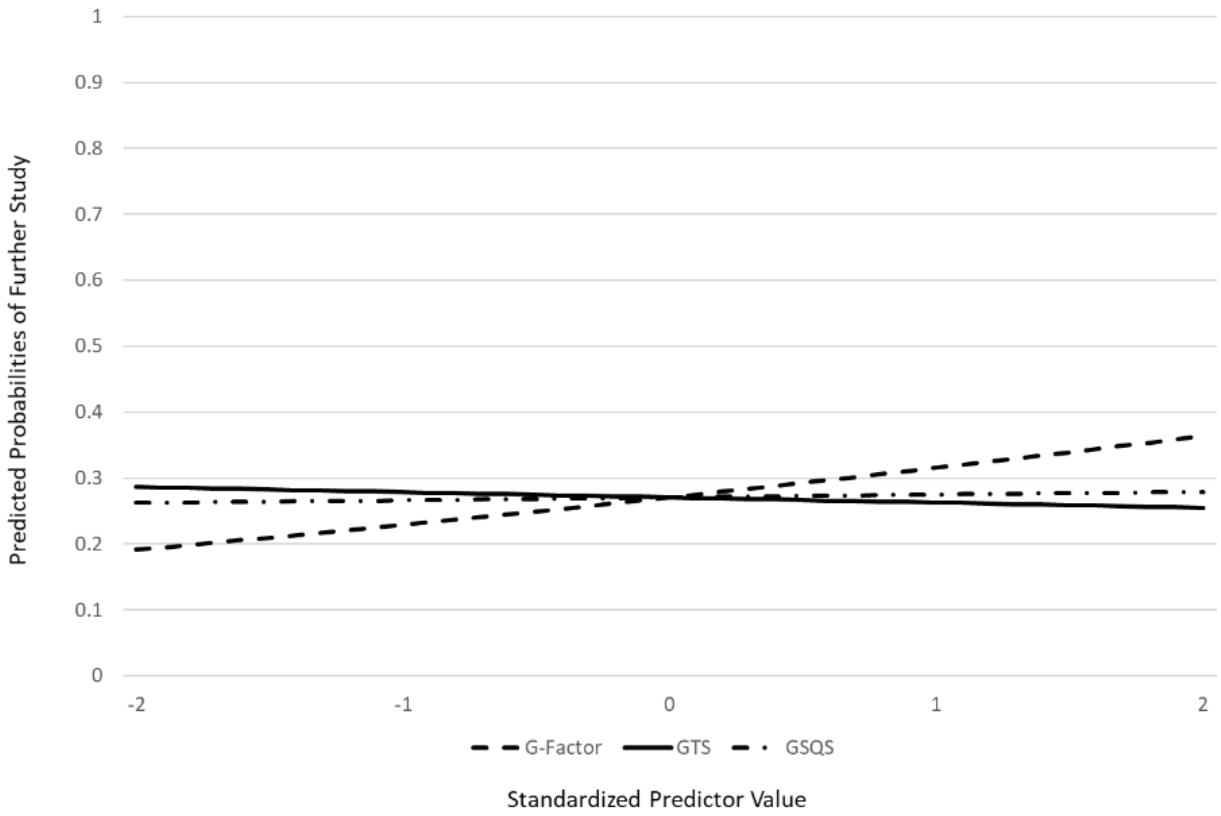


Figure 3. Plot of predicted probabilities for course satisfaction across a range of values for the (standardised) CEQ latent variables.