

Interweaving employability skills in ESL/EFL university direct entry integrated assessment tasks: A systematic review

ANGELA WINDSOR

University of Southern Queensland

ZOE HANCOCK

Southern Cross University

Integrated assessments are often used to measure English language readiness for university. As language assessment contexts are impacted by variables beyond the classroom, and technologised learners are increasingly striving to achieve “differentiated learning goals” (Peng et al., 2020, p. 88), the challenge is to design assessments that are more “person-centred” (Bensen, 2019, p. 60). One response is to interweave employability soft skills in English for academic purposes (EAP) direct entry program (DEP) integrated assessment. We conducted a systematic review of integrated assessment and soft skills assessment scholarship to inform integrated assessment design for supporting soft skills development in Australian university DEPs. The findings suggest that adopting a person-centred approach to the design and research of integrated assessments may more comprehensively support students’ adaptive communication capacities for 21st century ambitions.

Key words: integrated assessment, English for Academic Purposes (EAP), employability soft skills

Introduction

Direct entry programs (DEPs) provide second or foreign language (ESL/EFL) students with a pathway to university study (Roche & Booth, 2021) and so aim to facilitate a range of aspirations. For international students, gaining international qualifications is viewed as essential for securing future employment (Arkoudis et al., 2019; OECD, 2022) as English continues to maintain its lingua franca status, driving student mobility and education ambitions (Pillar & Bodis, 2022). For mobile ESL/EFL students who do not meet English language proficiency (ELP) requirements for entry to Australian university degrees, DEPs can help achieve these goals (Roche & Booth, 2021) as they

provide opportunities to enhance academic and ELP skills through an enculturation process (Dinneen, 2021).

An assessment instrument that has grown in popularity for measuring English language readiness for university study is the integrated assessment task (Knoch & Sitajalabhorn, 2013). These English language assessments are commonly found in EAP courses (Gokturk Saglam & Tsagari, 2022; Westbrook, 2023) and are considered to measure students' language skills more holistically (Plakans et al., 2018). They require synthesising reading and/or listening source input and communicating ideas in writing or speech in various combinations. In doing so, they demonstrate a range of academic skills such as summarising, paraphrasing, and referencing, in addition to their English proficiency standard (Haug, 2021). These characteristics of DEP integrated assessments support teachers to make important decisions regarding students' preparedness for future degree study.

Unlike standardised tests, DEP assessments are usually developed, validated, marked, and reported on within DEP contexts, making them artefacts that both emerge from but also influence interaction and activities in the instructed context (Wolfersberger, 2013). Therefore, integrated English assessment designers have a critical role to play that may require a mindset that is more inclusive and personalised, which reflects the pivot taking place more broadly in applied linguistics (Benson, 2019; Larsen-Freeman, 2018) as the field more adaptively responds to the dynamics of global mobility and education competitiveness (Arkoudis et al., 2019).

Person-centredness

One of the most cited reasons for using integrated assessment is authenticity (Plakans & Gebril, 2017). This is because in DEPs, integrated assessments usually reflect the type of tasks required in university programs, such as research-based essays, reports, oral presentations, notetaking, or group discussions. However, as DEP integrated assessment takes place in high-stakes environments, it is critical that construct validity (that assessments effectively assess what they claim to assess) underpins authenticity rationales so that improved student outcomes are actually realised, which has not always been the case for integrated writing (Plakans, 2015).

Target language use in DEP assessment

Incorporating a greater range of authentic target language use (TLU) domains beyond the academic may need to be considered to respond to the challenge of enhancing construct validity. For a long time, testing experts have emphasised the importance of considering TLU contexts for promoting authenticity (Hasrol, Zakaria, & Aryadoust, 2022). Messick (1996) argues that it is important that all variables relevant to the construct are included and measured in the assessment task. As assessment contexts

are porous (Larsen-Freeman, 2018), learners' targets for language use are more differentiated, interconnected, co-adaptive, and bespoke (Peng et al., 2020). Thus, assessments may need to more closely capture the learners' current and future TLU contexts, which is possible with the flexibility of integrated assessments.

Interweaving a broader range of TLU domains in integrated assessment beyond the future university context may promote person-centred integrated assessment practice. Ushioda (2011) suggests that orienting pedagogical activity towards more specific interests, goals, and ambitions to blend language assessment and life activities is critical for supporting motivation. In addition, rather than resulting only from explicit input made available in the classroom, language develops as students engage and interact with the resources and opportunities in their environment (Larsen-Freeman, 2023). Indeed, as integrated assessments closely correspond to teaching activity (Haug, 2021), only reflecting academic TLU in integrated assessments may now be considered detrimental as it limits learners' opportunities to develop and demonstrate the English skills required for participating in a complex world (Larsen-Freeman, 2018).

Soft skills in DEP assessment

The range of possible TLU contexts that could be interwoven with students' integrated assessments is vast. One possible TLU context that is likely to be a relevant part of students' identity building is soft skills for employment. Soft skills are those required for effective workplace participation beyond disciplinary knowledge or technical skill (Cotronei-Baird, 2020). Their development, which includes "complex problem solving, critical thinking, creativity, people management, cognitive flexibility, coordinating with others, emotional intelligence, judgement and decision making, service orientation, and negotiation" (Oxford University Press, 2020, p.8), is prioritised in higher education. Apart from the close association between English language skills and labour market entry (Humphreys, 2022), soft skills make students employable, and according to Cotronei-Baird (2020), employers' hiring decisions prioritise soft skills above technical or disciplinary expertise. In terms of research on their use in DEP integrated assessment, scholarship is rare. However, as language develops from meaningful use (Hiver et al., 2023), the diverse range of soft skills provides DEP test designers with consequentially valid opportunities for engaging students in authentic integrated English assessment tasks and supporting employability skills development from university pre-entry. Jackson and Bridgestock (2021) assert that employability development "is life-long, and life-wide" (p. 724), suggesting that introducing soft skills in DEPs may be just as important as developing them in degree studies.

The present study

The discussion above indicates that there is a benefit to extending the current approaches to DEP assessment to respond to differentiated learning goals. Embedding employability soft skills may be one way to address this issue. As far as we are aware, research investigating the embedding of employability within English as L2 integrated assessments is rare in the field. The aim of this paper is to address this gap through a systematic review of research in two disparate areas of scholarship, that is English as L2 integrated assessment and generic employability skills assessment in Australian higher education literature. Systematic reviews support researchers to uncover high quality research on an explicit topic or question to evaluate practice (Hiver et.al., 2022a). Our systematic review only considers studies that investigated integrated English language assessments and employability soft skills assessments authentically, or in-situ, to platform the holistic context of assessment. It is guided by the following research question and two sub-questions:

1. How can integrated assessment in university DEPs support the development of employability skills?
 - i. How are integrated English skills assessed in ESL/EFL university DEPs?
 - ii. How are generic employability skills in Australian university degree programs assessed?

Method

Literature identification

The search was conducted in EBSCOhost Megafile Ultimate, including Education Research Complete and Academic Search Ultimate, Scopus, ERIC, Taylor and Francis, Springer, Wiley Online Library, Informit, and Proquest One Academic databases. We used all databases to search the two distinct subject areas: DEP integrated assessment and soft skills assessment in Australian universities. We independently searched for studies that spanned a ten-year period from 2012 to 2022 that were published in English. Table 1 shows the range of terms used in a variety of search combinations. Systematic reviews are typically reproducible and comprehensive (Avenali et.al., 2023). When a systematic review topic is rare in the literature with disparate key themes, we recognise that it may be time-intensive. This may deter replication. We therefore recommend future systematic reviews limit the search to one or two key search terms. A total of 2,453 study titles and abstracts were independently imported by the researchers for screening into Covidence, and Covidence removed 716 duplicates. Covidence is a web-based collaboration software platform that streamlines the production of systematic and other literature reviews (Veritas Health Innovation, 2023). We used Covidence software to guide and record

the systematic review and discrepancy resolution process. Covidence was not used for data extraction.

Table 1.
Search Terms

Subject area 1:
"integrated assessment OR integrative" AND (population)
"reading to write" AND (population)
"reading to speak" AND (population)
"reading and listening to write" AND (population)
"reading and listening to speak" AND (population)
"authentic assessment" AND (population)
"source-based writing" AND (population)
"source-based speaking" AND (population)
"writing from sources" AND (population)
"speaking from sources" AND (population)
"multi-modal assessment" OR "multimodal assessment" AND (population)
(population) = (EAL OR "English as an additional language")/ (CALD OR "culturally and linguistically diverse") "international students" "ESL""TESOL" EAL/D" "EFL" "LBOTE OR language background other than English
Subject Area 2:
"soft skills" OR employability) AND (university OR "higher education") AND Australia AND assess*

Screening

To be included in this review, the following criteria were applied:

Subject area 1:

1. Population (see Table 1)
2. Non-standardised in-situ integrated assessment
3. DEP
4. University
5. Peer-reviewed academic journal article
6. In-situ empirical research

Subject area 2:

1. Non-discipline and non-technical based employability soft skills
2. Australian university
3. Peer reviewed academic journal article
4. In-situ empirical research

We excluded articles related to standardised test preparation courses and those that focused solely on rubric design and/or validation to the exclusion of test stakeholder experiences. We also excluded studies focused on work integrated learning (WIL) as these studies involve disciplinary employability skills. To begin, the authors independently assessed 1,737 title and abstracts against the inclusion criteria and met to discuss any discrepancies until full agreement was reached. 166 studies satisfied the initial screening, and the full texts of these studies were uploaded into Covidence. Next, the researchers independently screened each full text against the criteria and, also applied quality inclusion criteria. We used journal metrics to assess quality and only included studies that were published in Q1 or Q2 journals according to SCImago Journal Rank Indicator (SJR). Even though valuable insights are gained from articles in other journals and grey literature, peer review in Q1 and Q2 journals indicate a higher level of impact (Western Sydney University Library, 2023), and our aim at the outset was to build knowledge of best practice in recognition of the high stakes involved in DEP integrated assessment contexts. Similar decisions regarding high quality selection in the second language acquisition publications have been made by Zhang (2020). However, we acknowledge the limitation this has on the exhaustiveness of this review and recommend that a wider scope is included for more defined future systematic reviews. The authors met to discuss any inclusion and quality screening discrepancies until full agreement was reached. Twenty studies were eligible for inclusion in this review. The authors then independently conducted backward and forward searches based on the final twenty studies, but no additional studies were identified. See Figure 1 for an overview of the systematic review process.

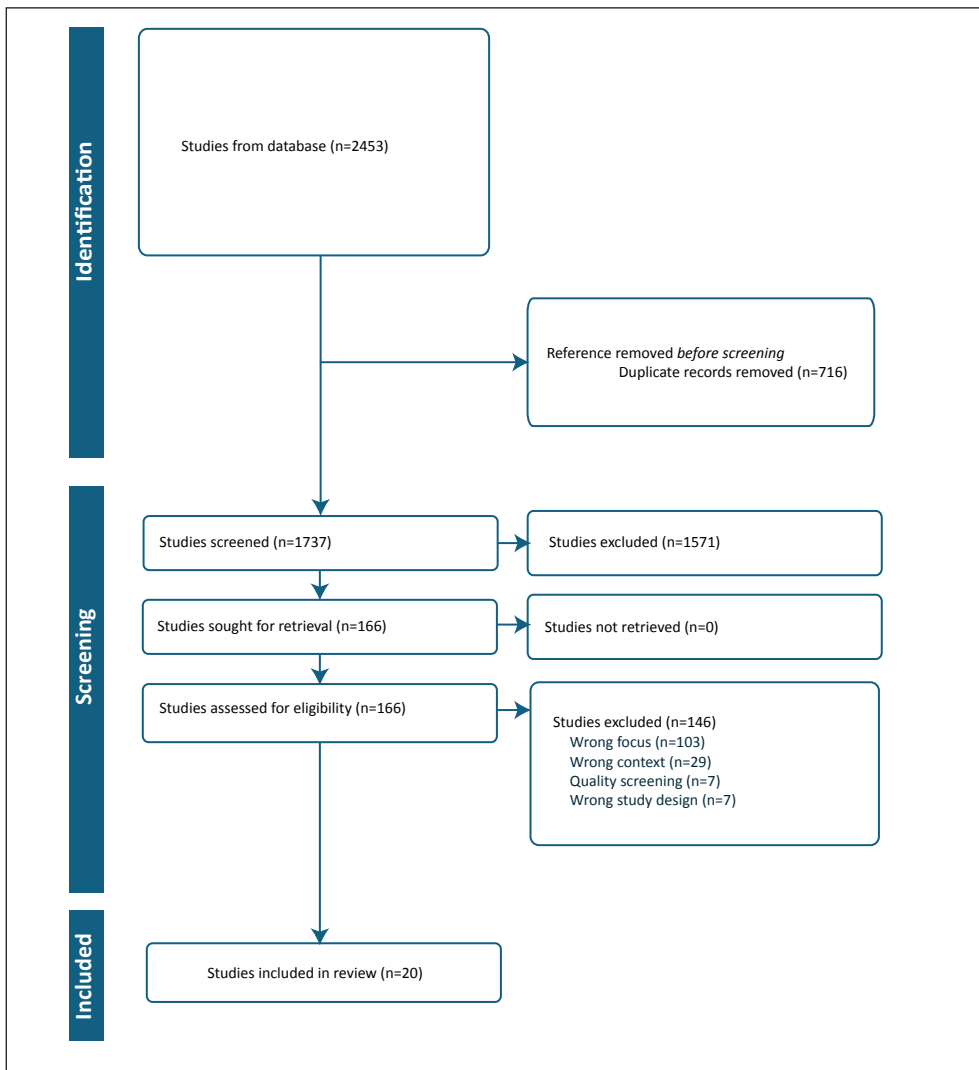


Figure 1. PRISMA Flowchart of systematic review process (Source: Veritas Health Innovation 2023)

Coding and data extraction

The authors met to discuss the descriptive categories of information required from the 20 papers for addressing the research questions, and data tables were created. Once full agreement was reached on the codes, data were then extracted independently. Finally, the authors met to systematically review the data extracts and reach 100% agreement. A third researcher independently extracted data from seven of the final 20 articles to validate the qualitative extracts to enhance the inter-rater reliability.

Findings

Integrated assessments in ESL/EFL DEPs

In terms of research question one, surprisingly, our extensive search of the literature found only seven studies that investigated integrated assessment in-situ (see Appendix A). Clearly, this is an emerging area of research that has the potential to extend insight into the role of context in integrated assessment performance. In terms of which skills were integrated, four of the seven studies investigated reading to write assessments, two focused on reading and listening to speak, and the remaining studies involved listening to write and reading and listening to write. Further evidence of prioritising academic reading and writing is also clear from the data describing the type of integrated assessment (see Appendix A). Apart from the need for further integrated assessment studies investigating in-situ performance in a range of speaking and listening genres, in-situ studies of DEP assessments that integrate all four macro skills were not found in the literature and highlights another significant gap in the literature for future research.

In terms of their nature, Appendix A outlines the key design and process characteristics. Detailed and multi-stage instructions are a feature of integrated assessments (see Appendix A) and indicates integrated assessment designers' efforts to enhance authenticity and focus on students' academic TLU. Finding detailed instructions in the data is not surprising as this reflects the number of stages and processes involved in preparing for common university assessment tasks that DEPs aim to build student capacity in. This also indicates that students should be able to manage employability soft skills assessment as these also would require detailed instructions. Evidence of the complexity of employability soft skills assessment is supported by the task data in Appendix B. Scaffolds are also a feature and are embedded into the steps and stages of the assessment tasks, and despite the small number of integrated assessments reviewed, a variety of scaffolds were found. Notetaking and brainstorming templates and background reading prior to assessment are integral in three of the studies, personal notes are embedded in two of the studies, and cloze notes, dictionaries, pre-writing questions, instruction reminders, and listening to test audio twice are scaffold artefacts or processes that also feature. However, it is interesting that only two of the studies involved students interacting with each other. As many soft skills involve interacting with others, addressing this gap in pedagogy and research seems an ideal way to not only build a range of diverse and important English communication skills and, therefore employability capability.

Furthermore, the data indicates that students require a wide variety of intertwined skills to complete assessments, with research, synthesis, and reading found to be key skill components. This also highlights the complexity of integrated English

constructs and, perhaps, explains the prevalence of integrated assessment studies whose purpose was to investigate rater and test moderation processes in the studies reviewed (see Appendix A). However, while it could be assumed that the nature of the academic tasks described in the seven studies involves evaluating students' critical thinking development and, to a certain extent, complex problem solving and creativity, other important soft skills are not explicitly assessed in these studies.

In terms of personalisation, it is not clear what choices and input students had regarding the features of the assessment or whether test designers considered student priorities. While it could be argued that relevance is evident in five of the studies in Appendix A as the topics either relate to the curriculum and/or are topically pertinent as human citizens, this argument views students as a population of learners rather than recognising their individuality and the potential role interweaving personal preferences into assessment could have on their language outcomes. Turning to how the integrated English tasks were assessed, two important features are evident. Firstly, each rubric assesses different skills, with English language and academic skills under focus. This reflects one of the advantages of DEP assessment design compared with standardised assessment, which is the opportunities they offer for adapting or personalising the assessment experience to learners and contexts. However, like the topics of assessment discussed previously, it is not clear to what extent the test designers used this opportunity to personalise rubrics to the DEP or course.

Soft skills assessment in Australian university degrees

To inform strategies for building soft skills capacities, 13 studies met the inclusion criteria for this purpose, and Appendix B provides an overview of types of soft skills assessment task. The research revealed there are a variety of ways to assess soft skills development (see Appendix B), and importantly, many of the studies provide strategies for interweaving and aligning this development within the curriculum, making these skills explicit to students, and building interactive communication skills through teamwork. Also, facilitated reflection and self-directed development assessment activities suggests that students' current capacities and/or needs were considered to facilitate soft skills development.

Turning to the features of these generic employability assessments, the purposes and contributions of the studies noted in Appendix B show that compared with integrated English assessments, employability soft skills assessments seem to be richer and interwoven with a more diverse range of holistic interactions, stakeholders, and activities, which appear to allow the assessment activity to reflect a wider context beyond the classroom. Aspects such as students, academics, university policies, disciplines, society and culture, along with students' affective variables,

interaction, and connection with future ambitions feature and provide strategic evidence for enhancing the person-centredness of integrated English assessments. The data also shows that a diverse range of soft skills are assessed for a range of purposes beyond communication skills and summative and placement assessments, respectively, including formative, personalised, and self-assessment functions. This range of skills and purposes contrasts with the evidence of assessment purpose for integrated assessment (see Appendix A) and highlights the pedagogical and research opportunities that interweaving employability skills into integrated assessment designs may provide for actively engaging students in real-world assessment tasks.

Discussion

This systematic review aimed to investigate how integrated assessments in DEPs could interweave employability soft skills development to support best practice in DEP integrated assessment. Through an in-situ authentic lens, we investigated how integrated assessments in DEPs are conducted and how Australian universities assess soft skills.

In terms of our first research question, the findings reveal that DEP integrated assessments are detailed, clearly staged, and scaffolded and they integrate a range of key academic skills. This sophistication means DEP integrated assessments are already conducive for incorporating employability soft skills development. However, it is clear that there is an opportunity for practitioners and researchers to more closely respond to student ambitions, such as gaining skills to secure future employment and enabling global mobility. Prioritising the person concurs with researchers' acknowledgments more generally (Benson, 2019; Dubreil & Thorne, 2017; Larsen-Freeman, 2018; Peng et al., 2022) regarding the need to see more of the person in the learner (Ushioda, 2011) and to realise the possibilities that language learning experiences could provide for both reflecting, supporting, and motivating students to navigate the variety of social and employment contexts they interact within. Achieving this would entail increasing the number of studies investigating in-situ integrated assessment activity across the different standard integrated assessment types in general. However, what may be more important is to extend tasks and scholarship into more diverse target language use contexts, in which employability skills is just one example, where definitions of success do not necessarily need to be defined according to proficiency level targets (Larsen-Freeman, 2018). It would seem that there is a natural role for DEP integrated assessment to incorporate soft skills development as DEP students are seeking to enter a variety of degrees. Therefore, focusing on soft skills in DEPs could be just as important as developing them during degree studies.

In terms of research question two, compared with DEP integrated assessments the

soft skills studies demonstrate a range of richer assessment experiences, assessment strategies, and assessment purposes that can be interwoven across curriculum, which also seem more personalised through self-assessment and self-directed assessment opportunities. Thus, they provide DEPs with a fruitful place to begin interweaving soft skills development into curriculum and integrated assessment design for ESL/EFL learners. This pedagogical recommendation mirrors researchers' calls for taking more deliberate transdisciplinary research action for extending impact and sophistication in applied linguistics (Hiver et al., 2022b). Furthermore, greater innovativeness supports the second language development literature where more social assessment designs are advocated (Larsen-Freeman, 2018).

CONCLUSION

DEP integrated assessments facilitate a range of ESL/EFL student aspirations. While they may be mostly associated with opportunities to demonstrate levels of academic English proficiency development, the inherent association between learning and assessment means that DEP integrated assessment activities are necessarily intertwined with a range of other student ambitions including the acquisition of an international education, future employment, and global mobility. DEP integrated assessments are not self-contained and isolated from the contexts in which they are interwoven. Rather, each student's assessment outcome is influenced by and are influential within the DEP higher education system. This systematic review has provided evidence that best practice in DEP integrated assessment recommends interweaving employability skills with academic skills development. In so doing, it suggests that DEP practitioners and researchers can innovatively extend integrated assessment practice and knowledge by adopting a person-centred perspective in order to closely meet the needs of each ESL/EFL DEP student. It should also be clear that adopting a person-centred lens works hand in hand with supporting students to reach their learning ambitions for 21st century success.

REFERENCES

- Arkoudis, S., Dollinger, M., Baik, C., & Patience, A. (2019). International students' experience in Australian higher education: Can we do better? *Higher Education*, 77, 799-813. <https://doi.org/10.1007/s10734-018-0302-x>
- Avenali, A., Daraio, C., Di Leo, S., Matteucci, G., & Nepomuceno, T. (2023). Systematic reviews as a metaknowledge tool: Caveats and a review of available options. *International Transactions in Operational Research*, 30(6), 2761–2806. <https://doi.org/10.1111/itor.13309>

- Bennett, D., & Ananthram, S. (2022). Development, validation and deployment of the employABILITY scale. *Studies in Higher Education, 47*(7), 1311-1325. <https://doi.org/10.1080/03075079.2021.1888079>
- Bennett, D., Knight, E., Dockery, A. M., & Bawa, S. (2020). Pedagogies for employability: Understanding the needs of STEM students through a new approach to employability development. *Higher Education Pedagogies, 5*(1), 340-359. <https://doi.org/10.1080/23752696.2020.1847162>
- Benson, P. (2019). Ways of seeing: The individual and the social in applied linguistics research methodologies. *Language Teaching, 52*(1), 60-70. <https://doi.org/10.1017/S0261444817000234>
- Cotronei-Baird, V. S. (2020). Academic hindrances in the integration of employability skills development in teaching and assessment practice. *Higher Education, 79*, 203-223. <https://doi.org/10.1007/s10734-019-00405-4>
- Crosthwaite, P., Boynton, S., & Cole III, S. (2017). Exploring rater conceptions of academic stance and engagement during group tutorial discussion assessment. *Journal of English for Academic Purposes, 28*, 1-13. <http://dx.doi.org/10.1016/j.jeap.2017.04.004>
- Crosthwaite, P. R., & Raquel, M. (2019). Validating an L2 academic group oral assessment: Insights from a spoken learner corpus. *Language Assessment Quarterly, 16*(1), 39-63. <https://doi.org/10.1080/15434303.2019.1572149>
- Cushing, D. F., Pennings, M., Willox, D., Gomez, R., Dyson, C., & Coombs, C. (2019). Measuring intangible outcomes can be problematic: The challenge of assessing learning during international short-term study experiences. *Active Learning in Higher Education, 20*(3), 203-217. <https://doi.org/10.1177/1469787417732259>
- Dinneen, C. (2021). Students' use of digital translation and paraphrasing tools in written assignments on direct entry English programs. *English Australia Journal, 37*(1), 40-51. <https://files.eric.ed.gov/fulltext/EJ1341751.pdf>
- Dubreil, S., & Thorne, S. L. (2017). Social pedagogies and entwining language with the world. In S. Dubreil, & S. L. Thorne (Eds.), *Engaging the world: Social pedagogies and language learning* (pp. 1-11). Cengage.
- Ghanbari, N., & Abdolrezapour, P. (2021). Using emotional intelligence in an EFL integrated writing assessment. *Studies in Educational Evaluation, 70*. <https://doi.org/10.1016/j.stueduc.2021.101017>
- Gokturk Saglam, A. L., & Tzagari, D. (2022). Evaluating perceptions towards the consequential validity of integrated language proficiency assessment. *Languages, 7*(1), 65. <https://doi.org/10.3390/languages7010065>

- Granado-Peinado, M., Cuevas, I., Olmos, R., Martin, E., Casado-Ledesma, L., & Mateos, M. (2023). Collaborative writing of argumentative syntheses by low-performing undergraduate writers: Explicit instruction and practice. *Reading and Writing, 36*, 909-936. <https://doi.org/10.1007/s11145-022-10318-x>
- Hains-Wesson, R., & Ji, K. (2020). Students' perceptions of an interdisciplinary global study tour: Uncovering inexplicit employability skills. *Higher Education Research & Development, 39*(4), 657-671. <https://doi.org/10.1080/07294360.2019.1695752>
- Hasrol, S.B., Zakaria, A., & Aryadoust, V. (2022). A systematic review of authenticity in second language assessment. *Research Methods in Applied Linguistics, 3*(1). <https://doi.org/10.1016/j.rmal.2022.100023>
- Hastie, C. R. (2018). 'TeamUP': An approach to developing teamwork skills in undergraduate midwifery students. *Midwifery, 58*, 93-95. <https://doi.org/10.1016/j.midw.2017.12.026>
- Haug, L. (2021). Introducing integrated language skills assessment at the language department of a Czech university. *Language Learning in Higher Education, 11*(1), 253-262. <https://doi.org/10.1515/cercles-2021-2010>
- Hembrough, T. (2020). Integrating reading-to-write strategies and pairing composition and reading courses for first-year, at-risk college students. *International Journal of Instruction, 13*(4), 177-196. <https://doi.org/10.29333/iji.2020.13412a>
- Hiver, P., Al-Hoorie, A. H., & Evans, R. (2022a). Complex dynamic systems theory in language learning: A scoping review of 25 years of research. *Studies in Second Language Acquisition, 44*(4), 913-941. <https://doi.org/10.1017/S0272263121000553>
- Hiver, P., Al-Hoorie, A., & Larsen-Freeman, D. (2022b). Toward a transdisciplinary integration of research purposes and methods for complex dynamic systems theory: Beyond the quantitative-qualitative divide. *International Review of Applied Linguistics in Language Teaching, 60*(1), 7-22. <https://doi.org/10.1515/iral-2021-0022>
- Hiver, P., Larsen-Freeman, D., Al-Hoorie, A. H., & Lowie, W. (2023). Complex dynamic systems and language education: A sampling of current research – Editorial. *International Journal of Complexity in Education, 4*(1), <https://doi.org/10.26262/ijce.v4i1.9476>
- Humphreys, P. (2022). The need for English language development in anglophone higher education settings: A case study from Australia. *Journal of English for Academic Purposes, 59*. <https://doi.org/10.1016/j.jeap.2022.101156>
- Jackson, D., & Bridgestock, R. (2021). What actually works to enhance graduate employability? The relative value of curricular, co-curricular, and extra-curricular learning and paid work. *Higher Education, 81*, 723-739. <https://doi.org/10.1007/s10734-020-00570-x>

- Jorre de St Jorre, T., & Oliver, B. (2018). Want students to engage? Contextualise graduate learning outcomes and assess for employability. *Higher Education Research and Development*, 37(1), 44-57. <https://doi.org/10.1080/07294360.2017.1339183>
- Kalfa, S., & Taksa, L. (2017). Employability, managerialism, and performativity in higher education: A relational perspective. *Higher Education*, 74, 687-699. <https://doi.org/10.1007/s10734-016-0072-2>
- Kim, H. J., & Kim, A. A. (2021). Examining the effectiveness of learning-oriented language assessment in second language reading-to-write tasks: Focusing on instructor feedback and self-assessment. *The Journal of Asia TEFL*, 18(4), 1250-1265.
- Knoch, U., & Sitajalabhorn, W. (2013). A closer look at integrated writing tasks: Towards a more focused definition for assessment purposes. *Assessing Writing*, 18(4), 300-308. <http://doi.org/10.1016/j.asw.2013.09.003>
- Kocatepe, M. (2021). Reconceptualising the notion of finding information: How undergraduate students construct information as they read-to-write in an academic reading class. *Journal of English for Academic Purposes*, 54. <https://doi.org/10.1016/j.jeap.2021.101042>
- Larsen-Freeman, D. (2018). Looking ahead: Future directions in, and future research into, second language acquisition. *Foreign Language Annals*, 51(1), 55-72. <https://doi.org/10.1111/flan.12314>
- Larsen-Freeman, D. (2023). Ten questions for Diane Larsen-Freeman. *English Australia Journal*, 39(1), 64-70. <https://www.englishaustralia.com.au/documents/item/1945>
- Leijten, M., Van Waes, L., Schrijver, I., Bernolet, S., Vangehuchten, L. (2019). Mapping master's students' use of external sources in source-based writing in L1 and L2. *Studies in Second Language Acquisition*, 41(3), 555-582. <https://doi.org/10.1017/S0272263119000251>
- Marttunen, M., & Kiili, C. (2022). Supporting university students' argumentative source-based writing. *Written Language and Literacy*, 25(2), 228-252. <https://doi.org/10.1075/wll.00068.mar>
- Mercer-Mapstone, L., & Kuchel, L. (2015). Teaching scientists to communicate: Evidence based assessment for undergraduate science education. *International Journal of Science Education*, 37(10), 1613-1638. <https://doi.org/10.1080/09500693.2015.1045959>
- Merkel, W. (2020). A case study of undergraduate L2 writers' concerns with source-based writing and plagiarism. *TESOL Journal*, 11(3). <https://doi.org/10.1002/tesj.503>
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241-256. <https://doi.org/10.1177/026553229601300302>

- OECD. (2022). What is the profile of internationally mobile students? In OECD, *Education at a Glance 2022: OECD Indicators* (1-460). OECD Publishing. <https://doi.org/10.1787/3197152b-en>
- Oxford University Press. (2020). *The Australian student voice on the soft skills needed for the future: And how universities can integrate these skills into their teaching*. Soft skills for graduate employability. <https://www.oup.com.au/higher-education/whitepaper/soft-skills-for-graduate-employability>
- Peng, H., Jager, S., Thorne, S. L., & Lowie, W. (2020). A holistic person-centred approach to mobile-assisted language learning. In W. Lowie, M. Michel, A. Rousse-Malpat, M. Keijzer & R. Steinkrauss (Eds.), *Usage-based dynamics in second language development* (pp. 87-106). Multilingual Matters.
- Peng, H., Jager, S., & Lowie, W. (2022). A person-centred approach to L2 learners' informal mobile language learning. *Computer Assisted Language Learning*, 35(9), 2148-2169. <https://doi.org/10.1080/09588221.2020.1868532>
- Pillar, I., & Bodis, A. (2022). Marking and unmarking the (non) native speaker through English language proficiency requirements for university admission. *Language in Society*, 53(1), 1-23. <https://doi.org/10.1017/S0047404522000689>
- Plakans, L. (2015). Integrated second language writing assessment: Why? What? How? *Language and Linguistics Compass*, 9(4), 159-167. <https://doi.org/10.1111/lnc3.12124>
- Plakans, L., Liao, L.-T., & Wang, F. (2018). Integrated assessment research: Writing-into-reading. *Language Teaching*, 51(3), 430-434. <https://doi.org/10.1017/S0261444818000149>
- Plakans, L., & Gebriel, A. (2017). Exploring the relationship of organisation and connection with scores in integrated writing assessment. *Assessing Writing*, 31, 98-112. <https://doi.org/10.1016/j.asw.2016.08.005>
- Qin, J., & Liu, Y. (2021). The influence of reading texts on L2 reading-to-write argumentative writing. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.655601>
- Rezaei, K., & Marandi, S. S. (2022). A socio-cognitive approach toward source-based writing instruction: Professors' perceptions and current constraints. *Journal of Language Horizons*, 6(2), 7-30. <https://doi.org/10.22051/lghor.2021.34247.1413>
- Roche, T., & Booth, S. (2021). A collaborative approach to assuring standards: Using the CEFR to benchmark university pathway programs' English language outcomes. *Language Teaching Research Quarterly*, 26, 18-38. <https://doi.org/10.32038/ltrq.2021.26.02>

- Ruge, G., & McCormack, C. (2017). Building and construction students' skills development for employability – reframing assessment for learning in discipline-specific contexts. *Architectural Engineering and Design Management*, 13(5), 365-383. <https://doi.org/10.1080/17452007.2017.1328351>
- Sarkar, M., Overton, T., Thompson, C. D., & Raynor, G. (2020). Academics' perspectives of the teaching and development of generic employability skills in science curricula. *Higher Education Research & Development*, 39(2), 346-361. <https://doi.org/10.1080/07294360.2019.1664998>
- Sotiriadou, P., Logan, D., Daly, A., & Guest, R. (2020). The role of authentic assessment to preserve academic integrity and promote skill development and employability. *Studies in Higher Education*, 45(11), 2132-2148. <https://doi.org/10.1080/03075079.2019.1582015>
- Sridharan, B., & Boud, D. (2019). The effects of peer judgements on teamwork and self-assessment ability in collaborative group work. *Assessment and Evaluation in Higher Education*, 44(6), 894-909. <https://doi.org/10.1080/02602938.2018.1545898>
- Sterba, S. K., & Bauer, D. J. (2010). Matching method with theory in person-oriented developmental psychopathology research. *Development and Psychopathology*, 22(2), 239-254. <https://doi.org/10.1017/S0954579410000015>
- Tarchi, C., Villalon, R., & Lamanda, E. (2022). Investigating the effect of actively-openminded thinking on source-based writing: A randomized control trial. *Thinking Skills and Creativity*, 46. <https://doi.org/10.1016/j.tsc.2022.101128>
- Uludag, P., & McDonough, K. (2022). Exploring EAP instructors' evaluation of classroom-based integrated essays. *TESOL Journal*, 13(3). <https://doi.org/10.1002/tesj.670>
- Ushioda, E. (2011). Language learning motivation, self and identity: Current theoretical perspectives. *Computer Assisted Language Learning*, 24(3), 199-210. <https://doi.org/10.1080/09588221.2010.538701>
- Veritas Health Innovation (2023). *Covidence systematic review software*, Veritas Health Innovation, Melbourne, Australia. Available at www.covidence.org
- Weigle, S. C., & Parker, K. (2012). Source text borrowing in an integrated reading/writing assessment. *Journal of Second Language Writing*, 21(2), 118-133. <https://doi.org/10.1016/j.jslw.2012.03.004>
- Westbrook, C. (2023). The impact of input format on written performance in a listening-into-writing assessment. *Journal of English for Academic Purposes*, 61. <https://doi.org/10.1016/j.jeap.2022.101190>
- Western Sydney University. (2023). *How do I measure a journal impact factor?* <https://answers.library.westernsydney.edu.au/faq/269743>

Wolfersberger, M. (2013). Refining the construct of classroom-based writing from readings assessment: The role of task representation. *Language Assessment Quarterly*, 10, 49-72. <https://doi.org/10.1080/15434303.2012.750661>

Xu, W., Zhao, P., Yao, Y., Pang, W., & Zhu, X. (2023). Effects of self-efficacy on integrated writing performance: A cross-linguistic perspective. *System*, 115. <https://doi.org/10.1016/j.system.2023.103065>

Zhang, X. (2020). A bibliometric analysis of second language acquisition between 1997 and 2018. *Studies in Second Language Acquisition*, 42, 199-222. <https://doi.org/10.1017/S0272263119000573>

Zhao, C. G., & Liao, L. (2021). Metacognitive strategy use in L2 writing assessment. *System*, 98. <https://doi.org/10.1016/j.system.2021.102472>

DOI: <https://doi.org/10.61504/JRNZ2577>

Angela Windsor has been a TESOL teacher for over 25 years. She holds a Master of Applied Linguistics (TESOL) (Macquarie, 2014) and a Master of Arts (Research) (Griffith, 2019). She is a PhD candidate at the Australian National University. Her interests include ESL writing and motivation development from a CDST perspective.

Angela.Windsor@unisq.edu.au

Zoe Hancock has been teaching English for more than 20 years. Commencing her teaching career in Japan, she returned to Australia and completed a Master of Applied Linguistics (UTS, 2007) and Master of Journalism (UQ, 2010). Her areas of interest include academic literacy development, teacher training, and spending time with her children.

Zoe.hancock@scu.edu.au

NATURE OF EAL DIRECT ENTRY INTEGRATED ASSESSME

Purpose and type of Integrated Assessment	Instructions	Topic	Written Genre
Summative Reading, listening to speak (Crossthaite et. al., 2017)	The student instructions were to conduct background readings on a topic supplied 72 hrs prior to the assessment. They develop a stance, brainstorm arguments, and identify ideas to evidence their argument in note-form on a single A4 page-sized template that they may use during the assessment. During the assessment, students discuss the topic in groups of five which are video recorded. Time: 72hrs for research & prepare 25-min discussion task	Supplied to students but not identified in study.	Note-taking template
Summative Reading, listening to speak (Croswaithe & Racquel, 2019)	Students conduct background readings on a topic supplied 72 hrs prior to the assessment. They develop a stance, brainstorm arguments, and identify ideas to evidence their argument in note-form on a single A4 page-sized template that they may use during the assessment. During the assessment, students discuss the topic in groups of five which are video recorded. Time: 72hrs with resources to prepare 25-min discussion task	Supplied to students but not identified in study.	Note-taking template
Summative Reading and listening to write (Gokturk Saglam & Tsagari, 2022)	Students use four reading texts and one listening text on the same topic and take notes. In the writing section of the test, students use their notes create a written text integrating information from the readings and the listening text. Time: Not given	Focus on themes from course content, such as the environment, culture, and globalisation.	Note-taking, synthesis writing; details not given
Placement Test Listening to write Reading to write (Haug, 2021)	Listening to write; Students listen twice to a 6-min-long lecture and take notes. Based on the talk and their notes, they write a persuasive text. Reading to write: Students read two articles complete a table with missing information. They use the information to compose a formal email. Time: 90mins (45min per task) 180-200 words per task	Listening to write: The effects of the overuse of antibiotics on human health. Reading to write. Laboratory internship	Note- taking, completing a table, persuasive text, formal email

INDEX A.

TITLES AND PURPOSE OF EMPIRICAL STUDIES 2013-2023

Spoken Genre	Readings (Number and details)	Number of listening texts (Number and details)	How assessed	Purpose of study
Group tutorial discussion	Academic sources found by students	N/A	Marking criteria assesses ability to explain academic concepts and stance (40%); ability to interact with others (30%); and ability to communicate comprehensibly and fluently (30%) and uses an in-house assessment scale. Not assessed by classroom teacher.	Understand teacher-raters' real-time conceptualisation of students' academic stance and engagement in a group oral assessment and the challenges.
Group tutorial discussion	Four academic sources provided to students	N/A	Marking criteria assesses ability to explain academic concepts and stance (40%); ability to interact with others (30%); and ability to communicate comprehensibly and fluently (30%), and uses an in-house assessment scale. Not assessed by classroom teacher.	Understand how errors, meta discourse and (dis)fluency features predict raters' marking of successful oral production in a group tutorial discussion.
n/a	Four texts of different lengths on the same topic; further details not given	One listening text; details not given	Not given.	Understand teacher and student perspectives on the consequential validity of an integrated tests and the viability of test-based decisions over time
n/a	Two x 300-words descriptive/informative texts (Flesch Reading Ease=42.6)	Six-minute lecture monologue created for the test (British neutral accent) played twice	Holistic rating scales based on ALTE Can Do Statements: task completion, incorporating ideas from the source and adding own, relevance of ideas, coherence and cohesion, and grammar and vocabulary.	Explore the validity and authenticity of an integrated assessment as an alternative to an independent assessment for determining direct entry.

Purpose and type of Integrated Assessment	Instructions	Topic	Written Genre
Summative Reading to write (Uludag & McDonough, 2022)	Students write one essay as part of a 3-hour final exam. Students are given a reading list relevant to the topic two weeks prior to the exam. The sources are discussed in class and students take notes using a template encouraging paraphrasing, noting key terms, and noting important quotations. In the exam, students select one essay topic and compose an essay supporting their opinion with ideas from their notes and students can use a paper-based English dictionary. Time: Readings provided two weeks before exam Essay written as part of 3-hour exam	The role of governments in reducing economic inequality (e.g., economic inequality, food banks, microloans).	Note-taking template; argumentative essay
Placement Test Reading to write (Weigle & Parker, 2012)	Students read two readings and answer eight short answer questions in writing then students write an argument essay which must include information from at least one of the readings as support for their argument.	Computers in Education and Globalisation	Argumentative essay
Placement Test Reading to write (Zhao & Liao, 2021)	Before writing the essay, three scaffolding pre-writing questions are used to facilitate attention to task requirements, planning, generating ideas and selecting of supporting evidence. Then test-takers read a short source text and write an essay in response to a prompt. Test instructions remind students to pay attention to assessment criteria while writing. After the test, students complete a metacognitive writing strategy questionnaire. Time: Not given	Familiar and relevant to students	Argumentative essay

Spoken Genre	Readings (Number and details)	Number of listening texts (Number and details)	How assessed	Purpose of study
n/a	Six readings including news reports and theme-based academic texts (900- 1,300 words)	n/a	Analytic rubric developed in house; three categories (i.e., content and organisation, grammar and vocabulary, and mechanics) with four score levels.	Investigate EAP instructors' orientation when assessing integrated writing tasks and the relationship between instructor ratings and textual measures of source use
n/a	Two x 200–300-word passages from authentic sources (e.g., magazines, newspapers) with opposing view	n/a	An analytic scale including four criteria worth ten marks each including content, organisation, language range complexity and language accuracy.	Investigate the extent of source text language borrowing in an integrated reading to write text.
n/a	One short text	n/a	In-house five-point holistic rating rubric capturing writing quality along four main dimensions: content, organisation, linguistic control, and sociolinguistic control.	Investigate the type of metacognitive strategies test takers report using in an integrated test and their relationship to test performance.

OVERVIEW OF EMPIRICAL STUDIES ON THE ASSESSMENT OF GENERIC EMP

Study Purpose	Type of soft employability skill assessed	Assessment Purpose
Support student's self-assessment of perceived employability. (Bennett & Ananthram, 2022)	Self-awareness, interactive, technological and digital literacy, problem-solving and decision-making, goal-directed behaviour, career identity and commitment, self-esteem, efficacy, ability and willingness to learn, understand the relevance of learning, emotional intelligence, career exploration and awareness, occupational mobility and ethical and responsible behaviour	Formative, self-assessment
Understand how STEM students perceive their future careers and employability compared with students in non-STEM fields of study using the EmployABILITY scale (Bennett et al, 2020)	Employability sub-scales utilised in the study: Goal-directed behaviour, career identity and commitment, ability and willingness to learn, perceived program relevance, career exploration and awareness, occupational mobility	Formative, self-assessment
Identify the hindrances to integrating employability skills into the university curriculum. (Cotronei-Baird, 2020)	Analysis, critical thinking, written, verbal and teamwork skills	Formative and summative assessment
Investigate assessment practices and outcomes possible during short-term study tours in internationalised curriculum and what form it takes. (Cushing et. al., 2019)	Confidence, independence, respect, imagination, ambition, self-awareness, empathy, critical thinking, and cultural competency	Formative and summative
Understand how interdisciplinary, global short-term study tours can enhance students' employability (Hains-Wesson & Ji, 2020)	Teamwork skills, and inexplicit employability skills such as managing complexity, developing agility and creativity	Summative assessment

DIX B.
LOYABILITY SKILLS IN AUSTRALIAN UNDERGRADUATE DEGREES 2013-2023

Context	Task	Contribution
10,193 students from multiple disciplines, age groups, year levels and study modes from over 15 Australian universities	Survey questions (employABILITY scale) was completed by students as required reading or available online. The scale is a self-measure of perceived employability designed to promote self-awareness of employability skills and strategies to improve independently. Personalised report based on results is provided to students and educators to follow up or intervene.	Provides a valid instrument for use in all years of study; assesses a combination of traits which relate to employability beyond academic discipline; understood by the non-expert and easily accessible by students
12,576 students enrolled in STEM and non-STEM disciplines at Australian universities	Survey questions (employABILITY scale) was completed by students as required reading or available online. The scale is a self-measure of perceived employability designed to promote self-awareness of employability skills and guide self-development. Personalised report based on results provided to students and educators to follow up or intervene.	Support students who need to enhance awareness of self-regulated behaviour for building capacity for flexible future careers.
Eight academics from accounting and economic disciplines teaching 1st and 3rd year subjects	Indirect assessment along with direct assessment of soft skills in; group work, essays, tutorial participation, data processing assignment and examinations	Reveal the impact of academic experience, position, and disciplinary boundaries on the opportunity academics have to make assessing employability a reality.
15 study tour leaders from 11 Australian universities	The tour leaders responded to survey questions on a range of assessment tasks including reflective writing through journals or essays, role playing, staff observations of students during the tour, student reflection and questionnaires before and after the tour.	Gain insight into the feasibility and practicality of assessing employability skills on short-term study tour programs.
114 in a range of majors including STEM, business, social sciences and health at an Australian university	Students were surveyed and interviewed on a study tour teamwork project investigating global issues and/or undertaking an industry-linked, group-based project for enhancing an industry client's needs	The role of study tour programs in promoting employability skills

Study Purpose	Type of soft employability skill assessed	Assessment Purpose
Develop a model (TeamUP) to enable educators to teach, develop and assess the development of teamwork skills across a degree (Hastie, 2018)	Fostering a team climate, project planning, facilitating teams, managing conflict and quality individual contribution to the group project.	Formative and summative
Evaluate the value of group work assignments for developing workplace teamwork skills (Kalfa & Taksa, 2017)	Teamwork skills; project & time mgt. skills, communication skills, negotiation skills	Summative
Understand what and how communication skills are assessed in undergraduate science assessment tasks (Mercer-Kapstone & Kuchel, 2015)	Communication skills	Not given
Understand how to align assessment with learning and early professional skill development (Ruge & McCormack, 2017)	Analysis and enquiry, problem solving, personal attributes, communication, working independently, working in groups	Personalised, formative, and summative assessment
Investigate science academics perspectives on how generic employability skills are developed and assessed and the associated challenges (Sarkar et al., 2020)	Leadership, flexibility or adaptability, commercial awareness, ICT, independent learning ability, verbal communication, use of own initiative, team-working, numeracy, information retrieval, time management and organisation, written communication, analytical and critical thinking, and problem-solving	Summative assessment
Investigates the potential for interactive oral examinations in enhancing authenticity, promoting employability skill development and improving academic integrity (Sotiriadou et. al., 2020)	Oral communication skills, professional identity and awareness, problem solving skills and identifying viable organisational solutions	Summative assessment

Context	Task	Contribution
21 midwifery students at a regional Australian University	Eight teamwork assignments of increasing difficulty, spread over the three years of the degree with opportunities for self and peer assessment of teamwork skills in addition to assessment by educator.	Provides a model for teaching, developing and assessing the social, emotional and practical behaviours required for effective teamwork.
13 business studies academics and 27 business alumni at an Australian university	Participants responded to questions about the effectiveness of group work assignments to promote relevant teamwork skills for the workplace during their studies.	Highlights the voices of university staff and alumni on the development and authenticity of teamwork skills assessment in degree programs, often absent in the literature.
Undergraduate science degrees comprising a range of core science courses from 4 research intensive Australia universities	Written documents containing instructions for 35 communication-style assessment tasks comprising online course profiles, assessment outlines, criteria rubrics, lecture and tutorial notes were categorised according to whether communication skills were explicit, implicit or absent in the instructions.	Provide recommendations to enhance the assessment of communication skills in science assessments.
Students and stakeholders associated with core courses in building and construction management degree program at an Australian university	Assessments across core courses were developed based on five educational design principles involving embedded assessment for employability, explicit statements about generic skills, active learner engagement, authentic learning experiences and personalised feedback and scaffolding.	Contributes new knowledge on constructive alignment and the use of assessment for learning and early professional skills development.
Academics from Australia and the UK coordinating 189 science units in a range of disciplines	Academics reported on which generic skills were developed and assessed in their units (189 science units in total)	Provides evidence for using alternative strategies for assessing generic employability skills due to the challenges faced by current university assessment practices.
107 undergraduate students Sport Mgt = 93 and International Mgt= 22 in a business degree in a large Australian university	Students completed three interconnected assessments in which the last assessment was based on an interactive oral task relevant to their course. Students then responded to questionnaires on their assessments focussing on the assessments' real-world relevance, the extent to which the assessment focused on completing a 'task', and its effectiveness in preparing the student for gaining employment.	Provides key characteristics and evidence for the design and use of oral interactive assessments in online and face-to-face settings to authentically assess employability skills.

Study Purpose	Type of soft employability skill assessed	Assessment Purpose
Investigates the relationship between self and peer judgements and collaborative groupwork outcomes (Sridharan & Boud, 2019)	Teamwork and self-assessment	Formative and summative
Understand what features of assessment increase students' understanding and engagement with graduate employability and employment preparation concepts and skills (St Jorre & Oliver, 2018)	Communication, teamwork, critical thinking, problem-solving, self-management, digital literacy and global citizenship	Formative and summative

Context	Task	Contribution
98 undergraduate and post-graduate students enrolled in business programmes at one Australian university	Collaborative group task (3-5 students' random allocation); Unique real-life project including three assessment submissions completed collaboratively over 11 weeks	Contributes knowledge about the use and effectiveness of self and peer feedback in collaborative projects
45 students enrolled in a wide range of disciplines at an Australian university	Students were interviewed about their perceptions and understanding of graduate capabilities and how to demonstrate their employability skills.	Provides insight into student preferences in terms of the development and assessment of employability in their programs of study.