2018

Challenges and strategies for assessing student workplace performance during work-integrated learning

Denise Jackson
Edith Cowan University

Challenges and strategies for assessing student workplace performance during work-integrated learning

Associate Professor Denise Jackson
Centre for Innovative Practice
School of Business and Law
270 Joondalup Drive
Joondalup
WA 6027
Australia
T: +61(0)8 6304 2794
d.jackson@ecu.edu.au
ORCID: 0000-0002-7821-3394

This study explores the challenges of assessing student workplace performance during work-integrated learning. It highlights the need for, yet difficulties with, combining positivist and constructivist assessments where workplace supervisors make evaluative judgements on performance yet students are also agents in their own assessment. It examines the ratings awarded by 163 workplace supervisors for 213 business undergraduates completing a work placement as part of their degree program in Western Australia. Students were rated on 17 capabilities associated with employability and results indicate, in alignment with previous studies, a tendency among supervisors to assign inflated marks across capabilities. The mean capability rating awarded to each student was significantly higher than their weighted course average, suggesting workplace supervisors mark more highly than academics in coursework units. To identify solutions to manage leniency bias, the study examined variations in supervisor ratings for a range of personal and contextual variables such as gender, organisation size, work area, and sector. Although supervisor ratings were inflated, they were consistent across the sample with variations recorded for only four capabilities in certain work areas. Reasons for leniency bias among workplace supervisors are explored in light of the findings and alternative approaches to evaluating student workplace performance are presented.

Keywords: work-integrated learning; work-based learning; supervisor evaluation; professional competence; workplace performance.

Work-Integrated Learning (WIL), also referred to as work-based learning and cooperative education, has seen significant growth in recent years. In Australia, this is catalysed by the National Strategy for WIL (Universities Australia et al. 2015) which highlights key areas to increase stakeholder participation in WIL. Forms of WIL may be broadly divided into two categories. 'Placement WIL' is where students are physically based in a professional setting for a defined period, such as internships, practicums, or work placements. The second comprises authentic learning experiences through, for example, simulations, virtual learning, consultancy-based work or client-based projects that do not require a student to be physically present in the workplace for any particular length of time. WIL's documented benefits vary depending on type and length yet include students gaining valuable insight into professional ideology and practice - assisting with career development learning, professional identity construction and becoming a critical practitioner - and applying and developing discipline-specific and non-technical skills (see Jackson 2016).

WIL is integral to many undergraduate degree programs, such as Education, Engineering and Health Sciences, which require practical skill application for professional accreditation purposes. In other disciplines, its growth may be attributed to industry calls for relevant work experience among new graduates (see, for example, Graduate Careers Australia (GCA) 2016). Quality WIL prepares students prior to their practice-based learning experience; makes explicit links between the practice-based setting and what is taught in the classroom through reflective activities; provides industry feedback to identify areas of weakness and strategies for improvement; encourages students to develop critical perspectives of work practices and how they can be improved; and incorporates reflection to consider personal strengths and career aspirations (Billett 2011; Smith 2012).

Despite its widely acknowledged benefits, there are known challenges to implementing WIL which include significant resourcing requirements; balancing high student demand for WIL opportunities against relatively lower levels of employer engagement; access and equity concerns with evidence of lower participation among international students and disadvantaged groups (Jackson et al. 2016) and managing the growing 'black market' to WIL in the form of self-organised, unpaid internships (Department of Employment 2016). A further challenge is designing and implementing assessment which, as with all aspects of a student's degree program, is constructively aligned to intended learning outcomes (Biggs and Tang 2003).

This study focuses on placement WIL and assumes its key purpose is to develop professional competence, achieved through the application of theory into practice in an authentic setting. Underpinned by constructivist approaches to assessment, there is strong argument for student involvement in the assessment process. This includes the development of assessment criteria, meaning the qualities by which they are assessed (Sadler 1987), which may involve their negotiation of individualised learning outcomes and performance criteria with academic coordinators and workplace supervisors (Rust 2007) and subsequent self-assessment of their achievements. Reflecting on performance and gauging their attainment of intended learning outcomes may be undertaken through, for example, learning journals, e-portfolios and structured reflections.

Some critique the sole use of student self-assessment as claims may not be verified and it may evaluate ability to articulate rather than actual performance (see McNamara 2013). It is commonly combined with an assessment conducted by the workplace supervisor of a positivist nature whereby the student's performance is measured against objective standards (Bloxham 2009), meaning the specified minimum achievement levels used to assess the quality of student

performance (Sadler 1987). Although there is variation across disciplines (Gonsalvez and Freestone 2007), this typically involves the completion of a criterion-referenced summative evaluation (Milne and Caldicott 2016). Wolf (2015) asserts that given the drive for higher education's development of work-ready graduates, benchmarking student workplace performance to industry expectations and standards is critical. Student performance in the workplace may differ greatly from that in the classroom and as workplace supervisors have access to students completing WIL and are able to observe their performance, it is important they are involved in the assessment process (Bernard and Goodyear 2013). In alignment, Gonsalvez and Freestone (2007) argue supervisor input is integral to WIL and is in fact mandated in some disciplines.

It appears, therefore, that involving both students and workplace supervisors in the assessment process – mediated by academic coordinators – is considered the ideal. Indeed the combination may facilitate the principles of quality assessment in WIL which include regular feedback among all parties (Rust 2007), utilising both formative and summative assessment pieces (Gonsalvez and Freestone 2007), and incorporating critical reflection (Hodges 2011). There are, however, long-held concerns for the reliability and validity of assessments conducted by workplace supervisors, largely relating to leniency bias - a tendency to mark very highly - and poor inter-rater reliability (Gonsalvez and Freestone 2007). Further, supervisors often resist direct involvement in the assessment process (McNamara 2013) and may not be appropriately skilled to assess accurately (Trede and Smith 2014). There can also be difficulties in incorporating supervisor marks into students' formal grades and learning outcomes in WIL can vary by context (Ferns and Zegwaard 2014), rendering standardised assessment tools problematic.

The study is driven by the tension created in combining assessments where workplace supervisors make evaluative judgements on performance yet students are also agents in their own assessment. Apparent disillusionment among WIL students, and difficulties for those responsible for WIL curricula, in the disparity between grades awarded by workplace supervisors and academics delivering WIL units also catalyses this examination of leniency bias and the identification of ways to more effectively combine positivist and constructivist epistemologies. The research objectives for this study were, therefore, to (i) examine evidence of leniency bias among workplace supervisors; (ii) identify any differences in the skill ratings assigned by workplace supervisors by personal and contextual characteristics; and (iii) recommend strategies for more valid and reliable assessment in WIL, based on the study's findings.

Research objectives were addressed using 213 workplace supervisor evaluations of business undergraduates in a range of placement contexts in Western Australia. Quantitative analysis explores the influence of supervisor characteristics and placement context on rating tendencies. The study extends discussion on challenges in assessing student workplace performance and ways to improve assessment in WIL (see, for example, Ferns and Zegwaard 2014; Wolf 2015), with a particular focus on improving reliability and better reconciling self-assessment and third-party assessment. It focuses on a less explored discipline, moving beyond the Health Sciences where most studies on the reliability of workplace supervisor evaluations are situated (Wolf 2015). Relevant literature is first reviewed, methodology outlined and results presented. A discussion of the results follows with strategies presented to improve workplace assessment in WIL.

What is assessed in WIL?

Assessment in WIL is focused on professional competence. Samples of student performance are used to assess competence, a process which Benett (1993) describes as 'what individuals are theoretically able to do and this ability is judged by the performance of what they actually do in particular circumstances' (87) and acknowledges a level of generalization is involved. Trede et al. (2015) assert assessment should extend beyond competence to include 'dispositions, student understanding and their professional reasoning that underpin their practice performance' (1003). In alignment, Smith (2014) believes assessment in WIL is focused on three areas: 'experience of the work-world; the development or refinement of skills; and the application of disciplinary knowledge in work contexts' (209).

The mix of capabilities used to gauge student workplace performance during WIL must reflect current industry perceptions of what constitutes professional competence and the complexities of professional practice (Trede and Smith 2014). Recent studies on the broader set of skills required of graduates to effectively navigate the changing world-of-work - such as communication, digital literacy, collaboration and project management - are useful (see, for example, Foundation for Young Australians 2016) and transdisciplinarity which captures knowledge production in the contemporary workplace (Walsh 2007). There is a dominating trend of assessing the development of employability skills during WIL (Ferns and Moore 2012) yet Richardson et al. (2013) argue there is 'a tendency to assess that which is easiest to assess and, in doing so, omit more detailed insights about student core employability skills (or lack thereof)' (28).

Forms of assessment in WIL

Ways of evaluating student performance in an authentic setting can be divided into two main categories: student self-assessment and assessment by a third party, typically the workplace supervisor.

Student self-assessment

Sadler (1989) asserts that students should know how their performance compares to ideal standards and therefore understand areas requiring improvement. Developing the ability to self-evaluate is, Sadler (2009) argues, an important graduate skill. Through self-assessment against standards, students will learn what quality workplace performance means and how it can be achieved (Nicol and Macfarlane-Dick 2006). Self-assessment may be particularly useful for exploring student exposure to the professional environment yet may not be the most valid means of measuring the achievement of learning outcomes (Smith 2014). Self-assessments may reflect that WIL students 'conceptualise good performance in terms of efficient completion of daily tasks and are unable to confidently judge their own performance' (Ibrahim et al. 2014, 417). Further, evidence suggests students favour third-party assessment as they find it difficult not to appear arrogant or too confident in self-assessment (Trede et al. 2015).

The role of reflection is widely supported in WIL and is a key point of differentiation from other forms of practical experience. Reflective activities and assessments encourage students to consider their personal strengths and weaknesses, develop career action plans and identify professional development pathways, and identify and reconcile differences between taught theory and industry practice (Billett 2011). Evidence-based learning provided by the student can come in many forms, including oral presentations, learning journals, structured reflections and e-portfolios. These draw on evidence of work completed, to attest student performance, yet provide 'only a snapshot of student performance rather than the entire domain of professional encounters' (McNamara 2013, 190). Student involvement in assessment

focused on deep, critical reflection and interpretation of the meaning of their experiences, encounters and behaviours may also capture the influential role of context on their learning and achievement.

Third-party assessment

Assessment by workplace supervisors or assessors can vary but includes work-based projects, observation, evaluation reports and simulation of practice (see, for example, Gonsalvez and Freestone 2007). As Norcini (2003) notes, third-party assessment can facilitate the assessment of multiple capabilities, augments feedback from industry practitioners and can draw on samples of actual practice. The importance of gathering supervisor input is acknowledged by many (Zegwaard, Coll, and Hodges 2003; Hodges 2011) with Bernard and Goodyear (2013) noting the supervisory relationship is focused on enhancing the professional growth of the student, monitoring their work, and raising the quality of new entrants to the profession. Industry feedback can facilitate deep learning in WIL (McNamara 2013), complemented by feed-forward which can help guide students on how to improve future performance (Knight 2006). It serves to clarify what is expected of a worker, helps them to monitor their own progress, enhances confidence in their work and raises aspirations and goals of what they can achieve, provides guidance on how to improve and can lead to enhanced workplace performance and career success (see Ibrahim et al. 2014; Nae, Moon, and Choi 2015).

Milne and Caldicott (2016) posit that most assessment in placement WIL requires the workplace supervisor to complete a summative evaluation which is criterion-referenced, largely related to employability skills and performance-based. This may be due to difficulties with the artificial feel of simulations or implementing observation in programs with large numbers of students (see McNamara 2013). This positivist form of assessment does not account

for context, assuming consistency across settings, time and raters (see Elton and Johnston 2002). McNamara (2013) encourages broadening third-party forms of assessment beyond a simple summative evaluation to collating additional forms of empirical evidence of student performance, such as observation or simulation.

Good practice principles in WIL assessment

As assessment helps students understand their strengths, weaknesses and how to improve their performance (Biggs and Tang 2003), feedback should be ongoing and regular through the WIL experience (Ibrahim et al. 2014). This highlights the importance of formative assessment which Zegwaard et al. (2003) argue can more easily focus on student growth and development rather than judgement and accountability. The value of feedback can vary with perceived quality, highlighting the need for accurate feedback which is of value to the student (Nae et al. 2015). Feedback should be provided by a supervisor who works closely with the student (Ibrahim et al. 2014), who has the expertise to provide quality feedback to students (Bernard and Goodyear 2013), and who has a commitment to the assessment for learning rather than compliance (Peach, Ruinard, and Webb 2014). Further, workplace supervisors must allow sufficient time to complete the summative evaluation process and the formal evaluation can be of a standardised format but should allow some level of personalisation (Ibrahim et al. 2014).

Zegwaard et al. (2003) posit a combination of self-assessment and third-party assessment is ideal for developing students who can reflect and can create a portfolio of their demonstrated abilities. They also acknowledge the role of supervisor input and suggest using negotiated placement outcomes and competency scales for assessment criteria as one example of their input. Ferns and Zegwaard (2014) also advocate industry involvement in defining authentic assessment criteria, arguing it would likely result in assessments which are more

closely aligned with what should and what could be feasibly assessed in students based in the workplace. Others (Jones et al. 2009; McNamara 2013) also promote students negotiating their intended learning outcomes with their host and academic coordinator, deepening stakeholder's understanding of assessment criteria (Peach et al. 2014). This, however, may be difficult in disciplines where students are dispersed across an array of contexts and in diverse fields. Delandshere (2001) argues that third-party assessment alone poses moral and ethical issues as the supervisor imposes judgement on a student without regard to their own reflection and self-assessment. Content validity would be more assured if 'multiple sources of evidence of learning are used as a basis for assessment (and self-assessment), since they would allow a broad range of key issues to be addressed' (Benett 1993, 85).

There has also been discussion on applying pass/fail grading criteria, rather than traditional academic scales, to workplace assessments although many criticise this as demotivating students and resulting in lower standards of performance (see Zegwaard et al. 2003; Reddan 2013). Hodges (2011) emphasises the importance of what he describes as a 'portfolio-based' approach to assessment in WIL which engages students in critical reflection through drawing on evidential sources of their performance, articulating their learning during the WIL experience and identifying weaknesses and developmental pathways.

Challenges in assessing WIL

The role of the supervisor

The popularity of summative evaluations by workplace supervisors is problematic for a number of reasons given significant evidence of poor reliability and bias (Gonsalvez et al. 2013). First, workplace supervisors typically rate students much more highly than academics (Milne and Caldicott 2016). This leniency bias, defined by Vinton and Wilke (2011) as 'the tendency to evaluate individuals more favorably than is warranted by their performance' (288), among

workplace supervisors is evidenced in many studies (see, for example, Gonsalvez and Freestone 2007; McGill, Van der Vleuten, and Clarke 2011; McNamara 2013). McGill and colleagues argue it could be attributed to low supervisor expectations, a ceiling effect among trainees, or little actual variation among trainees. Alternatively, supervisors may wish to avoid conflict caused by a student challenging their mark (Dudek, Marks, and Regehr 2005), supported by anonymous evaluations resulting in less leniency bias (Vinton and Wilke 2011). Wolf (2015) asserts that supervisors may also not wish to have to justify lower marks to academic coordinators while Dudek and colleagues state it could be caused by a lack of familiarity with grading processes and expectations or simply a tactic to retain a flow of WIL students into the organisation (Stone and McLaren 1999).

Wolf (2015) argues excessively higher supervisor marks 'may also be the consequence of the halo effect, allowing a student's commitment, enthusiasm or performance in other areas of competency to override the overall assessment' (1051). Halo bias may be due to assessors grading on potential, rather than actual performance, or them rating on one particularly outstanding aspect of performance, rather than considering all (see Wolf 2015). Elevated marks may cause students to develop an inflated sense of their own capabilities, preventing them for pursuing much-needed professional development (Gonsalvez and Freestone 2007), and could create unrealistic expectations of post-graduation employment.

To avoid leniency bias, Gonsalvez and Freestone (2007) recommend introducing multiple assessors to evaluate aspects of the WIL student's performance. There is, however, evidence of poor inter-rater reliability among workplace supervisors with significant variability in assigned scores for the same student, resulting in inconsistent outcome decisions (see Yeates et al. 2013). Interestingly, McGill et al. (2011) found that while there was significant variation

in the assignment of scores by assessors for different aspects of competence in a student, overall competence and communication were areas with acceptable levels of reliability and which could be assessed by the workplace supervisor with reasonable confidence.

Some workplace supervisors simply may not have the skills required for effectively assessing workplace performance (see McNamara 2013), perhaps due to a lack of formal training (Gonsalvez and Freestone 2007; Jones et al. 2009) although rater training has not proven to increase reliability in previous studies (see Govaerts et al. 2013; Yeates et al. 2013). A lack of attention by academic coordinators to guiding workplace supervisors to assess WIL students effectively may augment grade inaccuracies (see Milne and Caldicott 2016) and Benett (1993) notes that academics often assume industry practitioners are both familiar with the expected standards of student performance and have a frame of reference to assess them against. In accordance, McNamara (2013) observes 'it is questionable whether it is possible to ensure each supervisor has a consistent perception about what they are assessing and what standards are expected' (189) and offering training and information on the higher education institution's assessment policy would be useful. As Richardson et al. (2009) assert, 'universities often spend considerable resources in developing relationships with employers but fewer in ensuring that these employers are equipped and confident in fulfilling their role as workplace supervisor' (284).

Finally, workplace supervisors may find it difficult to give informal, formative feedback as this may be associated more with a mentoring role and in conflict with their need to complete a summative evaluation report (see Vinton and Wilke 2011). Workplace supervisors do not wish to inhibit student growth (Stark and Greggerson 2016) and may feel uncomfortable with applying standardised measures of behaviour in evaluations when they feel

their role is to foster personal development (Bogo et al. 2007). Yeates et al. (2013) assert that relationships between the supervisor and students can influence assessment judgement.

Inappropriate evaluation templates and rating scales

The inappropriate design of templates used to capture skill capabilities (see Milne and Caldicott 2016) and/or a poorly developed rating scales (McGill et al. 2011) raise concerns for validity and reliability. While the addition of behavioural anchors has improved reliability somewhat, restricting the scale range has not (see Yeates et al. 2013). Gonsalvez and Freestone (2007) argue more attention should be paid to clearly defining the capabilities included in formal evaluations and higher education institutions could review both the student's absolute ratings and their ratings relative to other peers in the WIL unit, although this has not appeared to reduce bias (Bushnell et al. 2011). McNamara (2013) believes the standardardisation of assessment criteria used by supervisors may improve quality and reliability yet this has not always proven successful (Govaerts et al. 2013) and is not always possible given the diversity within some disciplines. Sadler (2014) ponders the problematic use of a fragmented, analytic approach to assessment which operates a number of criteria and associated grading scales, such as excellent, good and weak. Here, the awarded mark comprises the sum of grades assigned to the different criteria with a lack of consideration for the assessor's sense of overall quality. He acknowledges the lack of anchorage in the use of criteria and grading scales, potentially leading to individual interpretation, subjectivity and bias.

Variations by context

In contrast to the classroom setting, the professional environment is diverse, particularly for students dispersed across different industries and sectors with individualised learning goals.

Delandshere (2001) laments the lack of consideration to the social context of learning, acknowledging that different knowledge will be developed in different communities and situated learning must be effectively captured by assessment. Knight (2006) asserts that where there is uncertainty in judgement – highly probable with multiple industry supervisors in a range of different professional contexts – "it would be prudent to treat them as local judgements—as honest, doubly-contexted judgements" (441). Assuming the traditional model of 'marker' and 'subject' are adopted, as opposed to actively involving the student in the assessment process, significant moderation may be required. While moderation is broadly justified for reasons relating to student confidence, staff development, and community building (Bloxham et al. 2016), its overarching purpose in WIL would be to increase the reliability of assessment through review of applied criteria and awarded marks to ensure equity and accountability.

As Knight (2006) argues, context influences assessment in two ways: "the context in which the achievement arose and the circumstances in which the judges judged" (435). Focusing on the latter, there are reported variations in supervisor assessments by context. Wolf (2015) explored leniency in workplace evaluations for Public Relation (PR) students and found differences by organisation type, namely small business, not-for-profit and firms with dedicated PR professionals. She found higher grades assigned in small businesses, attributed to the intimacy of the supervisor and student relationship leading to both halo and leniency bias. Also, given the size of the business, the assigned supervisor may not be an 'expert' in PR and thus less accurate in their rating of technical capabilities. Wolf also found high levels of leniency and halo bias in not-for-profit organisations which rely heavily on volunteers and therefore may not wish to jeopardise future placements by assigning relatively low marks. In addition, she asserted 'NFP organisations attract people with certain personality traits, who

may arguably be more inclined to provide positive reinforcement and opportunities for growth' (1055). Time pressures and the perceived purpose of assessment may also influence awarded ratings (Levy and Williams 2004).

Method

Participants

The study focuses on the assessment of 213 Business undergraduates completing an elective work placement as part of their degree program over four academic semesters. Students must complete at least one-half of their degree program before undertaking WIL and are recruited onto the program through interviews with the WIL team and host employer. Demonstration of sound work ethic is considered more important than formal academic grades, evidenced by recommendation from a relevant academic. While a credit course average or above among applicants is considered ideal, students with less are encouraged to apply. Placements attract course credit, are predominantly unpaid and require a minimum of 100 hours in the professional setting. The university assigns students to suitable placement opportunities based on cultural fit, personal interest and alignment of major/program to the opportunity's proposed learning outcomes. Placements therefore take place in diverse settings with tasks and outcomes unique to each student.

A rigorous preparation program, comprising both online and on-campus elements, is mandatory for WIL students and there is considerable liaison with workplace supervisors on how to manage student performance and any arising issues, the importance of formative feedback, and expectations regarding the summative evaluation. Supervisors are encouraged to discuss any concerns with student performance with the academic coordinator and must complete a mid-semester evaluation (which does not count towards the student's final mark) via email.

As recommended by Peach et al. (2014), students, hosts and academic coordinators enter into a tripartite arrangement and negotiate the student's intended learning outcomes with students taking responsibility for their own learning. This forms their first assessment, followed by a structured reflection at the mid-point of semester. An electronic skills portfolio is submitted at the conclusion of the semester, comprising i) a review of performance using self-ratings and associated commentary on the extent to which they achieved their defined learning outcomes, drawing on evidence from the workplace; (ii) career action plan; (iii) three minute video summarising their key achievements in the placement; and (iv) workplace supervisor evaluation with an associated 500-word response by the student. Assessment therefore combines self and third-party assessment pieces in both formative and summative format.

Of the 161 different supervisors, 60% were female and 16% were based in small organisations (less than 20 employees). Thirty of the 213 students sourced their own placement opportunity, mostly with their current employer, which required approval by the academic coordinator. Sixty one per cent of supervisors were based in the private sector, 24% in the public sector and 15% in not-for-profit sector. The work areas in which supervisors were based were Human Resource Management (20%), Accounting and Finance (28%), Marketing and Sales (15%), Tourism, Hospitality, Sports and Events Management (23%) and Management and related (14%).

Procedures

Data was collected over four academic semesters between April 2015 and October 2016. The aggregation of data over this two year time period was chosen in order to gather sufficient data to explore variations by supervisors' individual and contextual characteristics using quantitative techniques. A staged, action research approach, with analysis conducted each

semester and iterative developments to assessments and moderation processes implemented in the next semester cycle, was rejected given the dispersion of students - and their workplace supervisors – across such diverse areas.

The 161 different workplace supervisors of the 213 students undertaking the work placement during these four semesters were invited by email to share their summative evaluations for research purposes. Of the original sample of 219 students, six of their supervisors declined to participate and were removed from the sample. The summative workplace evaluation form is mandatory and requires supervisors to rate student performance in 17 capabilities broadly related to employability. The form also asks supervisors to provide open text comments on the extent to which intended outcomes were achieved; strengths and any areas for improvement; and whether the students possessed the skills and knowledge necessary to perform their duties and what others may be useful. As recommended by McNamara (2013), the supervisor report is not awarded a separate mark but embedded within the e-portfolio grade with associated criterion in the portfolio marking rubric. This serves to avert student queries on their mark and any conflict arising from this.

Students are provided with assessment guidelines which explain the purpose of the supervisor evaluation within their Portfolio and the e-portfolio marking criterion, labelled 'workplace performance', has five descriptors ranging from weak to excellent. These guidelines and associated rubric are discussed, and any questions clarified, with participating WIL students at both the induction and mid-semester learning sessions. Workplace supervisors are briefed by email – with any clarification where needed by phone – on the purpose of the evaluation and how to complete it at both the commencement and conclusion of the placement.

Measures

Supervisors were required to rate each of the 17 capabilities on a five-point scale: very poor (one), poor (two), average (three), good (four) and excellent (five), with an option for 'not applicable' if the placement did not offer the opportunity to demonstrate that particular capability. The capabilities derive from literature relating to attitudes, traits, competencies and skills considered important for the development of professional identity among students (see Jackson 2016). Jackson asserts that developing an affinity and alliance with one's intended profession, that 'sense' of feeling like a professional, is increasingly acknowledged as a key aspect of employability. Importantly, there is significant alignment with the graduate capabilities highly regarded by industry (see, for example, GCA 2016). The placement of students in areas oriented to their future career is considered important for facilitating the development of the defined capabilities.

Analysis

Evaluation of the scores assigned by workplace supervisors was undertaken in two ways. First, the five response categories were converted to a 100-point scale (1=0, 2=25, 3=50, 4=75, 5=100) and the mean score and distribution of ratings were recorded for all 17 capabilities. Second, a composite mean score was computed for the supervisor's percentage ratings for all capabilities for each student and compared to their course weighted average mark at the time of completing the placement. This forms a direct comparison between the workplace supervisor's assigned ratings and the marks awarded by a range of academics during the student's degree program. Variations in supervisor ratings were examined using a series of MANOVA for gender, organisation size, sector, work area, whether the placement was self-organised and their level of 'experience' in assessing students. A supervisor was considered 'experienced' if they had assessed at least one other student previously using the same evaluation form.

Results

The mean percentage ratings for the 17 capabilities are presented in Table 1, along with the frequency and proportion for each response category: very poor (0), poor (25), average (50), good (75) and excellent (100). Table 1 indicates workplace supervisors assign high ratings across all 17 capabilities with the mean percentage score for all but one equating to a Higher Distinction (80 to 100). 'Generates and suggests new ideas' achieved a mean percentage rating of a Distinction (70 to 80). Further, the proportion of students achieving either a 'good' (75) or 'excellent' (100) rating was above 85% for all capabilities other than, again, 'generates and suggests new ideas' where 76.2% of the supervisors rated good and above. Findings align with Gonsalvez and Freestone (2007) whose high percentage scores also suggested leniency bias among workplace supervisors.

 [Insert Table 1]

A composite mean score of the supervisor's percentage ratings for the 17 capabilities was computed for each student. The average across the 213 evaluations was 86.15% with a standard deviation of 11.49. Placing them into the same academic scales utilised in the Western Australian university, only 0.9% students were in the Fail category, 0.5% in Pass, 7.5% in Credit; 22.1% in Distinction and 69.0% in Higher Distinction. As with similar studies (Gonsalvez and Freestone 2007; Wolf 2015), an extremely high proportion of students received the equivalent to a Distinction or Higher Distinction in their workplace performance appraisal. Over the four academic semesters, only two students failed and one student achieved a Pass, adding support to workplace supervisors' reluctance to award anything other than high marks to placement students. Thirty three of the students were awarded an average rating of 100% across the 17 capabilities, a similar proportion to that recorded in Wolf's study.

There was no correlation reported between the composite mean score assigned by workplace supervisors and the student's achieved course weighted average mark when commencing the placement, $r(213) = .055$, p=.428. Further, on average, the workplace supervisor mean rating was 16.1% higher than the student's course weighted average mark, suggesting supervisors graded student performance more highly than academics in previous coursework units. Contingency analysis is presented in Table 2 and supports previous studies which assert more lenient marking among supervisors than academics (Gonsalvez and Freestone 2007; Wolf 2012, 2015). To compare, 69.0% of workplace supervisors marked their placement students as a Higher Distinction while only 10.7% of the same students had achieved the same average grade in their coursework.

[Insert Table 2]

A series of MANOVA (α=.05) was conducted to explore variations in supervisor ratings for the 17 capabilities by supervisor gender; the sector, work area and organisation size within which they and the student were based; whether the placement was organised by the student or the university; and supervisor prior experience in using the summative evaluation report. Unlike Wolf (2015) who found differences in supervisor ratings by organisation type and sector, the only significant interaction reported was for work area; λ=.619, $F(68, 689.068)$=1.318, p=.050, partial η^2=.113. Significant variations were recorded for four of the 17 capabilities (see Table 3). Tukey post-hoc analysis indicated that for the communication capability there was a significant difference between supervisors based in Accounting and Finance and Marketing Sales, with the former awarding significantly lower ratings (p=.012). This finding was replicated for the suggestion and generation of new ideas (p=.002); demonstrating self-awareness (p=.009); and exhibiting professional judgement and reasoning

ability (*p*=.010). For the latter capability, supervisors in Marketing and Sales also assigned significantly higher ratings than the Tourism grouping (*p*=.009).

[Insert Table 3]

Discussion

The elevated mean supervisor ratings, compared to the student's achieved course average, could be explained by them excelling in the practical application of skills rather than academic knowledge (Wolf 2015). The high marks could in fact be a true reflection of student performance, particularly as supervisors may be better placed to gauge workplace performance given their exposure to entry-level staff in the professional environment and close proximity to the student. Extant evidence of inflated performance ratings, however, and widespread employer declarations that new graduates perform inadequately in the workplace suggest it may not be unreasonable to interpret the consistently high ratings for all capabilities as leniency bias among supervisors. Within the study context, this may be occurring for a number of reasons.

First, actively encouraging students to discuss their evaluation with supervisors and the completion of a 500-word reflective response may adhere to good assessment practice yet could augment leniency bias. Supervisors may not wish to engage in negative dialogue surrounding marks, rather opting for a higher mark to avoid conflict and disappointment (Dudek et al. 2005). Second, as Bennett-Wimbush and Amstutz (2011) suggest, the work placement may be associated with 'cheap labour' and supervisors feel it appropriate to reward their students in the absence of financial compensation. Stone and McLaren's (1999) suggestion that high marks guarantee a flow of interns may be valid but excessive student demand for placements in

comparison with the limited supply of opportunities (Jackson et al. 2016) may render this irrelevant in this particular study .

Perhaps more likely is a lack of guidance for workplace supervisors on the summative assessment's purpose and frame of reference and the need to benchmark ratings to the standards expected by industry. While instructions are provided by the academic coordinator to ensure the evaluation is completed by the person who has most direct contact with the student and is best equipped to judge their performance, in addition to advising sufficient time should be allocated to the exercise, there is no formal training as such. The lack of difference between those who had previous experience with completing the evaluation report, and those who did not, tentatively suggests that repetition makes no difference. It may also reflect novice markers paying greater attention to instructions and the actual marking process (see Bloxham 2009). Finally, it is likely that the halo effect may be evident with supervisors' high marks reflecting the typically elevated levels of enthusiasm and gratefulness among WIL students.

Despite the inflated ratings, minimal variation among supervisor ratings from different backgrounds and work settings suggests there is some level of consistency in marking. Relatively high ratings from supervisors in Marketing and Sales prompt additional attention and action by academic coordinators. Anecdotally, obtaining relevant work experience in Marketing seems easier for undergraduates who regularly secure short-term opportunities to assist in the organisation and promotion of charity, sporting and/or community events. These students may therefore be more confident, self-aware and proficient in generating and suggesting new ideas, communication, and exhibiting professional judgement. Further, skill gaps in Accounting and Finance students in certain industry-required capabilities have been documented (see, for example, Yu and Churyk 2013).

Strategies going forward

Alleviating leniency bias

One approach could be to create rich avenues for formative and summative feedback from workplace supervisors yet confine these to a pass/fail mark. Students, however, crave assurance that their workplace performance is highly valued and integral to their achievements in WIL (Wolf 2015) so may not respond well to this. A further option is academic coordinators partnering with supervisors in workplace assessment through, for example, joint observation of actual or simulated practice or marking of the presentation of outcomes.

Another pathway for improvement could be developing a clearer frame of reference against which WIL students are observed, judged and assessed. An interpretivist approach which capitalises on the expert knowledge of workplace supervisors and their capacity to evaluate capability within a particular context may be useful. Here, there needs to be some shared consensus among supervisors on what is acceptable workplace performance, possibly drawing on expected knowledge and skills of a new graduate given the placement is in the latter stages of their degree (Bloxham 2009). Sadler (2014) suggests the use of standards which have been developed by 'recognised authority' such as accrediting and professional bodies. He asserts that standards are not average standards of performance but should be fixed points of reference for assessing student performance. Sadler argues "the performance of different cohorts of students could be compared, research on the effectiveness of teaching could be carried out, and general achievement levels in an academic program or institution could be mapped and evaluated longitudinally" (283).

There has been significant headway in developing fixed reference points for student achievement of graduate capabilities through consultative processes and the publication of industry threshold learning outcomes, accreditations standards and holistic rubrics by year level (see, for example Riebe and Jackson 2014). Further, supervisor evaluations can shift from numerical assessment and analytic grading to adopt a more holistic approach (see Sadler 2009) which may also incorporate open-response commentary. In essence, graduate capabilities could be criteria by which WIL students are evaluated but against a specified minimum level, expressed in absolute terms and developed through consultative processes with industry. The fixed, specified standard for each graduate capability should facilitate comparability of judgements of student performance. This would alleviate calls for moderating workplace assessor decisions by academics (Richardson et al. 2013) which would be resource-intensive and create difficulties if student performance was generalised based on one excerpt of practice.

Knight (2006) argues assessor judgement on student performance would become more trustworthy with familiarity with developed standards – rather than using descriptors with associated grading scales – and this is amplified further if assessment was undertaken by more than one workplace colleague/supervisor and training was provided on using the evaluation tool. Sadler (2014) maintains that standards-referenced judgements can still be consistent "of even single works at arbitrary times and places" (283), such as a work placement, yet also emphasises the need for a competent assessor. A valuable example of formalising supervisor training by the coordinating higher education institution is Trede et al.'s (2016) self-paced, online capacity building module which engages workplace supervisors in the learning process and provides key resources for their role as mentor and assessor. Further, it may be possible for higher education institutions to enrol workplace supervisors in formal graduate certificates focused on mentorship and performance management. Not only does this assure quality

supervision and assessment, the formal recognition will assist supervisors in their own professional development and enhance their engagement with WIL.

Academic coordinators are responsible for alerting supervisors to the importance of accurately assessing against specified standards. In tandem, defined industry standards should be clarified with WIL students. Reminding students that the purpose of WIL is to foster development and growth and they are unlikely to be 'perfect' in their professional preparedness, or equate to a graduate entrant at that point of time, may revise expectations of their achievement against specified standards. WIL plays a critical role in developing self-awareness, confidence and perceived employability and it is important to develop an assessment culture among students and supervisors which continues to foster this yet which also prepares for the realities of the contemporary workplace.

Reconciling self-assessment and third-party assessment

Despite success in integrating workplace learning and classroom learning through reflection, we need to find ways to reconcile self-assessment and third-party assessment and more seamlessly incorporate them into a student's final grade. Involving students in self-monitoring and appraisal of their own performance is critical (Sadler 2009) yet is more problematic in WIL than in traditional academic units as it is not always possible to provide students with exemplars and exposure to what constitutes quality work. Students may not have exposure to entry-level professionals against which they can benchmark themselves. Engaging with peers, who are based in similar workplace environments, in reflective activities and assessments which are focused on self-evaluating performance may develop student understanding of quality and expected performance levels.

Adopting a more constructivist approach, workplace supervisor evaluations could be adapted to incorporate both standards, aligned to threshold learning outcomes or other established criteria reflecting expected graduate capabilities, and new criteria which emerges during the WIL experience. The latter could derive from the individually negotiated learning outcomes in the early stages of the placement, allowing students to partially develop their own assessment criteria against which both parties could evaluate student performance. Again, evaluation of the achievement of standards can be combined with textual comments by both parties.

Another possibility is combining supervisor evaluation with student self-assessment in the summative piece. Students could create a portfolio of evidence, with associated commentary on achievement of each intended learning outcome. Included in this would be reflection on the capabilities developed and utilised during the process; issues encountered and managed; and experiences and learning in relation to professional ideology and career management. This, in turn, can be evaluated by their supervisor in the form of qualitative comments, or the selection of descriptor ratings to ease time pressures. An associated marking rubric would ensure the evaluation of skill mastery, areas for improvement, potential developmental pathways, alignment to professional standards, and dispositions and attitudes in in their approach to this particular outcome. Video-based commentary could be used by both students and supervisors to reduce time, given the time taken to appraise portfolios (Knight and Yorke 2008) and to align to graduate recruitment processes. The assessment process could be extended to a tripartite negotiation of the awarded portfolio mark, similar to that proposed by Cooper and Ord (2014), in a meeting with the student, workplace supervisor and academic coordinator at the conclusion of the placement.

This approach would motivate students to produce a quality skills portfolio which accurately captures their completed work. It would boost content validity through holistic review and evidence-based assessment of multiple learning outcomes using both third-party and self-assessment. Negotiated outcomes could have clearly defined parameters from the outset of WIL to include the application of discipline-specific skills and knowledge; operating successfully in the professional environment; and non-technical skill development. This ensures assessment is focused not only on activity and tangible output but also learning aligned to the demands of the contemporary workplace. The portfolio approach could be extended to incorporate the feedback of workplace peers, in line with 360 degree models, and provides rich industry feedback without quantitative marks. Supervisor training in assessing students remains important, including a frame of reference for benchmarking performance and fostering growth and development while providing realistic and useful feedback. An automated tool in an easily shared platform could relieve time pressures and may engage supervisors better with the feedback process.

Concluding remarks

The study enhances our understanding of the challenge of assessing student workplace performance during WIL in a less explored discipline. Despite implementing assessment design which incorporates both summative and formative formats; uses multiple points of assessment; combines third-party and assessment pieces; and prepares both students and supervisors on the purpose and nature of the assessments, there was evidence of leniency bias and concerns among students for the lack of formal recognition of industry feedback in their final grade. Interestingly, although supervisors assigned inflated ratings, these did not vary by personal characteristics and placement context variables, other than minor differences by work area (field of business).

The study highlights the need for educators to consider alternative approaches to evaluating student workplace performance and presents ways to improve the reliability of supervisor assessments, as well as combining the use of self-assessment and third-party assessment in WIL. The importance of, and ways to, better train supervisors in assessment are discussed. A portfolio approach which integrates industry feedback on defined learning outcomes that span professional identity construction and industry-required discipline-specific and non-technical capabilities is presented. In particular, the use of standards which reflect expected graduate capabilities and are developed through stakeholder consultation is emphasised. A shift away from positivist, summative evaluations to utilise methods which involve student and supervisor evaluation of both pre-defined standards and assessment criteria derived from individually negotiated learning outcomes are discussed.

There are limitations to the study which include the collection of data being confined to one institution and the business discipline although this is over four time periods (four academic semesters). Future research could include designing, implementing and evaluating the proposed portfolio approach which incorporates industry feedback on personalised learning outcomes. Second, a qualitative study of workplace supervisors to examine the difficulties experienced with summative evaluation reports – and suggestions for alternatives – would be beneficial. Finally, there would be significant value in extending the current study to explore reasons for leniency bias through surveying or interviewing placement supervisors.

References

Benett, Y. 1993. "The validity and reliability of assessments and self-assessments of work-based learning". *Assessment & Evaluation in Higher Education* 18 (2): 83-94.

Bennett-Wimbush, K., and M. Amstutz. 2011. "Characteristics and employer perspectives in undergraduate animal industry internships". *North American Teachers of Agriculture Journal* 55 (1):55-59.

Bernard, J., and R. Goodyear. 2013. *Fundamentals of clinical supervision* (5th ed.). Boston, MA: Pearson.

Biggs, J., and C. Tang. 2003. *Teaching for quality learning at university.* Maidenhead: Society for Research into Higher Education and Open University Press.

Billett, S. 2011. *Curriculum and pedagogical bases for effectively integrating practice-based experiences – final report*. Strawberry Hills, NSW: Australian Learning and Teaching Council.

Bloxham, S. 2009. "Marking and moderation in the UK: false assumptions and wasted resources". *Assessment and Evaluation in Higher Education* 34 (2): 209-220.

Bloxham, S., C. Hughes, and L. Adie. 2016. "What's the point of moderation? A discussion of the purposes achieved through contemporary moderation practices". *Assessment and Evaluation in Higher Education* 41 (4): 638-653.

Bogo, M., C. Regehr, P. Roxanne, and G. Regehr. 2007. "When values collide". *The Clinical Supervisor* 26 (1–2): 99–117.

Bushnell, J., P. Nicholson, R. Blackman, C. Allan, Y. Nasstasia, R. Knight, and C. Gonsalvez. 2011. "Where angels fear to tread? Leniency and the halo effects in practicum-based assessment of student competencies". Paper presented at the Australian Technology Network of Universities Conference, Perth, October 20-21.

Cooper, S., and J. Ord. 2014. Developing 'know how': a participatory approach to assessment of placement learning. *Journal of Vocational Education & Training* 66 (4): 518-536.

Delandshere, G. 2001. "Implicit Theories, Unexamined Assumptions and the Status Quo of Educational Assessment". Assessment in Education: Principles, Policy & Practice 8 (2): 113-133.

Department of Employment. 2016. *Unpaid work experience in Australia*. Canberra: Department of Employment.

Dudek, N., M. Marks, and G. Regehr. 2005. "Failure to fail: The perspectives of clinical supervisors". *Academic Medicine* 80 (10): S84–S87.

Elton, L., and B. Johnston. 2002. "Assessment in Universities: a critical review of research". LTSN Generic Centre. https://eprints.soton.ac.uk/59244/1/59244.pdf

Ferns, S., and K. Moore. 2012. "Assessing student outcomes in fieldwork placements: An overview of current practice". *Asia-Pacific Journal of Cooperative Education* 13 (4): 207-224.

Ferns, S., and K. Zegwaard. 2014. "Critical assessment issues in work-integrated learning". *Asia-Pacific Journal of Cooperative Education* 15 (3): 179-188.

FYA. 2016. *The new work mindset: 7 new job clusters to help young people to navigate the new work order*. Melbourne: Foundation for Young Australians.

GCA. 2016. *The report of the 2015 Graduate Outlook Survey*. Melbourne: Graduate Careers Australia.

Gonsalvez, C., and I. Freestone. 2007. "Field supervisors' assessments of trainee performance: Are they reliable and valid?" *Australian Psychologist* 42 (1): 23-32.

Gonsalvez, C., J. Bushnell, R. Blackman, F. Deane, V. Bliokas, K. Nicholson-Perry, and R.

Knight. 2013. "Assessment of psychology competencies in field placements: Standardized vignettes reduce rater bias". *Training and Education in Professional Psychology* 7 (2): 99-111.

Govaerts, M., M. Van de Wiel, L. Schuwirth, C. Van der Vleuten, and A. Muijtjens. 2013. "Workplace-based assessment: Raters' performance theories and constructs". *Advances in Health Sciences Education* 18 (3): 375-396.

Hodges, D. 2011. "The assessment of student learning in cooperative and work-integrated education". In *International handbook for cooperative & work-integrated education,* edited by R. Coll and K. Zegwaard, 53-62. Lowell, MA: World Association for Cooperative Education.

Ibrahim, J., A. MacPhail, L. Chadwick, and S. Jeffcott. 2014. "Interns' perceptions of performance feedback". *Medical education* 48 (4): 417-429.

Jackson, D. 2016. "Re-conceptualising graduate employability: The construction of pre-professional identity in the higher education landscape of practice". *Higher Education Research and Development* 35 (5): 925-939.

Jackson, D., D. Rowbottom, S. Ferns, and D. Mclaren. 2016. "Employer understanding of Work-Integrated Learning and the challenges of engaging in WIL opportunities". *Studies in Continuing Education* 39 (1): 35-51.

Jones, M., J. Jackson, E. Coiacetto, T. Budge, M. Coote, W. Steele, S. Gall, and M. Kennedy. 2009. *Generating Academic Standards in Planning Practice Education*. Strawberry Hills, NSW: Australian Learning and Teaching Council.

Knight, P. 2006. "The local practices of assessment". *Assessment and Evaluation in Higher Education* 31 (4): 435-452.

Knight, P., and M. Yorke. 2008. "Assessment close up: The limits of exquisite descriptions of achievement". *International Journal of Educational Research* 47 (3): 175-183.

Levy, P., and J. Williams. 2004. "The social context of performance appraisal: A review and framework for the future". *Journal of Management* 30 (6): 881–905.

McGill, D., C. Van der Vleuten, and M. Clarke. 2011. "Supervisor assessment of clinical and professional competence of medical trainees: a reliability study using workplace data and a focused analytical literature review". *Advances in Health Sciences Education* 16 (3): 405-425.

McNamara, J. 2013. "The challenge of assessing professional competence in work integrated learning". *Assessment & Evaluation in Higher Education* 38 (2): 183-197.

Nicol, D., and D. Macfarlane-Dick. 2006. "Formative assessment and self-regulated learning: A model and seven principles of good feedback practice". *Studies in Higher Education* 31 (2): 199–218.

Milne, L., and J. Caldicott. 2016. "Exploring Differences in Industry Supervisors' Ratings of Student Performance on WIL Placements and the Relative Importance of Skills: Does Remuneration Matter?" *Asia-Pacific Journal of Cooperative Education* 17 (2): 175-186.

Nae, E., H. Moon, and B. Choi. 2015. "Seeking feedback but unable to improve work performance? Qualified feedback from trusted supervisors matters". *Career Development International* 20 (1): 81-100.

Norcini, J. 2003. "ABC of teaching and learning in medicine work based assessment". *British Medical Journal* 326 (7392): 753–755.

Peach, D., E. Ruinard, and F. Webb. 2014. "Feedback on student performance in the workplace: The role of the workplace supervisors [Special issue]". *Asia-Pacific Journal of Cooperative Education* 15 (3): 239-249.

Reddan, G. 2013. "To Grade or Not to Grade: Student Perceptions of the Effects of Grading a Course in Work-Integrated Learning". *Asia-Pacific Journal of Cooperative Education* 14 (4): 223-232.

Richardson, J., B. Jackling, F. Kaider, K. Henschke, M. Kelly, and I. Tempone. 2009. "The role of industry supervisors in providing feedback to students as part of the assessment process in work integrated learning (WIL)". In *ATN Assessment Conference Proceedings,* 282-292. Melbourne: RMIT University.

Richardson, J., B. Jackling, K. Henschke, and I. Tempone, I. 2013. "Developing a collaborative model of industry feedback for work placement of business students". *Asia-Pacific Journal of Cooperative Education* 14 (1): 27-43.

Riebe, L., and D. Jackson. 2014. "The use of rubrics in benchmarking and assessing employability skills". *Journal of Management Education* 38 (3): 320-345.

Rust, C. 2007. "Towards a Scholarship of Assessment". *Assessment & Evaluation in Higher Education* 32 (2): 229–237.

Sadler, D. 1987. "Specifying and promulgating achievement standards". *Oxford Review of Education*, 13 (2): 191 - 209.

Sadler, D. 1989. "Formative assessment and the design of instructional systems". *Instructional Science* 18 (2): 119–44.

Sadler, D. 2009. "Transforming holistic assessment and grading into a vehicle for complex learning". In *Assessment, learning and judgement in higher education: A critical review*, edited by Gordon Joughin, 1-19. Netherlands: Springer.

Sadler, D. 2014. "The futility of attempting to codify academic achievement standards". *Higher Education* 67 (3): 273-288.

Smith, C. 2012. "Evaluating the quality of work-integrated learning curricula: A comprehensive framework". *Higher Education Research and Development* 31(2): 247-262.

Smith, K. 2014. "Assessment of student outcomes from work-integrated learning: Validity and reliability". *Asia-Pacific Journal of Cooperative Education* 15 (3): 207-221.

Stark, M., and K. Greggerson. 2016. "Differences in Perceptions of Supervisee Contribution: Supervisors' vs. Supervisees' Evaluations". *Journal of Counselor Preparation and Supervision* 8 (1): http://dx.doi.org/10.7729/81.1145.

Stone, W., and J. McLaren. 1999. "Assessing the undergraduate intern experience". *Journal of Criminal Justice Education* 10 (1): 171–83.

Trede, F., and M. Smith. 2014. "Workplace educators' interpretations of their assessment practices: a view through a critical practice lens". *Assessment & Evaluation in Higher Education* 39 (2): 154-167.

Trede, F., M. Mischo-Kelling, E. Gasser, and S. Pulcini. 2015. "Assessment experiences in the workplace: a comparative study between clinical educators' and their students' perceptions". *Assessment & Evaluation in Higher Education*, 40 (7): 1002-1016.

Trede, F., K. Sutton, and D. Maxwell. 2016. Get engaged! Growing WPL through a workplace supervisor capacity building module. Springvale: Australian Collaborative Education Network.

Universities Australia, BCA, ACCI, AIG, and ACEN. 2015. *National strategy on Work Integrated Learning in university education*. Canberra: Universities Australia.

Vinton, L., and D. Wilke. 2011. "Leniency bias in evaluating clinical social work student interns". *Clinical Social Work Journal* 39 (3): 288–295.

Walsh, A. 2007. "An exploration of Biggs' constructive alignment in the context of work-based learning". *Assessment & Evaluation in Higher Education* 32 (1): 79-87,

Wolf, K. 2012. "'Brooke has been an absolute angel' – leniency and the halo effect in industry based assessment of student competencies". Paper presented at the International Conference on Cooperative and Work-integrated Education, Istanbul, Turkey, June 20-21.

Wolf, K. 2015. "Leniency and halo bias in industry-based assessments of student competencies: a critical, sector-based analysis". *Higher Education Research & Development* 34 (5): 1045-1059.

Yeates, P., P. O'Neill, K. Mann, and K. Eva. 2013. "Seeing the same thing differently". *Advances in Health Sciences Education* 18 (3): 325-341.

Yu, C., and N. Churyk. 2013. "Are students ready for their future accounting careers?" *Global Perspectives on Accounting Education* 10: 1-15.

Zegwaard, K., R. Coll, and D. Hodges. 2003. "Assessment of workplace learning: A framework". *Asia-Pacific Journal of Cooperative Education* 4 (1): 10-18.

Table 1 Workplace supervisor capability ratings (*N*=213)

Capability	*N*	*M*	*SD*	< 50%		50%		75%		100%	
				N	%	N	%	N	%	N	%
Communicates effectively in a work environment	213	84.74	16.348	1	0.4	17	8.0	93	43.7	102	47.9
Works effectively with others	212	89.39	14.784	0	0	11	5.2	68	32.1	133	62.7
Pursues tasks and responsibilities with commitment and interest	213	89.44	16.821	0	0	12	5.6	61	28.8	139	65.6
Accepts and uses feedback in a constructive manner	211	89.10	16.537	0	0	10	4.8	67	31.9	133	63.3
Generates and suggests new ideas	211	76.18	20.523	3	1.4	47	22.4	93	44.3	67	31.9
Accepts responsibility and accountability for own tasks and actions	213	87.56	15.285	0	0	13	6.1	80	37.6	120	56.3
Shows initiative	212	82.43	19.794	3	1.4	25	11.8	85	40.3	98	46.5
Manages time effectively to achieve defined goals	213	85.45	16.090	1	0.5	15	7.0	91	42.7	106	49.8
Demonstrates self-awareness	211	84.00	16.958	1	0.5	21	10.0	90	42.7	99	46.8
Shows resilience	209	85.17	17.372	1	0.5	12	5.8	92	44.2	103	49.5
Upholds professional conduct, including following protocols, processes and dress codes	213	91.90	15.609	1	0.5	5	2.4	51	24.1	155	73.1
Exhibits technical expertise and knowledge at the expected level	208	81.25	18.307	2	1.0	20	9.7	105	50.7	80	38.6
Exhibits professional judgement and reasoning ability	208	84.01	16.110	1	0.5	16	7.7	98	47.1	93	44.7
Displays confidence in manner and approach	211	83.89	16.756	0	0	23	10.9	90	42.7	98	46.4
Demonstrates a sense of purpose and self-esteem	211	87.80	14.912	0	0	11	5.2	81	38.4	119	56.4
Able to apply their skills and knowledge in the work context	212	87.26	15.100	1	0.5	9	4.2	87	41.0	115	54.3
Shows interest in and commitment to professional development and future learning	207	91.43	14.220	1	0.5	7	3.4	54	26.1	145	70.0

Table 2 Percentage marks awarded by workplace supervisors and academics to the same student sample (*N=213*)

Mark categories (%)	Workplace supervisor		Average academic coursework	
	N	*%*	*N*	*%*
Fail (0 to 49)	2	0.9	2	0.9
Pass (50 to 59)	1	0.5	13	6.0
Credit (60 to 69)	16	7.5	98	45.7
Distinction (70 to 79)	47	22.1	79	36.7
Higher Distinction (80 to 100)	147	69.0	23	10.7

Table 3 Univariate analysis of workplace supervisor ratings by work area

Capability	df	MS	F	p	η^2
Communicates effectively in a work environment	4	789.206	3.043	.018	.060
Generates and suggests new ideas	4	1358.626	3.856	.005	.075
Demonstrates self-awareness	4	922.825	3.418	.010	.067
Exhibits professional judgement and reasoning ability	4	856.300	3.452	.009	.067