Evaluation of work-integrated learning: A realist synthesis and toolkit to enhance university evaluative practices

ELIZABETH J. COOK[1]
Edith Cowan University, Perth, Australia

Situated in the context of work-integrated learning (WIL), this paper aims to build the evaluative capacity of universities in response to an increasing need for evaluation in higher education. It contributes a realist synthesis of international peer-reviewed literature on university evaluation of WIL, which revealed no use of evaluation theory or approaches by the authors. In response, to support the enhancement of university evaluative practices, this paper offers a toolkit of evaluation theory and approaches, with examples relating to WIL, featuring an evaluation planning tool (RUFDATAE). RUFDATAE is demonstrated using a study from the realist synthesis, to highlight its relevance, usefulness and simplicity, or ease of use, for university stakeholders conducting any evaluation. This paper also contributes to recent scholarly debates about evaluation – how it is perceived and differs from research – suggesting evaluation could be considered as an extension of research.

Keywords: Evaluation, university, higher education, work-integrated learning, employability, realist synthesis

Situated in the context of work-integrated learning (WIL), this paper broadly aims to build the evaluative capacity of universities in response to an increasing need for evaluation in higher education. It offers a selection of evaluation theory and approaches (called the 'toolkit' in this paper) to address a knowledge gap identified through a realist synthesis of international peer-reviewed literature on university evaluation of WIL (empirical component). Notably, the researcher has practical and theoretical evaluation expertise, which informed the development of the toolkit and the assessment of the evaluation knowledge gap. A key feature of the toolkit is an evaluation planning tool (RUFDATAE), which is demonstrated using a study from the realist synthesis to highlight its relevance, usefulness and simplicity or ease of use for university stakeholders conducting any evaluation. Although this paper's commentary and examples relate to WIL, the evaluation toolkit is generalizable to any context and focus (i.e., system-level through to individuals). Moreover, since WIL is likely to be familiar to stakeholders associated with higher education, the toolkit may be easily shared and widely used, thus building collective evaluative capacity.

In this paper, WIL is defined as an educational approach that enables students to experience relevant and authentic work-based learning through engagement with industry and/or community partners as part of assessed university coursework (International Journal of Work-Integrated Learning, n.d.; Jackson, 2019). Importantly, particularly in the context of the COVID-19 pandemic, this definition includes workplace learning undertaken virtually or on-campus (e.g., consulting and projects).

The researcher's aims were to: (1) summarize and critically appraise the international peer-reviewed literature on the evaluation of WIL in university contexts; (2) highlight the gap in university evaluation skills and knowledge despite increasing need to evaluate; (3) offer a generalizable toolkit of evaluation theory and approaches to support the enhancement of university evaluative practices; and (4) contribute to recent scholarly debates (e.g., Gullickson, 2000; Wanzer, 2021) by attempting to clarify,

[1] Corresponding author: Elizabeth J. Cook, e.cook@ecu.edu.au

particularly for non-evaluator audiences, what evaluation is, and how it is perceived and differs from research.

The research questions were:

1. What peer-reviewed international research literature has been published about evaluating WIL in universities?
2. How are university researchers evaluating WIL, and are they using theory and approaches from the evaluation literature?
3. How might evaluation theory and approaches be applied to evaluate WIL?
4. What are the broader implications of these findings for universities?

The first two questions relate to the realist synthesis (review), which informed the latter two questions relating to the evaluation toolkit (theory and approaches). However, the paper is not structured in this order. Rather, the toolkit is intentionally split in two parts, so that important theory informs the realist synthesis, which then provides the steps for enacting changes to evaluative practice. The next section describes the socio-political context of relevance to this paper.

THE CONTEXT OF EVALUATION AND WIL IN HIGHER EDUCATION

Neoliberalism in higher education (Ball, 2008; Olssen & Peters, 2005; Tight, 2019) means that evaluation matters for universities. Governments expect universities to get graduates jobs despite growing numbers of graduates and a constrained labor market (Jackson, 2021). Accordingly, graduate employment outcomes feature in many university strategic plans (Jackson & Bridgstock, 2020) enacted via employability-focused policy and practices (Hewitt, 2020). Graduate employment rates are often used by universities as a proxy metric for employability. However, they are different phenomena, and employability is not an employment guarantee. In the context of careers, employability describes the process of lifelong and life-wide personal development (Jackson & Bridgstock, 2020) towards employment and other personally meaningful life outcomes. Both phenomena matter to individuals, governments and society (Billett et al., 2015; Dearing, 1997; Kinash, Crane, Judd & Knight, 2016; Organisation for Economic Co-operation and Development, 2016; Pennington & Stanford, 2019) and, therefore, need to be evaluated.

Commonly, a proportion of government funding for universities is contingent on graduate employment performance outcomes (e.g., Wellings et al., 2019) and, from 2021, Australian universities will also need to demonstrate performance in respect to WIL (Australian Government Department of Education, Skills and Employment, 2020). Increasingly, universities must produce comprehensive evidence of employability-related processes and graduate employment outcomes (Jongbloed & Vossensteyn, 2010; Organisation for Economic Co-operation and Development, 2017; Wellings et al., 2019; Williamson, 2019). This paper aims to support the enhancement of university evaluative practices in response to the increasing need for evaluation in higher education.

In relation to supporting and encouraging students to develop future-focused employability skills, knowledge and dispositions (Foundation for Young Australians, 2016; Australian Government National Skills Commission, 2020; Pearson, 2020; Smit et al., 2020; Strack et al., 2019) for the world of work (Brown et al., 2018; Hajkowicz, 2016; Pennington & Stanford, 2019), universities have found benefits for a range of initiatives (Jackson & Bridgstock, 2020). In particular, whole-of-institution curriculum-embedded approaches for achieving employability have been shown to achieve the greatest gains (Artess et al., 2017; Blackmore et al., 2016; Kinash, 2015; Kinash, Crane, & Judd, 2016)

and, among these, WIL has strong support (Blasko, et al., 2002; Jackson & Bridgstock, 2020; Kinash, Crane & Judd, 2016; Orrell, 2011).

Significant work has been done by universities and policy makers to clearly define WIL, including its various forms (e.g., Quality Assurance Agency, 2018; Tertiary Education Quality and Standards Agency, 2017; Universities Australia et al., 2015), which suited the progressively focused approach to the realist synthesis in this paper. Despite various challenges to WIL delivery (Doolan et al., 2019; Ferns & Zegwaard, 2014; Rook, 2017), a range of frameworks, validated scales, pedagogical resources and exemplars (Bandaranaike, 2018; Billett et al., 2020; Campbell et al., 2019; Cooper, et al., 2010; Nghia & Duyen, 2019; von Treuer et al., 2011; Winchester-Seeto, 2019) have enabled universities to implement WIL in ways that benefit students. However, the ways that universities highlight their actions towards achieving these outcomes to key stakeholders, could be improved.

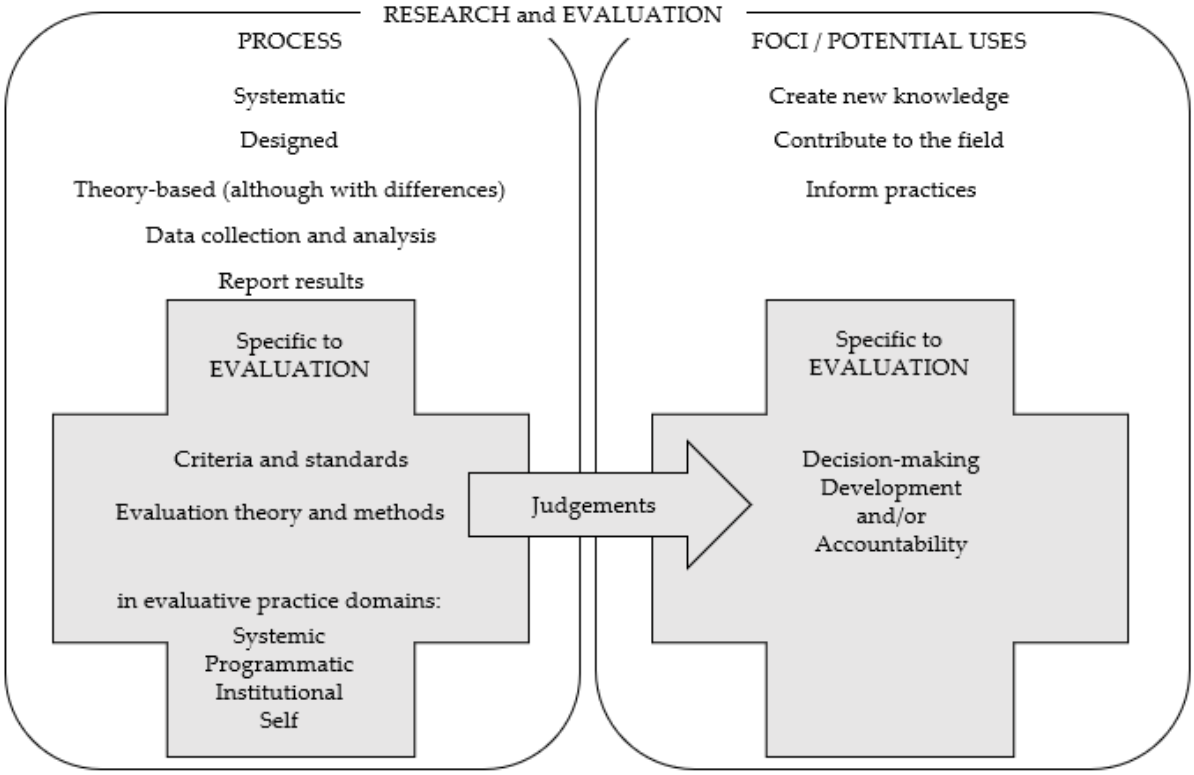
The next section is the first part of the toolkit. It provides foundational evaluation theory and approaches that were purposefully selected by the researcher, who has evaluation expertise, following the realist synthesis. This information will also assist readers to understand the categorization of the reviewed research provided in Appendix A.

EVALUATION THEORY AND APPROACHES

Recently, evaluation scholars have been discussing what evaluation is, how it should be defined and how it differs from research because, increasingly, a diverse range of stakeholders with differing perspectives and evaluation experience are undertaking evaluative work (Gullickson, 2020; Wanzer, 2021). In this paper, an inquiry is an evaluation or a research project depending on its purpose (Patton, 2002), and evaluation is defined as "the generation of a credible and systematic determination of merit, worth, and/or significance of an object through the application of defensible criteria and standards to demonstrably relevant empirical facts" (Gullickson, 2020, p. 4).

Evaluation is distinct from research in its use of criteria and standards to form judgments, which are used for decision-making, development and/or accountability purposes (Chelimsky, 1997; Gullickson, 2020). Gullickson's (2020) "expanded evaluation logic" (p. 3) clearly depicts the components of evaluation, including judgments and inherent reasoning. Building on Wanzer's (2021, p. 31) "possible relationships between evaluation and research", this paper offers a model of evaluation as an extension of research (signified by the plus symbols) with shared and distinctive features (Figure 1).

FIGURE 1: Model of evaluation as an extension of research.



Note. Adapted from "What is evaluation? Perspectives of how evaluation differs (or not) from research" by D. L. Wanzer, 2021, *American Journal of Evaluation,42*(1), p. 31 (https://doi.org/10.1177/1098214020920710). CC BY 4.0.

As a complex and situated social practice, evaluation involves stakeholder groups working together to purposefully gather, analyze and discuss observed evidence from relevant sources about the quality, worth and/or impact of delivery, development and/or policy (Saunders, 2006, 2011, 2012; Saunders et al., 2015). Evaluation occurs in four domains of social practice: systemic, programmatic, institutional and/or self (Saunders, 2011, 2012), as elaborated below with WIL examples.

Systemic evaluative practices are sector-wide (international, national and/or regional) and conducted for accountability, management, comparison and/or auditing purposes. In this domain, the criteria of merit are determined by external funders and/or accreditors. Works by Campbell et al. (2019), the United Kingdom's Quality Assurance Agency (2018), Venville et al. (2018) and Winchester-Seeto (2019) relate to systemic evaluation of WIL.

Programmatic evaluative practices, or WIL program evaluations as reviewed by Rowe et al. (2018) and Orrell (2011), are situated within a university's frameworks and conducted to assess the impacts, effects and value for money of specific interventions. The realist synthesis is focused on WIL program evaluation for its increasing relevance to the sector (e.g., Australian Government Department of Education, Skills and Employment, 2020) and association with time-limited funding. Challenges in WIL program evaluation (Rowe et al., 2018) may be reduced by constructing a logic model (Taylor-Powell & Henert, 2008) and/or theory of change (TOC) (Rogers, n.d.-b.) during the design stage.

A logic model is like a road map depicting the relationships between inputs (e.g., resources), activities, outputs, expected immediate and longer-term outcomes, and behavioral changes, in respect to an intervention (McLaughlin & Jordan, 1999). A TOC shows the theories and assumptions behind the expected changes due to an intervention (Taplin & Clark, 2012) and particularly assists in evaluating complex phenomena (Byrne, 2013), such as WIL.

Institutional evaluative practices are associated with internal quality standards, assurance and improvement (e.g., Palmer et al., 2018; Young et al., 2017). The evaluation criteria and standards should align to institutional policy. Campbell et al. (2019) offer a framework to guide quality evaluation of WIL by universities and the Tertiary Education Quality and Standards Agency (2017) outlines suitable evaluative criteria for universities to benchmark against. To evaluate WIL teaching quality, Smith's (2008) four quadrant model lists appropriate data sources relating to self, student learning, student experience and peers.

Self-evaluation includes judgments made by students, staff, supervisors and industry about the value, worth and/or impact of WIL experiences. Bandaranaike's (2018) WIL reflective practice framework has been shown as effective in assisting students to reflect on and articulate their progress in developing employability skills and autonomy. Self-evaluation was a key component in the designs of many of the studies included in the realist synthesis.

Like research, evaluation starts with questions relating to a particular focus, which inform the design. Table 1 (used in the analysis of the reviewed studies) provides examples of evaluation questions relating to different WIL foci.

TABLE 1: Typology of WIL evaluation.

Focus	Judgment	
	Formative (revise or change)	Summative (start, stop, continue or expand)
Needs assessment	How should delivery be adapted to meet the needs of specific student cohorts?	Is there sufficient need to expand the program?
Process	Do supervisors need more training to assure quality delivery?	Are sufficient numbers of international students participating to merit development of tailored supports?
Outcomes	How can the curriculum be revised to achieve improved outcomes?	Is the program achieving its goals to a sufficient extent to maintain funding?
Emphasized evaluative practice domain	Institutional	Programmatic

Note. Derived from *Program evaluation: Alternative approaches and practical guidelines* (p. 22), by J. L. Fitzpatrick, J. R. Sanders and B. R. Worthen (Eds.), 2004, Pearson Education; "Setting the scene: The four domains of evaluative practice in Higher Education" by M. Saunders, in In M. Saunders, P. Trowler, and V. Bamber (Eds.), *Reconceptualising evaluation in higher education: The practice turn* (pp. 1-17), 2011, McGraw-Hill

There are many different evaluation approaches that address specific evaluation questions and challenges (see Rogers, n.d.-a). Thus, it is important to note that the toolkit offered throughout this paper is quite specific and provides only foundational evaluation theory and approaches to spark interest in getting to know evaluation as a field. It features, RUFDATAE, which is a modified version of Saunders' approach (2000). RUFDATA was purposefully selected (by the researcher with evaluation expertise) for its simplicity, relevance (to the sector's needs) and appropriateness to initiate and guide "new evaluators into [and through] the evaluation planning process" (Saunders, 2000, p.7).

RUFDATAE (as a plural) emphasises that evaluation is a social practice involving collaboration, ethics and care. The acronym provides a simple framework of questions designed to prompt reflection and decision-making for effective evaluation planning at any level (i.e., systemic, programmatic, institutional and/or self). The questions are as follows (note the emphasised RUFDATAE letters, which form the acronym):

- What are the ***R***easons and purposes for the evaluation?
- How will the university ***U***se the evaluation?
- What are the evaluation ***F***oci?
- What ***D***ata and evidence should be collected and analyzed?
- Who is the ***A***udience?
- When should evaluation ***T***ake place?
- Whose ***A***gency will be required?
- What are the ***E***thical considerations?

There are no rules in applying RUFDATAE, nor any limitations. It can be used as a checklist (to ensure that all aspects of 'good' evaluation design have been attended to) in addition to a planning tool, and users can attend to questions as they please to suit their needs. The section titled Using RUFDATAE demonstrates how each question might be used to prompt thinking when planning an evaluation and uses an article from the realist synthesis as a worked example. The following sections are focused on the realist synthesis (review).

REVIEW METHODOLOGY

Realist synthesis is a review methodology that emerged from realist perspectives with the aim of determining what works, for whom, in what circumstances and why (Pawson, 2002). Lawarée et al. (2020) define realist synthesis as "an evaluation approach that combines an interest in the operation of interventions with an interest in their functioning in particular contexts" (p. 3). As such, the methodology is common in evaluation and evidence-based policy research (Klein Haarhuis & Niemeijer, 2009; Pawson, 2002, 2006; Pawson, et al., 2004) although new in higher education research.

This review (and paper) has two ontological and epistemological perspectives relating to evaluation (methodological contribution), and employability and WIL (substantiative contribution). From a realist standpoint, questions such as, what counts as employability practices? what counts as evaluation? and, what works, why and how? (Pawson & Tilley, 1997), guided the review and focused the research questions on the process of evaluation (i.e., how researchers are evaluating), as opposed to evaluation outcomes (i.e., findings).

The review approach was systematic, configurative and aggregative, but not exhaustive (Gough et al., 2012). While some researchers may, therefore, call this a systematised (Grant & Booth, 2009) or exploratory scoping review (Rumrill, et al., 2010), realist synthesis was preferred due to the broader realist (Pawson, 2002) aims of this paper (i.e., to determine what is working [the review] and could work [the toolkit] for university stakeholders evaluating WIL) as clearly presented.

METHOD

Microsoft Project was used to plan, note take and implement the review. The SALSA (Search, Appraisal, Synthesis, Analysis) framework, common in systematic reviews (Booth et al., 2012), was closely followed, as outlined below. Progressive focusing was used to gain insights into evaluative

practices associated with employability before focusing on WIL, meaning that the inclusion criteria were inductively derived.

Search Strategy

Keywords were tested in various combinations across several education and social sciences electronic databases. EBSCO, Informit and Scopus were chosen because they produced the greatest yield of relevant sources. The following search string was repeated in each database: ("higher education" OR college OR university OR tertiary OR institution) AND (curricul* OR course) AND (evalu* OR assess* OR judg* OR metric OR measur*) AND (employability OR "career development learning" OR "career education"). The search was limited to peer-reviewed journal articles in English and published since 1900. Additional inclusion and exclusion criteria (to narrow the focus) were applied in the Appraisal stage.

Appraisal

Screening was conducted in the citation management tool, Mendeley, using a system of folders. OneDrive folders for each database were also created to save a backup copy of each relevant article based on title and abstract. Unrelated articles were listed in Microsoft Excel but not downloaded. Duplicates were removed via the Mendeley menu option: Tools > Check for Duplicates. 446 peer-reviewed journal articles remained as summarized in Table 2.

TABLE 2: Database yield and number of accepted articles.

Database	Search period 2020	Yield	Accepted based on title and abstract
EBSCO (= British Education Index; ERIC; Education Administration Abstracts)	21-26 June	283	134
Informit (=A+Education)	15-20 June	196	114
Scopus	27 June-3 July	455	218
Total			466

All 466 articles were skim read and either excluded or included in the employability category. This theming process provided insights into the broader employability literature to situate the review and assure the relevance, credibility, and validity of the progressive focussing that followed. Table 3 lists the criteria in the order they were applied.

TABLE 3: Progressively focussed criteria.

Focus / SALSA stage	Included	Excluded	Reason
Broad / Search	English language	Other than English	English-speaking author.
	Peer-reviewed	Not peer-reviewed	Quality assurance.
	Journal articles	Grey literature; conference proceedings	Limit yield and assure quality.
	1900-current	Pre-1900	Exhaustive.
	Empirical and evaluative	Conceptual papers, case studies and reviews	Evaluation research focus.
	Employability skills focus	Single skill focus (e.g., teamwork or digital literacy)	Students need more than one employability skill to be employable.
	Embedded in curriculum	Extracurricular; co-curricular; 'bolt-on'	Align and limit yield; Embedded works best (Artess et al., 2017).
Narrow / Appraisal	Industry experiences; Work placements; Internships; Work-based learning (in this paper, collectively referred to as WIL, although the true definition encompasses forms listed in the exclusion column)	Capstones; mentoring; simulations; work experience preparation courses; teaching and nursing placements; consulting practicums / projects; problem-based / project-based learning (unless combined with work placement); service-learning; fieldwork; entrepreneurship / enterprise education; sustainable education	Teaching and nursing placements excluded because defined by professional accreditation. Other activities excluded to focus the review and limit yield. Selected the most published forms of WIL.
	Off campus	On campus	Selected forms happen off campus.
	Student performance or outcomes (perceived or real)	Student satisfaction, attitudes and/or expectations, and/or supervisor perspectives (unless in conjunction with the inclusion criterion); curriculum development	Limit yield and focus the review.
	Programmatic and self domains of evaluative practice	Solely systemic or institutional	Focus the review.
	Mentions evaluation or evaluate (word search within articles)	No mention of evaluation or evaluate	Examined how authors used these words.

The criteria resulted in 24 peer-reviewed journal articles, which were starred as 'Favorites' (Mendeley feature) to enable easy identification moving forward. During full read screening, 10 articles did not meet all criteria and were excluded as summarized below (Table 4). Table 5 shows the final 14 articles.

TABLE 4: Articles excluded in full read screening.

Author (Date)	Reason for exclusion
Bandaranaike & Willison (2015)	Exploratory research focused on student perceptions of learning outcomes in cognitive and affective domains, and whether students display emotional work-readiness.
Dollinger & Brown (2019)	Comparison of WIL types using case study examples.
Doolan et al. (2019)	Focused on stakeholder perspectives relating to implementation.
Jackson & Bridgstock (2020)	Compared student perspectives on the value of embedded, extra-curricular and co-curricular activities, and paid work.
Santiago (2009)	Focused on curriculum design to determine optimal work exposure for employment outcomes.
Samuel et al. (2018)	Focused only on student expectations and supervisor perspectives.
Simiyu et al. (2015)	Focused only on student expectations, experiences and attitudes.
Smith et al. (2016)	Developmental research to assure the validity and predictability of measures used to operationalize WIL curricula design and the concept of employment readiness.
Smith et al. (2019)	Focused on WIL curriculum design but, this time, to determine the optimal settings for student employability outcomes.
Zehr & Korte (2020)	Not embedded in curriculum, i.e., participants recruited through the careers service.

TABLE 5: Final 14 articles.

Citation	Title	Journal
Jackson (2013)	The contribution of work-integrated learning to undergraduate employability skill outcomes	Asia-Pacific Journal of Cooperative Education
Jackson (2015)	Employability skill development in work-integrated learning: Barriers and best practice	Studies in Higher Education
Jackson (2017)	Developing pre-professional identity in undergraduates through work-integrated learning	Higher Education
Jackson (2019)	Students' and their supervisors' evaluations on professional identity in work placements	Vocations and Learning
Jackson & Collings (2018)	The influence of work-integrated learning and paid work during studies on graduate employment and underemployment	Higher Education
Jackson, et al. (2019)	Enabling the transfer of skills and knowledge across classroom and work contexts	Vocations and Learning
Jackson & Wilton (2016)	Developing career management competencies among undergraduates and the role of work-integrated learning	Teaching in Higher Education
Nenzhelele (2014)	Employability through experiential learning course in open distance learning institution	Mediterranean Journal of Social Sciences
Rampersad (2020)	Robot will take your job: Innovation for an era of artificial intelligence	Journal of Business Research
Reddan (2015)	Enhancing students' self-efficacy in making positive career decisions	Asia-Pacific Journal of Cooperative Education
Reddan (2017)	Enhancing employability of exercise science students	Asia-Pacific Journal of Cooperative Education
Taylor & Hooley (2014)	Evaluating the impact of career management skills module and internship programme within a university business school	British Journal of Guidance and Counselling
Toledano-O'Farrill (2017)	Professional application projects: Work-based learning in the curriculum	Higher Education, Skills and Work-based Learning
Whelan & Reichelt-Brushett (2019)	Using internship placements to road test threshold learning outcomes for environment and sustainability	The Journal of Teaching and Learning for Graduate Employability

Synthesis

Categorical data were extracted from each article to Microsoft Excel and evaluative statements were found using a Control F search for evaluat* (i.e., evaluation, evaluate). Similarities and differences across the articles were noted and the analytical approach was finalized.

Analysis

In NVivo12, word frequency queries were generated to explore themes across the articles and produce a word cloud (query setting: 500 most frequent stemmed words a least 6 characters long). Drawing on the researcher's personal evaluation expertise, the evaluative language in each article was assessed to inform the context of this paper. The next section outlines the review findings that assisted the researcher to establish the sector's evaluation knowledge needs which, in turn, informed the components of the toolkit.

REVIEW FINDINGS

Table 6 provides descriptive statistics on the 14 articles.

TABLE 6. Descriptive statistics on the reviewed articles.

Broad category	Narrow category	Number of articles
Database*	Scopus	9
	EBSCO	3
	Informit	3
Location of research (University count)**	Australia	12
	United Kingdom	2
	New Zealand	1
	South Africa	1
	Mexico	1
Author of more than one included article	Jackson, D.	7
	Reddan, G.	2
Journal with more than one included article	Asia-Pacific Journal of Cooperative Education***	3
	Higher Education	2
	Vocations and Learning	2

Notes: *Reddan's (2015) article was sourced via EBSCO and Informit. **Two studies spanned more than one university. ***Presently 'International Journal of Work-Integrated Learning'.

Figure 2 shows the word cloud, which gives prominence to words appearing more frequently in the article texts and highlights that the concept of 'evaluation' was not a major focus for these authors, compared to other aspects of their research.

FIGURE 2. Word cloud showing evaluation was a lesser focus for these authors.

The authors used the word 'evaluation' to describe their research practices yet demonstrated only basic understandings of evaluation as summarized in Table 7. These university researchers were aware of the need to evaluate as demonstrated by their use of the word, sometimes to emphasize the importance of evaluation. However, they were not thinking or practicing as evaluators because they did not apply or reference any evaluation theory or approaches. It is appropriate that they referred to their works as a 'study', 'research' and/or 'investigation'.

TABLE 7. Summary of the use of words evaluation and/or evaluate in the articles.

Citation	Count of 'evaluat'	Examples (quotes/descriptions; not exhaustive)
Jackson (2013)	5	"Halo error is where participants consistently evaluate survey items" (p. 107). Employability Skills Framework includes 'evaluation' in descriptors of desirable behaviors (p. 105). Measures impact and involves student self-evaluation.
Jackson (2015)	12	Abstract: "Evaluation of WIL programs in enhancing skill development remains predominantly outcomes-focused" (p. 350). "prevailing labour market conditions must be considered when evaluating the impact of WIL" (p. 351). "lack of framework for systematically evaluating WIL curricula" (p. 351). Employability Skills Framework includes 'evaluation' in descriptions of desirable behaviors (p. 356). "students observed and evaluated" (p. 357).
Jackson (2017)	14	"Studies designed to evaluate" … (p. 850). "lack of evaluation and reflection" (p. 839). "Transitioning … will cause an individual to re-evaluate" (p. 839). Theme: Self-evaluation and reflection (p. 842). "skills in critical self-evaluation and reflective practise" (p. 844). "share responsibility with industry stakeholders in the development, monitoring and evaluation of PPI" (p. 837).
Jackson (2019)	18	Use of supervisor evaluation reports and students' self-evaluations to investigate impacts. Evaluations in the Title.
Jackson & Collings (2018)	3	Abstract: "the study evaluates the influence" (p. 405). "evaluate institutional data … lack of empirical evidence … and evaluation" (p. 405).
Jackson, et al. (2019)	4	"evaluation of performance" (p. 463). "generic skills form key criteria for employers' evaluations of student preparedness for the workplace" (p. 467).
Jackson & Wilton (2016)	5	Research question two: "evaluate the role of WIL in the development of undergraduate career management competencies" (p. 267). "evaluation of the variations … using MANOVA" (pp. 273-274). Student survey item: "Evaluate how personal priorities may impact upon future career options" (p. 275). "students rated themselves more highly on evaluating how" … (p. 276).
Nenzhelele (2014)	4	Abstract: "important to evaluate" (p. 1602). "learning is the result of … evaluation and reflection of these experiences" (p. 1602). "ensure that the courses attain [objectives] … done by continuous evaluation" (p. 1610).
Rampersad (2020)	9	"Further research is needed to quantitatively evaluate the impact" (p. 70). "confirmatory factor analysis and hypothesis testing was undertaken to evaluate" (p. 70). "normality was evaluated" (p. 70).

Citation	Count of 'evaluat'	Examples (quotes/descriptions; not exhaustive)
		"Reliability was evaluated using coefficient alpha" (p. 71). "scales were evaluated" (p. 71). Evaluation is a dimension in the factor analysis (p. 72). "Evaluating skill levels in WIL students helps educators … inform corrective action" (p. 72).
Reddan (2015)	2	"research is now being focused on the evaluation of counselling interventions designed to increase career decision-making self-efficacy" (p. 292). "scale scores can be utilized to evaluate the effectiveness of educational and career interventions" (p. 293).
Reddan (2017)	9	"A flexible framework that provides a process for discussion, reflection, action and evaluation is essential" (p. 26). "written evaluation of the day's program" (p. 29). Evaluate / evaluation included multiple times in the instrument used.
Taylor & Hooley (2014)	3	Evaluating in Title. Abstract: "This study evaluates" (p. 487). "literature highlights a paucity of empirical, evaluative or illuminative research" (p. 489).
Toledano-O'Farrill (2017)	5	Evaluation is the first of four keywords. Project management perspective: "we have a general and practical vision of the whole process of consulting, diagnosis, change proposal, validation, implementation, evaluation and closure" (pp. 28-29). "Student performance and learning … formally evaluated … evaluation … departs from a conventional academic perspective" (p. 31).
Whelan & Reichelt-Brushett (2019)	10	Threshold Learning Outcome: "thinking critically and creatively in designing and evaluating" … (p. 38). "provided the ability to evaluate" (p. 39). "the host evaluates" (p. 46). "evaluated the performance of students" (p. 46).

Appendix A summarizes the reviewed research and demonstrates the use of the typology of evaluation (introduced in Table 1) to transition readers towards evaluative thinking. The next section demonstrates the RUFDATAE framework using Taylor and Hooley's article as a worked example.

USING RUFDATAE TO PLAN AN EVALUATION

The preceding review demonstrated that there is a gap in evaluation knowledge among international researchers who say they are evaluating WIL. Thus, in addition to the first part of the toolkit (i.e., in the Evaluation theory and approaches section), this section demonstrates how RUFDATAE might be used to guide evaluation planning throughout the sector. Each component of RUFDATAE is explained, with examples drawn from Taylor and Hooley's (2014) study. Their article was specifically chosen for its clarity, broad sector relevance and similarity to evaluation.

Reasons and Purposes for the Evaluation

This component prompts consideration of the internal and external drivers, and the evaluation aims and questions. As outlined in Table 1, reasons may relate to outcomes, needs and/or developments. For example, Taylor and Hooley (2014) described expectations from stakeholders (needs/outcomes focus) and an intention to compare the impact of different interventions (development focus) as their reasons for evaluating. Internal drivers may include institutional policy and/or the needs and perspectives of staff (implementers and performance managers), students (customers), senior executive (decision-makers) and administration (performance reporters and acquitters). External drivers are context-dependent, often in response to relevant legislation and literature.

Uses of the Evaluation

Where possible, institutional and programmatic evaluations should be designed to meet the requirements of related systemic evaluation. That is, use formal internal evaluations as evidence in external performance reporting and acquittal. This is achieved through use practices (Saunders, 2012) and by: (1) designing internal evaluations and reports, which align to university strategy and policy, and external priorities, policy and reporting requirements (as determined by government); and (2) coordinating evaluation reporting to coincide with institutional committee meeting cycles to enable review by internal stakeholders prior to external submission.

Taylor and Hooley (2014) used their evaluation to publish and, in doing so, contributed evidence in support of their WIL approach. They also used the findings to inform university practices (i.e., monitoring and development of the program) and most likely reported the outcomes to strategic decision-makers to inform future delivery and ensure the sustainability of the program. Evaluations can also be used to promote uptake and/or delivery, celebrate successes and/or recognize stakeholder contributions.

Foci

Referring again to Table 1 (i.e., needs assessment, outcomes and/or development), foci may be to refine data collection, determine program needs, impacts, outcomes and/or effects, inform decision-making, planning and delivery, and/or identify enablers and barriers to implementation and uptake. Although Taylor and Hooley's (2014) main focus was to compare the impacts of different interventions on graduates' employment outcomes (rates and levels of employment achieved), they also sought graduates' perspectives on the effectiveness of various recruitment supports.

Data and Evidence

Qualitative and/or quantitative data should be considered and selected based on relevance. Quantitative data may focus on measuring the uptake and/or impact (numbers or proportions) of a program with or without comparing groups and/or external benchmarking. Detailed analyses may be required at various levels, e.g., cohort, course, department, campus and/or university. Taylor and Hooley's (2014) study was purely quantitative (i.e., national and institutional graduate surveys and descriptive statistical analysis).

Qualitative data, gathered via surveys, focus groups, interviews and/or observations, may provide valuable insights and, when triangulated with quantitative data, can enrich evaluation outputs. Qualitative data may focus on participation (i.e., reasons, preferences, experiences and/or benefits) and/or implementation (i.e., perceived benefits and/or impacts of training and/or the intervention).

Audience

Evaluation outputs (i.e., design and content) and modes of dissemination (i.e., email, web content, documents; formal or informal) will depend on the audience and the urgency and/or need to facilitate change. Internally, audiences may include any stakeholder interested or invested in the evaluation (e.g., senior executives, implementing staff and students). External audiences may include current and/or prospective collaborators (e.g., associations, business, industry and/or government departments). Taylor and Hooley's (2014) audience included readers of the article and internal university stakeholders.

Timing

Evaluation for internal monitoring purposes should coincide with program delivery. A semesterly and/or yearly reporting schedule may be suitable. Evaluation for external purposes would usually require annual reporting based on a year of data. Using a yearly planner can assist with meeting due dates.

Agency

This component prompts consideration of collaborative and participatory evaluation approaches. Determine who will prepare, collect and analyze the data, and produce and disseminate the evaluation outputs.

Ethics

Like research, data collection and use must be respectful, protect the anonymity and privacy of participants and abide by relevant institutional and national privacy and ethics legislation. Therefore, determine whether ethics clearance is required before proceeding.

CONCLUSION

This paper summarized and critically appraised the international peer-reviewed literature on the evaluation of WIL in university contexts and revealed a gap in the evaluation skills and knowledge of authors of this literature, despite an increasing need to evaluate in higher education. In response, to support the enhancement of university evaluative practices, this paper offers a generalizable toolkit of evaluation theory and approaches, and clarifies, particularly for non-evaluator audiences, what

evaluation is and how it might be viewed as an extension of research. It also promotes ethical and social practice approaches to evaluation achieved through coordinated and collaborative work at any level.

Most importantly, this paper highlights that, in the critical realm of WIL and employability, university stakeholders are not embracing theory and approaches from the discipline of evaluation to the detriment of the sector's evaluative capacity and subsequent outputs. Although there are guiding principles for evaluation (e.g., Campbell et al., 2019), it appears from this research that staff do not yet apply these and/or understand how to evaluate. It is recommended that universities learn evaluation basics (theory and techniques), particularly in respect to criteria and design. Until that time, the sector is in danger of lackluster evaluative practices and lost opportunities to demonstrate its worth and impact, which is a risk in this metricized and performance-measured world; one that can be easily avoided through a focus on evaluation.

Now that this paper has demonstrated the value of evaluation, how might the sector begin to transition towards evaluative thoughts and actions?

To begin, here are some key considerations:

1. Consider merit, judgments, criteria, significance and use in and of evaluation, i.e., engage in evaluative thinking.
2. Choose what makes sense and is possible given the evaluation context.
3. Look for opportunities to transform evaluation, particularly when evaluating complex, changing systems.
4. When evaluating, be aware of broader occurrences and events to identify if, when and how an evaluation might need to adapt in response to changes in systems.

This paper's limitations are associated with the realist synthesis and include: (1) the exclusion of grey and other literature; (2) exclusion of hand and citation searching and reference checking; and (3) publishing lag impacting the currency of the review. Further, the review was conducted by one researcher, which some may consider to be a limitation, however, this ensured that the research was consistent and articulate in its systematized approach and rigor.

Future research might review improvements made to university evaluative practices and move towards developing a suite of workable systemic evaluation frameworks that meet the needs of universities and governments, whilst enabling institutional and international benchmarking.

Implications for universities include the need to review and enact improvements to institutional and programmatic evaluation and associated practices, which could be used to directly meet systemic evaluation requirements. Governments could collaborate with universities to design explicit evaluation frameworks that include well-considered examples of qualitative and quantitative metrics to guide how universities should meet their clearly defined objectives. While universities have no control over external factors, they can enhance their evaluative practices to produce unassailable justifications for their actions and use of public funds. Improved evaluation and communication of outputs and outcomes will help to highlight the powerful value and impact of universities to societies and people's lives.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the support of the Department of Educational Research, Lancaster University, U.K., and thanks Professor Murray Saunders, Lancaster University and Dr. Kylie Stevenson, Edith Cowan University for their helpful feedback on different versions of the draft article.

REFERENCES

Artess, J., Hooley, T., & Mellors-Bourne, R. (2017). *Employability: A review of the literature 2012-2016*. Higher Education Academy.

Australian Government Department of Education, Skills and Employment. (2020). *National priorities and industry linkage fun: Consultation paper*

Australian Government National Skills Commission. (2020). *A snapshot in time: The Australian labour market and COVID-19*

Ball, S. (2008). *The education debate*. Policy Press.

Bandaranaike, S. (2018). From research skill development to work skill development. *Journal of University Teaching & Learning Practice*, *15*(4), Article 7.

Bandaranaike, S., & Willison, J. (2015). Building capacity for work-readiness: Bridging the cognitive and affective domains. *Asia-Pacific Journal of Cooperative Education, 16*(3), 223–233.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Baxter Magolda, M. (1998). Developing self-authorship in young adult life. *Journal of College Student Development, 39*(2), 143–156.

Blackmore, P., Bulaitis, Z. H., Jackman, A. H., & Emrullah, T. (2016). *Employability in higher education: A review of practice and strategies around the world*. Pearson.

Billett, S., Choy, S., Dymock, D., Smith, R., Henderson, A., Tyler, M., & Kelley, A. (2015). *Towards more effective continuing education and training for Australian workers*. NCVER.

Billett, S. R., Orrell, J. E., Jackson, D., & Valendia-Forrester, F. (Eds.). (2020). *Enriching higher education students' learning through post-work placement interventions*. Springer.

Blasko, Z., Little, B., & Woodley, A. (2002). *UK graduates and the impact of work experience.* Higher Education Funding Council for England

Booth, A., Papaioannou, D., & Sutton, A. (2012). *Systematic approaches to a successful literature review*. Sage.

Bransford, J., & Schwartz, D. (1999). *Rethinking transfer: A simple proposal with multiple implications*. Sage.

Brown, J., Gosling, T., Sethi, B., Sheppard, B., Stubbings, C., Sviokla, J., Williams, J., Zarubina, D., & Fisher, L. (2018). *Workforce of the future: The competing forces shaping 2030*. PwC.

Byrne, D. (2013). Evaluating complex social interventions in a complex world. *Evaluation, 19*(3), 217-228. https://doi.org/10.1177%2F1356389013495617

Campbell, M., Russell, L., McAllister, L., Smith, L., Tunny, R., Thomson, K., & Barrett, M. (2019). *A framework to support assurance of institution-wide quality in work-integrated learning*. ACEN.

Chelimsky, E. (1997). Thoughts for a new evaluation society. *Evaluation*, *3*(1), 97–109. https://doi.org/10.1177/135638909700300107

Cooper, L., Orrell, J., & Bowden, M. (2010). *Work integrated learning: A guide to effective practice*. Routledge.

Dearing, R. (1997). *Higher education in the learning society*. National Committee of Inquiry into Higher Education.

Dollinger, M., & Brown, J. (2019). An institutional framework to guide the comparison of work- integrated learning types. *Journal of Teaching and Learning for Graduate Employability*, *10*(1), 88–100.

Doolan, M., Piggott, B., Chapman, S., & Rycroft, P. (2019). The benefits and challenges of embedding work integrated learning: A case study in a university education degree program. *Australian Journal of Teacher Education*, *44*(6), 91–108. https://doi.org/10.14221/ajte.2018v44n6.6

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

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2004). *Program evaluation: Alternative approaches and practical guidelines* (3rd ed.). Pearson Education.

Foundation for Young Australians. (2016). *The new basics: Big data reveals the skills young people need for the new work order*.

Gough, D., Thomas, J., & Oliver, S. (2012). Clarifying differences between review designs and methods. *Systematic Reviews*, *1*(28), 1–9. https://doi.org/10.1186/2046-4053-1-28

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, *26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Gullickson, A. M. (2020). The whole elephant: Defining evaluation. *Evaluation and Program Planning*, *79*, Article 101787. https://doi.org/10.1016/j.evalprogplan.2020.101787

Hajkowicz, S., Reeson, A., Rudd, L., Bratanova, A., Hodgers, L., Mason, C., & Boughen, N. (2016). *Tomorrow's digitally enabled workforce: Megatrends and scenarios for jobs and employment in Australia over the coming twenty years*. CSIRO.

Hewitt, R. (2020). *Getting on: Graduate employment and its influence on UK higher education*. Higher Education Policy Institute.
International Journal of Work-Integrated Learning. (n.d.). *Defining work-integrated learning*. https://www.ijwil.org/defining-wil
Jackson, D. (2013). The contribution of work-integrated learning to undergraduate employability skill outcomes. *Asia-Pacific Journal of Cooperative Education*, *14*(2), 99–115.
Jackson, D. (2015). Employability skill development in work-integrated learning: Barriers and best practice. *Studies in Higher Education*, *40*(2), 350–367. https://doi.org/10.1080/03075079.2013.842221
Jackson, D. (2017). Developing pre-professional identity in undergraduates through work-integrated learning. *Higher Education*, *74*(5), 833–853. https://doi.org/10.1007/s10734-016-0080-2
Jackson, D. (2019). Students' and their supervisors' evaluations on professional identity in work placements. *Vocations and Learning*, *12*(2), 245–266. https://doi.org/10.1007/s12186-018-9207-1
Jackson, D. (2021). The changing nature of graduate roles and the value of the degree. *Journal of Higher Education Policy and Management*, *43*(2), 186-197. https://doi.org/10.1080/1360080X.2020.1777634
Jackson, D., & Bridgstock, R. (2020). What actually works to enhance graduate employability? The relative value of curricular, co-curricular, and extra-curricular learning and paid work. *Higher Education*, *81*(4), 723-739. https://doi.org/10.1007/s10734-020-00570-x
Jackson, D., & Chapman, E. (2012). Non-technical competencies in undergraduate business degree programs: Australian and UK perspectives. *Studies in Higher Education*, *37*(5), 541–567. https://doi.org/10.1080/03075079.2010.527935
Jackson, D., & Collings, D. (2018). The influence of work-integrated learning and paid work during studies on graduate employment and underemployment. *Higher Education*, *76*(3), 403–425. https://doi.org/10.1007/s10734-017-0216-z
Jackson, D., Fleming, J., & Rowe, A. (2019). Enabling the transfer of skills and knowledge across classroom and work contexts. *Vocations and Learning*, *12*(3), 459–478. https://doi.org/10.1007/s12186-019-09224-1
Jackson, D., & Wilton, N. (2016). Developing career management competencies among undergraduates and the role of work-integrated learning. *Teaching in Higher Education*, *21*(3), 266–286. https://doi.org/10.1080/13562517.2015.1136281
Jongbloed, B., & Vossensteyn, H. (2010). Keeping up performances: An international survey of performance-based funding in higher education. *Journal of Higher Education Policy and Management, 23*(2), 127-145. https://doi.org/10.1080/13600800120088625
Kinash, S. (2015). *8 ways to enhance your students' graduate employability*. Bond University
Kinash, S., Crane, L. H., & Judd, M-M. (2016). *Good practice report: Nurturing graduate employability in higher education*. Australian Government Office for Learning and Teaching.
Kinash, S., Crane, L. H., Judd, M-M., & Knight, C. (2016). Discrepant stakeholder perspectives on graduate employability strategies. *Higher Education Research and Development, 35*(5), 951–967. https://doi.org/10.1080/07294360.2016.1139555
Klein Haarhuis, C. M., & Niemeijer, E. (2009). Synthesizing legislative evaluations: Putting the pieces together. *Evaluation, 15*(4), 403–425. https://doi.org/10.1177/1356389009341897
Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall.
Kumar, A. (2007). *Personal, academic & career development in higher education: soaring to success*. Routledge.
Law, B., & Watts, A. G. (1977). *Schools, careers and community*. CIO Pub. for the General Synod Board of Education.
Lawarée, J., Jacob, S., & Ouimet, M. (2020). A scoping review of knowledge syntheses in the field of evaluation across four decades of practice. *Evaluation and Program Planning*, *79*, Article 101761. https://doi.org/10.1016/j.evalprogplan.2019.101761
Leong, F. T. L., & Barak, A. (Eds.). (2001). *Contemporary models in vocational psychology*. Lawrence Erlbaum.
McLaughlin, J. A., & Jordan, G. B. (1999). Logic models: A tool for telling your program's performance story. *Evaluation and Program Planning*, *22*(1), 65–72. https://doi.org/10.1016/S0149-7189(98)00042-1
Nenzhelele, T. E. (2014). Employability through experiential learning course in open distance learning institution. *Mediterranean Journal of Social Sciences*, *5*(20), 1602–1612. https://doi.org/10.5901/mjss.2014.v5n20p1602
Nghia, T. L. H., & Duyen, N. T. M. (2019). Developing and validating a scale for evaluating internship-related learning outcomes. *Higher Education*, *77*(1), 1–18. https://doi.org/10.1007/s10734-018-0251-4
Organisation for Economic Co-operation and Development. (2016). *Enhancing employability*: *Report prepared for the G20 Employment Working Group with inputs from the International Monetary Fund*.
Organisation for Economic Co-operation and Development. (2017). *Benchmarking higher education system performance: Conceptual framework and data*.
Olssen, M., & Peters, M. A. (2005). Neoliberalism, higher education and the knowledge economy: From the free market to knowledge capitalism. *Journal of Education Policy, 20*(3), 313–345. https://doi.org/10.1080/02680930500108718
Orrell, J. (2011). *Good practice report: Work-integrated learning*. Australian Learning and Teaching Council.
Palmer, S., Young, K., & Campbell, M. (2018). Developing an institutional evaluation of the impact of work-integrated learning on employability and employment. *International Journal of Work-Integrated Learning*, *19*(4), 371–383.
Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Sage.
Pawson, R. (2002). Evidence-based policy: The promise of `realist synthesis'. *Evaluation*, *8*(3), 340–358. https://doi.org/10.1177/135638902401462448
Pawson, R. (2006). *Evidence-based policy a realist perspective*. Sage.

Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2004). *Realist synthesis: An introduction*. ESRC Research Methods Programme, University of Manchester.
Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Sage.
Pearson. (n.d.). *The future of skills*. https://futureskills.pearson.com/research/#/homescreen
Pennington, A., & Stanford, J. (2019). *The future of work for Australian graduates: The changing landscape of university-employment transitions in Australia*. Australia Institute Centre for Future Work.
Peterson, G., Sampson, J., & Reardon, R. (1991). *Career development and services: A cognitive approach*. Pacific Grove; Brooks/Cole.
Quality Assurance Agency for Higher Education. (2018). *UK quality code, advice and guidance: Work-based learning*
Rampersad, G. (2020). Robot will take your job: Innovation for an era of artificial intelligence. *Journal of Business Research*, *116*, 68–74. https://doi.org/10.1016/j.jbusres.2020.05.019
Reddan, G. (2015). Enhancing students' self-efficacy in making positive career decisions. *Asia-Pacific Journal of Cooperative Education*, *16*(4), 291–300.
Reddan, G. (2017). Enhancing employability of exercise science students. *Asia-Pacific Journal of Cooperative Education*, *18*(1), 25–41.
Rogers, P. (n.d.-a). *Approaches*. Better Evaluation. https://www.betterevaluation.org/en/approaches
Rogers, P. (n.d.-b). *Develop programme theory / theory of change*. Better Evaluation. https://www.betterevaluation.org/en/rainbow_framework/define/develop_programme_theory
Rook, L. (2017). Challenges implementing work-integrated learning in human resource management university courses. *Asia-Pacific Journal of Cooperative Education*, *18*(3), 199–212.
Rowe, A. D., Nay, C., Lloyd, K., Myton, N., & Kraushaar, N. (2018). Telling your story of work-integrated learning: A holistic approach to program evaluation. *International Journal of Work-Integrated Learning*, *19*(3), 273–285.
Rumrill, P. D., Fitzgerald, S. M., & Merchant, W. R. (2010). Using scoping literature reviews as a means of understanding and interpreting existing literature. *Work*, *35*(3), 399–404. https://doi.org/10.3233/WOR-2010-0998
Samuel, G., Donovan, C., & Lee, J. (2018). University–industry teaching collaborations: A case study of the MSc in Structural Integrity co-produced by Brunel University London and The Welding Institute. *Studies in Higher Education*, *43*(4), 769–785. https://doi.org/10.1080/03075079.2016.1199542
Santiago, A. (2009). Impact of sandwich course design on first job experience. *Asia-Pacific Education Researcher*, *18*(2), 205–217.
Saunders, M. (2000). Beginning an evaluation with RUFDATA: Theorizing a practical approach to evaluation planning. *Evaluation, 6*(1), 7–21. https://doi.org/10.1177/13563890022209082
Saunders, M. (2006). The 'presence' of evaluation theory and practice in educational and social development: Toward an inclusive approach. *London Review of Education*, *4*(2), 197–215. https://doi.org/10.1080/14748460600855526
Saunders, M. (2011). Setting the scene: The four domains of evaluative practice in Higher Education. In M. Saunders, P. Trowler, & V. Bamber (Eds.), *Reconceptualising evaluation in higher education: The practice turn* (pp. 1-17). McGraw-Hill.
Saunders, M. (2012). The use and usability of evaluation outputs: A social practice approach. *Evaluation, 18*(4), 421–436. https://doi.org/10.1177/1356389012459113
Saunders, M., Sin, C., & Dempster, S. (2015). Evaluative research in higher education policy analysis. In J. Huisman & M. Tight (Eds.), *Theory and method in higher education research Vol.1*, (pp 49–69). Emerald. https://doi.org/10.1108/s2056-375220150000001003
Simiyu, R. R., Okaka, F. O., & Omondi, P. (2015). Geography students' assessment of internship experience at a Kenyan university. *Journal of Geography in Higher Education*, *39*(3), 343–355. https://doi.org/10.1080/03098265.2015.1038780
Smit, S., Tacke, T., Lund, S., Manyika, J., & Thiel, L. (2020, June 20). *The future of work in Europe.* McKinsey Global Institute. https://www.mckinsey.com/featured-insights/future-of-work/the-future-of-work-in-europe
Smith, C. (2008). Building effectiveness in teaching through targeted evaluation and response: Connecting evaluation to teaching improvement in higher education. *Assessment and Evaluation in Higher Education*, *33*(5), 517–533. https://doi.org/10.1080/02602930701698942
Smith, C., Ferns, S., & Russell, L. (2014). *Impact of work integrated learning on student work-readiness*. http://hdl.voced.edu.au/10707/337518
Smith, C., Ferns, S., & Russell, L. (2016). Designing work-integrated learning placements that improve student employability: Six facets of the curriculum that matter. *Asia - Pacific Journal of Cooperative Education*, *17*(2), 197–211.
Smith, C., Ferns, S., & Russell, L. (2019). Placement quality has a greater impact on employability than placement structure or duration. *International Journal of Work-Integrated Learning*, *20*(1), 15–29.
Strack, R., Kaufman, E., Kotsis, A., Sigelman, M., Restuccia, D., & Taska, B. (2019, September 12). *What's trending in jobs and skills*. Burning Glass Technologies. https://www.burning-glass.com/research-project/bcg-trending-jobs-skills/
Taks, M., Tynjala, P., Toding, M., Kukemelk, H., & Venesaar, U. (2014). Engineering students' experiences in studying entrepreneurship. *Journal of Engineering Education. 103*(4), 573–598. https://doi.org/10.1002/jee.20056
Taplin, D. H., & Clark, H. (2012). *Theory of change basics: A primer on theory of change*. ActKnowledge.
Taylor, A. R., & Hooley, T. (2014). Evaluating the impact of career management skills module and internship programme within a university business school. *British Journal of Guidance and Counselling*, *42*(5), 487–499. https://doi.org/10.1080/03069885.2014.918934

Taylor, K. M., & Betz, N. E. (1983). Applications of self-efficacy theory to the understanding and treatment of career indecision. *Journal of Vocational Behavior*, *22*(1), 63–81. https://doi.org/10.1016/0001-8791(83)90006-4

Taylor-Powell, E., & Henert, E. (2008). *Developing a logic model: Teaching and training guide*. University of Wisconsin-Extension, Program Development and Evaluation.

Tertiary Education Quality and Standards Agency. (2017, October 11). *Guidance note: Work integrated learning*. https://www.teqsa.gov.au/latest-news/publications/guidance-note-work-integrated-learning

Tight, M. (2019). The neoliberal turn in higher education. *Higher Education Quarterly, 73*(3), 273–284. https://doi.org/10.1111/hequ.12197

Toledano-O'Farrill, R. (2017). Professional application projects: Work-based learning in the curriculum. *Higher Education, Skills and Work-Based Learning*, *7*(1), 21–34. https://doi.org/10.1108/HESWBL-07-2016-0047

Universities Australia, ACCI, AiGroup, Business Council of Australia, & ACEN. (2015). *National strategy on work integrated learning in university education*.

Venville, A., Lynch, B., & Santhanam, E. (2018). A systematic approach to the evaluation of the student experience in work-integrated learning. *International Journal of Work-Integrated Learning*, *19*(1), 13–21.

von Treuer, K., Sturre, V., Keele, S., & Mcleod, J. (2011). An integrated model for the evaluation of work placements. *Asia-Pacific Journal of Cooperative Education*, *12*(3), 196–204.

Wanzer, D. L. (2021). What is evaluation? Perspectives of how evaluation differs (or not) from research. *American Journal of Evaluation, 42*(1), 28-46. https://doi.org/10.1177/1098214020920710

Wellings, P., Black, R., Craven, G., Freshwater, D., & Harding, S. (2019). *Performance-based funding for the Commonwealth Grant Scheme*. Australian Government Department of Education Skills and Employment.

Whelan, M., & Reichelt-Brushett, A. (2019). Using internship placements to road test threshold learning outcomes for environment and sustainability. *The Journal of Teaching and Learning for Graduate Employability*, *10*(2), 36–49.

Williamson, B. (2019). Policy networks, performance metrics and platform markets: Charting the expanding data infrastructure of higher education. *British Journal of Educational Technology, 50*(6), 2794-2809. https://doi.org/10.1111/bjet.12849

Wilton, N. (2011). Do employability skills really matter in the UK graduate labour market? The case of business and management graduates. *Work, Employment and Society*, *25*(1), 85–100. https://doi.org/10.1177/0950017010389244

Winchester-Seeto, T. (2019). *Quality and standards for work integrated learning*. Australian Council of Deans of Science Teaching and Learning Centre.

Young, K., Palmer, S., & Campbell, M. (2017). Good WIL hunting: Building capacity for curriculum re-design. *Journal of Teaching and Learning for Graduate Employability, 8*(1), 215. https://doi.org/10.21153/jtlge2017vol8no1art670

Zehr, S. M., & Korte, R. (2020). Student internship experiences: Learning about the workplace. *Education + Training*, *62*(3), 311–324. https://doi.org/10.1108/ET-11-2018-0236

APPENDIX A: Summary of the reviewed research and categorization according to the evaluation typology (Table 1).

Citation	Research questions (paraphrased)	WIL activity	Design / theoretical framework	Participants	Location	Findings	Evaluation typology	
							Evaluative practice domains	Underlying judgment / focus
Jackson (2013)	(1) Impact on employability skills. (2) Variations in skill outcomes by demographic, background and/or placement characteristics.	Work placement	Quantitative. Employability skills framework (adapted from Jackson & Chapman, 2012).	131 undergraduates; any study year. Business; Law; Management; Education; Health; Science; Engineering.	Public university; Australia.	WIL significantly improved students' perceived ability to perform all ten employability skills. Study background and demographic characteristics produced minor variations in skill outcomes. For six skills, more time on placement associated with greater confidence in performing skills, and often also associated with greater performance outcomes.	Programmatic, Institutional and Self	Formative / Needs assessment, Process and Outcomes Summative / Outcomes
Jackson (2015)	(1) Importance of classroom and placement learning/assessment activities for employability skill development. (2) Factors impeding skill development and performance.	Work placement	Quantitative and qualitative. Employability skills framework (as above).	132 undergraduates; any study year. Business; Law; Education; Health; Science; Engineering.	Public university; Australia.	Students' perceptions of the classroom and placement activities that are important and effective in developing employability skills broadly aligned to WIL best practice principles. Problems experienced in performing skills during placement can be overcome by good design.	Programmatic, Institutional and Self	Formative / Needs assessment, Process and Outcomes
Jackson (2017)	(1) Influence on Pre-Professional Identity (PPI) development.	Work placement	Qualitative phenomenological study (students' structured reflections).	105 undergraduates; final years. Business.	Public university; Australia.	Work placements, particularly learning and assessment activities of reflection and critical appraisal of experience and current practices, are	Programmatic, Institutional and Self	Formative / Needs assessment, Process and Outcomes

Citation	Research questions (paraphrased)	WIL activity	Design / theoretical framework	Participants	Location	Findings	Evaluation typology	
							Evaluative practice domains	Underlying judgment / focus
	(2) Strategies to improve PPI development.		PPI stages assessed using self-authorship framework (Baxter Magolda, 1998).			important preparation for graduate employment, and differentiate WIL from extra-curricular work experiences. Work placements can offer a valuable platform for fostering identity construction and sensemaking of an intended profession through observation, questioning and interacting with experienced practitioners.		
Jackson (2019)	In respect to developed Professional Identity (PI) capabilities: (1) Differences in student/ supervisor perspectives. (2) Differences based on individual characteristics. (3) Changes. (4) Challenges experienced.	Work placement	Quantitative and qualitative. Framework of 17 PI capabilities (as per Jackson, 2017).	212 undergraduates; 161 workplace supervisors. Business.	Public university; Australia.	Supervisors and students reported improvements to professional identity capabilities during work placements. Students broadly agreed on strengths and weaknesses, yet supervisors were more generous with ratings. International students recorded lower mean ratings on capabilities related to confidence, communication and teamwork.	Programmatic, Institutional and Self	Formative / Needs assessment, Process and Outcomes

Citation	Research questions (paraphrased)	WIL activity	Design / theoretical framework	Participants	Location	Findings	Evaluation typology	
							Evaluative practice domains	Underlying judgment / focus
Jackson & Collings (2018)	(1) Influence of WIL on graduate employment and under-employment. (2) Influence of paid employment during studies on graduate employment and underemployment. (3) Graduate perspectives on inhibitors/enablers to employability/ employment.	Practical work experience and/or placement	Quantitative. No theoretical framework.	Two graduate samples: 628 domestic (2013); 237 domestic and international (2015). Completed an undergraduate degree with WIL and/or paid work in final year. Excluded Nursing and Education.	Western Australian university	Participation was not associated with increased full-time employment rates but there was evidence suggesting it may lead to higher quality and more relevant employment, short and long term. Paid employment in final undergraduate year produced higher full-time employment rates but had little effect on underemployment.	Programmatic, and Institutional	Formative / Needs assessment, Process and Outcomes
Jackson, et al. (2019)	(1) Extent to which students perceive they transferred skills/knowledge across classroom and work settings. (2) Inhibitors/ barriers to transfer. (3) Strategies facilitating transfer.	Workplace-based experience	Mixed method. Transfer or preparation for future learning (Bransford & Schwartz, 1999)	151 undergraduates and postgraduates; 24 industry supervisors. Second half of degree. Completed at least 100 hours of WIL (typically unpaid) in current or previous	Three universities: Western Australia (N=97), New South Wales (N=7) and New Zealand (N= 47).	Students practice skill and knowledge transfer but do it more during less complex, discipline-specific tasks than generic ones. WIL augments transfer and certain program and work characteristics can enhance students' confidence and capabilities in transfer including, for example: pre-placement preparation; rigorous screening to appropriately match students to placement;	Programmatic, Institutional and Self	Formative and Summative / Needs assessment, Process and Outcomes

Citation	Research questions (paraphrased)	WIL activity	Design / theoretical framework	Participants	Location	Findings	Evaluation typology	
							Evaluative practice domains	Underlying judgment / focus
				semester. Business; Sociology; Sport/Recreation.		creating collaborative workplace environments; and workplace supervisors' familiarity of students' coursework requirements.		
Jackson & Wilton (2016)	(1) Extent of career management competencies. (2) Impact on competency development. (3) Variation in competencies by individual characteristics.	Work placement	Quantitative and qualitative. DOTS (decision-making; opportunity awareness; transition learning; self-awareness) career management framework (Peterson et al., 1991).	2 samples. 480 students, presumably undergraduates, 1st - 3rd year, Business.	Two 'vocationally focused' universities: UK (N=136) and Australia (N=344).	Students considered themselves reasonably proficient in career management, yet variations existed across DOTS dimensions depending on whether students participated in WIL, the nature of their experiences, and other study and employment characteristics. Placements fostered students' self-awareness, opportunity awareness, decision-making and transition learning the most; and job search strategies and understandings of the labor market the least. Without concurrent employment, placements benefited the development of career management competencies.	Programmatic, Institutional and Self	Formative / Needs assessment and Outcomes Summative / Outcomes
Nenzhelele (2014)	(1) Impact on employability.	Work placement	Quantitative. Kolb (1984) cycle of experiential	97 students in 'experiential learning course'.	University of South Africa (largest	On average, 85% of students agreed the course improved employability skills.	Programmatic and Self	Summative / Outcomes

Citation	Research questions (paraphrased)	WIL activity	Design / theoretical framework	Participants	Location	Findings	Evaluation typology	
							Evaluative practice domains	Underlying judgment / focus
	(2) Influence of demographics.		learning. Employability skills (Wilton, 2011).	Diploma of Administrative management.	Open Distance Learning university in Africa; Public).	Highest levels of agreement were reported for gains in spoken communication (96.9%), basic computer literacy (94.9%) and written communication skills (90.6%). Gains in advanced computer skills had the lowest level of agreement.		
Rampersad (2020)	Factors influencing development of 'innovation' via WIL.	Work placement involving project-based learning	Quantitative. Develops a theoretical model for driving innovation through confirmatory factor analysis. Based on skills concepts (Taks et al., 2014) and experiential learning pedagogical theory.	111 undergraduates. Course level and discipline unknown (presumably Engineering).	Mid-sized Australian university.	Critical thinking, problem solving, communication and teamwork skills significantly impacted the development of innovation skills and students' perceptions of capabilities in these skills increased post-placement.	Programmatic and Self	Summative / Outcomes
Reddan (2015)	Effects on career self-efficacy and self-confidence in making positive career decisions.	Work experience placement and career development learning	Quantitative. Self-efficacy theory (Bandura, 1997; Leong & Barak, 2001). SOAR Model of Career Development Learning (Kumar,	15 undergraduates (entire cohort), 2nd year, Exercise Science.	Griffith University, Gold Coast campus. Public Australian university.	WIL and career development learning increased students' confidence, knowledge of specific occupations, goal selection, planning and problem-solving capabilities, and awareness	Programmatic, Institutional and Self	Summative / Outcomes Formative / Needs assessment and Outcomes

Citation	Research questions (paraphrased)	WIL activity	Design / theoretical framework	Participants	Location	Findings	Evaluation typology	
							Evaluative practice domains	Underlying judgment / focus
			2007). Career Decision Self-Efficacy Scale – short-form (Taylor & Betz, 1983).			of personal strengths and weaknesses related to employability.		
Reddan (2017)	(1) Impact on employability. (2) Student perceptions on course aspects that changed their abilities.	Fieldwork placement, workshop and related career development assignments	Quantitative. Six dimensions of employability (Smith, et al., 2014). SOAR as a pedagogical tool (Kumar, 2007).	8 undergraduates (entire cohort), 2nd year, Exercise Science.	Griffith University, Gold Coast campus.	Pre-post comparison of work readiness showed improvements for all employability dimensions except informed decision-making. Students rated placements as having a greater impact on abilities than workshops and assignments. However, all three course components contributed to the development of workplace competencies. Students indicated the course increased awareness of employability strengths and weaknesses, and knowledge of specific occupations.	Programmatic and Self	Summative / Outcomes
Taylor & Hooley (2014)	Impact and efficacy of career management skills module and/or internship programme on graduate employment and	Industrial placement and/or career management skills module comprised of lectures, seminars	Mixed methodology informed by cited studies. DOTS model (Law & Watts, 1977): self-awareness, opportunity	Business undergraduates. Part 1 (national employment survey): three graduate sample groups over two years: 73	UK university.	Structured work experiences associated with improved ability of to secure employment in 'graduate level' jobs within six months of graduation. Graduates who completed the module also had higher	Programmatic, Institutional and Self	Formative / Needs assessment and Process Summative / Outcomes

Citation	Research questions (paraphrased)	WIL activity	Design / theoretical framework	Participants	Location	Findings	Evaluation typology	
							Evaluative practice domains	Underlying judgment / focus
	level of employment.	and practical exercises.	awareness, decision-making and transition into employment.	completed module and placement; 110 only completed module, no placement; 460 no module or placement. Part 2 (perceptions – surveys): 61 graduates; 24 final year and 22 placement year students (response rate 26.10%).		rates of employment post-graduation. Students who completed both had the greatest employment success.		
Toledano-O'Farrill (2017)	None articulated. Uncover self-reported learning outcomes via analysis of assessment task.	Placements and intervention projects for clients	Qualitative. No theoretical framework as such but social service perspectives underpin the embedded activity.	37 teams of undergraduates. Business. Across two years (student numbers unknown).	Mexican private university with university-wide Professional Application Projects (PAPs).	PAP mostly successful based on students' self-reported learning outcomes.	Programmatic and Self	Summative / Process and Outcomes
Whelan & Reichelt-Brushett (2019)	(1) Assess suitability of Threshold Learning Outcomes (TLO) as an instrument	Internship program	Quantitative. Australian TLO statements.	33 undergraduates (3rd and 4th years); 14 host-supervisors; 10 teaching staff	Southern Cross University; Australian public university.	Students exceeded expectations of host-supervisors for all TLOs. Hosts expected graduates to be capable performers and rated overall performance of	Programmatic, Institutional and Self	Formative and Summative / Needs assessment

Citation	Research questions (paraphrased)	WIL activity	Design / theoretical framework	Participants	Location	Findings	Evaluation typology	
							Evaluative practice domains	Underlying judgment / focus
	to quantify graduate employability. (2) Quantify TLO performance expectations of employers and whether students meeting expectations.			(2015). 2016 graduate survey data. Environmental Science.		interns as capable. Teaching staff rated performance lower than students and host-supervisors. Results indicated that the degree met the needs of industry and graduates seeking professional work in the discipline.		and Outcomes