3-11-2021

# Implementing programmatic assessment transforms supervisor attitudes: An explanatory sequential mixed methods study

Janica Jamieson
*Edith Cowan University*

Margaret Hay

Simone Gibson

Claire Palermo



10.1080/0142159X.2021.1893678
This is an Accepted Manuscript of an article published by Taylor & Francis in MEDICAL TEACHER on 11/03/2021, available online:http://www.tandfonline.com/10.1080/0142159X.2021.1893678.
Jamieson, J., Hay, M., Gibson, S., & Palermo, C. (2021). Implementing programmatic assessment transforms supervisor attitudes: An explanatory sequential mixed methods study. *Medical Teacher, 43*(6), 709-717.
https://doi.org/10.1080/0142159X.2021.1893678

## Abstract

**Introduction:** Programmatic assessment (PA) is an increasingly popular approach to competency-based assessment (CBA), yet evaluation evidence is limited. This study aimed to identify and explore supervisor attitudes before and after implementing a novel PA using a sequential explanatory mixed methods design. In phase one, a survey was used to identify supervisor perspectives on work-based placements, PA and CBA. Survey results were then applied to develop focus group questions to further explore supervisor attitudes. **Results:** PA was found to improve supervisor-student relationships by removing high-stakes assessment decisions and creating greater capacity for feedback and teaching, leading to a productive learning environment. Assessment was perceived as an important role and supervisors wanted to feel valued and heard within PA. Trust was conceptualised as a triad between supervisor, student and university, and enabled supervisors to engage with PA which was important for success. Supervisor learning of PA was experiential and often supported by students, highlighting the need for hands-on training. **Conclusion:** Participants reported a high level of agreement with PA and CBA principles which may have made them amenable to educational change. Further research is needed to explore the experience of all stakeholders and to understand how worldviews and culture influence assessment initiatives.

## Introduction

Competency-based education (CBE) is the predominate approach to training and certifying health professionals (Holmboe 2015). The emergence of CBE has initiated a transformation in competency-based assessment (CBA) with a move away from individual instruments towards a holistic programmatic assessment (PA) approach (Iobst and Holmboe 2020).

PA recognises that single instruments are inherently limited when used to assess the complex phenomenon of competence. PA utilises a suite of high-quality instruments to longitudinally collect data on student performance and drive meaningful learning (*assessment for learning*). Assessment is conceptualised on a spectrum of increasing stakes proportional to outcomes. Low-stakes assessments are regular feedback-rich moments that illuminate student development and support progression. Low-stakes assessment data are purposefully combined to provide an information-rich picture of the learner which is used to inform high-stakes assessment decisions (van der Vleuten et al. 2012).

PA engages a range of stakeholders including students, workplace supervisors, mentors/ advisors, and assessment committees, each with a critical role (Meeuwissen et al. 2019; Rich et al. 2020). Workplace supervisors have historically held responsibility for both teaching and high-stakes assessment decisions. This creates a conflict of interest compromising the quality of teaching provided to students and can trivialise assessment. PA seeks to purposefully separate supervisors from high-stakes moments and reorientate their role to teaching and feedback, captured in low-stakes assessment moments (van der Vleuten et al. 2012).

Despite increasing popularity, PA is not a panacea for CBA challenges. PA is a complex intervention and its application is unique to the individual setting. Implementation is shaped by contextual parameters such as institutional procedures, time and financial resources,

stakeholder worldviews, and supervisor engagement, which challenge the translation of principles into practice (Wilkinson and Tweed 2018; Schuwirth and van der Vleuten 2019; Torre et al. 2020). This complexity necessitates robust evaluation to understand the intricacies and mechanisms that enable successful transference of knowledge to practices (Haji et al. 2013; Wilkinson and Tweed 2018; Torre et al. 2020). Emerging research suggests underdeveloped stakeholder relationships and shared understanding may impede PA implementation. An inclusive and collegial approach, which supports supervisors, is needed to empower all stakeholders and overcome implementation barriers (Schut et al. 2021).

Supervisors are highly variable in how they conceptualise and enact their role within PA. This variability is derived from individual worldviews, mentoring approaches, and assessment perceptions, which moderate their practices and acceptance of educational change (de Jonge et al. 2017; Schut et al. 2018; Meeuwissen et al. 2019; Castanelli et al. 2020; Schut et al. 2020). Supervisors are uncomfortable documenting negative feedback (Acai et al. 2019; Rich et al. 2020), find progress decisions emotionally burdensome (Castanelli et al. 2020), and have difficulty adjusting to increased student autonomy (de Jonge et al. 2017; Meeuwissen et al. 2019; Schut et al. 2020).

Given the important role of supervisors within PA, an understanding of their reaction to the implementation of PA is valuable learning. This mixed methods study aimed to identify and explore the attitudes of supervisors involved in the implementation of a novel PA.

## Methods

In 2015, a working group comprised of workplace supervisors and university academics designed a novel PA using a participatory research approach for the 20-week work-based placement in a dietetics program in Western Australia. The novel PA replaced the non-PA approach to CBA which was based on historical antecedent (Jamieson et al. 2017). The

placement occurred in the final 6 months of the two-year course and was undertaken full-time in three settings (clinical, food service and primary prevention) with students supervised by practitioners (workplace supervisors). Students undertook authentic learning tasks with performance captured in low-stakes assessments (e.g., self-reflection; performance appraisals completed by the student, supervisors, and peers; practice outcomes evidence; client perspective). With university guidance, students aggregated low-stakes assessments into a portfolio for the high-stakes assessment. University staff held responsibility for the high-stakes assessment decision in consideration of national professional competency standards. The competencies articulate standards for collaborative nutrition experts with professional attributes and communication skills who provide evidence-based nutrition care (Palermo et al. 2016). All stakeholders, including students, participated in student-led competency development meetings which provided support and where necessary, remediation. At completion of the placement, students are entrusted to independently implement nutrition interventions that address determinants of nutrition and health to enhance outcomes across multiple settings (Begley et al. 2020). Prior to PA implementation, supervisors and students were provided training including CBA literature; PA purpose and goals; role changes, expectations and duties; and the assessment system. The PA was implemented in 2016 and a two-phase explanatory sequential mixed method design was applied to evaluate the PA focusing on workplace supervisor attitudes (Supplementary Figure 1).

In phase one (quantitative), an online survey was administered to the same cohort of supervisors, before (February 2016 and 2017) and after (July 2016 and 2017) PA implementation, to determine baseline attitudes and ascertain change. The 2016 phase one data were then used to inform the phase two (qualitative) methods (Creswell 2014). In phase two (July 2016 and 2017), focus groups were held with supervisors, recruited from the same cohort as phase one, to further explore the identified attitudes. Focus groups were conducted until the

researchers were satisfied that there was adequate breadth and depth of data to answer the research question (Malterud et al. 2016). As such, data collection occurred in both years to maximise sample size. Phase one and two results will be presented separately then integrated in the discussion for interpretation as is typical for sequential mixed method designs (Creswell 2014).

The research was positioned in a constructivist paradigm. This framing influenced the study design by seeking to understand supervisor interactions and interpretations with the PA, enabling their perspective to contribute to the evaluation. One author (JJ) is an academic staff member who coordinated the placement and had existing relationships with many of the supervisors. JJ was a co-researcher and member of the working group to design the PA (Jamieson et al. 2017) and has a thorough understanding of the PA. The author (JJ) adopted a critical reflexive position to recognise and mitigate potential biases resulting from this insider position (Berger 2015). In addition, focus group analysis were undertaken and verified by a second author (CP) who was an independent external expert for the working group (Jamieson et al. 2017) and has a understanding of the PA. CP, MH and SG have experience in health professional assessment and mixed methods research providing an additional lens to data interpretation. The study received ethics approval (Edith Cowan University 12549 and Monash University 8925) and all participants provided written informed consent for both phases.

***Phase one (quantitative)***

Survey statements were derived from literature (van der Vleuten et al. 2012), working group priorities (Jamieson et al. 2017), and, with permission, a survey determining dietetic supervisor attitudes (Nasser et al. 2014). The survey had four sections: (i) identification code for matching participant responses and demographic data (Table 1); (ii) 14 statements exploring attitudes towards placement; (iii) 13 statements exploring attitudes towards CBA;

and (iv) 10 statements exploring attitudes towards PA. A purposeful mix of positive and negative statements were used to avoid participant acquiescence (Rattray and Jones 2007). Participants responded to the same statements before and after engaging with the PA. Level of agreement for each survey statement was indicated using a five-point Likert scale with open-text responses following each section to capture additional information. Researchers collected validity evidence through expert review (CP, SG and MH). Modifications were made to the question content, style and structure to improve usability and address the research question. The survey was designed and administered using Qualtrics™ (Provo, UT).

Supervisors engaging with the PA for the first time were emailed the survey link before and after the placement. A reminder email was sent after one week to maximise response rates. Demographic data for participants is presented as frequency or mean ± SD (range). The statement response data were entered into SPSS™ (IBM, version 26) for analysis with non-parametric tests applied. Descriptive results for statements from the two points (before and after) are given as percent agreement, representing both 'agreed' and 'strongly agreed'. Participant responses were matched and analysed using the Wilcoxon signed-rank test. The Benjamini-Hochberg procedure, which controls for the false discovery rate, was applied to raw p-values, in consideration of the multiple tests conducted with a small sample. The q-values are presented with <0.05 considered statistically significant (Benjamini and Hochberg 1995). Effect size was determined using Cohen's r equation for non-parametric methods ($r= z/\sqrt{N}$). An r value of 0.1 was considered a small effect, 0.3 medium, and 0.5 large (Fritz et al. 2012). Open-text survey responses were imported into NVivo™ 12 (QSR International, version 12) and analysed using the framework developed for phase two data by one author (JJ). A second author (CP) reviewed and confirmed the analysis.

***Phase two (qualitative)***

Phase two utilised focus groups to further explore the attitudes identified by supervisors in phase one. The 2016 phase one results were descriptively analysed by one author (JJ) and were used to inform the development of phase two focus group questions, along with the working group priorities (Jamieson et al. 2017) and literature (van der Vleuten et al. 2012). Questions explored the experience of, and attitudes towards, the PA, gave comparison to previous non-PA experiences, considered the credibility and dependability of the PA approach, and applicability to differing student needs (Supplementary Table 1).

Focus group participants were workplace supervisors recruited from the phase one cohort who may or may not have completed the phase one survey. An email invitation to participate was circulated to supervisors at each placement site. Where possible, focus groups were held at participants' place of employment to maximise attendance. Two focus groups were conducted using either telephone or videoconference for non-metropolitan participants. Participant demographic information was collected. Focus groups were between 30 and 90 minutes and were conducted by the first author. Focus groups were audio-recorded and the author took notes (Barbour 2007). Focus group data was transcribed verbatim and analysed according to the framework analysis method (Gale et al. 2013). Two authors (JJ and CP) inductively open coded three focus groups by hand. The authors then compared codes and agreed upon an analytical framework by grouping codes into categories. One author (JJ) developed an initial framework with 14 codes and 71 sub-codes, each with a definition, which was reviewed by the second author (CP) with minor changes. The framework and all transcribed focus groups data were entered into NVivo$^{TM}$ for subsequent analysis. One author (JJ) applied the initial framework to three focus group transcripts using the indexing method which was reviewed by a second author (CP). Minor changes were made to the framework with the clarification of one sub-code and the addition of three new sub-codes. The final framework contained 14 codes and 68 sub-codes and was applied by one author (JJ) to all

remaining transcriptions. NVivo$^{TM}$ was used to produce a framework matrix which was exported to Microsoft Excel. The framework matrix contained the coded text for each focus group across all codes and sub-codes. This data were summarised and synthesised initially by one author (JJ) to develop themes through iterative readings of the data (both coding and focus group transcripts) with references to the study aim. The resulting themes were then reviewed by the second author (CP) for agreement, with minor changes.

## Results

### *Phase one (quantitative)*

The surveys were emailed to 169 supervisors at 13 sites (n=144 across 13 sites in 2016 and n=25 across 4 sites in 2017). For 2016 and 2017 combined, the before had 50 responses (response rate 30%) and 36 for the after survey (response rate 21%). Twenty respondents were matched using the unique identification code. Two respondents were removed as they had not engaged with the novel PA, leaving eighteen participants for analysis (n=17 in 2016 and n=1 in 2017) (Supplementary Figure 2). Four participants self-identified as working group members involved in designing the PA. Demographic and statement responses for these participants were descriptively compared to non-working group respondents with no observed differences, therefore their data were included in analyses. Demographic data and statement responses for all participants were descriptively compared to the matched subset with no observed differences, indicating those included in analyses were representative of the total sample.

Demographic data for matched participants is given in Table 1. Participants reported supervising students across clinical (n=11), primary prevention (n=5) and food service (n=2) settings. Ten (56%) had attended PA training with no observed differences in statement responses to those who did not attend, when compared descriptively. None of the statements achieved statistically significance (Table 2).

[Table 1]

[Table 2]

***Phase two (qualitative)***

Seven focus groups were held in 2016 (n=25 supervisors) and two in 2017 (n=7 supervisors) with an average of 4 participants per focus group (range 1 – 5). This sample reflected 19% of practitioners at 13 placement sites (n=144 across 13 sites in 2016 and n=25 across 4 sites in 2017) (Table 1). Participants reported supervising students across clinical (n=26), primary prevention (n=6) and food service (n=4) settings. Seventeen (53%) had attended PA training. One participant was a member of the working group to design the PA.

Five themes were identified that explored supervisor attitudes towards assessment (Table 3), after the implementation of PA: (i) relationship transformation; (ii) valuing voices; (iii) student-led approach; (iv) fit-for-purpose assessment; and (v) importance of practice and training.

*(i) Relationship transformation.* Implementing PA facilitated a transformation in the supervisor-student relationship, shifting from an assessment focus to supporting student development, creating an enjoyable and productive teaching environment. Supervisors described high-stakes assessment decisions as burdensome and impeding student support. Removing responsibility for high-stakes assessment transformed their role from authority figure to supportive mentor. Supervisors described feeling comfortable and relaxed within their new role, enabling them to cultivate a productive learning environment. While student performance appraisal was perceived as being innate and fundamental, removal of high-stakes

decisions facilitated greater opportunity for direct performance observation, constructive feedback and teaching to support student development.

The transformation of the supervisors' role created an open relationship with students. Supervisors perceived students to be comfortable and less anxious regarding their performance. Students were described as shifting from passive to active individuals holding greater responsibility for their learning with increased confidence communicating needs. This reduced the reliance and demand on supervisors. The change in supervisor role had a positive impact on student development and promoted improved performance and competence development.

*(ii) Valuing voices*. Supervisors wanted their opinions of student performance to be heard and valued within any assessment approach, including PA. Student assessment was considered a significant responsibility and supervisors were exceptionally concerned that their opinion would not be incorporated and considered in high-stakes decisions determined by the university. Supervisors were previously (in the non-PA approach) uncomfortable having sole responsibility for assessment decisions, which created stress and was perceived a conflict of interest. Ultimately, supervisors wanted their opinion of student performance to be recognised and meaningfully contribute to high-stakes assessment, particularly for underperforming students.

Supervisors described a change to their understanding of best practice assessment. They recognised their opinion as a singular interpretation of student performance and articulated that CBA was underpinned by multiple opinions collected over time. This facilitated an acceptance of their role within PA and alleviating concerns of not being heard or valued.

Supervisors initially described inconsistency in assessment and interpretation of competency standards between themselves, which they perceived to compromise the rigour and quality. They noted the variability could be confusing for students. Reflecting on this, supervisors began to place value on assessment that collected multiple opinions over time and

perceived the university in a position to holistically assess a student based on multiple perspectives. Importantly, supervisors needed to trust that the university valued and consider their opinion when making assessment decisions.

*(iii) Student-led approach.* Supervisors articulated seven benefits to a student-led assessment approach: empowering students; providing insight into student reasoning and decision-making; facilitating meaningful performance discussion; promoting a collaborative, transparent and open relationship; enhancing student self-reflection; driving meaningful and rich assessment as students critically reflected on their experiences; and reducing the burden and responsibility for supervisors. Supervisors recognised that a student-led approach better prepared students for the workforce through the development of independence, initiative, responsibility, time management and workload prioritisation. These attributes and skills were considered fundamental to the workplace but were challenging to assess. Supervisors noted that supervisor-led assessment methods did not facilitate the development of such skills within students.

*(iv) Fit-for-purpose assessment*. PA was perceived to facilitate meaningful feedback and discussion of student performance, which led to an accurate and global assessment by the university. Supervisors valued fit-for-purpose and efficient approaches to communicate their perceptions of student performance whilst minimising the burden. Fit-for-purpose was conceptualised as a context-sensitive approach to assessment that achieved the PA purpose. Supervisors expressed concern that within a student-led PA, students could selectively present favourable assessment moments to mask performance issues. However, as supervisors became familiar with the assessment principles, in that multiple opinions were collected over time using transparent methods, they recognised students would have difficulty being covert.

Supervisors agreed that PA enabled robust assessment decisions as it utilised fit-for-purpose instruments which had relevant criteria and facilitated feedback and discussion

inclusive of all stakeholders. Documented evidence of performance was of a high standard and the collection of multiple opinions over time was perceived to overcome individual assessor bias. Some supervisors reported that the assessment would be accurate for "good" students but were hesitant to extend this to underperforming students. This apprehension arose from a reliance on "gut-feeling" when undertaking CBA which invoked vulnerability, exacerbated by underperforming students. This inhibited supervisors from commenting on the applicability of PA for underperforming students without personal experience.

Supervisors perceived the PA to facilitate meaningful and targeted performance feedback that was holistic rather than focussed on individual skills and attributes, promoting student confidence and development. The PA was described as facilitating open and transparent dialogue between all stakeholders that supported the unpacking of performance in a collaborative and non-confrontational manner. This led to early identification and intervention of issues before escalation.

Supervisors described the PA as flexible and adaptable to the setting, enhancing authenticity. This alleviated the need for supervisors to create artificial learning moments to "tick-boxes" which had occurred previously. Supervisors were divided on the inclusion of qualitative (narrative) and quantitative assessment methods. Most preferred the inclusion of narrative as it provided detail and flexibility when describing competence. It also provided an opportunity to holistically summarise student performance and facilitated discussion. However, some supervisors wanted more quantitative (lists and tick boxes) as this was perceived to be measurable and objective. The previous non-PA model was described as being frustrating, inflexible, repetitive, confusing, and having unclear criteria. The novel PA alleviated these issues and was considered efficient, reducing the burden, which was important for busy supervisors.

*(v) Practice and training are important.* Initially, supervisors felt uncertain using the PA despite formal training being offered prior to implementation. They articulated that experience and practice engaging the PA accelerated their learning, experience and confidence. As such, they highly valued on-the-ground support. Students were an important learning asset as they were often more experienced and confident than supervisors, and therefore could provide guidance.

## Discussion

This study applied an explanatory sequential mixed method to identify and explore the attitudes of supervisors involved in the introduction of a novel PA. The results demonstrate that a successfully implemented PA can transform the role of supervisors and their comprehension of assessment practices, facilitating a feedback-rich environment for students which optimises teaching, learning and assessment. All components of PA, as previously proposed (van der Vleuten et al. 2012), are required for success.

The supervisor-student relationship is a powerful influence on student learning and development (Carless et al. 2012; Massie and Ali 2016; Pitkänen et al. 2018), extending beyond the student to caring for clients (Kilminster and Jolly 2000). Supervision encompasses the complex dual role of teaching and assessment (Pitkänen et al. 2018; O'Connor et al. 2019) which can be a conflict of interest and inflate already subjective performance interpretations, compromising the credibility of assessment (van der Vleuten et al. 2012; Trede and Smith 2014; Bacon et al. 2017; Lockyer et al. 2017). This dual role is stressful, creating tension which compromises the supervisor-student relationship (Trede and Smith 2014; Cantillon et al. 2019; Castanelli et al. 2020). In the PA model implemented and examined in this paper, high-stakes assessment decisions were intentionally removed from supervisors and their role was focused onto teaching and feedback-rich low-stakes assessments. This was found to reduce supervisor

burden and improved their responsiveness to meeting student needs, creating a productive learning environment.

Separating supervisors from high-stakes assessments can enhance student learning and assessment credibility (van der Vleuten et al. 2012) but risks uncoupling those producing low-stakes assessment evidence (supervisors) from those using the evidence for high-stakes decisions (university-based committees), which may compromise congruence (Rich et al. 2020). PA needs to operate within an inclusive systems approach, with robust dialogue between all stakeholders, allowing for an understanding of individual roles, the overall system, and how assessment decisions occur (Acai et al. 2019; Rich et al. 2020; Schut et al. 2021). Our research supports this, as we found supervisors sought to understand how their interpretations of student performance fit within the PA, particularly how they were incorporated into high-stakes assessment decisions. Supervisors wanted their opinions to have value within the PA as they perceived assessment as an innate and important role. Underlying this was a need for supervisors to trust the PA and those making high-stakes decisions. Trust within assessment practices has been conceptualised between the supervisor and student, with a mutual responsibility being held by both (de Jonge et al. 2017). Our research extends this concept to a triad including the supervisor, student and university. Trust within assessment, achieved through open dialogue, allows supervisors to engage with PA which is vital for success.

Implementing PA led to a positive change in supervision practices and comprehension which aligned with best practice (Lockyer et al. 2017). This is an important finding, as engaging supervisors in formal CBA training is challenging (Massie and Ali 2016) and may not influence behaviour (de Jonge et al. 2017). Although effective supervisor training techniques are reported (Lockyer et al. 2017), participation can be minimal and instead supervisors rely on experience and senior colleagues to develop skills (Spencer 2003; Massie and Ali 2016; Lockyer et al. 2017). Although offered, we observed poor supervisor engagement with formal training when

introducing our PA. Rather, learning was experiential, at times provided by students, and occurred through ongoing engagement with the PA leading to understanding, acceptance, and confidence. This reflects the need for supervisor training to be purposefully embedded and longitudinal (Massie and Ali 2016; Lockyer et al. 2017) and exemplifies that all stakeholders are learners requiring regular feedback (Schut et al. 2021).

Assessment practices are shaped by an individual's worldview which has significant implications for student learning and development (de Jonge et al. 2017; Cantillon et al. 2019; Meeuwissen et al. 2019). Supervisors engaging with our PA described a highly favourable transformation in assessment attitudes and practices. The worldviews of supervisors were not explored in our research but may have influenced participation in the voluntary research and acceptance of PA. This is supported by the quantitative results which indicate a high level of agreement with principles of CBA and PA both before and after the implementation of the PA. Research by Meeuwissen et al. (2019) suggests supervisors with a philosophy favouring a reflective and student-led approach may be amenable to PA. This highlights the need for further research to unpack the complex interplay between worldviews, assessment, and culture, in order to understand supervisor acceptance or resistance to assessment change (Torre et al. 2020).

Our research is limited by the exclusion of other stakeholders including students and university. Further research is warranted to explore the experience and role of all stakeholders. Student attributes, such as intrinsic motivation and self-directedness in learning, promote feedback engagement and may favour PA (Sargeant et al. 2010; de Jonge et al. 2017). The small sample size and the single context limits generalisability and transference of results. However, the mixed method approach and rich context description seek to overcome this by allowing readers to understand the PA and data, which can be adapted to their own settings (Torre et al. 2020). The propinquity of first author (JJ) to the PA may have led to an

unconscious bias in data analysis. Adoption of a critical reflexive position and critical oversight by co-researchers (CP, SG and MH), who were outsiders to the PA, sought to identify and mitigate such biases (Berger 2015). The small sample size further limits the quantitative analysis and calculation of statistical significance and effect size (Bakker et al. 2019; Kraft 2020).

## Conclusion

This mixed method study provides valuable insight into the acceptability of PA by workplace supervisors. It highlights the potential for PA to transform the supervisor role which enhances the supervisor-student relationship and supports quality feedback for student development. Importantly, supervisors need to understand the PA and trust that their interpretations of student performance will be valued and considered. The supervisors in our study demonstrated an inclination for the principles of CBA and PA which may have made them amenable to change. Further research is warranted to unpack the impact of supervisor worldviews and explore the experience of other stakeholders.

## Acknowledgements

The authors wish to acknowledge Dr Jonny Lo for statistical analysis guidance.

## References

Acai A, Li SA, Sherbino J, Chan T. 2019. Attending Emergency Physicians' perceptions of a programmatic workplace-based assessment system: The McMaster Modular Assessment Program (McMAP). Teach Learn Med. 31(4):434-444.

Bacon R, Holmes K, Palermo C. 2017. Exploring subjectivity in competency-based assessment judgements of assessors. Nutr Diet. 74(4):357-364.

Bakker A, Cai J, English L, Kaiser G, Mesa V, Van Dooren W. 2019. Beyond small, medium, or large: points of consideration when interpreting effect sizes. Educ Stud Math. 102(1):1-8.

Barbour R. 2007. Qualitative research kit: doing focus groups. 2nd ed. London: Sage.

Begley A, Bird A, Palermo C. 2020. Developing national conceptual understanding to describe entry-to-practice dietetics competence. J Nutr Educ Behav. 52(4):351-358.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc. 57(1):289-300.

Berger R. 2015. Now I see it, now I don't: researcher's position and reflexivity in qualitative research. Qual Res. 15(2):219-234.

Cantillon P, Dornan T, De Grave W. 2019. Becoming a clinical teacher: identity formation in context. Acad Med. 94(10):1610-1618.

Carless SA, Robertson K, Willy J, Hart M, Chea S. 2012. Successful postgraduate placement experiences: what is the influence of job and supervisor characteristics? Aust Psychol. 47(3):156-164.

Castanelli DJ, Weller JM, Molloy E, Bearman M. 2020. Shadow systems in assessment: how supervisors make progress decisions in practice. Adv Health Sci Educ. 25(1):131-147.

Creswell J. 2014. Research design: qualitative, quantitative, and mixed methods approaches. 4th ed. Los Angeles (CA): Sage.

Dalton M, Davidson M, Keating J. 2011. The Assessment of Physiotherapy Practice (APP) is a valid measure of professional competence of physiotherapy students: a cross-sectional study with Rasch analysis. J Physiother. 57(4):239-246.

de Jonge LPJWM, Timmerman AA, Govaerts MJB, Muris JWM, Muijtjens AMM, Kramer AWM, van der Vleuten CPM. 2017. Stakeholder perspectives on workplace-based performance assessment: towards a better understanding of assessor behaviour. Adv Health Sci Educ. 22(5):1213-1243.

Fritz CO, Morris PE, Richler JJ. 2012. Effect size estimates: current use, calculations, and interpretation. J Exp Psychol. 141(1):2-18.

Gale N, Heath G, Cameron E, Rashid S, Redwood S. 2013. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol. 13(1):117-124.

Haji F, Morin MP, Parker K. 2013. Rethinking programme evaluation in health professions education: beyond 'did it work?'. Med Educ. 47(4):342-351.

Holmboe SE. 2015. Realizing the promise of competency-based medical education. Acad Med. 90(4):411-413.

Iobst WF, Holmboe ES. 2020. Programmatic assessment: the secret sauce of effective CBME implementation. J Grad Med Edu. 12(4):518-521.

Ivankova NV. 2014. Implementing quality criteria in designing and conducting a sequential QUAN → QUAL mixed methods study of student engagement with learning applied research methods online. J Mix Method Res. 8(1):25-51.

Jamieson J, Jenkins G, Beatty S, Palermo C. 2017. Designing programmes of assessment: a participatory approach. Med Teach. 39(11):1182-1188.

Kilminster SM, Jolly BC. 2000. Effective supervision in clinical practice settings: a literature review. Med Educ. 34(10):827-840.

Kraft MA. 2020. Interpreting effect sizes of education interventions. Educ Researcher. 49(4):241-253.

Lockyer J, Carraccio C, Chan M-K, Hart D, Smee S, Touchie C, Holmboe ES, Frank JR. 2017. Core principles of assessment in competency-based medical education. Med Teach. 39(6):609-616.

Malterud K, Siersma VD, Guassora AD. 2016. Sample size in qualitative interview studies: guided by Information Power. Qual Health Res. 26(13):1753-1760.

Massie J, Ali JM. 2016. Workplace-based assessment: a review of user perceptions and strategies to address the identified shortcomings. Adv Health Sci Educ. 21(2):455-473.

Meeuwissen SNE, Stalmeijer ReE, Govaerts M. 2019. Multiple-role mentoring: mentors' conceptualisations, enactments and role conflicts. Med Educ. 53(6):605-615.

Nasser R, Morley C, Cook S, Coleman J, Berenbaum S. 2014. Dietitians' perceptions of precepting: knowledge, skills, attitudes, barriers, and training. Can J Diet Pract Res. 75(1):7-14.

O'Connor A, Cantillon P, Parker M, McCurtin A. 2019. Juggling roles and generating solutions: practice-based educators' perceptions of performance-based assessment of physiotherapy students. Physiotherapy. 105(4):446-452.

Palermo C, Conway J, Beck EJ, Dart J, Capra S, Ash S. 2016. Methodology for developing competency standards for dietitians in Australia. Nurs Health Sci. 18(1):130-137.

Pitkänen S, Kääriäinen M, Oikarainen A, Tuomikoski A-M, Elo S, Ruotsalainen H, Saarikoski M, Kärsämänoj T, Mikkonen K. 2018. Healthcare students' evaluation of the clinical learning environment and supervision: a cross-sectional study. Nurs Educ Today. 62:143.

Rattray J, Jones MC. 2007. Essential elements of questionnaire design and development. J Clin Nurs. 16(2):234.

Rich JV, Fostaty Young S, Donnelly C, Hall AK, Dagnone JD, Weersink K, Caudle J, Van Melle E, Klinger DA. 2020. Competency-based education calls for programmatic assessment: but what does this look like in practice? J Eval Clin Pract. 26(4):1087-1095.

Sargeant J, Armson H, Chesluk B, Dornan T, Eva K, Holmboe E, Lockyer J, Loney E, Mann K, van der Vleuten C. 2010. The processes and dimensions of informed self-assessment: a conceptual model. Acad Med. 85(7):1212-1220.

Schut S, Driessen E, Tartwijk J, van der Vleuten C, Heeneman S. 2018. Stakes in the eye of the beholder: an international study of learners' perceptions within programmatic assessment. Med Educ. 52(6):654-663.

Schut S, Heeneman S, Bierer B, Driessen E, Tartwijk J, van der Vleuten C. 2020. Between trust and control: teachers' assessment conceptualisations within programmatic assessment. Med Educ. 54(6):528-537.

Schut S, Maggio LA, Heeneman S, van Tartwijk J, van der Vleuten C, Driessen E. 2021. Where the rubber meets the road - An integrative review of programmatic assessment in health care professions education. Perspect Med Educ. 10:6-13.

Schuwirth LWT, van der Vleuten CPM. 2019. Current assessment in medical education: programmatic assessment. J Appl Test Tech. 20(S2):2-10.

Spencer J. 2003. The clinical teaching context: a cause for concern. Med Educ. 37(3):182-183.

Torre DM, Schuwirth LWT, van der Vleuten CPM. 2020. Theoretical considerations on programmatic assessment. Med Teach. 42(2):213-220.

Trede F, Smith M. 2014. Workplace educators' interpretations of their assessment practices: a view through a critical practice lens. Assess Eval High Educ. 39(2):154-167.

van der Vleuten C, Schuwirth L, Driessen E, Dijkstra J, Tigelaar D, Baartman L, van Tartwijk J. 2012. A model for programmatic assessment fit for purpose. Med Teach. 34(3):205-214.

Wilkinson T, Tweed M. 2018. Deconstructing programmatic assessment. Adv Med Educ Pract. 9:191-197.

**Practice points**

- PA can be successful implemented and is acceptable to supervisors.
- PA positively transforms supervisor practices creating a feedback-rich environment for students.
- Separating supervisors from high-stakes assessment is achievable, reducing stress and role conflict.
- Supervisors need to feel heard and trust their interpretations are valued.
- Supervisors are learners who gain confidence be engaging with PA, reflecting the need for individualised training.

**Notes on contributors**:

**Janica Jamieson**, Grad Dip Diet (Hons), is a lecturer in Nutrition and Dietetics at Edith Cowan University and a doctoral student at Monash University.

**Professor Margaret Hay**, BBSc (hons), Grad Dip Epi, GCHPE, PhD, MAPS, is the Head, Monash Centre for Professional Development and Monash Online Education in the Portfolio of the Deputy Vice-Chancellor (Education) at Monash University.

**Associate Professor Simone Gibson**, Grad Dip Diet, Grad Cert Health Prof Ed, PhD, is the Director of Education for the School of Clinical Sciences.

**Associate Professor Claire Palermo**, BSci, MNutrDiet, MPH, Grad Cert Health Prof Ed, PhD, is the Director of the Monash Centre for Scholarship in Health Education and Associate Professor in the Department of Nutrition, Dietetics and Food in the Faculty of Medicine, Nursing and Health Sciences.

Table 1. Demographic characteristics of phase one (QUAN) and phase two (Qual) participants.

| **Characteristic** | **Phase one (n=18)** | **Phase two (n=32)** |
|---|---|---|
| *Age* (years) | 33±10 (23–58) | 35±11 (24-65) |
| *Gender:* | | |
| Female | 18 (100%) | 30 (94%) |
| Male | 0 (0%) | 2 (6%) |
| *Time since graduation* (years) | 11±10 (2–37) | 10±8 (2-31) |
| *Geographical location:* | | |
| Metro or large urban (>100,000 people) | 15 (83%) | 28 (88%) |
| Regional centre | 3 (17%) | 3 (9%) |
| Rural or remote | 0 (0%) | 1 (3%) |
| *Work load:* | | |
| Full-time | 15 (83%) | 28 (88%) |
| Part-time | 3 (17%) | 3 (9%) |
| Casual | 0 (0%) | 1 (3%) |
| *Self-reported student supervision experience (Dalton et al. 2011):* | | *n=31* |
| No experience | 2 (11%) | 2 (6%) |
| Some experience | 1 (6%) | 3 (10%) |
| Average experience | 8 (44%) | 12 (39%) |
| Above average experience | 5 (28%) | 8 (26%) |
| Very experienced | 2 (11%) | 6 (19%) |

Table 2. Statement analysis for matched participants

| **Statements** | **Agreement[a] *n(%)*** | | **Z score[b]** | **p-value[b]** | **q value[c]** | **Effect size[d]** |
|---|---|---|---|---|---|---|
| | **Before** | **After** | | | | |
| Thinking back on your pre-PA experience of supervision, please rate the following statements (n=15): | | | | | | |
| I enjoy supervising students | 15(100) | 15(100) | 1.000 | 0.317 | 0.555 | 0.18 |
| I feel confident in my ability to teach dietetic skills and knowledge to students^ | 14(93) | 15(100) | 2.236 | **0.025** | 0.177 | 0.41 |
| It is important that students are closely supervised to ensure patient safety | 13(87) | 12(80) | -0.333 | 0.739 | 0.913 | -0.06 |
| I receive sufficient training in student supervision^ | 10(67) | 14(93) | 2.640 | **0.008** | 0.116 | 0.48 |
| I enjoy assessing student performance | 9(60) | 9(60) | 0.000 | 1.000 | 1.000 | 0.00 |
| I feel confident in my decisions when assessing student performance | 13(87) | 14(93) | 1.890 | 0.059 | 0.206 | 0.35 |
| I feel confident in using the tools/ forms to assess student performance | 10(67) | 10(67) | 1.604 | 0.109 | 0.254 | 0.29 |
| I receive sufficient training on how to make student assessment decisions | 9(60) | 11(73) | 0.879 | 0.380 | 0.590 | -0.16 |
| I feel valued for my contribution to student supervision | 14(93) | 14(93) | 1.890 | 0.059 | 0.206 | 0.35 |
| I feel comfortable allowing students to make mistakes | 12(80) | 11(73) | -0.333 | 0.739 | 0.913 | -0.06 |
| *Supervising students is/ was stressful* | 3(20) | 3(20) | 0.054 | 0.957 | 1.000 | 0.01 |
| *I experience burnout when supervising students* | 2(13) | 2(13) | 0.276 | 0.783 | 0.913 | 0.05 |
| *Supervising students is/ was time consuming* | 13(87) | 11(73) | -1.633 | 0.102 | 0.254 | -0.30 |
| Supervising students adds to my job satisfaction | 15(100) | 15(100) | 1.000 | 0.317 | 0.555 | 0.18 |
| CBA for placement should (n = 18): | | | | | | |
| Use multiple tools to assess student performance | 17(94) | 15(83) | -1.613 | 0.107 | 0.694 | -0.27 |
| Include the perspectives of multiple assessors | 17(94) | 18(100) | 0.302 | 0.763 | 0.870 | 0.05 |
| Use one standardised and validated assessment tool for the dietetics profession | 11(61) | 12(67) | 0.595 | 0.552 | 0.851 | 0.10 |
| Ensure that assessors receive training | 18(100) | 18(100) | 0.447 | 0.655 | 0.851 | 0.07 |
| Separate the roles of teacher and assessor | 14(78) | 16(89) | 0.250 | 0.803 | 0.870 | 0.04 |
| Engage students in active learning | 18(100) | 18(100) | 0.447 | 0.655 | 0.851 | 0.07 |
| Engage students in self-reflection | 18(100) | 18(100) | 0.447 | 0.655 | 0.851 | 0.07 |
| Encourage students to take responsibility for their assessment | 17(94) | 17(94) | 0.577 | 0.564 | 0.851 | 0.10 |
| *Be the sole responsibility of supervisors who observe the student on placement* | 4(22) | 0(0) | -1.807 | 0.071 | 0.694 | -0.30 |
| Encourage students to led the assessment process | 13(72) | 16(89) | 1.261 | 0.207 | 0.851 | 0.21 |
| Include opportunity for regular and meaningful feedback | 18(100) | 18(100) | 0.816 | 0.414 | 0.851 | -0.14 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Include quantitative components | 16(89) | 17(94) | 0.447 | 0.655 | 0.851 | 0.07 |
| Include qualitative components | 18(100) | 18(100) | 0.000 | 1.000 | 1.000 | 0.00 |
| To what extend do you agree with the following statements regarding CBA (n = 18)? | | | | | | |
| Assessment during placement should be standardised | 16(89) | 15(83) | -1.000 | 0.317 | 0.793 | -0.17 |
| The expertise of those completing the assessment tool are more important than the tool itself | 4(22) | 6(33) | 0.465 | 0.642 | 0.802 | 0.08 |
| Judgement is an important part of assessment | 11(61) | 12(67) | 0.546 | 0.585 | 0.802 | 0.09 |
| Objectivity is essential for assessment | 17(94) | 16(89) | -0.707 | 0.480 | 0.802 | -0.12 |
| Subjectivity in assessment is avoidable | 4(22) | 7(39) | 0.465 | 0.642 | 0.802 | 0.08 |
| A good assessment tool should allow any supervisor to perform high quality, reliable assessment | 18(100) | 16(89) | -1.134 | 0.257 | 0.793 | -0.189 |
| *Students should not be aware of assessment processes except their final grade* | 1(6) | 1(6) | 1.342 | 0.180 | 0.793 | 0.227 |
| Assessment during placement is difficult to standardise as it occurs in the 'real world' | 8(44) | 7(39) | -0.351 | 0.726 | 0.806 | -0.06 |
| Assessment processes should be transparent to all stakeholders | 18(100) | 18(100) | 2.000 | 0.046 | 0.455 | 0.33 |
| Assessment should occur as a continuum across placement settings (e.g., food service, community and public health, and clinical) | 18(100) | 15(83) | 0.000 | 1.000 | 1.000 | 0.00 |

*Italicise indicate statements with negative phrasing*

[a] 'Agreed' and 'strongly agreed' are collapsed and presented as 'agreed'

[b] Wilcoxon-signed ranked test

[c] Benjamini Hochberg (False Discovery Rate 0.05)

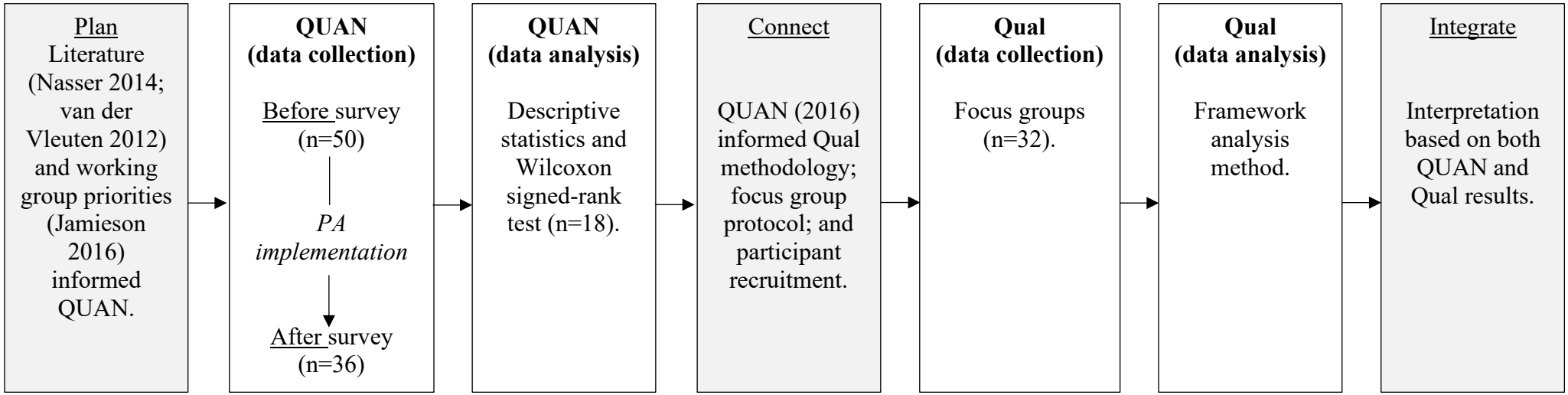
[d] Cohen's effect size

^ Statistically significant using Wilcoxon signed-rank test ($p < 0.05$) but not using Benjamini Hochberg

Table 3. Qualitative themes and quotations

| **Theme** | **Participant quotation** |
|---|---|
| Relationship transformation | *"…before it was more that you were an assessor and now I feel like now I am just their* [the student's] *teacher…that's what that model is trying to achieve so that is definitely why* [what] *I tried to embrace while we were doing it…"* (supervisor 1, FG3)<br><br>*"…we are teachers, not assessors…it's partly us letting go of the role of assessors" (supervisor 1, FG1)* |
| Valuing voices | *"They're* [supervisors] *feeling like their role as the assessor is being taken away and it's all in the hands of…* [the university] *and the student. And that maybe your* [the supervisor's] *opinions won't end up being taken on board* [in high-stakes assessments]*, or if there are problems, it might get overridden, because your role wasn't the assessor; it was just the teacher." (supervisor 2, FG8)*<br><br>*"… as long as the students understand that when they come on placement, the placement supervisor's feedback is also important for their final* [high-stakes assessment] *decision…"* (supervisor 4, FG1)<br><br>*"…it's not just this one person* [supervisor] *who they* [students] *might not necessarily get along with…there are a lot of other people involved in the final* [high-stakes assessment] *decision. I think that is extremely beneficial…"* (supervisor 3, FG5)<br><br>*"You* [the university] *have a very holistic view because you've seen their academic work as well as what's been fed back to you from multiple points. We're only seeing one aspect* [of student performance]*, so I think you* [the university] *having a broader view, and an experience with the student, I think* [it] *is a good thing."*(supervisor 3, FG2) |
| Student-led approach | *"… it puts the onus on the student and…they can't ignore responsibility for their learning and they have to proactively go and seek certain things* [learning or assessment tasks]*….That's…great for us because we not only have to manage the student, but manage your normal caseload…With the student having to manage* [assessment]*, that's the real world…"* (supervisor 2 , FG5)<br><br>"*It prepares them* [student] *for the workplace where they have to take responsibility for their own actions… just having this process and encouraging them to do that, I think is a step in the right direction*." (supervisor 2, FG5) |
| Fit-for-purpose assessment | *"…students can do that* [be covert]. *They can get their best educations…they can pick their best patient…and the best nurse that gives them the lovely feedback…but if our* [supervisor] *feedback is reflecting…half the patients they're doing well and half they're not doing well…Then the university can look at that and go, okay, so they've picked their best ones."* (supervisor 1, FG2)<br><br>*"I think it really focused on the students' strengths and weaknesses instead of just going through and ticking certain boxes. So that if a* |

| | |
|---|---|
| | *student was doing really well, you could really focus and give them positive reinforcement and…if the student was struggling in a particular area, you could focus your constructive criticism in that one area."* (supervisor 1, FG9)<br><br>*"I think* [it] *is a much better process definitely than just the previous forms that were used and ticking off boxes…* [the student] *might get this number of boxes ticked at the start and then at the end you get all your boxes ticked. But apart from a few comments it doesn't really capture that journey. Whereas this process* [the PA]*…you see them…grow along the level of competency…And by setting those learning goals and having to discuss a bit more and show more evidence as to why they are where they are, it just captures that* [competency] *much better."* (supervisor 2, FG6) |
| Importance of practice and training | *"It was definitely doing something new and different. At times it took a little bit more to get my head around it* [PA] *and I guess that was me just wanting to make sure we were doing the right things…definitely very student-led and they had a good handle on everything so they guided us through…"* (supervisor 1, FG9) |



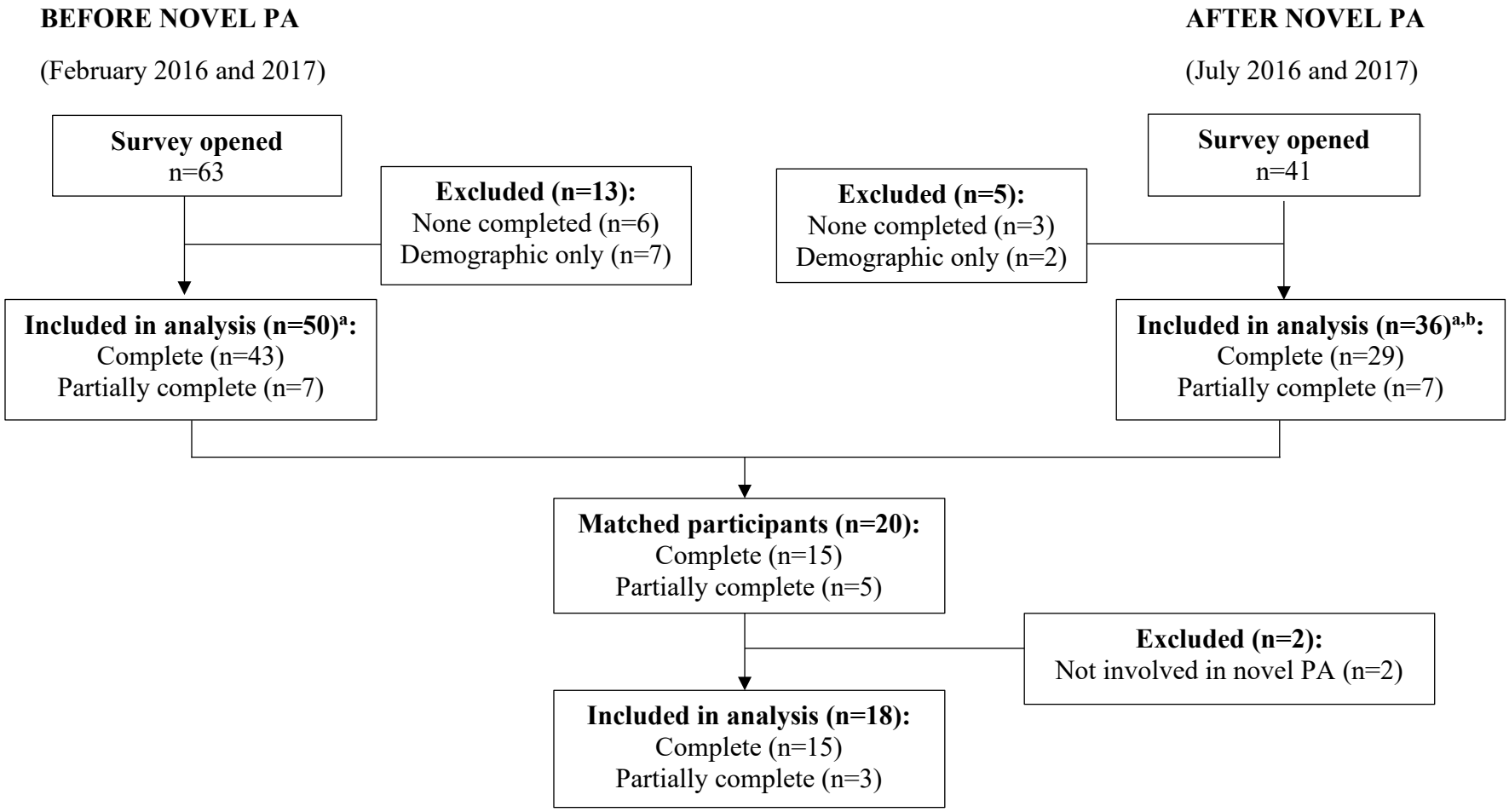

Supplementary Figure 1. Explanatory sequential mixed methods study design and procedure, adapted from Ivankova (2014)

Supplementary Table 1. Focus group questions and influences.

| **Focus group questions** | **Influenced by:** | | |
|---|---|---|---|
| | **Phase one (QUAN)** | **Working Group[1]** | **Literature[2]** |
| Thinking back on your experience of placement in 2016/2017 what thoughts, feelings and words come to mind? | | | |
| How does the new assessment model compare to the previous model? | ✓ | ✓ | |
| Thinking about *other* dietetics training program in which you have been involved, how does the new placement model compare? | | | |
| What does the term *student led* learning mean to you? How do you relate this concept, *student led*, to the new assessment model? | ✓ | ✓ | ✓ |
| Has there been a change in your student supervision role with the new assessment model? OR<br>The new assessment model aims to separate the roles of teacher and assessor. What are your thoughts on this approach? OR<br>Has there been a change in the way you spend your time during student supervision with the new assessment model? | ✓ | ✓ | ✓ |
| The literature states that placement assessment models should enable accurate decisions to be made regarding student performance. What are your thoughts on the ability of the new assessment model to allow accurate decisions to be made about student performance? What aspects of the assessment model allow these decisions to be made? AND/ OR<br>What needs to change in the assessment model to allow these decisions to be made? | ✓ | ✓ | ✓ |
| We conducted an online survey with placement supervisors and found agreement with the concept of assessment occurring as a *continuum across the range of placement settings*, for example, food service, community nutrition and clinical. What does a *continuum of assessment across the range of placement settings* mean to you? How do you relate this concept to the new assessment model? | ✓ | ✓ | |
| The same survey found agreement with the statement "some students are easier to supervise than others". What are your thoughts on this statement? | ✓ | | |
| The survey also asked about the stress related to supervising students and found varying results. In your experience, what factors contribute to a stressful supervision experience for a supervisor? | ✓ | | |

[1] Jamieson et al. (2017)

[2] van der Vleuten et al. (2012)




Supplementary Figure 2. Phase one (quantitative) survey response rate.

[a] One participant completed the before survey in both 2016 and 2017 (and did not complete an after survey in either year), for analysis the 2017 before survey was considered an after survey response.
[b] One participant completed the before survey twice and the after survey once in 2017, for analysis the second before survey response was removed.