Title: Measuring exposure to bullying and harassment in health professional students in a clinical workplace environment: Evaluating the psychometric properties of the Clinical Workplace Learning NAQ-R scale.

Abstract

Background

Instruments that measure exposure to bullying and harassment of students learning in a clinical workplace environment (CWE) that contain validity evidence are scarce. The aim of this study was to develop such a measure and provide some validity evidence for its use.

Method

We took an instrument for detecting bullying of employees in the workplace, called the Negative Acts Questionnaire – Revised (NAQ-R). Items on the NAQ-R were adapted to align with our context of health professional students *learning* in a CWE and added two new factors of sexual and ethnic harassment. This new instrument, named the Clinical Workplace Learning NAQ-R, was distributed to 540 medical and nursing undergraduate students and we undertook a Confirmatory Factor Analysis (CFA) to investigate its construct validity and factorial structure.

Results

The results provided support for the construct validity and factorial structure of the new scale comprising five factors: workplace learning-related bullying (WLRB), person-related bullying (PRB), physically intimidating bullying (PIB), sexual harassment (SH), and ethnic harassment (EH). The reliability estimates for all factors ranged from 0.79 to 0.94.

Conclusion

This study provides a tool to measure the exposure to bullying and harassment in health professional students learning in a CWE.

Introduction

Bullying and harassment of students in health professional education is a significant, ongoing and widespread problem, with evidence from medicine (Fnais et al. 2014), nursing (Clarke et al. 2012), dentistry (Rowland et al. 2010), physiotherapy (Whiteside et al. 2014), and pharmacy (Knapp et al. 2014). In addressing bullying and harassment in a clinical workplace environment (CWE), it is likely to be useful to examine the extent of the exposure – both to establish the extent, and to also measure the impact of any intervention. In health professional education, the prevalence of bullying and harassment is highly variable ranging from 6.3% (Wolf et al. 1998) to 87.4% (Owoaje et al. 2012) in medical students; 45.1% (Ferns and Meerabeau 2008) to 90% (Foster et al. 2004) in nursing students; 34.6% in dentistry students (Rowland et al. 2010); and 25% in physiotherapy students (Stubbs and Soundy 2013). Reasons for this variability are due to the different definitions of what constitutes bullying and harassment, along with different instruments used to measure the phenomena (Einarsen et al. 2011).

A significant amount of learning in health professional education occurs in the CWE, which may account for students in the health professions reporting more bullying and harassment than other higher education students (Rautio et al. 2005). Health professional students are also placed in a setting which is not just a learning environment, but also a working environment delivering healthcare to patients who are in need of care and treatment. It is in this environment that learning about the profession interacts with the provision of the health service, and is also where abuse of health professional students occurs. This is illustrated in a study by Rees et al. (2015) on clinical workplace abuse narratives of students in a variety of health professional education institutions, where many examples of verbal and physical abuse along with sexual and ethnic harassment were described.

When looking at addressing bullying and harassment of health professional students, the leaders of health professional education institutions first need to ask the questions: do we have bullying and harassment occurring at our institution when students are learning in CWE's? If so, to what extent? And lastly, what specific forms of bullying and harassment are occurring? Answering these questions is key to informing the development of a response to bullying and harassment of health professional students. Furthermore, as effective interventions are developed, for example the Creating a Positive Learning Environment (CAPLE) initiative (Gamble Blakey et al. 2019a; Gamble Blakey et al. 2019b; Gamble Blakey et al. 2019c), we need reliable and valid measures to determine impact.

To answer these questions, reliability and validity evidence is needed for an instrument used to measure the exposure to bullying and harassment of health professional students in a CWE. To our knowledge, no studies have been published explicitly demonstrating an instrument's validity and psychometric properties measuring the prevalence and type of bullying and harassment (that includes sexual and ethnic harassment), in health professional education specifically within a CWE. In order to address this issue, we took the following approach to develop such an instrument.

The NAQ-R

An instrument already developed to investigate bullying in the workplace is the Negative Acts Questionnaire – Revised (NAQ-R). The questionnaire has been previously researched to provide validity and reliability evidence for its use and is widely used in measuring exposure to workplace bullying of employees (Einarsen et al. 2009). The NAQ-R contains 22 questions with three factors: work-related bullying, person-related bullying, and physically intimidating bullying.

Previous literature has illustrated that sexual and ethnic harassment are significant factors associated with the bullying and harassment of health professional students (Dineen et al. 1997; Richardson et al. 1997; White 2000; Rautio et al. 2005; Witte et al. 2006; Wilkinson et al. 2006; Garbin et al. 2010; Premadasa et al. 2011; Rees and Monrouxe 2011; Bruce et al. 2015; Rees et al. 2015), however, the NAQ-R does not include behaviours associated with these factors. Therefore, it was determined appropriate to include two additional factors of sexual and ethnic harassment, to provide a more comprehensive account of bullying and harassment.

Purpose of the study

We have called the modified questionnaire the Clinical Workplace Learning NAQ-R scale. Specifically, this study attempts to answer the research question of:

To what extent are the factors present in the modified version of the questionnaire, applicable to a CWE; and what is the effect of adding two new factors on the validity of the instrument? To this end, the psychometric properties of the Clinical Workplace Learning NAQ-R scale were evaluated.

Method

Participants and Procedures

The Clinical Workplace Learning NAQ-R questionnaire was distributed to all undergraduate medical students who were in their clinical years (years 4-6) at the University of Otago's six-year medical degree, and all undergraduate nursing students in their final two years of the Otago Polytechnic's Bachelor of Nursing degree. Years 4-6 of the medical curriculum entails the immersion of medical students in different CWEs for their learning, predominately hospital and general practice learning environments. Year 2 of the nursing degree is when nursing students begin to learn in CWEs, with final year students (year 3) spending the largest amount of time learning in CWEs. The CWEs for nursing students consist of Primary Healthcare settings, hospitals and residential care facilities, with the predominant amount of clinical learning conducted in the hospital environment.

For the medical students, hardcopy questionnaires for year 4 and 5 were administered during whole class sessions. Year 6 students are more geographically dispersed so on-line questionnaires were used. For the nursing students hardcopy questionnaires were distributed during a whole class teaching session. Ethical approval of the study was obtained from the University of Otago Human Ethics Committee and by the Otago Polytechnic Ethics Committee.

Theoretical underpinnings of the original NAQ-R

In the original NAQ-R bullying is defined as *"a situation in which hostile and aggressive actions are systematically directed at one or more persons in such a way that they are stigmatized and victimized"* (Mikkelsen and Einarsen 2001, p.394). Additionally, describing the dimensions of what constitutes bullying in the workplace in which the NAQ-R is situated is necessary to understand its construction. Bullying in the workplace constitutes negative

and unwanted behaviours (Einarsen et al 2011) along with "...evolving and often escalating hostile workplace relationships rather than discrete and disconnected events and is associated with repetition (frequency), duration (over a period of time), and patterning (of a variety of behaviours involved) as its most salient features" (Einarsen, 2009, p.25).

The NAQ-R contains three factors: work-related bullying, person-related bullying, and physically intimidating bullying. Work-related bullying consists of behaviours targeted at an individual's working role and activities such as being given unreasonable deadlines, or meaningless tasks (Einarsen et al 2011). Person-related bullying consists of behaviours that are targeted at the individual themselves for example, spreading gossip or rumours about you, having insulting or offensive remarks made about your person (Einarsen et al 2011). As stated by Einarsen et al 2011, the behaviours associated with person-related bullying "...are by and large, independent of the work organisation" (p.13). Physically intimidating bullying is associated with behaviours targeting the individual with explicit acts of physical aggression or violence or threats of violence (Einarsen et al 2011).

Instrument development

Demographic data collected in the survey included age, ethnicity, sex and sexual orientation. Sexual orientation was included purposively as previous literature indicates that students who are in a minority regarding sexual orientation are more likely to be bullied and/or harassed (Przedworski et al. 2015).

To develop the instrument, we undertook the following processes. The original concept underpinning the NAQ-R (bullying and harassment of employees in a workplace) was modified in order to fit our context (bullying and harassment of students learning in a CWE). Items used in the original NAQ-R were then edited to align with our new context. Finally, we

added two new factors of sexual and ethnic harassment. Both processes were conducted while maintaining the original instrument's behavioural design.

Modifying the concept of bullying and harassment in the workplace

The NAQ-R consists of a three factor model of workplace related bullying, person-related bullying, and physically intimidating bullying. The NAQ-R was originally designed in the context of workplace bullying and harassment of *employees* and therefore, the items in the questionnaire related to the definition of bullying and harassment in the context of *working* as an *employee*. In order to make the NAQ-R effective for measuring exposure of bullying and harassment of *students learning* in a CWE we undertook the following conceptual modifications.

Firstly, we modified the existing definition of bullying used for the NAQ-R of *"a situation in which hostile and aggressive actions are systematically directed at one or more persons in such a way that they are stigmatized and victimized"* (Mikkelsen and Einarsen 2001, p.394) to fit our context *"a situation in which hostile and aggressive actions are systematically directed at one or more students in such a way that they are stigmatized and victimized in a clinical workplace learning environment."*

Then, we edited specific items in the original questionnaire so they would align with this new modified concept of bullying and harassment, from *employees working,* to *students learning* in a CWE. Two statements that did not fit the context of students learning in a clinical environment were removed. One statement from the work-related bullying factor: 'Pressure not to claim something to which by right you are entitled (e.g. sick leave, holiday entitlement, travel expenses)', as it relates to the role of being employed which is not relevant to students. The second statement that was omitted was from the person-related bullying factor: 'Practical

jokes carried out by people you don't get along with'. As health professional students generally move around different clinical workplaces to experience different areas of healthcare practice, we concluded that the behaviour indicated in the statement may occur less frequently. We also re-analysed the remaining 20 statements in the original NAQ-R associated with its three factors, and re-worded 10 statements to align with our modified concept. The changes to the wording of particular items can be seen in Table 1:

--- Insert Table 1 here ---

Two statements obtained from previous literature, were also added to more accurately and thoroughly reflect the context of experiencing bullying and harassment by students learning in a clinical environment. Items added were: Being assigned work for punishment rather than for educational value; and having learning opportunities blocked or withheld by others. These statements were added to the workplace related bullying factor and was re-named the workplace learning-related bullying factor to fit our new modified concept, (see Appendix, Table A1, for all Clinical Workplace Learning NAQ-R Scale Items).

Adding two new factors

Two new factors were introduced in our modified version of the questionnaire: sexual harassment and ethnic harassment (Appendix, Table A1).

To include sexual harassment in our CWE context we adopted Till's definition of 'academic sexual harassment' as the *"use of authority to emphasise the sexuality or sexual identity of a student in a manner which prevents or impairs that student's full enjoyment of educational*

benefits, climate, or opportunities" (Till 1980, p.7). The items used for this factor that aligned with this definition drew on the general sexual harassment literature such as the Sexual Experiences Questionnaire (SEQ) (Fitzgerald et al. 1995); along with questionnaires used in the bullying and harassment literature focused on professional working environments (Crebin et al. 2015); and health professional students (Sheehan et al. 1990; Baldwin et al. 1991; Uhari et al. 1994; Wilkinson et al. 2006; Woolley et al. 2006; Rowland et al. 2010; Clarke et al. 2012; Knapp et al. 2014; Whiteside et al. 2014; AAMC 2018).

We adopted the conceptualisation of ethnic harassment as described by Schneider et al, defined as *"threatening verbal conduct or exclusionary behaviour that has an ethnic component and is directed at a target because of his or her ethnicity"* (Schneider et al. 2000, p.3). The items that were used for this factor aligning with this definition were derived from a combination of the Ethnic Harassment Experiences scale (EHE) (Schneider et al. 2000); and the bullying and harassment literature in the workplace (Keashly 1998; Einarsen, et al. 2011; Crebin et al. 2015) and of health professional students (Sheehan et al. 1990; Baldwin et al. 1991; Uhari et al. 1994; Wilkinson et al. 2006; Woolley et al. 2006; Rowland et al. 2010; Clarke et al. 2012; Knapp et al. 2014; Whiteside et al. 2014; AAMC 2018).

Maintaining instrument design

Two main aspects in the design of the original NAQ-R ensures all items are written using specific behavioural statements, and there are no definitions given about or words mentioning "bullying" or "harassment" when participants undertake the questionnaire (Einarsen et al. 2009). For example, the questionnaire asks participants how often they have been subjected to the following negative acts and gives a list of the terms such as 'Having insulting or offensive remarks made about your person (i.e. habits and background), your attitudes or

your private life', as opposed to asking how the participant feels about the behaviour. Asking about specific behaviours without providing terms or definitions assists in minimising misinterpretations by participants, specifically when developing an instrument investigating incidence and/or prevalence (Arvey and Cavanaugh 1995), such as is the focus of this research. Taking this behavioural approach to the design of the instrument *"is considered to provide a more objective estimate of exposure to bullying behaviours than self-labelling approaches, as respondents' need for cognitive and emotional processing of information would be reduced"* (Einarsen et al. 2009, p. 27). Therefore, when modifying the original NAQ-R, we kept the same approach of using behavioural terminology and not mentioning the words "bullying" or "harassment".

The frequency rating of behaviours in the original NAQ-R was kept, because we determined that it would still fit our context of students learning in a CWE. For this measurement, participants are asked to rate how often they experienced the behaviours listed on a 1-5 point Likert scale where 1 = *Never*, 2 = *Now and then*, 3 = *Monthly*, 4 = *Weekly*, 5 = *Daily*.

While in the NAQ-R, participants are asked to rate their experiences over the past 6 months, we modified this to 8 months for the current questionnaire, as this gives the students the opportunity to reflect on their experiences for the majority of their clinical year.

In summary, in light of these modifications the newly formed 31-item Clinical Workplace Learning NAQ-R scale has been developed to measure health professional students' exposure to negative inter-personal interactions in a CWE. It is comprised of five hypothesised factors: workplace learning-related bullying (WLRB), person-related bullying (PRB), physically intimidating bullying (PIB), sexual harassment (SH), and ethnic harassment (EH) (the Clinical Workplace Learning NAQ-R questionnaire is located in Appendix A2).

Statistical Analyses

We developed and examined the psychometric characteristics of the Clinical Workplace Learning NAQ-R scale using a structural equation modelling approach to explore the validity evidence. A Structural Equation Modelling (SEM) approach, specifically Confirmatory Factor Analysis (CFA), was carried out using MPlus 7 (Muthén and Muthén, 2012). We conducted the data analyses in two stages. Initially, we examined the data for outliers and missing cases. Then, the factorial structure of the Clinical Workplace Learning NAQ-R scale was examined by means of Confirmatory Factor Analysis (CFA). To provide support for the validity evidence for the hypothesised factor structure of the Clinical Workplace Learning NAQ-R scale, we investigated and compared the goodness-of-fit of different competing models as suggested by Noar (2003) and Strauss and Smith (2009).

Confirmatory Factor Analysis (CFA)

The hypothesized model (five-factor) model was compared to three other alternative competing models. The hypothesized model included three factors reported by Einarsen et al. (2009) as well as two new factors that are sexual harassment (SH) and ethnic harassment (EH) all of which are related to each other. The competing models included: a) a one-factor (unidimensional) model that assumed all manifest variables loaded on a single factor, b) a three-factor model that suggests WLRB, PRB, and PIB items loaded on a single factor, and c) a second-order (higher-order) factor model with the five scale factors subordinating to a single second-order factor.

One-factor model means that what we are measuring is a unidimensional construct and university students are not differentiating the bullying and harassment factors. Evidence for a

three-factor model indicates that all three factors reported by Einarsen et al. 2009 are not distinct from each other. Support for second-order model would suggest that these related five factors can be accounted for by an underlying higher order construct. Support for the hypothesized five-factor correlated (oblique) model would suggest that medical and nursing students differentiate between the five bullying and harassment factors that are related to each other.

Since the data were ordered-categorical, the weighted least squares mean and variance adjusted (WLSMV) estimation procedure was used for CFA analyses. The WLSMV is a robust estimation technique that is suggested for modelling ordinal data (Flora and Curran 2004; Brown 2006). The consequences of treating ordinal responses as continuous which may lead to reporting inaccurate results are well-established in the literature (Lubke and Muthén 2004).

A number of different indices were examined to compare the different models and evaluate model-data fit (Cheung and Rensvold 2002; Fan and Sivo 2005, 2007). Each of these measures reflects a different aspect of model fit and may not perform equally well under different types of model conditions (Fan and Sivo 2007). Thus, it is important to use multiple indices rather than relying on one measure (Hair et al. 2010). Indices reported in this study included the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), the Tucker-Lewis index (TLI), and the Weighted Root Mean Square Residual (WRMR). The chi-square (χ^2) values were also reported but not used for model fit decisions as this statistic and its significance are inflated with large sample sizes. The commonly accepted cut-offs for 'acceptable' or 'good' fit (Browne and Cudeck 1992; MacCallum et al.1996; Hu and Bentler 1999; Yu 2002; Hair et al. 2010) included: a non-significant chi-square (χ^2), RMSEA with values < .08 indicating an acceptable fit and values < .05 indicating a good fit, CFI and TLI with values > .90 being indicative of reasonable fit and

values > .95 indicating a good fit, and WRMR with values being close to 1. The limitations of coefficient alpha (α) as a measure of reliability estimate is well documented in the literature (Sijtsma 2009; Teo and Fan 2013). Therefore, using polychoric correlations, we calculated and reported McDonald's omega (ω) (McDonald 1999) as a better estimate of reliability.

Results

Participant and demographic information

A total of 428 from an eligible 852 medical students completed the questionnaire giving a response rate of 50% (428/852). A total of 69 nursing students in year 2 and 43 year 3 completed the questionnaire, from an eligible 212 nursing students, a giving a response rate of 53% (112/212). The questionnaire took approximately 5 minutes to complete.

Therefore, 540 medical and nursing students completed the questionnaire giving an overall response rate of 51% (540/1064). Of the participants, 65.2% (n = 352) were females and 34.4% (n = 186) were males. Only one student did not report their sex and one student identified themselves as transgender. Mean age was 23.7 years (range 19-53 years, SD = 4.35). Self-reported ethnic composition was reported as; New Zealand European (67.8%), Māori (9.4%), Chinese (8.3%), Indian (3.0%), Samoan (1.3%), Cook Island Māori (0.7%), Tongan (0.4%), other ethnicities (20%), and not stated (0.7%). Because individuals can be of more than one ethnicity, these totals are greater than 100%. Sexual orientation was reported as heterosexual (91.7%), homosexual (3.3%), bisexual (2.4%), other (1.5%), and not stated (1.1%).

Descriptive Statistics

No univariate outliers were identified to have an effect on the results. The proportion of missing cases for each item was trivial ranging from mostly zero to one percent. The Expectation Maximization (EM) algorithm which assumes that observations are missing at random (MAR) was utilized to impute the missing cases rather than listwise deletion. The means and standard deviations for the five factors of the Clinical Workplace Learning NAQ-R scale are summarized in Table 2.

--- Insert Table 2 here ---

Factor means ranged from 1.15 to 1.44, suggesting that most students endorsed 'never' or 'now and then' with the statements. However, examination of the item means most of which ranged from 1.00 to 5.00 revealed that there were some students who had been subjected to negative acts on a daily basis. The standard deviations ranged from 0.28 to 0.44 indicating that the dispersion of responses for each factor were somewhat similar.

Confirmatory Factor Analysis (CFA)

The goodness-of-fit measures of hypothesized and alternative models are summarized in Table 3.

--- Insert Table 3 here ---

Evaluation of the unidimensional model revealed that this model was not representing the sample data sufficiently. The RMSEA, CFI, TLI and WRMR values did not meet the commonly acceptable fit criteria. The chi-square statistic and fit indices (RMSEA= 0.045; CFI= 0.955; TLI= 0.955; and WRMR=1.232) suggested that the hypothesized five-factor correlated model provided the best model fit with the data. The three-factor and second-order models had also good fit close to the hypothesized model. However, the chi-square difference test (DIFFTEST) between the second-order and the hypothesized model indicated that adding a higher-order dimension significantly worsened the fit. Thus, the hypothesized model was retained as the model of best fit with the five factors of workplace learning-related bullying (WLRB), person-related bullying (PRB), physically intimidating bullying (PIB), sexual harassment (SH), and ethnic harassment (EH).

For this model, factor correlations and reliabilities are provided in Table 4.

--- Insert Table 4 here ---

The correlations between the Clinical Workplace Learning NAQ-R factors ranged from 0.48 to 0.95. The highest correlations were found between workplace learning-related bullying

(WRB), person-related bullying (PRB), and physically intimidating bullying (PIB). These results were closely consistent with findings of the Einarsen et al. (2009) study. However, we observed moderate correlations between the new factors sexual harassment (SH), ethnic harassment (EH) - and other factors indicating that the evidence of discriminant validity improved with the addition of SH and EH factors.

All of the omega reliability estimates were greater than the recommended level (0.70). Standardized factor loadings for the hypothesized model are provided in Table 5.

--- Insert Table 5 here ---

All the unstandardized factor loadings were significant. Standardized factor loadings for the hypothesized model ranged from .63 to .96 providing support for convergent validity. All of the items were strong indicators of the factors they were related to.

In summary, the CFA analyses and reliability estimates provided support for the validity evidence of the factorial structure of the Clinical Workplace Learning NAQ-R.

Discussion

The aim of this research was to construct and provide validity evidence for a self-report instrument to measure health professional students' experience of bullying and harassment in a clinical workplace learning environment. Previous research looking into the prevalence and type of bullying and harassment that health professional students face were from a variety of instruments lacking validity evidence to support the intended purpose.

This study has the potential to contribute to the literature on measuring exposure to bullying and harassment of health professional students by providing: empirical evidence through supporting the validation of a new instrument; an instrument that is specific to a clinical workplace learning environment; and an instrument that contains more relevant factors associated with literature on bullying and harassment of health professional students.

Using an instrument (NAQ-R) that was developed for a different context (workplace bullying among employees), and with evidence supporting its validity, we modified it to investigate if it would fit in our context of (clinical workplace learning environments for students, not employees). The original NAQ-R contained three factors related to bullying in the workplace: workplace learning-related bullying, person-related bullying and physical related bullying. We modified these three factors and also added two new factors not in the original NAQ-R related to sexual harassment and ethnic harassment, because of the reported prevalence of these types of experiences by students learning in a CWE. Our analyses provide evidence for the validity of these two new factors.

We suggest naming this instrument the Clinical Workplace Learning NAQ-R scale, as this acknowledges the significant body of work by Einarsen et al. (2009) in developing and analysing the psychometric properties of the original NAQ-R designed for workplace bullying of employees.

The results indicated that the original three factors of the NAQ-R that we modified to fit a clinical workplace learning (workplace learning-related bullying, person-related bullying and physical related bullying) had relatively high factor correlations. This indicates the modifications we made to the original NAQ-R (Table 1) to reflect the new context of measuring exposure to clinical workplace learning bullying, did not change the strength of the associations among factors we borrowed from the original NAQ-R.

The magnitude of the factor loadings indicated that all items were strong indicators of the factors they were related to. The factor loadings of the items of two new factors of sexual harassment and ethnic harassment were even higher than the original NAQ-R items which provided further convergent validity evidence. The correlations between the two novel factors and the original ones were moderate which revealed discriminant validity evidence. Both convergent and discriminant validity are important components of construct validity. The CFA analysis suggests that adding these two new factors support that medical and nursing students differentiate between the five factors that are also inter-related to each other for the overall construct of clinical workplace learning bullying and harassment.

Implications

Having a comprehensive five factor model that includes sexual and ethnic harassment among the bullying behaviours experienced provides a more comprehensive instrument that aligns with the definition of bullying offered earlier, and more accurately reflects the varied negative experiences of health professional students learning in the clinical workplace described in the literature (Dineen et al. 1997; Richardson et al. 1997; White 2000; Rautio et al. 2005; Witte et al. 2006; Wilkinson et al. 2006; Garbin et al. 2010; Premadasa et al. 2011; Rees and Monrouxe 2011; Bruce et al. 2015; Przedworski et al. 2015; Rees et al. 2015).

During the development of modifying the original NAQ-R and adding the two factors of sexual harassment and ethnic/racial harassment we were mindful of trying to keep the instrument short. The final questionnaire is a 31 item instrument, which is only nine items longer the original NAQ-R questionnaire. The length of a questionnaire is important to consider, as practically implementing long questionnaire in large organisations, or student groups can be difficult to administer, and run the risk of larger attrition of participant responses. When implemented, the 31 item questionnaire takes participants approximately 5 minutes to complete.

The Clinical Workplace Learning NAQ-R scale, as a five factor model, would be useful for health professional education institutions who would like to measure their students' exposure to bullying and harassment in clinical workplace learning environments. Developing a scale specifically for student learners in the workplace and extending the original NAQ-R to a five factor model could assist institutions in identifying particular problematic areas (if any) that their students maybe experiencing. For example, are students experiencing bullying behaviours that reflect the workplace learning aspects of their student learning role (e.g. being asked to do something above their level of competence); or are they experiencing more person-related bullying (e.g. Being ignored or excluded from the clinical team); or experiencing sexual harassment (e.g. inappropriate physical contact); or ethnic harassment (e.g. made derogatory comments about your racial or ethnic group). Identifying these specific areas of bullying and harassment would significantly benefit institutions in planning any interventions to address the negative behaviours experienced by health professional students.

Limitations

The questionnaire was delivered to only two health professional groups (medical students and nursing students) yet these two groups would represent the largest health professional groups in New Zealand. Although these two groups occupy a variety of settings that include and represent various primary and secondary and community clinical environments, we acknowledge they may not exactly mirror all health professional CWEs. Additionally, using these groups in our study could also be viewed as a strength, given many validation studies only include more homogeneous populations.

Moreover, even though we view the statements used in the Clinical Workplace Learning NAQ-R scale are generic enough to apply to many clinical workplace learning settings, further testing to look at how the instrument works with other health professional student groups to confirm this would be worthwhile.

We also understand that using this instrument design of measuring only behaviours and their frequency does not provide answers to other specific questions institutions may be wanting. For example, this method does not examine what Einarsen et al. (2009, p.40) describes as 'who did what to whom'. However, this issue could be addressed by adding a self-labelling method. For example, after administering the Clinical Workplace Learning NAQ-R scale, a definition of bullying and harassment is offered to participants and then asked if they view themselves as victims according to this definition and to describe what happened to them and by who (Einarsen et al. 2009).

Although the five-factor model yielded the best fit, having acceptable fit for three-factor and second-order models suggests that there is still room for improvement on the psychometric properties of the NAQ-R scale. We agree with Einarsen et al. (2009) that even though the original three dimensions of reported workplace bullying can be distinguished, "yet they do

not discriminate well between different types of bullying behaviours, suggesting co-occurrence of these different types of bullying (p.31)". Also, results from the second-order model may support that idea that NAQ-R constructs are correlated reflecting the presence of a more general construct at a higher conceptual level and can be considered together to create a composite score. In this research, we wanted to maintain the integrity and structure of the original NAQ-R as much as possible but future research may consider improving the constructs by refining the item wording or shortening the scale especially for the WLRB and PRB factors.

The Clinical Workplace Learning NAQ-R was developed based in a New Zealand cultural context which shares some similarities with Einarsen's et al. (2009) Anglo-American context, yet also maintains its own cultural context. The literature reports many behaviours that are similar between many cultural contexts in relation to bullying and harassment behaviours at medical and nursing schools. However, it would be pertinent to assume that there would be different beliefs, values and practices specific to certain cultures that may inform the concepts of bullying and harassment. This would influence the wording of the statements used in the instrument along with the meaning that is attributed to the statements as well. Therefore, further work needs to be conducted in making the instrument applicable in different cultural contexts. Finally, future validity research could explore the how sensitive the scores are to change over time, for example following an intervention.

Conclusion

In this paper we have shown the development of the Clinical Workplace Learning NAQ-R scale and analysis of its factor structure and provided supporting validity evidence for its use. The Clinical Workplace Learning NAQ-R scale is a quickly administered instrument in order to measure exposure to bullying and harassment experienced by health professional students in a clinical workplace learning environment. Its structure may assist health professional

leadership to obtain vital information into the negative experiences students may be facing, including what specific experiences may be occurring more frequently than others in regards to the bullying and harassment of their students. In turn, this information may assist in developing specific interventions to target the particular experiences faced by health professional students learning in a CWE.

Acknowledgments

The authors would like to thank all of the Associate Deans of Medical Education for Otago Medical School who assisted with the data collection for this questionnaire, as well as Faculty in The School of Nursing at Otago Polytechnic. We would also like to thank Dr Ella Iosua and Michel de Lange for their initial advice on the analysis of this paper. We would also like to thank the students for their time and effort in completing this questionnaire.

References

Arvey RD, Cavanaugh MA. 1995. Using surveys to assess the prevalence of sexual harassment: some methodological problems. J Soc Issues. 51:39-52.

Association of American Medical Colleges (AAMC). 2018. Medical School Graduation Questionnaire. https://www.aamc.org/download/490454/data/2018gqallschoolssummaryreport.pdf Accessed November 2018.

Baldwin DC Jr, Daugherty SR, Eckenfel EJ. 1991. Student perceptions of mistreatment and harassment during medical school: A survey of ten United States schools. West J Med. 155:140-145.

Brown, T. 2006. Confirmatory factor analysis for applied research. Guildford, New York.

Browne MW, Cudeck R. 1992. Alternative ways of assessing model fit. Sociol Methods Res. 21:230–258.

Bruce AN, Battista A, Plankey MW, Johnson LB, Marshall MB. 2015. Perceptions of gender-based discrimination during surgical training and practice. Med Educ Online. 20:25923.

Cheung GW, Rensvold RB. 2002. Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equ Modeling. 9:233–255.

Clarke C, Kane D, Rajacich D, Lafreniere K. 2012. Bullying in undergraduate clinical nursing education. J Nurs Educ. 51:269-276.

Crebbin W, Campbell G, Hillis DA, Watters DA. 2015. Prevalence of bullying, discrimination and sexual harassment in surgery in Australasia. ANZ J Surg. 85:905-909.

Dineen M, Cole M, Dally RJ, Dinned S, Gaughran F. 1997. Stress in medical students. Ir Med J. 90:184-185.

Einarsen S, Hoel H, Notelaers G. 2009. Measuring exposure to bullying and harassment at work: Validity, factor structure and psychometric properties of the Negative Acts Questionnaire-Revised. Work Stress. 23:24-44.

Einarsen S, Hoel H, Zapf D, Cooper C, ed. 2011. Bullying and Harassment in the Workplace: Developments in Theory, Research, and Practice (2nd ed.). Boca Raton: CRC Press Taylor & Francis Group.

Fan X, Sivo SA. 2005. Sensitivity of fit indexes to misspecified structural or measurement model components: Rationale of two-index strategy revisited. Struct Equ Modeling. 12:343–367.

Fan X, Sivo SA. 2007. Sensitivity of fit indices to model misspecification and model types. Multivariate Behav Res. 42:509–529.

Ferns T, Meerabeau L. 2008. Verbal abuse experienced by nursing students. J Adv Nurs. 61:436-444.

Fitzgerald LF, Gelfand MJ, Drasgow F. 1995. Measuring sexual harassment: Theoretical and psychometric advances. Basic Appl Soc Psych. 17: 425-445.

Flora DB, Curran PJ. 2004. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. Psychol Methods. 9:466-491.

Fnais N, Soobiah C, Chen M, Lillie E, Perrier L, Tashkhandi M, Straus S, Mamdani M, Al-Omran M, Tricco A. 2014. Harassment and discrimination in medical training: A systematic review and meta-analysis. Acad Med. 89:817-827.

Foster B, Mackie B, Barnett N. 2004. Bullying in the Health Sector: A Study of Bullying of Nursing Students. N Z J Employ Relat. 29:67-83.

Gamble Blakey A, Smith-Han K, Anderson L, Collins E, Berryman E, Wilkinson, T. 2019. Interventions addressing student bullying in the clinical workplace: a narrative review. BMC Med Educ. 19:220.

Gamble Blakey A, Smith-Han K, Anderson L, Collins E, Berryman E, Wilkinson, T. 2019. "They cared about us students:" learning from exemplar clinical teaching environments. BMC Med Educ. 19:119.

Gamble Blakey A, Smith-Han K, Anderson L, Collins E, Berryman E, Wilkinson, T. 2019. (2019c). It's 'probably the teacher!' A strategic framework for clinical staff engagement in clinical student bullying intervention. BMC Med Educ. 19:116.

Garbin C, Zina L, Garbin A, Moimaz S. 2010. Sexual harassment in Dentistry: prevalence in dental school. J Appl Oral Sci. 18:447-452.

Hair JF, Black WC, Babin BJ, Anderson RE. 2010. Multivariate data analysis (7th ed.). Englewood Cliffs: Prentice Hall.

Hu L, Bentler PM. 1998. Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. Psychol Methods. 3:424–453.

Keashly L. 1997. Emotional Abuse in the Workplace. J Aggress Maltreat Trauma. 1(1), 85-117.

Knapp K, Shane P, Sasaki-Hill D, Yoshizuka K, Chan P, Vo T. 2014. Bullying in the clinical training of pharmacy students. Am J Pharm Educ. 78:117.

Lubke GH, Muthén BO. 2004. Applying multigroup confirmatory factor models for continuous outcomes to Likert scale data complicates meaningful group comparisons. Struct Equ Modeling. 11:514-534.

MacCallum RC, Browne MW, Sugawara HM. 1996. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods. 1:130.

McDonald R. 1999. Test Theory: a Unified Treatment. Mahwah, NJ: Lawrence Erlbaum Associates.

Mikkelsen EG, Einarsen S. 2001. Bullying in Danish work-life: Prevalence and health correlates. Eur J Work Organ Psychol. 10: 393-413.

Muthén LK, Muthén BO. 2012. Mplus: Statistical analysis with latent variables: User's guide. Los Angeles, CA: Muthén & Muthén.

Noar, SM. 2003. The role of structural equation modeling in scale development. Struct Equ Modeling. 10:622–647.

Owoaje ET, Uchendu OC, Ige OK. 2012. Experiences of mistreatment among medical students in a university in south west Nigeria. Niger J Clin Pract. 15:214-219.

Premadasa IG, Wanigasooriya NC, Thalib L, Ellepola ANB. 2011. Harassment of newly admitted undergraduates by senior students in a Faculty of Dentistry in Sri Lanka. Med Teach. 33:556-563.

Przedworski JM, Dovidio JF, Hardeman RR, Phelan SM, Burke SE, Ruben MA, Perry S, Burgess DJ, Nelson DB, Yeazel MW, Knudsen JM, van Ryn M. 2015. A Comparison of the Mental Health and Well-Being of Sexual Minority and Heterosexual First-Year Medical Students: A Report from the Medical Student CHANGE Study. Acad Med. 90:652-659.

Rautio A, Sunnari V, Nuutinen M, Laitala M. 2005. Mistreatment of university students most common during medical studies. BMC Med Educ. 5:36.

Rees CE, Monrouxe LV, Ternan E, Endacott R. 2015. Workplace abuse narratives from dentistry, nursing, pharmacy and physiotherapy students: a multi-school qualitative study. Eur J Dent Educ. 19:95-106.

Rees CE, Monrouxe LV. 2012. "A morning since eight of just pure grill": a multischool qualitative study of student abuse. Acad Med. 86:1374-1382.

Richardson DA, Becker M, Frank RR, Sokol R.J. 1997. Assessing Medical Students' Perceptions of Mistreatment in Their Second and Third Years. Acad Med. 72:728-730.

Rowland ML, Naidoo S, AbdulKadir R, Moraru R, Huang B, Pau A. 2010. Perceptions of intimidation and bullying in dental schools: A multi-national study. Int Dent J. 60:106-112.

Schneider KT, Hitlan, RT, Phanikiran R. 2000. An examination of the nature and correlates of ethnic harassment experiences in multiple contexts. J Appl Psychol 85:1:3-12

Sheehan KH, Sheehan DV, White K, Leibowitz A, Baldwin DC. 1990. A pilot study of medical student abuse: Student perceptions of mistreatment and misconduct in medical school. JAMA. 263:533-537.

Sijtsma K. 2009. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. Psychometrika, 74:107-120.

Strauss ME, Smith GT. 2009. Construct validity: Advances in theory and methodology. Annu Rev Clin Psychol. 5:1–25.

Stubbs B, Soundy A. 2013. Physiotherapy students' experiences of bullying on clinical internships: an exploratory study. Physiotherapy. 99:178-180.

Teo T, Fan X. 2013. Coefficient alpha and beyond: Issues and alternatives for educational research. Asia-Pac Educ Res. 22:209-213.

Till, F. 1980. Sexual harassment: A report on the sexual harassment of students. Washington. DC: National Advisory Council on Women's Educational Programs.

Uhari M, Kokkonen J, Nuutinen M, Vainionpas L, Rantala H, Lautala P, Vayrynen M. 1994. Medical student abuse: An international phenomenon. JAMA. 271:1049-1051.

White GE. 2000. Sexual harassment during medical training: The perceptions of medical students at a university medical school in Australia. Med Educ. 34:980-986.

Whiteside D, Stubbs B, Soundy A. 2014. Physiotherapy students' experiences of bullying on clinical internships: a qualitative study. Physiotherapy. 100:41-46.

Wilkinson T, Gill DJ, Fitzjohn J, Palmer CL, Mulder RT. 2006. The impact on students of adverse experiences during medical school. Med Teach. 28:129-135.

Witte FM, Stratton TD, Nora LM. 2006. Stories from the field: Students' descriptions of gender discrimination and sexual harassment during medical school. Acad Med. 81:648-654.

Wolf T, Scurria PL, Webster MG. 1998. A four-year study of anxiety, depression, loneliness, social support, and perceived mistreatment in medical students. J Health Psychol. 3:125-136.

Woolley DC, Paolo AM, Bonaminio GA, Moser SE. 2006. Student treatment on clerkships based on their specialty interests. Teach Learn Med. 18(3):237-243.

Yu CY. Evaluating cut-off criteria of model fit indices for latent variable models with binary and continuous outcomes [Doctoral dissertation]. 2002. University of California Los Angeles.

Tables

Table 1

Re-worded Statements from the NAQ-R Used in the Clinical Workplace Learning NAQ-R scale

Original statement in NAQ-R	Wording change/omission*	Rationale for change
Work-related bullying factor		
Someone withholding information which affects your performance	Someone withholding information which affects your *learning*	To fit our context of students learning in a workplace environment
Being ordered to do work below your level of competence	Being ordered to do *tasks above* your level of competence	This reflects the importance of students learning tasks to be within their level of competence in order to keep themselves safe, and their patients safe during the learning process.
Having your opinions ignored	Having your opinions *and views* ignored	Aides understanding of the statement to broaden it out to more than just having your 'opiniated comments' ignored, but whatever your contribution happens to be at the time ignored.
Being given tasks with unreasonable deadlines	Being given tasks with unreasonable *or impossible targets or* deadlines	Aides understanding of the statement.
Person-related bullying factor		
Being humiliated or ridiculed in connection with your work	Being humiliated or ridiculed in connection with your *learning*	To fit our context of students learning in a workplace environment
Having key areas of responsibility removed or replaced with more trivial or unpleasant tasks	Having key areas *of your student role* removed or replaced with more trivial or unpleasant tasks	Highlights significant change in what the student should be doing at the level they are currently at, which changes as they progress.
Being ignored or excluded	Being ignored or excluded *from the clinical team*	To fit our context of students learning in a workplace environment

Having insulting or offensive remarks made about your person, attitudes or your private life	Having insulting or offensive remarks made about your person *(i.e. habits and background)*, attitudes or your private life	Aides understanding of the statement.
Hints or signals from others that you should quit your job	Hints or signals from others that you should *quit studying your profession*	To fit our context of students learning in a workplace environment
Persistent criticism of your errors or mistakes	Persistent criticism of your *work and effort*	To make sure students didn't get this confused with patient safety literature (where errors etc. have specific definitions).

*Change/addition of wording shown in italics

Table 2

Descriptive Statistics of Clinical Workplace Learning NAQ-R Scale Factors

	Number of items	*M*	*SD*	*Min*	*Max*
Workplace learning-related bullying (WLRB)	8	1.44	0.44	1.00	3.75
Person-related bullying (PRB)	11	1.43	0.43	1.00	4.09
Physically intimidating bullying (PIB)	3	1.18	0.35	1.00	3.67
Sexual harassment (SH)	5	1.15	0.28	1.00	3.20
Ethnic harassment (EH)	4	1.15	0.38	1.00	4.00

Table 3

Confirmatory Factor Analysis of Alternative Models

Model	χ^2	*df*	RMSEA	CFI	TLI	WRMR
Alternative one-factor (unidimensional)	1995.244*	434	0.082	0.845	0.834	2.136
Alternative three-factor	906.221*	431	0.045	0.953	0.949	1.266
Hypothesized five-factor	878.706*	424	0.045	0.955	0.951	1.232
Alternative second-order	906.289*	429	0.045	0.953	0.949	1.280

Note. RMSEA= Root Mean Square Error of Approximation; CFI= Comparative Fit Index; TLI= Tucker-Lewis Index; WRMR= Weighted Root Mean Square Residual.

*$p<.01$

Table 4

Clinical Workplace Learning NAQ-R Scale Factor Correlations and Reliabilities

	WLRB	PRB	PIB	SH	EH
Workplace Learning-Related Bullying (WLRB)	-				
Person-Related Bullying (PRB)	0.95	-			
Physically Intimidating Bullying (PIB)	0.80	0.89	-		
Sexual Harassment (SH)	0.59	0.54	0.52	-	
Ethnic Harassment (EH)	0.48	0.58	0.49	0.48	-
Reliability (ω)	0.89	0.93	0.79	0.90	0.94

Table 5

Standardized factor loadings of the Clinical Workplace Learning NAQ-R

Item	Standardized estimate
WLRB 1	0.65
WLRB 2	0.63
WLRB 3	0.72
WLRB 4	0.75
WLRB 5	0.69
WLRB 6	0.71
WLRB 7	0.80
WLRB 8	0.70
PRB 1	0.75
PRB 2	0.66
PRB 3	0.71
PRB 4	0.69
PRB 5	0.73
PRB 6	0.73
PRB 7	0.83
PRB 8	0.65
PRB 9	0.82
PRB 10	0.83
PRB 11	0.80
PIB 1	0.79
PIB 2	0.76
PIB 3	0.76
SH 1	0.91
SH 2	0.92
SH 3	0.65
SH 4	0.66
SH 5	0.93
EH 1	0.93
EH 2	0.96
EH 3	0.91
EH 4	0.91

APPENDICIES

Table A1 - Clinical Workplace Learning NAQ-R *Scale Items*

Workplace Learning-Related Bullying (WLRB)	
WLRB 1	Someone withholding information which affects your learning
WLRB 2	Being ordered to do tasks above your level of competence
WLRB 3	Having your opinions and views ignored
WLRB 4	Being given tasks with unreasonable or impossible targets or deadlines
WLRB 5	Excessive monitoring of your work
WLRB 6	Being exposed to an unmanageable workload
WLRB 7	Being assigned work for punishment rather than for educational value
WLRB 8	Having learning opportunities blocked or withheld by others
Person-Related Bullying (PRB)	
PRB 1	Being humiliated or ridiculed in connection with your learning
PRB 2	Having key areas of your student role removed or replaced with more trivial or unpleasant tasks
PRB 3	Spreading of gossip and rumours about you
PRB 4	Being ignored or excluded from the clinical team
PRB 5	Having insulting or offensive remarks made about your person (i.e. habits and background), your attitudes or your private life
PRB 6	Hints or signals from others that you should quit studying your profession
PRB 7	Repeated reminders of your errors or mistakes
PRB 8	Being ignored or facing a hostile reaction when you approach
PRB 9	Persistent criticism of your work and effort
PRB 10	Having allegations made against you
PRB 11	Being the subject of excessive teasing and sarcasm
Physically Intimidating Bullying (PIB)	
PIB 1	Being shouted at or being the target of spontaneous anger
PIB 2	Intimidating behaviour such as finger-pointing, invasion of personal space, shoving, blocking/barring the way
PIB 3	Threats of violence of physical abuse or actual abuse
Sexual Harassment (SH)	
SH 1	Sexually explicit or offensive jokes
SH 2	Sexual slurs

SH 3	Questions or insinuations about your sexual or private life
SH 4	Inappropriate physical contact
SH 5	Unwanted sexual advances
Ethnic Harassment (EH)	
EH 1	Told jokes about your racial or ethnic group
EH 2	Made derogatory comments about your racial or ethnic group
EH 3	Used racial or ethnic slurs to describe you
EH 4	Made racist comments (for example, says people of your ethnicity aren't very smart or can't do the job)

Questionnaire A2 - The Clinical Workplace Learning NAQ-R

Thank you for taking the time to participate in this survey. We are collecting data on your experience with interacting with particular negative behaviours during your clinical learning years.

The first set of questions collects demographic data. The second set of questions look at how often you may have experienced particular negative behaviours during this year. Some of these questions are sensitive in nature, but they are important to ask for XXX to get sense of your experience with these behaviours during your clinical learning years.

We are <u>not</u> collecting any identifying data (i.e. your name or student ID number).

Demographic information:

1. Age:

2. Are you (please tick): Male?___ Female?___ Other? (please describe)_____

3. What is your sexual orientation (please tick)?:

___Heterosexual

___Bisexual

___Homosexual

___Other (please describe)_____

4. Ethnicity (question taken from the New Zealand 2013 Census):

Which ethnic group do you belong to?

Tick the box or write in the spaces which apply to you

___ New Zealand European

___ Maori

___ Samoan

___ Cook Island Maori

___ Tongan

___ Niuean

___ Chinese

___ Indian

___ other such as *DUTCH, JAPANESE, TOKELAUAN*. Please state:

Please read the following directions for completing the rest of the survey: Since the start of the year, how often have you been subjected to the following negative acts? Please check the boxes that best corresponds with your experience from the beginning of this year (there are no right or wrong answers to this questionnaire):

Please tick the appropriate box.

Negative Act	*Never*	*Now and then*	*Monthly*	*Weekly*	*Daily*
1. Someone withholding information which affects your learning					
2. Being humiliated or ridiculed in connection with your learning					
3. Being ordered to do tasks above your level of competence					
4. Having key areas of your student role removed or replaced with more trivial or unpleasant tasks					
5. Spreading of gossip and rumours about you					
6. Being ignored or excluded from the clinical team					
7. Told jokes about your racial or ethnic group					
8. Having insulting or offensive remarks made about your person (i.e. habits and background), your attitudes or your private life					
9. Being shouted at or being the target of spontaneous anger					
10. Made derogatory comments about your racial or ethnic group					
11. Intimidating behaviour such as finger-pointing, invasion of personal space, shoving, blocking/barring the way					
12. Repeated reminders of your errors or mistakes					
13. Sexual slurs					
14. Made racist comments (for example, says people of your ethnicity aren't very smart or can't do the job)					
15. Inappropriate physical contact					
16. Being ignored or facing a hostile reaction when you approach					
17. Persistent criticism of your work and effort					

Negative Act	*Never*	*Now and then*	*Monthly*	*Weekly*	*Daily*
18. Used racial or ethnic slurs to describe you					
19. Having your opinions and views ignored					
20. Being exposed to an unmanageable workload					
21. Sexually explicit or offensive jokes					
22. Being given tasks with unreasonable or impossible targets or deadlines					
23. Having allegations made against you					
24. Excessive monitoring of your work					
25. Being the subject of excessive teasing and sarcasm					
26. Questions or insinuations about your sexual or private life					
27. Threats of violence of physical abuse or actual abuse					
28. Being assigned work for punishment rather than for educational value					
29. Unwanted sexual advances					
30. Hints or signals from others that you should quit studying your profession.					
31. Having learning opportunities blocked or withheld by others					