Journal of Intellectual & Developmental Disability

ISSN: 1366-8250 (Print) 1469-9532 (Online) Journal homepage: http://www.tandfonline.com/loi/cjid20

Measuring practice leadership in supported accommodation services for people with intellectual disability: Comparing staff-rated and observational measures

Emma Bould, Julie Beadle-Brown, Christine Bigby & Teresa Iacono

To cite this article: Emma Bould, Julie Beadle-Brown, Christine Bigby & Teresa Iacono (2016): Measuring practice leadership in supported accommodation services for people with intellectual disability: Comparing staff-rated and observational measures, Journal of Intellectual & Developmental Disability, DOI: 10.3109/13668250.2016.1259466

To link to this article: http://dx.doi.org/10.3109/13668250.2016.1259466

Published online: 23 Nov 2016.

JOURNAL OF INTELLECTUAL & DEVELOPMENTAL DISABILITY, 2016
http://dx.doi.org/10.3109/13668250.2016.1259466

Taylor & Francis
Taylor & Francis Group

BRIEF REPORT

OPEN ACCESS

Measuring practice leadership in supported accommodation services for people with intellectual disability: Comparing staff-rated and observational measures

Emma Bould[a], Julie Beadle-Brown[a,b], Christine Bigby[a] and Teresa Iacono[a]

[a]Living with Disability Research Centre, La Trobe University, Bundoora, Australia; [b]Tizard Centre, University of Kent, Canterbury, UK

ABSTRACT

Background Studies incorporating staff-rated or observational measures of practice leadership have shown that where practice leadership is stronger, active support is better implemented. The study aim was to compare measures of practice leadership used in previous research to determine the extent of their correspondence.

Method A subset of data from a longitudinal study regarding 29 front-line managers working across 36 supported accommodation services in Australia was used. An observed measure of practice leadership, based on an interview and observation of a front-line manager, was compared with ratings of practice leadership completed by staff. The quality of active support was rated after a 2-hour structured observation.

Results Correlations between staff-rated and observed measures were nonsignificant. Only the observed measure was correlated with the quality of active support.

Conclusions This study provides evidence to support using an observational measure of practice leadership rather than reliance on staff ratings.

KEYWORDS

intellectual disability; practice leadership; observed; staff-rated; quality of support; supported accommodation

Introduction

The quality of staff support and outcomes in supported accommodation services for people with intellectual disability are very variable (Mansell & Beadle-Brown, 2012; Mansell, Beadle-Brown, & Bigby, 2013). A large body of evidence has shown that the quality of active support is associated with higher levels of service user engagement and better quality of life outcomes, but that embedding this staff practice in organisations has been difficult (Mansell & Beadle-Brown, 2012). Recent evidence suggests a positive association between the strength of practice leadership shown by the front-line managers and the quality of active support (Beadle-Brown, Bigby, & Bould, 2015). Practice leadership was defined by Mansell, Beadle-Brown, Ashman, and Ockenden (2004) as practices implemented by front-line manager of a staff team to develop and maintain good staff support for service users. Mansell et al. (2004) operationalised these practices according to five domains that address (a) focusing leadership on staff support for service users' quality of life, (b) allocation and organisation of staff support to meet service users' needs and wants, (c) coaching staff in good practice through feedback and modelling of good practice, (d) regular reviewing of staff practices on an individual basis, and (e) reviewing the extent to which staff teams are enabling service users to be actively engaged in meaningful activities and relationships during regular team meetings. Similarly, Deveau and McGill's (2016) study emphasised the importance of practice leaders directly observing and monitoring staff practice and inadequacies of relying on paperwork-based quality assurance mechanisms. An issue faced in researching practice leadership is how best to measure its strength.

Beadle-Brown et al. (2014) developed a 12-item measure of staff perceptions of practice leadership in order to determine the extent to which it was provided in supported accommodation services. They found that practice leadership accounted for improved quality of active support over time, and it was more effective in the context of good general management. A concern raised by the authors was that staff perceptions might be limited by their experience and expectations of managers. Hence they argued for an observed measure of practice leadership to supplement that based on staff perceptions alone. Beadle-Brown et al. (2015) argued further limitations of a staff-rated report measure were bias caused by social desirability and different interpretations of questions. These limitations prompted development of an

CONTACT Emma Bould e.bould@latrobe.edu.au

observational measure of practice leadership, completed by a trained researcher, which combines an interview with the front-line manager on the perception of their role, a brief interview with any staff on shift at the time where possible, a review of paperwork associated with practice leadership, and an observation of front-line managers working in the service they manage. Using this observational measure, they provided further evidence to support Beadle-Brown et al.'s (2014) conclusion that practice leadership by front-line managers was related to the level of active support, and, in turn, outcomes for service users. Although both methods of measuring practice leadership have limitations, they have not been evaluated for their comparative accuracy in assessing the strength of practice leadership.

Previous research by Higgins (2010), in which no correlation was found between paper-based records completed by support staff, and observed service user engagement and quality of support suggests that accurate measurement might rely on direct observations. Certainly, Mansell (2011) argued for the use of observational data in assessing and improving the quality of services. The aim of this study was to compare two methods of measuring practice leadership: staff-rated and observational. We addressed two research questions:

(1) Is there correspondence between staff-rated and observational measures of practice leadership used in previous research?
(2) Do either practice leadership measures correspond with the quality of support provided to service users?

Method

Design

The data reported in this paper were collected from September 2012 to June 2013 and are drawn from a longitudinal study following a cohort of supported accommodation services provided by nine organisations across three states in Australia. The study was approved by the La Trobe University Human Research Ethics Committee.

The study was primarily a naturalistic observation study, although some level of intervention was provided for each organisation: First, each year, organisations received a report based on data collected for the study that described how their organisation was performing with regard to active support, with recommendations on what they needed to focus on to maintain and improve practice. Second, each year, the research team provided training on various topics related to active support; for example, training in person-centred active support and training in practice leadership.

Participants and settings

The nine organisations, of varying sizes, are funded by their respective state governments to deliver supported accommodation services. One organisation is a regional branch of a government department and the others are not-for-profit agencies. Aggregated data from one year of data collection is reported in respect of 29 front-line managers who worked across 36 services (seven worked across two services). Further information on the characteristics of the people and services involved in the study are reported below. Each of the nine organisations had expressed the intention to implement active support across all intellectual disability services, but was at different stages in the implementation.

Measures

Staff-rated practice leadership index

Staff perceptions of the practice leadership they experienced from the front-line manager were determined from 12 items within an adapted version of the Staff Experiences and Satisfaction Questionnaire (SESQ; Beadle-Brown, Gifford, & Mansell, 2003). The SESQ also includes questions on staff characteristics, training, and experience, as well as knowledge of active support, person-centred planning, involvement of senior management, and other motivational structures (i.e., job satisfaction, feelings about the line manager).

A staff-rated practice leadership index (staff-rated index) score was calculated using the methods described by Beadle-Brown et al. (2014). In brief, the score is calculated from a series of questions shown in Table 1, with either 5- or 6-point Likert scales (of frequency, usefulness, or satisfaction) or binary yes/no responses. For the purpose of the practice leadership index, all Likert scales were recoded into binary yes/no responses (see Table 1). This staff-rated index is focused on three domains of practice leadership: coaching, supervision, and team meetings. The maximum total score for the staff-rated index is 12. A percentage score was calculated overall and for each of the three domains. It was not possible to assess test–retest agreement for this measure during this study, but Cronbach's alpha at the staff level indicated that internal consistency was acceptable at 0.682, and comparable to that obtained by Beadle-Brown et al. (2014; 0.641). Beadle-Brown et al. (2014) did not report internal consistency for each of the three domains because the measure had been designed to provide an overall index of reported practice

Table 1. Summary of the components of the staff-rated practice leadership indexes.

Components	Staff-rated index original	Staff-rated index expanded
Coaching		
1. 1 if manager observed staff practices at least monthly	√	√
2. 1 if the main focus of observation was on how they support service users	√	√
3. 1 if manager usually modelled good support	√	√
4. 1 if manager usually gave feedback	√	√
5. 1 if the manager gave feedback that was at least usually useful	√	√
6. 1 if manager was approachable if they had a problem	X	√
7. 1 if manager's advice on a problem was at least helpful	X	√
Supervision		
1. 1 if supervision included discussion of user activities/active support	√	√
2. 1 if the main topic of supervision was not on admin/paperwork	√	√
3. 1 if supervision occurred at least monthly	√	√
4. 1 if supervision was rated as at least usually useful	√	√
5. 1 if the manager gives constructive feedback in supervision	X	√
Team meetings		
1. 1 if supporting service users is discussed	X	√
2. 1 if the main focus of team meetings was user participation/quality of life/active support	√	√
3. 1 if team meetings held at least monthly	√	√
4. 1 if team meetings at least usually useful	√	√
Allocating staff		
1. 1 if manager usually explains about tasks for day-to-day work	X	√
2. 1 if work time is rarely wasted through manager's lack of planning	X	√
3. 1 if manager explains what is expected from staff	X	√

Note. √ = included in the tool; X = not included in the tool.

leadership, with a limited number of items overall, making reliability in any one domain difficult to establish.

Expanded staff-rated practice leadership index

The staff-rated index includes only three of the five domains of the observed measures of practice leadership. We drew on other items in the SESQ to develop an expanded staff-rated index to address the domain of allocating staff. Only one question of the SESQ addresses manager focus; as a result, it was insufficient for inclusion as a domain in the expanded staff-rated practice leadership index. In addition, the manager focus item was very much about staff perceptions of their managers' focus, making it a more subjective item than the others for this measure. As with the original staff-rated index, the expanded index was calculated from a series of questions with either 5- or 6-point Likert scales (of frequency, usefulness, or satisfaction) or binary yes/no responses, with all Likert scales recoded into binary yes/no responses (see Table 1).

The number of items related to each domain were as follows: coaching, seven; supervision, five; team meetings, four; and allocating staff, three. The maximum total score for the expanded staff-rated index was 19. A percentage score was calculated overall and for each of the four domains. The Cronbach's alpha at the staff level for the four domains was 0.725 (at service level it was 0.826), which is higher than the Cronbach's alpha we obtained with the original practice leadership index at staff level of 0.682, and the 0.641 obtained by Beadle-Brown et al. (2014). As was the case for the original staff-rated index, the only individual domain with reasonable reliability was coaching (Cronbach's alpha 0.744 at staff level and 0.788 at service level).

Observed measure of practice leadership

The observed measure of practice leadership combines data obtained from interviews with front-line managers, if feasible a brief interview with any staff on shift at the time, a review of paperwork associated with practice leadership (e.g., minutes of team and supervision meetings), and a 30–60 min observation of front-line managers working in the service they manage. Beadle-Brown et al. (2015) demonstrated that the measure had good internal consistency and acceptable interrater reliability when used for 46 front-line managers. For the current study, we drew on data from 29 of the same 46 front-line managers used by Beadle-Brown et al. (2015). Interviews with front-line managers conducted by researchers lasted approximately one hour and were digitally recorded. Detailed field notes were written as soon as possible after each visit, which, together with the interview recording, were used to score the five domains on the measure. The maximum possible score for each domain was 5. Internal consistency of the observed measure of practice leadership at the service level (i.e., aggregating data from houses within a service) was high (Cronbach's alpha = 0.869).

Service user needs and characteristics

A measure of service user needs and characteristics was obtained by questionnaires completed by a keyworker or another member of staff who knew the individual well. These questionnaires included the short form of

the Adaptive Behavior Scale (SABS) Part I (Hatton et al., 2001) and the Aberrant Behavior Checklist (ABC; Aman, Burrow, & Wolford, 1995). Additional questions addressed characteristics such as gender, date of birth, and the presence of other disabilities. The reliability and validity of the ABS (from which the SABS was drawn) and ABC have been evaluated and reported as acceptable by their authors. A full-scale score for Part I of the Adaptive Behaviour Scale can be estimated from the SABS using the formula provided by Hatton et al. (2001), and has been reported here.

Quality of active support

The Active Support Measure (ASM; Mansell & Elliott, 1996; Mansell, Elliott, & Beadle-Brown, 2005) is a measure of the quality of active support provided by disability support staff and was completed after a 2-hour observation of service users to measure engagement in meaningful activities and relationships using the Engagement in Meaningful Activity and Relationships observational measure (EMAC-R; Mansell & Beadle-Brown, 2005). The ASM includes 15 items focusing on opportunities for service user involvement in activities and the skills with which staff provided and supported those opportunities. Each item is rated on a scale of 0 (poor, inconsistent support/performance) to 3 (good, consistent support/performance). The maximum possible total score was 45. Further details of both measures, including administration and interrater reliability, have been reported previously (Beadle-Brown, Hutchinson, & Whelton, 2012; Mansell et al., 2013). Interrater reliability data for the observations reported in this study were not available.

Procedure

Once consent was gained, service user questionnaires were sent to each service, with a request that they be completed by a staff member who knew each individual well, and returned to the research team using the prepaid envelopes provided. The staff questionnaires were mailed to front-line and more senior managers associated with each service, who were asked to give a copy to each consenting member of staff, along with a prepaid envelope for posting back directly to the research team. A researcher visited each service to conduct the observations using the EMAC-R, at the end of which the ASM was completed for each service user. The observation to complete the EMAC-R and the observed measure of practice leadership were conducted on different days and by different researchers, who were blind to the other results. As a result, each service had two visits (the exception being when a front-line manager worked across more than one service, in which situation only one interview/observation was conducted in one of those services), usually within one to two months of each other; exceptions were for six services, for which circumstances resulted in a longer interval of two and a half to three months.

Analysis

The criterion for inclusion of data in the analysis was a minimum of two staff surveys returned for a service (or, in the case of the 10 front-line managers, for the two services they worked across). Of the 46 front-line managers interviewed, sufficient staff surveys were returned for only 29 (response rate range: 40–100%). For the 17 front-line managers for whom there were insufficient staff surveys (response rate range: 0–25%) to meet the inclusion criterion, data from all consenting staff in the service(s) in which the manager worked were removed from the analysis (service user, $n = 55$; staff, $n = 15$; observed practice leadership data, $n = 17$). There were no significant differences between surveys included and those excluded in terms of the quality of active support (Mann–Whitney $z = -0.339$, $p = .735$), the level of ability of the service users supported (Mann–Whitney $z = -0.277$, $p = .782$), or the observed practice leadership scores overall (Mann–Whitney $z = -0.343$, $p = .732$), and for each of the five domains: manager focus (Mann–Whitney $z = -0.285$, $p = .776$), allocating staff (Mann–Whitney $z = -0.342$, $p = .732$), coaching (Mann–Whitney $z = -0.498$, $p = .618$), supervision (Mann–Whitney $z = -0.677$, $p = .499$), and team meetings (Mann–Whitney $z = -0.444$, $p = .657$). There were also no significant differences between surveys included and those excluded in terms of overall staff satisfaction (Mann–Whitney $z = -1.516$, $p = .129$) or the proportion of staff that received training in active support (Mann–Whitney $z = -1.828$, $p = .068$).

Analysis was completed at the levels of service user, staff, and service to determine the correspondence between the staff-rated and observed measures of practice leadership. For the service-level analysis, service user and staff data were aggregated across the service or services in which a front-line manager worked.

In light of the relatively low Cronbach's alpha scores across some domains of the expanded staff-rated index, a predominantly descriptive analysis was completed. Correlation analysis was used to explore the relationships between (a) the staff-rated and observed measures of practice leadership, and (b) each of these practice leadership measures and the ASM. In addition, differences between the staff-rated and observed measures of practice leadership were examined using Fisher's exact tests. Significance of correlations and Fisher's exact tests are reported at $p < .01$,

and Cohen's (1988) guidelines were used to report effect sizes where appropriate. Univariate linear regression was used to determine the extent to which the SABS score and measures of practice leadership accounted for variation in the ASM score.

Results

Description of participants and settings

The 29 practice leaders whose data were included worked across 36 services (seven worked across two services), and data were collected from 134 service users and 135 support workers. Just over half (52%) of the service users included in the analysis at the service-user level were male, and the average age was 42 years (range: 16–75). The sample comprised people with a range of abilities, but overall they were relatively able (ABS 135, range: 22–260), 35% were nonverbal, 31% had a physical disability, and 17% (range: 0–78%) were rated by staff as showing some form of challenging behaviour on the ABC.

For the ASM measure, the total score out of the maximum possible (45) was converted to a percentage. The average percentage score across observational groups was 48 (range: 2–97%).

At the staff level, 30% of the sample was male and 66% aged 46 years or older. Sixty-eight percent of the staff had been working in services for people with intellectual disability for 5 or more years, and 50% had been working in their current service for 5 or more years (only 15% for more than 10 years). In terms of qualifications, 53% had a Certificate 3, 4, or 5 in disability, community, social services, and health management (obtained at registered training organisations, such as Technical and Further Education institutions), 75% had been trained in active support, and 50% of staff said the training had been in the last 12 months.

Level of practice leadership

At the staff level, the mean percentage score on the staff-rated index was 50% (range: 8–92%) and the mean percentage score on the expanded index was 55% (range: 4–82%). There was a high and significant correlation between the two measures ($\rho = .867$, $n = 135$, $p < .001$). At service level, using the two aggregated staff-rated indexes, the correlations were also high and significant ($\rho = .819$, $n = 29$, $p < .001$). According to Cohen's (1988) guidelines for interpreting coefficients (see Dunst & Hamby, 2012; Lipsey & Wilson, 2001), these are both large effect sizes.

To allow for comparison between the two versions of the staff-rated index with the observed measure of practice leadership, all three measures at the aggregated service level were first converted to percentage scores, and then grouped according to low (0–50%) and high (51–100%). Table 2 provides the scores at service level for both staff-rated indexes and the observed measure of practice leadership, and percentage of practice leaders with a high percentage score on each measure (i.e., those over 50%). These percentage scores have been used for the remainder of the analysis.

Comparing staff-rated and observed measures

In light of the high correlation between the staff-rated indexes, and because the original version addressed three rather than four of the domains, we excluded the original index from further analysis. Spearman correlation between the overall expanded staff-rated index and the overall observed practice leadership measure was low and not significant ($\rho = -.034$, $n = 29$, $p = .861$), with negligible effect size (Cohen, 1988). Similarly, the correlations between the expanded index and observed practice leadership measure for the domains of coaching ($\rho = .101$, $n = 29$, $p = .603$), supervision ($\rho = .117$, $n = 29$, $p = .546$), team meetings ($\rho = .061$, $n = 29$, $p = .754$), and allocating staff ($\rho = .31$, $n = 29$, $p = .875$) were low and not significant. The effect sizes (Cohen, 1988) ranged from medium (allocating staff), to small (coaching and supervision), to negligible (team meetings).

Fisher's exact tests were used to compare the number of front-line managers with low (0–50%) versus high (51–100%) scores on the expanded staff-rated index and on the observed practice leadership measure. The expanded staff-rated index resulted in more services being rated as having high practice leadership on the allocating staff and team meeting domains and on the overall score, whereas the observed practice leadership measure resulted in more services rated as having high practice leadership on the coaching and supervision domains. However, there were no significant associations between the expanded staff-rated index and the observed measure, overall or across any of the four domains. Wilson's (2001) effect size calculator was used to convert the p value from the Fisher's exact test to Cohen's d, and the effect size for these associations were small overall (0.222) and for the domain of team meetings (0.331); and negligible for the domains of allocating staff (0.147), coaching (0.144), and supervision (0.135; Cohen, 1988).

Table 2. Summary data for overall and component scores on the staff-rated practice leadership indexes and the observed measure of practice leadership.

	Measures					
	Staff-rated index					
	Original		Expanded		Observed	
Components	*M* (Range)	High scoring	*M* (Range)	High scoring	*M* (Range)	High scoring
Overall across components	47% (14–73%)	67%	61% (28–80%)	90%	49% (20–92%)	41%
Coaching						
Overall	37% (0–100%)	28%	53% (15–86%)	46%	46% (20–100%)	62%
1. Manager watches you more than monthly	43% (0–100%)	41%	43% (0–100%)	41%	n/a	n/a
2. Focus of observations on involving service users in activities	42% (0–100%)	31%	42% (0–100%)	31%	n/a	n/a
3. Manager usually shows you how to work well with people	31% (0–100%)	21%	31% (0–100%)	21%	n/a	n/a
4. Manager usually tells you how well you are doing	42% (0–100%)	31%	42% (0–100%)	31%	n/a	n/a
5. Feedback is helpful	60% (0–100%)	59%	60% (0–100%)	59%	n/a	n/a
6. When you have a problem with someone you can support you feel able to approach your manager	n/a	n/a	76% (0–100%)	83%	n/a	n/a
7. Advice on problem-solving is helpful	n/a	n/a	77% (0–100%)	79%	n/a	n/a
Supervision						
Overall	62% (25–100%)	48%	59% (13–90%)	59%	44% (20–80%)	62%
1. Supporting service users/active support is discussed in supervision meetings	90% (0–100%)	90%	90% (0–100%)	90%	n/a	n/a
2. Main topic of discussion is not paperwork	86% (0–100%)	90%	86% (0–100%)	90%	n/a	n/a
3. Supervision is at least monthly	27% (0–100%)	21%	27% (0–100%)	21%	n/a	n/a
4. Supervision is usually meaningful and useful	52% (0–100%)	48%	52% (0–100%)	48%	n/a	n/a
5. Manager usually gives you constructive feedback	n/a	n/a	48% (0–100%)	38%	n/a	n/a
Team meetings						
Overall	63% (0–100%)	66%	74% (44–92%)	97%	55% (20–100%)	38%
1. Supporting service users is discussed	n/a	n/a	93% (0–100%)	97%	n/a	n/a
2. Main topic is supporting service users' quality of life	46% (0–100%)	39%	46% (0–100%)	39%	n/a	n/a
3. Team meetings held at least monthly	85% (0–100%)	86%	85% (0–100%)	86%	n/a	n/a
4. Usually meaningful and useful	64% (0–100%)	62%	64% (0–100%)	62%	n/a	n/a
Allocating staff						
Overall	n/a	n/a	58% (39–74%)	69%	50% (20–100%)	59%
1. Manager usually lets you know about tasks for your day-to-day work	n/a	n/a	70% (0–100%)	72%	n/a	n/a
2. Work time is rarely wasted through your manager's lack of planning	n/a	n/a	8% (0–50%)	0%	n/a	n/a
3. You know what your manager expects of you	n/a	n/a	94% (33–100%)	97%	n/a	n/a
Manager focus						
Overall	n/a	n/a	n/a	n/a	52% (20–80%)	41%

Note. n/a = not applicable to the tool.

Predicting active support

Further Spearman correlations were conducted to determine which of the practice leadership measures was a better predictor of active support. The only significant correlation, when controlling for adaptive behaviour (SABS), was between ASM and the overall observed practice leadership measure ($\rho = .547$, $p < .01$, large effect size according to Cohen, 1988).

As expected from the correlations, an exploratory univariate regression analysis showed that the only factors that contributed significantly to the variance in the ASM were the observed practice leadership measure and SABS (48%). The SABS explained 25% of the variance on the ASM score, and the observed practice leadership measure explained the remaining 23%, $F(2, 28) = 11.79$, $p < .001$; R square = 0.690; adjusted R square = 0.476.

Discussion

The findings from this study provide support for the importance of practice leadership in supporting staff to provide high-quality support in the form of person-centred active support, which in turn improves outcomes for the service users supported. However, it has to be acknowledged that practice leadership does not explain all of the variability in the quality of support provided (as also reported in Beadle-Brown et al., 2014, 2015). This finding may reflect the relatively low levels of practice leadership observed (as well as the relatively low levels of active support). It is also likely to reflect, as argued by Mansell and Beadle-Brown (2012), that although practice leadership is an important factor in improving and maintaining the quality of staff support, it is not the only factor required. The nature and quality of training staff had received and organisational factors related to the motivational structures in place for staff are likely to account, at least to some degree, for the unexplained variance in the quality of active support, but these have yet to be directly investigated.

Previous research has shown that the level of ability of the people supported also influences the quality of support received from staff (e.g., Felce, Lowe, Beecham, &

Hallam, 2000; Felce, Lowe, & Jones, 2002; Netten et al., 2010). If active support is well implemented, then staff should be providing the right amount of the right type of support for the person to compensate for the level of disability or the presence of complex needs. When completing the ASM, observers take into account the severity of disability, and hence there should be no or little relationship between the score and level of ability – the support provided should be matched to the needs of the individual. It has been reported elsewhere (Mansell & Beadle-Brown, 2012), however, that when active support is not yet fully implemented, then service users who are more able do get the better support. The irony of this pattern is that service users who are more able do not need such skilled support, but rather the right type of support is often standing back and encouraging rather than intervening. In contrast, service users who have more severe disability will need more skilled, direct support. In the current study, the finding that level of ability (as measured on the SABS) did contribute to variance in active support, along with the relatively low ASM scores, indicate that implementation of active support had not yet been fully achieved in many of these services.

A possible further indication that active support was yet to be fully embedded in the services was the low return rate for surveys; that is, one third of services did not have enough staff surveys returned to be included in the analysis. It may be that staff are not fully engaging with active support or that there is low morale within staff teams. Still, the finding of no differences between the services in which staff did and did not return surveys in terms of the quality of support, level of practice leadership, the ability of the service users, overall satisfaction, and training in active support provide evidence that the sample we have was at least comparable to the other services in the study.

In terms of the aim of this study to explore how well staff-rated measures compared to observational measures of practice leadership, our main finding was that staff-rated corresponded poorly with an observational measure. Although some caution is needed when interpreting our findings given the small sample size, the results do indicate that accuracy about the extent of practice leadership within a service relies on the use of direct strategies, including observation of practice, interviews of the practice leaders themselves, and document review. The strength of the observational measure may lie in the fact that it triangulates information from reports by front-line managers and front-line staff with observations of leadership practice, thereby providing a means to verify self-reported actions and expectations of staff processes and practices articulated in organisational policies and procedures.

The finding that the measure of staff perceptions of practice leadership does not relate as strongly to the quality of support as the observed measure is consistent with previous research (Beadle-Brown et al., 2014, 2015). Further, it corroborates concerns expressed by Mansell (2011) about the unreliability of proxy reports. Mansell (2011) discussed problems in using reports by support staff about aspects of quality of life for people with disability, noting that they tend to paint a more positive picture than is evident from direct observation, perhaps reflecting a reluctance to say anything that could be perceived as being critical of the service in which they are employed. Higgins (2010) also reported a lack of correspondence between staff-rated, manager-rated, and observed measures of quality of support and levels of engagement, with staff ratings tending to be more positive than observed ratings. Beadle-Brown et al.'s (2014) staff-rated index focused on three of the five domains (coaching, supervision, and team meetings) of practice leadership and was reported at the overall level. However, to address the domain of allocating staff we drew on other items in the SESQ to develop an expanded staff-rated index, and to allow for a more direct comparison between staff self-reports with the observed measure of practice leadership we calculated both an overall score and a score for each of the four domains. Certainly in our findings, on some domains that could be directly compared across measures, self-reports tended to yield higher ratings than did the direct observations, suggesting that either staff perceived practices to be better than they were, or they were reluctant to report otherwise. The reverse pattern was evident for two domains: staff supervision and a focus on active support in team meetings. The observational measure indicated that these did occur more often than may have been reported by staff, perhaps reflecting a failure of staff to recognise their managers' interactions as providing supervision, and to remember the content of meetings, respectively. It should be noted that ratings on these two domains drew more on managers' reports about what they do in team meetings and supervision, than on observations, and as such these could potentially be biased to some extent by the social desirability of managers. The potential for such bias can be reduced through triangulation from different sources, providing a strong rationale for using the observational measure.

Although staff-rated measures may not necessarily provide accurate data on practice leadership, Beadle-Brown et al. (2014) argued the importance of staff views on the basis that if staff do not think they are receiving practice leadership, for example, by their

practice being observed, and through coaching and supervision, then the impact of practice leadership is likely to be limited. On the other hand, it cannot be assumed that if staff report the presence of these elements they are being implemented effectively. Similarly, there are limitations to the use of the observational measure, primarily in terms of the time and resources needed to collect the data to make the ratings. First, in order to ensure reliability, those completing the observations need to be trained and experienced in recognising good (and poor) practice. Second, the observational measure requires researchers to travel to and visit the service, although once there all the data needed for the practice leadership ratings can be collected in approximately half a day. In this study, because the measure was new, we felt it was important to keep practice leadership independent from the measures of quality of staff support and user outcomes. This strategy required two visits to each setting, by different researchers, to collect all the data needed. A study is currently underway in the UK in which the same researcher will collect all required data within each service, which will allow exploration of potential efficiencies and disadvantages.

On the basis of the argument that it would be more efficient to obtain reliable data that can be used to determine the quality of practice leadership from staff rather than through direct observations, especially in large-scale studies, one option worth considering is the revision of the staff-rated measure. The inclusion of more questions perhaps in different formats might increase its reliability. In this study, the staff-rated measure was part of a larger questionnaire that explored many other elements of staff experience and knowledge, but having a longer version of the practice leadership section could yield more useful data.

An expanded staff-rated measure alone would not, however, deal with staff issues in rating practice leadership. It is likely that in order to accurately rate practice leadership, staff need to know conceptually and from their own experience what it should look like. As such, training for direct support workers might usefully include expectations about practice and practice leadership. Finally, although observations may provide only a snapshot of staff practices and the lived experience of those living within a service, Mansell and Beadle-Brown (2011) and Mansell (2011) found that it was possible to obtain a valid picture of the overall level of service quality from even short observations. Although some might argue that observations inevitably are intrusive, as they require a researcher to be present in a person's home, many years of observational research have shown that it is possible to collect such data without negative impact on the people supported (Mansell, 2011). Beadle-Brown, Hutchinson, and Whelton (2008), Ashman, Ockenden, Beadle-Brown, and Mansell (2010), and Mansell and Beadle-Brown (2012) each demonstrated ways that observations can manageably be incorporated into quality monitoring processes. These studies have demonstrated that key to the successful use of this strategy is having observers who are skilled in observational techniques while being sensitive to indicators that someone might be unhappy with their presence.

Conclusion

This study demonstrated little correspondence between staff-rated and an observational measure of practice leadership, and a correlation between the observed measure of practice leadership and the quality of active support only. In light of these findings, the benefits of using observational techniques to measure staff practice and service quality must be weighed against their limitations, and resource and skill requirements. Most commonly, however, in practice, self-reported measures are used both for staff practice and service quality, through a range of data-collection strategies, including paper-based recording by staff, interviews and surveys with people with intellectual disability, reports from families, or reports from staff and managers. Not only might these approaches be limited in terms of accuracy, but also their use may result in the exclusion of the perspective of people with more severe intellectual disability and communication difficulties. In light of these concerns, there is a strong argument for investment in the use of observational measures to gauge the extent and quality of practice leadership to ensure quality support for service users.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Emma Bould http://orcid.org/0000-0003-3108-2072
Christine Bigby http://orcid.org/0000-0001-7001-8976

References

Aman, M. G., Burrow, W. H., & Wolford, P. L. (1995). The aberrant behavior checklist-community: Factor validity and effect of subject variables for adults in group homes. *American Journal on Mental Retardation*, *100*, 283–292.

Ashman, B., Ockenden, J., Beadle-Brown, J., & Mansell, J. (2010). *Person-centred active support: A handbook*. Brighton, UK: Pavilion.

Beadle-Brown, J., Bigby, C., & Bould, E. (2015). Observing practice leadership in intellectual and developmental disability services. *Journal of Intellectual Disability Research*, *59*, 1081–1093. doi:10.1111/jir.12208

Beadle-Brown, J., Gifford, J., & Mansell, J. (2003). *Staff Experiences and Satisfaction Questionnaire (Learning Disability)*. Canterbury, UK: Tizard Centre.

Beadle-Brown, J., Hutchinson, A., & Whelton, B. (2008). A better life: The implementation and effect of person-centred active support in the Avenues Trust. *Tizard Learning Disability Review*, *13*(4), 15–24. doi:10.1108/13595474200800034

Beadle-Brown, J., Hutchinson, A., & Whelton, B. (2012). Person-centred active support: Increasing choice, promoting independence and reducing challenging behaviour. *Journal of Applied Research in Intellectual Disabilities*, *25*, 291–307. doi:10.1111/j.1468-3148.2011.00666.x

Beadle-Brown, J., Mansell, J., Ashman, B., Ockenden, J., Iles, R., & Whelton, B. (2014). Practice leadership and active support in residential services for people with intellectual disabilities: An exploratory study. *Journal of Intellectual Disability Research*, *58*, 838–850. doi:10.1111/jir.12099

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Deveau, R., & McGill, P. (2016). Practice leadership at the front line in supporting people with intellectual disabilities and challenging behaviour: A qualitative study of registered managers of community-based, staffed group homes. *Journal of Applied Research in Intellectual Disabilities*, *29*, 266–277. doi:10.1111/jar.12178

Dunst, C. J., & Hamby, D. W. (2012). Guide for calculating and interpreting effect sizes and confidence intervals in intellectual and developmental disability research studies. *Journal of Intellectual & Developmental Disability*, *37*, 89–99. doi:10.3109/13668250.2012.673575

Felce, D., Lowe, K., Beecham, J., & Hallam, A. (2000). Exploring the relationships between costs and quality of services for adults with severe intellectual disabilities and the most severe challenging behaviours in Wales: A multivariate regression analysis. *Journal of Intellectual & Developmental Disability*, *25*, 307–326. doi:10.1080/13668250020019593

Felce, D., Lowe, K., & Jones, E. (2002). Staff activity in supported housing services. *Journal of Applied Research in Intellectual Disabilities*, *15*, 388–403. doi:10.1046/j.1468-3148.2002.00130.x

Hatton, C., Emerson, E., Robertson, J., Gregory, N., Kessissoglou, S., Perry, J., … Hillery, J. (2001). The Adaptive Behavior Scale-Residential and Community (Part I): Towards the development of a short form. *Research in Developmental Disabilities*, *22*, 273–288. doi:10.1016/S0891-4222(01)00072-5

Higgins, I. (2010). *How should the delivery of active support be monitored?* Unpublished manuscript. Canterbury, UK: Tizard Centre.

Lipsey, M. W., & Wilson, D. B. (2001). *Applied Social Research Methods Series: Vol. 49. Practical meta-analysis*. Thousand Oaks, CA: Sage.

Mansell, J. (2011). *Structured observational research in services for people with learning disabilities* (SSCR methods review, 10). London, UK: NIHR School for Social Care Research.

Mansell, J., & Beadle-Brown, J. (2005). *Engagement in meaningful activity and relationships: An observational measure*. Canterbury, UK: Tizard Centre.

Mansell, J., & Beadle-Brown, J. (2011). Estimating activity duration by momentary time-sampling of part or all of the day. *Journal of Applied Research in Intellectual Disabilities*, *24*, 489–494. doi:10.1111/j.1468-3148.2011.00629.x

Mansell, J., & Beadle-Brown, J. (2012). *Active support: Enabling and empowering people with intellectual disabilities*. London, UK: Kingsley.

Mansell, J., Beadle-Brown, J., Ashman, B., & Ockenden, J. (2004). *Person-centred active support: A multi-media training resource for staff to enable participation, inclusion and choice for people with learning disabilities*. Brighton, UK: Pavilion.

Mansell, J., Beadle-Brown, J., & Bigby, C. (2013). Implementation of active support in Victoria, Australia: An exploratory study. *Journal of Intellectual & Developmental Disability*, *38*, 48–58. doi:10.3109/13668250.2012.753996

Mansell, J., & Elliott, T. E. (1996). *Active Support Measure*. Canterbury, UK: Tizard Centre.

Mansell, J., Elliott, T. E., & Beadle-Brown, J. (2005). *Active Support Measure (Revised)*. Canterbury, UK: Tizard Centre.

Netten, A., Beadle-Brown, J., Trukeschitz, B., Towers, A.-M., Welch, E., Forder, J., … Alden, E. (2010). *Measuring the outcomes of care homes: Final report* (PSSRU Discussion Paper 2696/2). Canterbury, UK: Personal Social Services Research Unit, University of Kent.

Wilson, D. B. (2001). Practical meta-analysis effect size calculator [Computer software]. Retrieved from http://www.campbellcollaboration.org/resources/effect_size_input.php