Research in Developmental Disabilities 94 (2019) 103477

Contents lists available at ScienceDirect

Research in Developmental Disabilities

journal homepage: www.elsevier.com/locate/redevdis

# Factors associated with increases over time in the quality of Active Support in supported accommodation services for people with intellectual disabilities: A multi-level model

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

[a] Living with Disability Research Centre, La Trobe University, Australia
[b] Department of Occupational Therapy, Monash University, Australia
[c] Tizard Centre, University of Kent, United Kingdom

ARTICLE INFO

Number of reviews completed is 3.

*Keywords:*
Implementation
Engagement
Active Support
Practice leadership
Training
Supported accommodation
Multilevel model

ABSTRACT

*Background:* Disability support organisations have embraced Active Support, but it has proved difficult to embed in services.
*Aims:* This study aimed to identify the factors associated with increases over time in the quality of Active Support.
*Method:* Data were collected on the predicted variable of the quality of Active Support, and predictor variables of service user, staff and service characteristics, including practice leadership, and composition and size of services from 51 services in 8 organisations over 2–7 time points. Data were analysed using multi-level modelling.
*Results:* There was significant linear change in Active Support scores (group mean centered at the organisational level) over time. Individuals with lower support needs received better Active Support and those with higher support needs experienced greater increases over time. Stronger practice leadership and more staff with training in Active Support were significant predictors of the quality of Active Support. Larger services with seven or more individuals and where there was a very heterogeneous mix of individuals were associated with lower quality of support.
*Conclusions:* Ensuring strong practice leadership, and staff training in Active Support that emphasises the principle of adapting support to each individual's level of ability and preferences are key to delivering high levels of Active Support.

## What this paper adds

This paper is the first to use a longitudinal multi-level modelling design to add new knowledge of what service providers should focus on to improve the quality of Active Support to the people they support. The study identifies the importance of all staff being trained in Active Support, and being skilled in applying the principle of tailoring support to an individual's level of ability. Training should emphasise how strategies required to support engagement of people with more severe disabilities differ from those for people with lower support needs. Organisations must ensure that staff receive good front-line practice leadership: this means coaching, regular individual feedback based on observation, ensuring staff are clear about their role on shift, have team meetings in which quality of practice is discussed and staff are frequently reminded about having a focus on the quality of life of the people they support. Organisations should ensure their services are small with six or less residents, and the mix of residents in terms of support needs does

* Corresponding author at: Living with Disability Research Centre, La Trobe University, Bundoora, Victoria 3086 Australia.
*E-mail address:* c.bigby@latrobe.edu.au (C. Bigby).

https://doi.org/10.1016/j.ridd.2019.103477
Received 20 March 2019; Received in revised form 17 August 2019; Accepted 19 August 2019
Available online 05 September 2019

not include people with a wide range of support needs.

## 1. Introduction

Since the deinstitutionalisation reforms of the 1970s, shared supported accommodation has been the dominant form of residential support for adults with intellectual disabilities who do not live at home with their family in Australia, the United Kingdom (UK), United States (US) and Scandinavia. A body of research has demonstrated variability in the quality of life outcomes for people with intellectual disabilities living in this type of service (for review see Mansell & Beadle-Brown, 2012). The research exploring propositions about factors thought to account for this variability was reviewed by Bigby and Beadle-Brown (2018). They concluded that the strongest evidence pointed to the quality of staff support, and, in turn, staff use of an approach known as Active Support, with emerging evidence about the influence of service culture and front-line leadership.

Active Support is a practice in which an enabling relationship is utilised to facilitate the engagement of people with intellectual disabilities in meaningful activities and social relationships (Mansell & Beadle-Brown, 2012). In a systematic review and meta-analysis of studies, Flynn et al. (2018) synthesised evidence of the effectiveness of Active Support in improving the quality of staff support to people with intellectual disabilities, which lead to "significant increases in the amount of time residents spent engaged in all types of activities at home" (p. 994). Although some studies have shown that Active Support also has a positive effect on reducing depressive symptoms and challenging behaviour, and increasing adaptive skills, choice and community participation, Flynn et al.'s review failed to show convergence on the significance of change associated with any of these factors. Active Support has been argued to reduce challenging behaviour, rather than being a singular approach to supporting people with challenging behaviour, it is one component of comprehensive interventions, such as Positive Behaviour Support (McGill et al., 2018; Ockendon, Ashman, & Beadle-Brown, 2014).

### 1.1. Embedding Active Support in services

Active Support has been adopted widely in the UK and Australia as part of the practice framework for shared supported accommodation services (services) and is also beginning to be used in other contexts, such as schools and programs to support community participation. Organisational claims about Active Support implementation have not always been reflected in staff practices, with difficulties demonstrated across studies in embedding it in services and maintaining quality over time (Flynn et al., 2018; Qian, Tichá, & Stancliffe, 2017). For example, in an Australian study, only one of six organisations that had included Active Support as part of their practice frameworks for more than five years was found to be delivering consistently good Active Support (Mansell, Beadle-Brown, & Bigby, 2013).

Three types of factors are thought to influence the successful implementation of Active Support: staff training, staff motivation and management commitment (Mansell, Beadle-Brown, Whelton, Beckett, & Hutchinson, 2008). Flynn et al. (2018) concluded from a systematic review of 10 studies about experiences of Active Support that the strongest, yet still tentative, evidence was the impact on implementation of the type of training (i.e., advantages of combined classroom and in-situ), lower staff-to-resident ratios, working in larger services (but only to a maximum of eight residents), and organisational leadership and management support for staff. Not included in the Flynn et al. review were two studies not specifically focused on experiences of implementing Active Support but investigating practice leadership which has been identified as a factor associated with the quality of Active Support. Beadle-Brown, Bigby, and Bould (2015) found services where practice leadership was stronger provided significantly better active support ($z = 2.540\ p = .01$, $n = 46$), and Bould, Beadle-Brown, Bigby, and Iacono (2018) found that when a practice leader was present in a service, the quality of Active Support was significantly higher, for service users with both more severe ($t\ (42) = 4.241$, $p = .001$) and less severe ($t\ (76) = 3.513$, $p = .001$) intellectual disability compared to when the practice leader was absent.

### 1.2. Limitations of previous studies

Few studies have investigated the quality of Active Support over a period longer than 12 months or included more than one organisation. Although organisations in Mansell et al.'s (2013) study had been implementing Active Support longer than five years, the design was cross sectional rather than longitudinal. Furthermore, research into the variables associated with Active Support has been limited by the use of single level linear regression, where all variables have been treated equally, regardless of where they sit within an organisational hierarchy with variably shared dependencies at each level. Hence, assigning mean scores obtained for staff working within a service on certain measures to every individual service user within the same service ignores statistical problems arising from data dependence within clusters (e.g., service users and services). In statistical terms, this error increases the chance for Type 1 errors (i.e., rejecting a true null hypothesis) arising from underestimating parameter estimates and their standard errors (Raudenbush & Bryk, 2002; Snijders & Bosker, 1999).

In an exploration of many of the previously identified variables associated with good Active Support (Bigby, Bould, Iacono, Kavangh & Beadle-Brown, in press) used multi-level modelling (MLM), which improves estimation of parameters because the variance of the dependent variable is partitioned into the hierarchical structure of the data (Raudenbush & Bryk, 2002). Similar to Mansell et al.'s (2013) study, however, it was a cross sectional rather than longitudinal design, with data from one time point only for each of 134 services managed by 14 organisations. The model developed by (Bigby et al., in press) enabled identification of Active Support quality predictors at three levels: (1) the individual service user - greater adaptive behaviour; (2) the service - stronger practice leadership, higher percentage of staff trained in Active Support, and fewer than seven service users with limited

heterogeneity; and (3) the organisation - smaller number of services managed and longer period implementing Active Support. However, the model did not show significant associations between quality of Active Support and variables previously explored relating to staff characteristics, such as qualifications, experience, and attitudes, and organisational hygiene, such as job satisfaction, role clarity and conflict.

### 1.3. Present study

The aim of the present study was to explore whether the factors identified in the literature as affecting the implementation, and thereby the quality of Active Support, predict increases in the quality of Active Support over time and identify whether these are similar to those found to be associated with the quality of Active Support at a single point in time by (Bigby et al., in press). The data reported were drawn from a large-scale study of services in Australia that commenced in 2009.

## 2. Materials and method

### 2.1. Design

The study was a repeated measures longitudinal design. The same data were collected at baseline and then intervals of 12–18 months, over periods of 2–7 years from the same 51 services managed by eight organisations. The predicted variable was the quality of Active Support, with predictor variables of service user, staff, service and organisational characteristics. The study began with six organisations, with others joining as it progressed. Thus, not all organisations participated for the same period of time: data were collected at baseline and six subsequent time points for one organisation, five for two organisations, three for two organisations, two for one organisation, and one for two organisations. Rather than calendar years or years in the study, we refer to data collection at baseline and subsequent *time points*, reflecting that a similar trajectory of repeated measures occurred for each organisation, but did not coincide with the same calendar year for all organisations.[1]

The La Trobe University Human Research Ethics Committee (HREC) approved the study. Staff and service users with capacity provided their own consent; for service users without consent capacity, permission was obtained from a person who usually made decisions for them, typically a next-of-kin or senior staff member of the service.

### 2.2. Participants and settings

The eight participating not-for-profit organisations operated in five different Australian states. As Table 1 shows, organisations had been implementing Active Support for periods ranging from 1 to 14 years, and managed from 5 to 34 services.

Depending on the size of the organisation, all or a sub-set of services were included in the study. Inclusion criteria were services providing 24 -h support for 1–12 service users ($M = 4.8$) living in an ordinary house dispersed among other houses in the community, and having at least one service user consenting to participate. Table 2 shows the number of services and service users with intellectual disabilities who were included at baseline and each subsequent time point for each organisation. Changes in service users' place of residence or ill health that prevented observational data collection at a scheduled time caused numbers to fluctuate. Consequently, the data set is unstructured, in that service users have an unequal number of observation data collection points. Change in the quality of Active Support was measured for two time points for just under half (46%) of service users, three time points for 29 (18%) service users, four time points for 32 (19.9%) service users, five time points for 16 (9.9%) service users, six time points for five (3.1%) service users, and all seven time points for only five (3.1%) service users.

### 2.3. Measures

#### 2.3.1. Service user characteristics

Data about service users were collected through a staff-completed audit questionnaire, which included questions about gender, date of birth and other disabilities present. It also included the short form of the Adaptive Behavior Scale (SABS) Part 1 (Hatton et al., 2001) to determine level of support needs and two specific items to indicate general receptive and expressive communication skills. The Aberrant Behavior Checklist (ABC) (Aman, Burrow, & Wolford, 1995) was used to measure level of challenging behaviour. The full-scale score for Part 1 of the Adaptive Behaviour Scale (ABS) was estimated from the SABS using the method described by Hatton et al. (2001). Authors of these measures have reported them to have acceptable reliability and validity.

#### 2.3.2. Staff experiences and satisfaction

The Staff Experiences and Satisfaction Questionnaire (SESQ) (Beadle-Brown, Gifford, & Mansell, 2005) was used in an adapted form. The elements used in this study included Section A items addressing staff demographics and training; Section B items regarding experiences at work in terms of satisfaction, role clarity and conflict and staff perception of management; and a shortened 13-item version of the original Section D scale on attitudes towards people with intellectual disabilities. The scales used are described in detail

[1] For example, baseline data for seven organisations were collected in 2009/10, and the collection was repeated at time point 2 which was in 2011/12, and time point 3 was 2013 and so on; for organisations that joined the study in 2015, baseline was 2015 and time point 2 was 2016.

**Table 1**
Size of each organisation and number of years since Active Support first implemented at time of 2017 data collection.

| Organisation | Services | Service users (total) | Years of Active Support |
|---|---|---|---|
| 1 | 5 | 21 | 8 |
| 2 | 15 | 28 | 14 |
| 3 | 5 | 18 | 13 |
| 4 | 34 | 155 | 12 |
| 5 | 7 | 29 | 5 |
| 6 | 10 | 62 | 5.5 |
| 7 | 7 | 42 | 1 |
| 8 | 31 | 142 | 1 |

**Table 2**
Number of services, consenting service users (SUs) and staff surveys from each organisation included in the analysis at each time point.

| Time Point | Frequencies | Organisation | | | | | | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | |
| Baseline | Services | 5 | 3 | 2 | 6 | 7 | 6 | 6 | 2 | **37** |
| | SUs | 16 | 8 | 6 | 23 | 22 | 20 | 29 | 6 | **130** |
| | Staff | 25 | 11 | 7 | 42 | 36 | 21 | 26 | 9 | **177** |
| 2 | Services | 5 | 3 | 2 | 5 | 6 | 6 | 5 | 2 | **34** |
| | SUs | 18 | 9 | 7 | 11 | 19 | 21 | 27 | 6 | **118** |
| | Staff | 20 | 12 | 8 | 32 | 24 | 21 | 22 | 11 | **150** |
| 3 | Services | 5 | 5 | 3 | 0 | 6 | 7 | 0 | 0 | **26** |
| | SUs | 18 | 14 | 8 | 0 | 21 | 25 | 0 | 0 | **86** |
| | Staff | 22 | 22 | 11 | 0 | 33 | 42 | 0 | 0 | **130** |
| 4 | Services | 5 | 7 | 5 | 6 | 0 | 0 | 0 | 0 | **23** |
| | SUs | 17 | 15 | 12 | 25 | 0 | 0 | 0 | 0 | **69** |
| | Staff | 19 | 26 | 18 | 28 | 0 | 0 | 0 | 0 | **91** |
| 5 | Services | 5 | 7 | 5 | 0 | 7 | 7 | 0 | 0 | **31** |
| | SUs | 19 | 18 | 18 | 0 | 25 | 29 | 0 | 0 | **109** |
| | Staff | 18 | 28 | 21 | 0 | 32 | 26 | 0 | 0 | **125** |
| 6 | Services | 2 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | **8** |
| | SUs | 6 | 0 | 0 | 14 | 0 | 0 | 0 | 0 | **20** |
| | Staff | 9 | 0 | 0 | 23 | 0 | 0 | 0 | 0 | **32** |
| 7 | Services | 2 | 5 | 4 | 0 | 0 | 0 | 0 | 0 | **11** |
| | SUs | 6 | 9 | 13 | 0 | 0 | 0 | 0 | 0 | **28** |
| | Staff | 10 | 22 | 15 | 0 | 0 | 0 | 0 | 0 | **47** |

in Mansell et al. (2008, pp 401-402) along with reliability and validity from that study. Cronbach's alpha for the 13-item attitude scale from a large scale evaluation involving over 550 staff was 0.856

### *2.3.3. Staff-to-resident ratio*

A researcher completed a proforma at the time of observation in the service, recording information about the number of residents living in the service and of staff on duty.

### *2.3.4. Practice leadership*

The Observed Measure of Practice Leadership developed by Beadle-Brown et al. (2015) was used to measure the quality of five elements: (1) the focus, overall, on the quality of life of all service users; (2) the allocation and organisation of staff to provide the support people need; (3) the extent of coaching, observing, modelling and giving feedback; (4) performance reviews with individual staff during supervision; and (5) performance reviews of teams during team meetings. This measure has been shown to be a reliable and valid measure with good internal consistency, inter-rater reliability and construct validity (Beadle-Brown et al., 2015). To complete the measure, a researcher interviewed the front-line manager, reviewed paperwork associated with practice leadership and observed within the service for 15–30 minutes, then rated each of the elements on a five-point scale (anchored by 1 = no or almost no evidence of the element being in place and 5 = excellent – could not really improve on this element). A mean score was calculated from scores summed across the elements, and represented the overall strength of practice leadership in a service. The measure was implemented by five researchers; each had been trained by one of the authors and conducted at least two visits with one other trained observer before collecting data alone.

### *2.3.5. Quality of Active Support*

The predicted variable was the quality of Active Support, determined using the Active Support Measure (ASM) (Mansell, Elliott, & Beadle-Brown, 2005). The authors of the ASM have reported the measure to have acceptable reliability and validity, with a Cronbach

alpha over 0.9 in most studies (see for example, Beadle-Brown, Hutchinson, & Whelton, 2012; Mansell, Beadle-Brown, Macdonald, & Ashman, 2003). The measure was completed according to guidelines provided by Mansell et al. (2005) and the fourth author was involved in developing the measure, preparing the guidance and training the researchers in the current study. The ASM comprises 15 items addressing the quality of staff support to individual service users and their skills in enabling them to be engaged in meaningful activities and relationships. A researcher completes the ASM at the end of 2 h of observation. During the observation, detailed notes are taken about the type of activities and nature of the contact observed. These notes are then used to rate the ASM items immediately after the observation. Each of the 15 items are rated on a scale of 0 (poor, inconsistent support) to 3 (good, consistent support), in line with the scoring guidelines provided by Mansell et al. (2005), with tallies across items converted to a percentage. The total possible raw score for each observed service user is 45, unless the two items relating to challenging behaviour are scored as 'Not applicable' (i.e. the service user was not observed to display challenging behaviour), in which case the maximum possible score is 39. A percentage score of 66.66 is considered a good level of Active Support (Mansell & Beadle-Brown, 2012).

Observations were conducted by a team of 13 observers over the course of the study, including the first author, who, along with the second and fourth author trained the others. Inter-observer reliability determined for each of the 15 items of the ASM within the group of seven observers involved in the 2009 data collection (and the two observers in 2010) was 60% on average (range 29–98%, *number of paired scores* = 24). Kappa was on average .32 (this low score was explained, in part, by low occurrences of Active Support). Reliability on the ASM was not conducted for data collected in 2012 because all observations were by one observer (first author), who was involved in all years of the study. In 2014, there was 84% average agreement across three observers (range 73–100%, $n = 15$), and average Kappa was .61 (range .21–.80). In 2015, 2016 and 2017, for four observers, there was 66% average agreement (range 55–100%, $n = 10$); 58% (range 30–100%, $n = 10$) and 87% (range 69–100%, $n = 26$) respectively. Kappa was on average .55 (range .20–.100); .51 (range .29–.100) and .73 (range .53–.100), respectively. Although across the years, agreement was found to be low for some ASM items, paired T-Tests showed there were no significant differences for the overall ASM scores (range $p = .271$–.385).

### 2.3.6. Procedures

For each organisation, an audit database was created and sent to a contact person from each organisation, with instructions to (a) complete the coding of service users identified within the database; (b) distribute questionnaires for all service users in the organisation, with requests for a staff member who knew the individual well to complete and return to the contact person; (c) remove the service users name on each questionnaire and leave only a unique code from the database; and (d) return completed audit questionnaires to the research team in the pre-paid envelopes provided.

Each service was also sent staff questionnaires with a request that they be distributed to staff, including the front-line manager. Each staff member was invited to individually complete and return a questionnaire in a pre-paid envelope. A researcher visited each service, usually between 4:00-6:00 pm on a weekday to conduct the observation and complete the ASM. On another day, a researcher visited the service and completed the Observed Measure of Practice Leadership. Every year, therefore, two visits were made to each service within a 2–4 month period, with the exception of services that shared a front-line manager with another participating service, in which case the measure was completed during a second visit to only one of these services.

### 2.3.7. Analyses

Data were entered into IBM SPSS 24. Descriptive statistics and correlational analysis were conducted, with Cohen's (1988) guidelines used to report correlation effect sizes. For each service user, the percentage ASM score was calculated, and an Adaptive Behaviour Scale (ABS) Part 1 score derived from the short adaptive scale as described by Hatton et al. (2001). The ABS scores were recoded initially into two groups below 151 and 151 and above. This rough cut off has been used in other studies to indicate more or less severe disabilities (see Mansell et al., 2013). However, preliminary analysis indicated that there appeared to be a lower cut off of an ABS score of 80 and one organisation only supported people below that cut off. As such the ability grouping was revised to include ABS score less than 80, ABS score between 81–150, or ABS score of 151 and above. The ABS groups were aggregated to the service level representing the number of ABS groups supported by that service. For example, in a service with four service users with two scoring an ABS of less than 80 and two scoring 81–150, the total number of ABS groups was two. Also, at the service level, a mean practice leadership score was calculated across the five elements of the Observed Measure of Practice Leadership for each service(s) in which the front-line manager worked. The unique codes from each organisation derived from the audit database were used to ascertain the total number of service users in each service, which were grouped into two categories: 1–6 and 7 + . This cut off is based on earlier studies by Tøssebro (1995) and Flynn et al. (2018). The aggregated data for ability group, practice leadership score and size of setting were assigned to all the individual service users within the same service(s).

For the staff questionnaires, data were included in the analysis only if at least three staff surveys were returned for a service at each time point. Individual staff data on job satisfaction, role clarity and conflict, perception of practice leadership, quality of senior management, attitudes towards people with intellectual disabilities and training in Active Support were aggregated to the service level through a mean score for each service, and subsequently assigned to all the individual service users within the same service. In terms of staff training, although staff reported whether or not they had training in Active Support, few answered the question about the type of training (i.e. classroom and or in-situ); hence, type of training was not included in the analysis. Table 2 shows the number of staff at baseline and each time point included in this analysis. Due to missing data, two service users were excluded from the analyses, but no services were excluded, resulting in final totals in the Multi-level modelling (MLM) of 194 service users from 51 services. The fluctuations shown in Table 2 reflect differences in the number of year's organisations (and their services) were involved in the study.

Taking clustering into account, the data structure had five levels: baseline and six subsequent time points (level 1) nested within

194 individual service users (level 2) nested within 51 services (level 3), nested within eight organisations (level 4), nested within five states (level 5). Because relatively few services were distributed across eight organisations and five states, the ASM scores were group mean centred (Organisation Mean –ASM Score = ASM GMC) at the organisational level. Organisations are nested within states, so this approach resulted in a response variable ICC of zero at both the organisation (level 4) and state (level 5) levels (Enders & Tofighi, 2007).

MLM regression was implemented using the MLwiN program (Version 3.02; Charlton, Rasbash, Browne, Healy, & Cameron, 2017). In light of the size and structure of the data set, and the ICC of zero at levels 4 and 5, the variance required partitioning at three levels (Rodriguez, 2007), thus the Deviance Information Criterion (DIC) (Spiegelhalter, Best, Garlin, & van der Linde, 2002) statistics for model comparison was calculated using the Markov chain Monte Carlo (MCMC) (Browne, 2017). Using this approach, any decrease in the goodness of fit diagnostic, the DIC, indicated a better model. All models were estimated using non-informative priors (Browne, 2004) with a burn-on of 1000 and 20,000 iterations to allow each model to converge on the correct posterior distribution, and collect sufficient independent samples from the posterior distribution to permit a good estimate. An initial null model was estimated, which computes an intraclass correlation coefficient (ICC): that is, the expected (population) correlation between two randomly chosen elements in the same group (Hox, 2010). Using a bottom-up approach, a series of multi-level models were then built (Hox, 2010; Raudenbush & Bryk, 2002). It was necessary to specify starting values for the level 3 variance (set to 1) and covariance (set to 0) prior to fitting the model with MCMC to ensure the variance-covariance matrix was positive definite. The fully adjusted model was:

$$\begin{aligned}\text{ASM GMC Score}_{ijk} = {} & \beta_{ijk} + \text{Time Point}_{ijk} + \text{ABS Score}_{ijk} + \text{Service Size (base, 6 or less service users)}_{ijk} \\ & + \text{Time Point*Service Size}_{ijk} + \text{Number of ABS Groups}_{ijk} + \text{Mean Practice Leadership}_{ijk} \\ & + \text{Staff with Active Support Training}_{ijk} + e_{ijk}\end{aligned}$$

In this model, $i$ refers to the time point, $j$ to the service users and $k$ to the services. $\beta_{ijk}$ refers to the grand mean (i.e., average Active Support GMC score across the seven time points from 194 individuals, across 51 services) and $e_{ijk}$ to a random effect.

Significant results were reported for estimates that were more than twice their estimated empirical standard error. All predictors were grand mean centred (in order that the intercept be centred around the mean of the sample) to facilitate interpretation of the intercepts and slopes, and because the influence of the service level was of primary interest (Enders & Tofighi, 2007).

## 3. Results

Descriptive statistics for the 194 service users included in the analysis along with the Active Support scores for all time points are presented in Table 3. As can be seen from this table, over the time points, service user mean age increased from 40 to 48 years, the proportion of males decreased, and there was some variability in ABS score, but less so in ABC scores. The proportion of service users who were non-verbal was relatively stable at around a third, but decreased to 18% at the final time point (perhaps accounted for by a drop-in sample size). Of particular note in Table 3 is the increase in the means for the ASM, which remained above the level of 66.66% considered to be good Active Support from time point 5. At most time points, the range in ASM scores tended to be large, but became much narrower at time point 6, with a higher low score indicating an overall shift up for all service users represented. Hence, ASM scores became less variable, particularly over the final two time points.

Relationships among the predictor levels included in the final model were examined using correlational analyses (Spearman) at the overall (across all seven time points) service user level and are shown in Table 4. The largest correlation with the quality of Active Support (ASM Score GMC), according to Cohen's (1988) guidelines, was the level of adaptive behaviour (ABS) ($\rho = .317$, $n = 560$, $p < .001$), with a medium effect (Table 4).

**Table 3**
Characteristics of service users at baseline and each time point and the quality of Active Support.

| Variable | | Time Point | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Baseline | 2 | 3 | 4 | 5 | 6 | 7 |
| | *N/n* | *130* | *118* | *86* | *69* | *109* | *20* | *28* |
| Age (years) | *M* | 40 | 42 | 42 | 41 | 44 | 44 | 48 |
| | Range | 18 to 76 | 18 to 77 | 19 to 78 | 17 to 66 | 19 to 81 | 28 to 66 | 29 to 70 |
| Males | | 53% | 48% | 43% | 41% | 42% | 25% | 39% |
| ABS score | *M* | 140 | 134 | 133 | 155 | 139 | 129 | 155 |
| | Range | 22–260 | 24–251 | 22–260 | 34–251 | 22–263 | 36–244 | 88–216 |
| ABC total score | *M* | 29 | 25 | 23 | 31 | 22 | 44 | 26 |
| | Range | 0–103 | 0–93 | 0–110 | 0–110 | 0–110 | 3–87 | 0–81 |
| Non-verbal[1] | | 32% | 37% | 36% | 26% | 34% | 35% | 18% |
| ASM score | *M* | 45 | 55 | 57 | 61 | 67 | 69 | 75 |
| | Range | 7–98 | 10–92 | 13–92 | 8–92 | 18–100 | 18–87 | 54–92 |

[1] Refers to the service users who did not use speech, but relied on non-speech modes of informal (e.g. gestures, facial expressions, body language, posture) or formal (e.g., signs, picture symbols) means of communication.

**Table 4**
Spearman correlations between predictor variables.

| | ABS Score | Service size (SU) | Number of ABS groups | Mean Practice Leadership | Staff with Active Support Training |
|---|---|---|---|---|---|
| ASM Score GMC | .317[b] | −.135[a] | −.200[b] | .190[b] | .174[b] |
| ABS Score | | .102[a] | −.103[a] | −0.049 | −.022 |
| Service size - Total number of SUs | | | .262[b] | .043 | −.145[b] |
| Number of ABS groups | | | | .027 | .011 |
| Mean Practice Leadership | | | | | .007 |

[a] Correlation significant at the 0.05 level (2-tailed).
[b] Correlation significant at the 0.01 level (2-tailed).

Tables 5 and 6 show the modelling results as parameter (beta) coefficients and their standard errors, along with the model-fitted diagnostic DIC. One purpose of longitudinal three-level modelling was to assess service (level 3) influences on the individual average (level 2) change over time (level 1). The first model shown in Table 5, therefore, provides an unconditional three level model, which enabled three ICCs (see Hoffman, 2015; Snijders & Bosker, 2012) to be calculated to assess the influence of service characteristics on change in the quality of Active Support (ASM score GMC) over time. The proportion of total variable variation that occurred between ASM GMC scores across all the services was 16% (73.47/ [375.24 + 1.51 + 73.47]). This level 3 ICC estimate can be interpreted as the expected correlation between two ASM GMC scores drawn completely at random from any time point, from two service users within the same service. An alternative level 3 ICC was 98% (73.47 / [73.47 + 1.51]), interpreted as the expected correlation between the mean (i.e., averaged across the repeated measurements) ASM GMC scores from two service users drawn completely at random from the same service. The level 2 ICC was 17% ([1.51 + 73.47] / [375.24 + 1.51 + 73.47]), interpreted as the expected correlation between two repeated measurements sampled from the same service user.

In the second model shown in Table 5, time point was centred at baseline and fixed to establish the average change in the ASM score over time across all 194 service users nested within 51 services. Fig. 1 shows that, on average, there was a linear increase in ASM scores over time. The average ASM GMC score at baseline for all service users nested within all services was −10.06%, and the constant expected rate of change in the score was 4.68 percentage points per time point after baseline. However, the first question to address was how to best model and quantify the change in ASM GMC score over time, as with repeated measurements, T - 1 fixed effects, and T - 2 random effects could be needed to accurately model change (Snijders & Bosker 1999). As such we looked at polynomial trend components (i.e. linear, quadratic, and cubic) and the DIC indicated linear-only (6512.601) more accurately modelled the change, as there was a better model fit compared to quadratic linear (6513.209), or linear, quadratic, and cubic polynomial components (6513.518). Furthermore, when the quadratic and cubic effects of time point were allowed to vary randomly across service users within services at Level 2 and across services at Level 3 the models failed to converge. This, as Peugh and Heck (2017) stated, is likely to be due to attempting to "over-fit" the model.

In light of the repeated measurements collected from service users (within services) over time, differences in rates of change (using linear polynomial trend) between service users were determined. In the third model shown in Table 5, time point was allowed to vary across individual service users at level 2. The average ASM GMC score and expected linear change remained relatively unchanged. In contrast, the level 1 residual variance decreased by 13% and the average ASM GMC score variation across services decreased 3%, while the variation in the mean ASM GMC scores across service users within services (level 2) at baseline increased from that of the previous model due to estimating random effects at level 2, rather than variation being fixed, as in the previous model. Results show that linear increases in scores on the ASM GMC differed significantly across all service users within all services. ASM GMC scores for 95% of service users within all services changed, on average, by differences from −1.59 and 11.07, around a mean increase of 4.74 points per time point after baseline. Furthermore, the intercept/slope covariance estimate (−29.41) was

**Table 5**
Parameter (beta) estimates of the multi-level models and deviance information criterion (MCMC) on the effect of time on Active Support Score GMC.

| | Model 1 (S.E) | Model 2 (S.E) | Model 3 (S.E) | Model 4 (S.E) |
|---|---|---|---|---|
| **Fixed parameters** | | | | |
| Intercept | 1.44 (0.44) | −10.06 (1.81) | −10.16 (1.91) | −9.86 (2.53) |
| Data collection time point – Centred at Baseline | | 4.68 (0.39) | 4.74 (0.46) | 4.67 (0.79) |
| **Random effects** | | | | |
| Level 1 (Time Point): Residual | 375.24 (20.04) | 307.18 (16.67) | 267.55 (16.95) | 251.20 (15.28) |
| Level 2 (Individual SU's): Intercept | 1.51 (3.66) | 0.240 (4.86) | 94.14 (32.41) | 8.35 (11.55) |
| Level 2 (Individual SU's): Covariance | | | −29.41 (9.68) | −61 (1.72) |
| Level 2 (Individual SU's): Slope | | | 10.43 (3.36) | 0.80 (0.50) |
| Level 3 (Services): Intercept | 73.47 (22.56) | 99.99 (27.85) | 97.00 (27.48) | 257.33 (70.26) |
| Level 3 (Services): Covariance | | | | −58.92 (19.09) |
| Level 3 (Services): Slope | | | | 21.46 (6.46) |
| **Deviance Information Criterion (DIC)** | 6656.320 | 6512.601 | 6478.248 | 6410.465 |
| **Change in DIC** | | 143.719 | 34.353 | 67.783 |

All estimates are significant at 0.05 probability level or smaller.

**Table 6**
Parameter (beta) estimates of the multi-level models and deviance information criterion (MCMC) with the level 2 and level 3 predictor variables.

| | Model 5 (S.E) | Model 6 (S.E) | Model 7 (S.E) | Model 8 (S.E) |
|---|---|---|---|---|
| **Fixed parameters** | | | | |
| Intercept | −9.89 (2.33) | −8.495 (2.54) | −10.13 (2.43) | −5.69 (2.44) |
| Time point – Centred at Baseline | 4.67 (0.80) | 4.50 (0.79) | 5.15 (0.75) | 4.76 (0.99) |
| Individual (Level 2) predictors | | | | |
| ABS Score | 0.12 (0.02) | 0.12 (0.02) | 0.12 (0.02) | 0.12 (0.02) |
| Service level (Level 3) predictors | | | | |
| Service size - Total number of SUs (6 or less base) | | −15.51 (4.91) | 7.21 (7.41) [N.S] | 14.76 (7.94) [N.S] |
| Time Point X Service size - Total number of SUs (6 or less base) | | | −9.70 (2.55) | −10.32 (3.40) |
| Number of ABS groups in the service (1 or 2 ABS groups base) | | | | −10.09 (3.49) |
| Mean practice leadership | | | | 4.18 (1.11) |
| Staff with training in Active Support | | | | 5.61 (2.20) |
| **Random effects** | | | | |
| Level 1: Residual | 236.67 (13.50) | 234.41 (13.29) | 232.70 (13.29) | 227.31 (15.61) |
| Level 2: Intercept | 2.74 (3.39) | 2.45 (3.11) | 2.62 (3.47) | 2.73 (3.71) |
| Level 2: Covariance | −0.45 (0.85) | −0.43 (0.88) | −0.43 (0.90) | −0.46 (0.92) |
| Level 2: Slope | 0.63 (0.38) | 0.63 (0.40) | 0.64 (0.38) | 0.76 (0.48) |
| Level 3: Intercept | 212.92 (61.62) | 253.16 (70.84) | 218.13 (62.56) | 185.05 (61.05) |
| Level 3: Covariance | −46.81 (17.42) | −58.73 (19.34) | −46.45 (16.39) | −63.99 (22.91) |
| Level 3: Slope | 21.21 (6.64) | 21.86 (6.61) | 17.57 (5.45) | 30.57 (10.11) |
| **Deviance Information Criterion (DIC)** | 6357.57 | 6350.727 | 6343.939 | 4706.53 |
| **Change in DIC** | 52.895 | 6.843 | 6.788 | 1637.409 |

All estimates, except where indicated by [NS] are significant at 0.05 probability level or small.

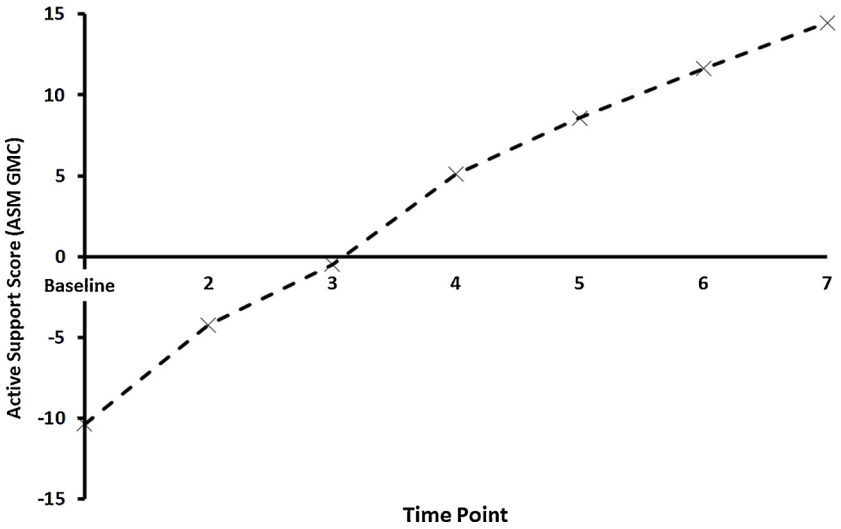


**Fig. 1.** Change in Active Support score over time.

significant, and, together with significant and positive slope for linear time (4.74), shows that the rate of increase in ASM GMC scores over time was, on average, slower for service users within services who had higher ASM GMC scores at baseline.

To determine whether there were differences in rates of change across services, time point was allowed to vary randomly both across service users and services in the fourth model, shown in Table 5. The significant intercept variance (257.33) remained, but significant variation over time on ASM scores across services (21.46) were found. ASM GMC scores for 95% of services changed on average by −4.41 and 13.75, around a mean increase of 4.67 points per time point after baseline. Furthermore, the intercept/slope covariance estimate at level 3 (−58.92) was significant, and together with significant and positive slope for linear time (4.67), shows that the rate of increase in ASM GMC scores over time was, on average, slower for services with higher ASM GMC scores at baseline. The results of model 4, however, show significant linear change in ASM GMC scores occurred over time, and the linear increase varied randomly across both service users within services and across services.

In the fifth model shown in Table 6, service user ABS scores were added to the three-level analysis as a level 2 predictor variable (as a fixed effect) and the main effect was significant (0.12). At baseline, service users with lower support needs, as measured by the ABS, showed, on average, ASM GMC scores that were 0.12 points higher than scores for service users with greater support needs. The time point by ABS score was not significant (-0.017), indicating that the change per time point for service users with lower support needs beyond baseline was not significantly different to service users with higher support needs. When the ABS main effect was

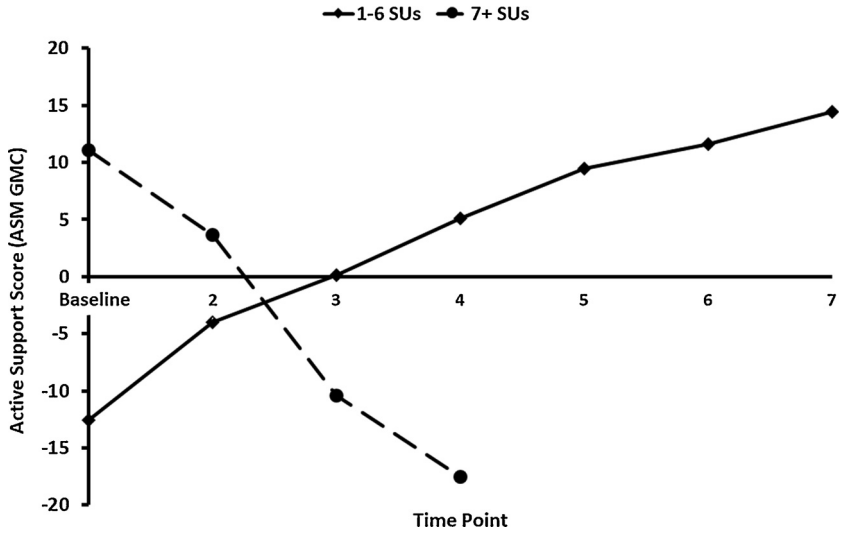


**Fig. 2.** Two-way interaction between time point and service users in a service.

allowed to vary randomly at level 3, the result was not significant, which can be interpreted to mean that the main effect of ABS on the ASM GMC score did not vary significantly across services.

No other level 2 predictors were shown to have a significant effect on the ASM GMC score (i.e., age, challenging behaviour). Therefore, in the sixth model shown in Table 6, one level 3 predictor variable was added, resulting in a main effect of service size – Total number of service users in a service: at baseline, services with seven or more service users showed, on average, ASM GMC scores that were 15.51 points lower than the scores for services with 1–6 service users. There was, however, a significant interaction effect between time point and service size, as shown in model 7 in Table 6: change in ASM GMC scores beyond baseline for services with 7 or more people was, on average, 10.32 points lower than services with 1–6 people. This significant interaction is presented graphically in Fig. 2. However, the significant interaction resulted in the main effect becoming non-significant.

The final model shown in Table 6 included additional level 3 predictors. Services with higher practice leadership scores and a higher percentage of staff who had received training in Active Support had higher ASM GMC scores. Conversely, as shown in Fig. 3, in services with a great deal of heterogeneity amongst service users (i.e., ABS scores falling within each of the three ABS groups), ASM GMC scores were on average 10.09 points lower than the scores in services with only one or two ABS groups. No other level 3 predictors or interactions were shown to have a significant effect on the ASM GMC score. However, the inclusion of the additional level 3 predictors led to the main effect of number of service users in a service becoming positive, after having been negative, and significant (in Model 6). This finding is not accurately reflected in the graph of the significant two-way interaction between time point and number of services users in a service, in Fig. 2, most likely because only 4/51 services had seven or more services users living together.

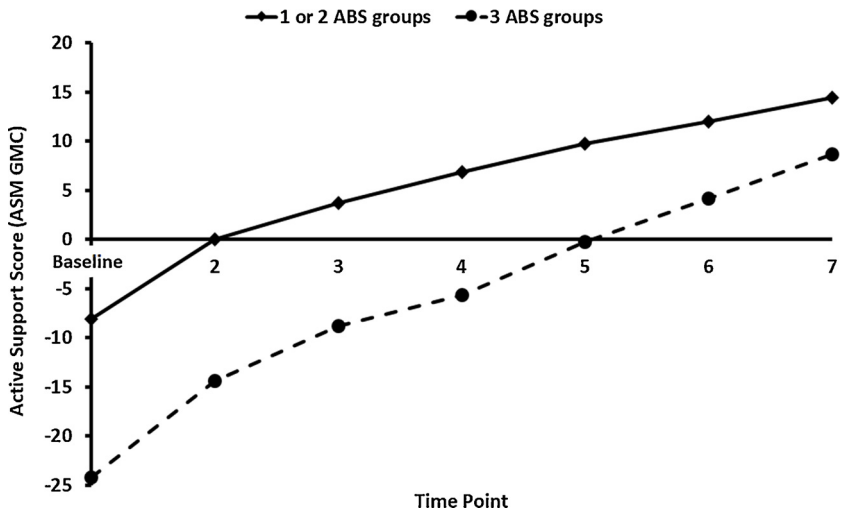


**Fig. 3.** Main effect for number of ABS groups in a service.

A pseudo-$R^2$ effect size estimate was calculated using a three-step process to ascertain the overall effect size modelled by the main and interaction effects (Hox & Roberts, 2011; Peugh & Heck, 2017). First, we obtained model-predicted ASM GMC scores for all 194 residents using the intercept and slope estimates, as follows.

$$\begin{aligned}\text{ASM GMC Score}_{ijk} = &-5.69 + 4.76\,(\text{Time Point}_{ijk}) + 0.12\,(\text{ABS Score}_{ijk}) + 14.76\,(\text{Service size} - \text{Total number of SUs}_{ijk}) \\ &- 10.32\,(\text{Time Point*Service Size}_{ijk}) - 10.09\,(\text{Number of ABS Groups}_{ijk}) + 4.68\,(\text{Mean Practice Leadership}_{ijk}) \\ &+ 5.61\,(\text{Staff with Active Support Training}_{ijk})\end{aligned}$$

Second, we obtained Pearson correlations between the model-predicted ASM GMC scores in step 1 ($\text{ASM GMC Score}_{Predictedijk}$) with the observed ASM GMC scores ($\text{ASM GMC Score}_{Observedijk}$) in the dataset. Third, we squared the resulting Person Correlation [$r_{Predicted,\ Observed}$] $^2$) to produce the pseudo-$R^2$ value. Results showed that $R^2_{Pseudo} = [.5962]^2 = .36$; 36% of the variance in ASM GMC scores across the $I = 7$ time points for all $J$ residents, across all $K$ services was modelled by the main and interaction effects. However, the equation above contains just fixed effect estimates, not the random effect estimates, and because of the choice of centreing, the predicted ASM GMC scores are based on the assumption that the individual represented by a mean ABS score is living in a service where the practice leader scores at the sample mean on the practice leader measure, and in which the percentage of staff trained in Active Support is at the sample mean. Consequently, the $R^2_{Pseudo}$ value obtained is considered a conditional effect size (Peugh & Heck, 2017).

### 3.1. Summary of results

Scores on the ASM GMC increased, on average, from baseline to the other time points, and Fig. 1 showed that a linear-only trend captured that change. Further, the linear increase in ASM GMC scores, was, on average, slower for individuals within services and across services with higher ASM GMC scores at baseline. In terms of the influence of individual and service variables on ASM GMC scores, greater levels of adaptive behaviour (ABS score), higher practice leadership scores, and a higher percentage of staff who had received training in Active Support were significant predictors of higher levels of Active Support; conversely, services with high heterogeneity amongst service users' (3 ABS groups) were associated with lower quality of support. There was also a significant two-way interaction on ASM GMC scores and service size, such that increases in ASM GMC scores occurred only in services with 1–6 service users, and not in those with 7 or more (Fig. 2). The Psuedo-$R^2$ estimate was used as an overall effect size and this indicated that the main and interaction effects in the model account for 36% of the variance in ASM scores.

## 4. Discussion

These findings confirm those from previous cross-sectional studies, which have used either MLM (Bigby et al., in press) or regression analysis (Mansell & Beadle-Brown, 2012), that show higher levels of adaptive behaviour are predictive of the quality of Active Support. A novel finding was the increase in Active Support over time, regardless of level of adaptive behaviour, with the rate of change not differing significantly across service users according to support needs. It would seem, however, that it is more challenging for staff to support people with more severe impairments. Consistent use of Active Support is positively associated with increased service user engagement in meaningful activities and relationships, which in turn is central to various domains of quality of life such as personal development, social inclusion, physical and emotional well-being. Accordingly, this study provides further evidence that service providers are failing to realise the full potential of Active Support to increase the quality of life of service users with severe intellectual disabilities (Bigby et al., in press).

The difficulty that staff have in tailoring support to individual needs was evident from our finding that greater heterogeneity amongst services users, such that there are more than two ability groups (as per ABS categories) in a service, is negatively associated with both the quality of Active Support and its rate of improvement over time. This association was also found in the cross-sectional data by (Bigby et al., in press). It may be particularly challenging for staff working in services in which service users' impairments range across the full spectrum from profound to mild. This range of support needs requires staff to switch between intensive hand-over-hand assistance to support engagement, to standing back to give time for more able people to complete tasks, or creating opportunities for them to engage in more complex tasks. These findings affirm the need to address apparent skill shortfalls among staff in tailoring Active Support to each individual service users' needs, and in tailoring Active Support practice to people with more severe or profound levels of intellectual disability identified by (Bigby et al., in press).

Findings from the present study strengthen evidence from previous studies of variables at the service level (stronger practice leadership, higher percentage of staff with Active Support training) being associated with higher levels of Active Support (Mansell et al., 2008; Bigby et al., in press). In addition, they reveal for the first time, using an observational measure of practice leadership, the association between these variables and improvements in Active Support over time, as well as its rate of change. Accordingly, the findings add to the growing body of evidence about the significance of both front-line practice leadership and staff training in Active Support to the quality of staff support. In contrast, other staffing-related variables that previously have been proposed (qualifications, experience, attitudes, satisfaction, role clarity, role conflict, and perceptions of the quality of leadership) were explored but not found to be significant predictors of the quality of Active Support over time. The implications of similar findings from the cross-sectional data were discussed by (Bigby et al., in press), who suggested the need for service delivery organisations to focus staff training on tailoring Active Support according to service users' impairment levels, and to tackle staff motivation through development of strong

front-line leadership. Due to missing data from staff about the type of training they had received, these results do not directly address whether training should be classroom or in-situ. However, given that a core element of practice leadership is in situ coaching, present study findings lend further support to previous studies that have demonstrated the advantages of a combination of classroom and in-situ Active Support training (see Flynn et al., 2018).

The negative association between service size (more than six service users) and quality of Active Support found in the cross-sectional study by (Bigby et al., in press) was also found in this longitudinal study, while further demonstrating that Active Support scores appear to increase in smaller services over time and decrease slightly over time in larger services. Combined, the studies support evidence from Flynn et al. (2018) and Tøssebro (1995) about the importance of service size to quality of life outcomes. It may be that six service users is the maximum threshold number, beyond which staff experience difficulties in providing consistent Active Support.

The finding that the mean ASM GMC scores increased at each time point after baseline, reaching a peak at the last time point support the previous finding of a positive association between the quality of Active Support and length of time since its implementation in an organisation (Bigby et al., in press). A caveat to interpreting this finding is that the rate of increase in ASM GMC scores was not uniform, but rather slower in services with higher mean ASM GMC scores at baseline. Perhaps it is unsurprising that most increase is likely for services with more scope for change. Importantly too, there was not necessarily a relationship between time points and the period over which an organisation had been implementing Active Support. Although time since Active Support had been implemented was not measured directly, there were indications that over time, staff may become more skilled in catering to diverse needs: first there was reduced variability in Active Support scores at the last two time points, and second there was a trend suggesting a reduced difference across services with more versus less variability in service user ability levels (see Fig. 3).

### *4.1. Limitations and directions for further research*

Culture, which has been repeatedly suggested as influencing staff practice (see Bigby & Beadle-Brown, 2018), was not included as a variable because when the study commenced, there were no reliable measures appropriate for the specific context of services for people with intellectual disabilities. Since that time, such a measure has been developed (Humphreys, 2018; Humphreys, Bigby, Bould, & Iacono, under review) and we recommend it for further research of the type reported here.

Organisational level factors were not explored, despite their influence on the quality of Active Support, as suggested by Bigby, Bould, and Beadle-Brown (2019) and findings by (Bigby et al., in press). Such factors include the presence of and strategies for monitoring practice quality and recruitment practices, including position descriptions, selection criteria, and induction. Inclusion of an additional level of variables requires a four-level approach, which is beyond the scope of published sources that assist researchers analysing longitudinal data using two levels (e.g., Raudenbush & Bryk, 2002; Snijders & Bosker, 2012) or three-levels (Peugh & Heck, 2017). A four or five-level model would require a sample of organisations and states significantly larger than the eight and five respectively in the present study. Nevertheless, in order to address the potential for inferential errors arising from ignoring organisational, and in turn, state level variance, the ASM scores were group mean centred at the level of the organisation.

The longitudinal design of this study was a strength given evidence about the fragile and variable nature of the quality of Active Support over time (Bigby et al., 2019; Flynn et al., 2018; Qian et al., 2017). However, although these findings revealed the factors that predicted improvements in Active Support over time, the size of the data set and the nature of the data meant they did not answer the question about factors that predict sustained good levels. The small number of services delivering good Active Support in the earlier years of the study precluded reaching a sub-sample of such services of sufficient size to include in the analysis. The fact that organisations joined the study over time was reflective of growing positive concern about how to best support service users with intellectual disabilities to live active and socially engaged lives. Slow accumulation of data over many years may provide one strategy for achieving the sample size needed to explore the complex interaction of multiple level influences on the quality of such support experienced by individual service users, and thereby account more fully for the variance in Active Support. Another strategy proposed is use of web-based shared repositories of data from international studies. Such an approach would require agreement across research groups on measures, an aim that would seem to be increasingly achievable in light of the accumulating research on Active Support.

## 5. Conclusions

These findings provide compelling evidence for the significance of practice leadership and staff training to the quality of Active Support, and the presence of these two factors are strong indicators of service quality. The implication is that service providers need to ensure that all staff receive strong practice leadership and training in Active Support. The importance of these factors to the provision of quality support will need to be carefully factored into future funding schemes for users of shared supported accommodation. Furthermore, the size of services is an important consideration, and services should support no more than six people if good levels of Active Support are to be provided and sustained over time.

## Funding

This study was funded with support from the Australian Research Council Linkage Grant LP13100189 and partner disability services.

## Declaration of Competing Interest

The authors declare that there are no conflicts of interest.

## Acknowledgements

Thanks are extended to the disability services participating in this study, and to research assistants Louise Phillips, Samuel Murray, Emma Caruana, Lincoln Humphreys, Rosa Solá Molina and Andrew Westle for support with data collection.

## 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* 293–292.

Beadle-Brown, J., Bigby, C., & Bould, E. (2015). Observing practice leadership in intellectual disability services. *Journal of Intellectual Disability Research, 59*, 1081–1093.

Beadle-Brown, J., Gifford, J., & Mansell, J. (2005). *Staff experience and satisfaction questionnaire (Learning disability).* Canterbury, UK: Tizard Centre.

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.

Bigby, C., & Beadle-Brown, J. (2018). Improving quality of life outcomes in supported accommodation for people with intellectual disability: What makes a difference? *Journal of Applied Research in Intellectual Disabilities, 31*, e182–e200.

Bigby, C., Bould, E., & Beadle-Brown, J. (2019). Implementation of active support over time in Australia. *Journal of Intellectual & Developmental Disability, 44*, 161–173.

Bigby, C., Bould, E., Iacono, T., Kavangh, S., Beadle-Brown, J. (in press). Factors that predict good Active Support in services for people with intellectual disabilities: A multilevel model. *Journal of Applied Research in Intellectual Disabilities*.

Bould, E., Beadle-Brown, J., Bigby, C., & Iacono, T. (2018). The role of practice leadership in active support: Impact of practice leaders' presence in supported accommodation services. *International Journal of Developmental Disabilities, 64*, 75–80. https://doi.org/10.1080/20473869.2016.1229524.

Browne, W. J. (2004). An illustration of the use of reparameterisation methods for improving MCMC efficiency in crossed random effects models. *Multilevel Modelling Newsletter, 16*(1), 13–25.

Browne, W. J. (2017). *MCMC estimation in MLwiN v3.00.* Centre for Multilevel Modelling, University of Bristol.

Charlton, C., Rasbash, J., Browne, W. J., Healy, M., & Cameron, B. (2017). *MLwiN version 3.02.* Centre for Multilevel Modelling, University of Bristol.

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

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*, 121–138.

Flynn, S., Totsika, V., Hastings, R. P., Hood, K., Toogood, S., & Felce, D. (2018). Effectiveness of Active Support for adults with intellectual disability in residential settings: Systematic review and meta-analysis. *Journal of Applied Research in Intellectual Disabilities, 31*, 983–998.

Hatton, C., Emerson, E., Robertson, J., Gregory, N., Kessissoglou, S., Perry, J., et al. (2001). The adaptive behavior scale-residential and community (part I): Towards the development of a short form. *Research in Developmental Disabilities, 22*, 273–288.

Hoffman, L. (2015). *Longitudinal analysis: Modeling within-person fluctuation and change.* New York, NY: Routledge.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York: Routledge.

Hox, J. J., & Roberts, J. K. (Eds.). (2011). *Handbook of advanced multilevel analysis*. New York, NY: Routledge.

Humphreys, L. (2018). *Does organisational culture matter in group homes for people with intellectual disabilities*Unpublished PhD thesis. Melbourne, Australia: La Trobe University. http://hdl.handle.net/1959.9/565789.

Humphreys, L., Bigby, C., Iacono, T, & Bould, E. (under review). Development and psychometric evaluation of the Group Home Culture Scale. *Journal of Applied Research in Intellectual and Developmental Disabilites*.

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

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.

Mansell, J., Beadle-Brown, J., Macdonald, S., & Ashman, B. (2003). Resident involvement in activity in small community homes for people with learning disabilities. *Journal of Applied Research in Intellectual Disabilities, 16*, 63–74.

Mansell, J., Beadle-Brown, J., Whelton, R., Beckett, C., & Hutchinson, A. (2008). Effect of service structure and organisation on staff care practices in small community homes for people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities, 21*, 398–413.

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

McGill, P., Vanono, L., Clover, W., Smyth, E., Cooper, V., Hopkins, L., et al. (2018). Reducing challenging behaviour of adults with intellectual disabilities in supported accommodation: A cluster randomized controlled trial of setting-wide positive behaviour support. *Research in Developmental Disabilities, 81*, 143–154.

Ockendon, J., Ashman, B., & Beadle-Brown (2014). Active support – Fundamental to positive behaviour support. *Research and Practice in Intellectual and Developmental Disabilites, 1*, 98–107.

Peugh, J. L., & Heck, R. H. (2017). Conducting three-level longitudinal analyses. *The Journal of Early Adolescence, 37*, 7–58.

Qian, X., Tichá, R., & Stancliffe, R. J. (2017). Contextual factors associated with implementing active support in community group homes in the United States: a qualitative investigation. *Journal of Policy and Practice in Intellectual Disabilities, 14*, 332–340. https://doi.org/10.1111/jppi.12204.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, California: Sage.

Rodriguez, G. (2007). Multilevel generalized linear models. Chapter 9. In J. De Leeuw, & E. Meijer (Eds.). *Handbook of multilevel analysis* (pp. 335–376). New York: Springer.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* London: Sage.

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* London, England: SAGE.

Spiegelhalter, D. J., Best, N. G., Garlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion and rejoinder). *Journal of the Royal Statistical Society Series B, 64*, 583–639.

Tøssebro, J. (1995). Impact of size revisited: Relation of number of residents to self-determination and deprivatization. *American Journal on Mental Retardation, 100*, 59–67.