

Measuring palliative care self-efficacy of intellectual and developmental disability staff using Rasch models

Original Article

Cite this article: Kim J, Gray JA (2024). Measuring palliative care self-efficacy of intellectual and developmental disability staff using Rasch models. *Palliative and Supportive Care* **22**, 146–154. <https://doi.org/10.1017/S1478951522001833>

Received: 27 June 2022
Revised: 31 October 2022
Accepted: 23 December 2022

Keywords:

Palliative care self-efficacy; Training; Rasch analysis; Intellectual and developmental disability; Instrument

Author for correspondence: Jinsook Kim, School of Health Studies, Northern Illinois University, Wirtz Hall 254, DeKalb, IL 60115, USA. Email: jkim4@niu.edu

Jinsook Kim, PH.D., M.P.H., D.D.S.  and Jennifer A. Gray, PH.D., M.P.P.

School of Health Studies, Northern Illinois University, DeKalb, IL, USA

Abstract

Objectives. The objectives of this study were to evaluate the psychometric properties of a palliative care self-efficacy instrument developed for intellectual and developmental disability (IDD) staff using Rasch analysis and assess the change in palliative care self-efficacy between 2 time points using Rasch analysis of stacked data.

Methods. Staff from 4 nonprofit IDD services organizations in a US Midwestern state ($n = 98$) answered 11 questions with Likert-style responses at baseline and 1-month follow-up post training. Rasch analysis was performed to examine rating scale structure, unidimensionality, local independence, overall model fit, person and item reliability and separation, targeting, individual item and personal fit, differential item functioning (DIF), and change in palliative care self-efficacy between 2 time points.

Results. The rating scale structure improved when 5 response categories were collapsed to 3. With the revised 3 response categories, the instrument demonstrated good psychometric properties. Principal components analysis of Rasch residuals supported the assumption of unidimensionality. Model fit statistics indicated an excellent fit of the data to the Rasch model. The instrument demonstrated good person and item reliability and separation. Gender-related DIF was found in 1 item, and work tenure-related DIF in 3 items. Overall palliative care self-efficacy improved between 2 time points.

Significance of results. Rasch analysis allowed for a more thorough examination of this palliative care self-efficacy instrument than classical test theory and provided information on rating scale structure, targeting, DIF, and individual persons and items. These recommendations can improve this instrument for research and practical contexts.

Introduction

Palliative care is a holistic model of care that focuses on the alleviation of suffering and improving quality of life by meeting the physical, social, psychological, cultural, and spiritual needs of people with chronic and life-limiting illnesses and their families (World Health Organization 2015). Palliative care involves multiple aspects including communication, end-of-life (EOL) decisions, care before and after death, information about social services, and cultural competency (Cha et al. 2020; Hahn and Cadogan 2011). Self-efficacy is defined as “people’s judgments of their capabilities to organize and execute courses of action required to attain designated types of performances” (Bandura 1986, 391). Since self-efficacy is a domain-specific concept, palliative care self-efficacy refers to an individual’s belief in his or her own ability to be competent in providing palliative care (Adriaansen et al. 2005).

Individuals with intellectual and developmental disability (IDD) demonstrate limitations in cognitive functioning and practical life functioning skills (including activities of daily living, occupational skills, money management, and social skills) relatively early in life (American Association of Intellectual and Developmental Disabilities 2023). As individuals with IDD are living longer (Coppus 2013; World Health Organization and The World Bank 2011), more are experiencing chronic health conditions, and the potential need for palliative care services increases (García-Domínguez et al. 2020; Kinnear et al. 2018).

Staff serving people with IDD have demonstrated gaps in knowledge and efficacy in providing palliative care (Cartlidge and Read 2010; Fahey-McCarthy et al. 2009; Ng and Li 2003; Ryan et al. 2010). There is a need for staff to develop a wide range of skill sets, encompassing communication, recognizing spiritual, social, and emotional needs and cultural competence, understanding legal issues, and bereavement care (Friedman et al. 2008; Lord et al. 2017; Reynolds et al. 2008; Stein 2008; Watters et al. 2012; Wiese et al. 2014). To our knowledge, no instruments exist that measure IDD staff’s self-efficacy in provision of palliative care. A variety of instruments have been used to measure the self-efficacy and knowledge of health professionals providing palliative care (Frey et al. 2011; Friesen and Andersen 2019; Karacsony et al. 2015). These include instruments for physicians

(Fox 2007; Nakazawa et al. 2009), medical students and residents (Billings et al. 2009; Buss et al. 2005; Mulder et al. 2009), nurses (Desbiens and Fillon 2011; Dobie et al. 2016; Edwards et al. 2020; Lange et al. 2011; Moura Minosso et al. 2017; Nakazawa et al. 2009; Phillips et al. 2011; Shipman et al. 2008; Wilkinson et al. 2008), care assistants (Dryden and Addicott 2009; Fox 2007; Jenkins et al. 2010; Phillips et al. 2011; Resnick 2008), and other health professionals (Jenkins et al. 2010; Lange et al. 2011).

Palliative care self-efficacy instrument for intellectual and developmental disability staff

We created a palliative care self-efficacy instrument to assess the level of confidence of IDD staff regarding palliative care provision. We call it Palliative Care Self-Efficacy Instrument for Intellectual and Developmental Disability Staff (PCSE-IDD). The instrument was used in evaluating the effect of an online training program on palliative care self-efficacy of IDD staff but can be used separately as well to gauge workers' palliative care self-efficacy cross-sectionally. The online palliative care training developed by the authors was delivered using Internet-connected tablets. The 2-hour training included modules on the overview of palliative care, legal and ethical issues, cultural diversity and competency, communication with people with IDD, symptom management, EOL care and logistics after death, bereavement and grief of people with IDD, and staff grief and coping strategies (Kim and Gray 2021). Training content and the palliative care self-efficacy instrument items were based on literature from the IDD and non-IDD fields (Adriaansen et al. 2005; Bekkema et al. 2015; Desbiens and Fillon 2011; Fahey-McCarthy et al. 2009; Fox 2007; Hahn and Cadogan 2011; Kirkendall and Waldrop 2013; McCarron et al. 2010; McEvoy et al. 2012; Nakazawa et al. 2009; Ng and Li 2003; Phillips et al. 2011; Ryan et al. 2011a, 2011b, 2010; Stein 2008; Tuffrey-Wijne et al. 2007; Wark et al. 2014; Wittenberg-Lyles et al. 2014) as well as the results of need assessments with IDD staff (Gray and Kim 2020; Kim and Gray 2018). The effect of the training was evaluated using one-group pretest–posttest design (Kim and Gray 2021).

Palliative care self-efficacy was measured using a question “How confident are you regarding the following?” for 11 items with 5 Likert-style responses of *not at all* (1), *slightly* (2), *somewhat* (3), *moderately* (4), and *a lot* (5). The self-efficacy questions reflect the subjective nature of the concept, which relies on the perception of the participants. The 11 items addressed response to ethical problems, advance directives, communicating with people with different background, informing bad news to people with IDD, recognizing pain, nonmedical pain management, patient care before death, post-death tasks, identifying grieving behavior of people with IDD, helping people with IDD recover after a loss, and managing own grief (Table 1).

Rasch modeling

The Rasch model is a psychometric model that shows what should be expected in response to test or questionnaire items if the outcome measure from the test or questionnaire is an interval measure (Tennant and Conaghan 2007). The Rasch model incorporates a method for ordering persons (e.g., from a sample of IDD staff) according to their ability (e.g., where they stand in terms of palliative care self-efficacy) and ordering items according to their difficulty (e.g., which level of palliative care self-efficacy each item represents) (Bond et al. 2021, 11).

Table 1. Items in the palliative care self-efficacy instrument

Item no.	Item text
1	How to respond to ethical problems
2	Utilizing advance directives
3	Communicating with people with different background (e.g., race, ethnicity, religion, and LGBT)
4	Telling people with IDD about bad news (e.g., serious illness, death, and loss)
5	Recognizing different types of pain
6	Reducing pain with nonmedical techniques
7	Caring for patients before they die
8	What to do when a client dies
9	Identifying behaviors of people with IDD who are grieving
10	Helping people with IDD recover after a loss
11	Managing your own experience of grief

Rasch modeling supplements classical test theory and overcome limitations of classical test theory approaches (Tennant and Conaghan 2007). First, using Rasch models, individual persons and items can be examined while classical test theory focuses on the summary of items (e.g., sum or average of response scores) (Anselmi et al. 2015). For example, when a paired *t*-test or effect size statistics are used to measure change by an intervention, only the overall change of the sample can be assessed. Information at the individual level, such as who responded to the intervention better, is not available in classical test theory (Anselmi et al. 2015). Second, Rasch analysis enables the transformation of a raw ordinal score into a linear, interval-level variable if ordinal data fit to Rasch model expectations (Chang et al. 2014). It is problematic to treat raw scores and values assigned to ordinal response categories as interval data without validating that the measure is interval (Grimby et al. 2012). The linear transformation in the Rasch model allows for valid use of mathematical operations and parametric analysis and provides a truer depiction of a trait (Chang et al. 2014). Third, in measuring change over time, Rasch models ensure the invariance of the instrument across time points (Bond et al. 2021, 203). Additionally, Rasch models provide a variety of psychometric tools and information that enable examination of key properties of composite measures including dimensionality, rating scale structure, differential item functioning (DIF), person and item separation and reliability, and targeting (Tennant and Conaghan 2007).

Rasch models to measure change

Rasch models provide solutions to challenges in measuring change. The primary interest of evaluation research is to determine whether an intervention improved persons' ability. Persons are expected to show changes between Time 1 and Time 2 because of an intervention. The challenge is how to measure persons and items in a common frame of reference encompassing different time points so that the measurement of change has an unambiguous numerical representation and a substantive meaning (Anselmi et al. 2015). This can be accomplished by estimating a Rasch model on stacked data. Stacked analysis produces a set of item measures that are consistent at both time points and a person measure for each individual

	Item 1	Item 2	Item 3	...
Data of Person 1 at Time 1				
Data of Person 2 at Time 1				
Data of Person 3 at Time 1				
...				
Data of Person 1 at Time 2				
Data of Person 2 at Time 2				
Data of Person 3 at Time 2				
...				

Fig. 1. Conceptual visualization of stacked data.

at each time point (Anselmi et al. 2015). Stacking data is done by appending the observations for the post intervention below the baseline (before intervention) observations (Figure 1). As a result, the data file contains twice as many cases as the persons being measured. In addition to measuring persons and items in a common frame of reference encompassing 2 time points, stacked analysis allows for tracking differential impacts of an intervention on persons (Wright 2003).

This study aimed to evaluate the psychometric properties of a newly created palliative care self-efficacy instrument developed for staff who care for people with IDD using Rasch analysis and assess the change in palliative care self-efficacy between 2 time points using Rasch analysis of stacked data.

Methods

Participants and procedures

Participants were recruited from 4 nonprofit community-based organizations serving people with IDD in a US Midwestern state. Informed consent was obtained before conducting the training and associated evaluations. The university institutional review board reviewed the research protocol, including the consent process (ORC #HS17-0189). The majority of the participants were female (89%) and White (58%) or African American (35%). Approximately 37% were in rural areas. On average, participants were 39 years old (20–67) and worked for 10 years in the field (1–30).

Before taking the training, participants answered questions about palliative care knowledge and palliative care self-efficacy as well as demographic and work-related information. Participants took a posttest about palliative care knowledge upon completing the training. One month after the training, participants were invited via email to complete an online follow-up survey that included items on palliative care self-efficacy. Palliative care self-efficacy was measured 1 month post training instead of upon completing the training to allow participants to have time to apply what they have learned in the training to their work for a month. Among 132 training participants, 98 answered questions on palliative care self-efficacy at both baseline and 1-month follow-up. Participants who completed both the baseline and follow-up survey ($n = 98$) had a higher education level than those who completed the baseline survey only ($n = 34$) ($p < 0.005$, Fisher's exact test), although 2 groups were similar in other characteristics.

Analysis

This study used baseline ($n = 98$) and follow-up data ($n = 98$) on palliative care self-efficacy stacked together. The sample size 98 was enough for Rasch modeling. For stable results, the sample size needs to be at least 6 multiplied by the number of items (Mundfrom

et al. 2005), which would be a minimum of 66 persons, considering that our instrument comprised 11 items.

Data were analyzed using WINSTEPS computer software version 4.8.0.0 (Linacre 2021b). The rating scale model was used because all items had the same number of response categories and category values (Bond et al. 2021, 97). Rasch analysis was performed to examine rating scale structure, unidimensionality, local independence, overall model fit, person and item reliability and separation, targeting, individual item and personal fit, DIF, and change in palliative care self-efficacy between 2 time points.

Before the main analysis, local dependency across time points was investigated to determine whether use of stacked analysis was appropriate. If the person measures estimated in the stacked analysis with anchors are not significantly different from those estimated in the stacked analysis, the effect of local dependency is considered negligible (Anselmi et al. 2015). In the stacked analysis with anchors, the data from Time 2 were analyzed first to obtain Rasch measures. Then, the data from Time 1 were analyzed by anchoring the item measures to the estimates from the analysis with Time 2 data. There was no evidence of significant local dependency according to t statistics. The t statistic comparing the person measures obtained in the stacked analysis with and without anchors ($df = 20$ by $df = 2k - 2$, where k is the number of items) was not significant for any person. When there is no evidence of significant local dependency, measures from either approach (i.e., anchored or nonanchored) can be used (Anselmi et al. 2015). We used the stacked analysis without anchors in this study.

Results

Rating scale structure

Analysis of the instrument with original 5 response categories indicated that the response category structure was not optimal. The instrument did not meet the criterion of having at least 10 observations per response category (Linacre 2002). There were fewer than 10 observations per response category in 10 of 11 items specifically in category 1 or 2. As presented in the left section of Table 2, the spacing between Andrich thresholds of categories 1 and 2 (first threshold) and categories 2 and 3 (second threshold) was only 0.45 logits, which was less than the minimum of 1.40 logits (Bond et al. 2021, 226). It indicated that categories 2 and 3 were too similar or close to be separate categories. The observed average measures and thresholds were in linear order, and fit statistics were within the acceptable range between 0.60 and 1.40.

Visual examination suggested similar pictures. Figure 2 displays the category characteristic curves for the instrument with 5 response categories (Figure 2a) and 3 response categories (Figure 2b). The curves represent the probability of respondents selecting any particular category for the difference between any person ability and item difficulty estimates (Bond et al. 2021, 226). The x axis represents person ability levels relative to the item's difficulty in logits (Bond et al. 2021, 112). The y axis represents the expected probability of selecting any given category. For the original 5 response categories (Figure 2a), the curves for category 2 and category 3 did not display a distinct peak. The highest probability of category 2 and category 3 was less than 0.5, which is the minimum probability peak value for a category to be considered functional (Bond et al. 2021, 112). It indicated that 2 adjacent categories (category 2 and category 3) represented too narrow intervals on the

Table 2. Rating scale structure of instruments with 5 and 3 response categories ($n = 196$)

5 response categories					3 response categories				
Category	Observed average	Infit MNSQ	Outfit MNSQ	Andrich threshold	Category	Observed average	Infit MNSQ	Outfit MNSQ	Andrich Threshold
1	-0.45	1.16	1.31	None	1	-1.32	0.94	0.97	None
2	-0.20	0.93	0.94	-1.61	2	0.00	0.97	1.00	-1.10
3	0.44	0.85	0.85	-1.16	3	1.22	1.07	1.07	1.10
4	1.49	0.87	0.99	0.29					
5	2.52	1.13	1.08	2.48					

MNSQ, mean square residual statistic.

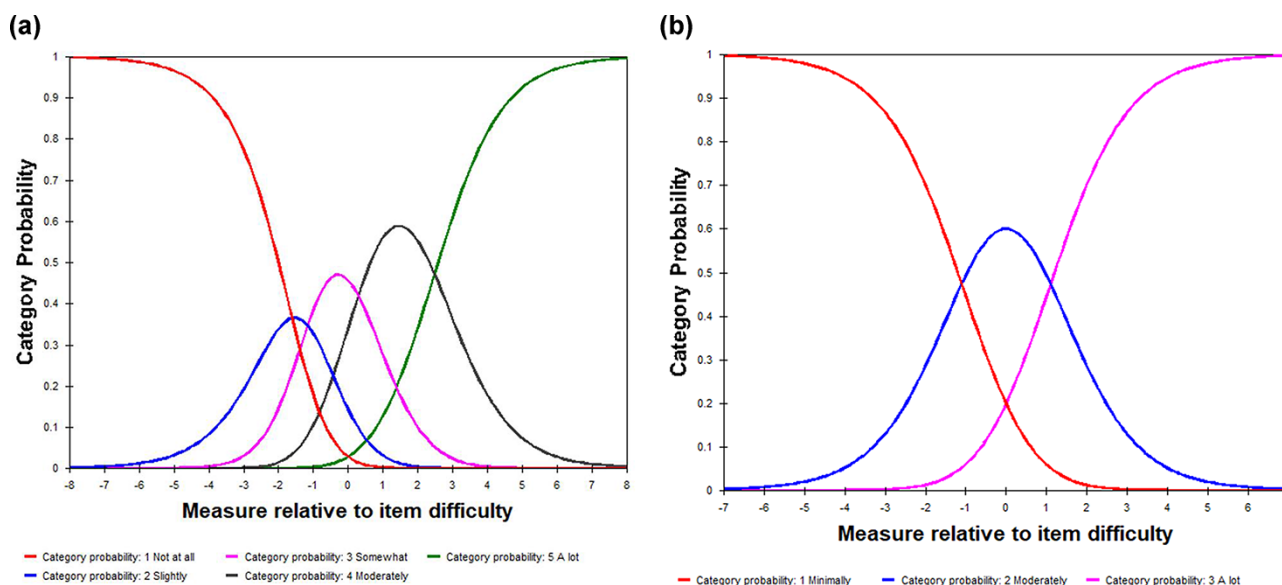


Fig. 2. Category characteristic curves. (a) Instrument with 5 response categories. (b) Instrument with 3 response categories.

latent trait, and one of the response categories was rarely chosen by respondents (Linacre 2002). These results together pointed to the need to collapse categories 1 to 3 into one category.

The revised instrument with 3 response categories (minimally, moderately, and a lot) indicated improvement to the original instrument. All 3 criteria for rating scale functioning (Linacre 2002) were met: (1) there were more than 10 observations per rating category, (2) measures increased linearly with each category, and (3) mean square residual statistics were between 0.60 and 1.40. The right section of Table 2 presents the structure calibrations for a revised instrument with 3 response categories. Both observed average measures and thresholds were in linear order, and the infit and outfit mean square residual statistics were between 0.60 and 1.40. The spacing between thresholds was greater than 1.40 logits. Figure 2b demonstrates an alternative where original response categories 1, 2, and 3 have been collapsed into one, resulting in 3 response categories. The category characteristic curves (Figure 2b) look like a range of hills with visible peaks for each response category, which is expected in a scale with ordered thresholds and adequate spacing between thresholds (Bond et al. 2021, 226; Linacre 2021a, 375). Since the instrument with 3 response categories showed an optimal response category structure, results from this point will be about analysis of the instrument with 3 response categories.

Unidimensionality, local independence, and model fit

Principal components analysis of Rasch residuals indicated that the Rasch dimension explained 42.3% of the total variance with an eigenvalue of 8.06. The variance in the first contrast had an eigenvalue of 1.97, suggesting that the residuals were random noise, given it was less than 2.0 (Bond et al. 2021, 255; Linacre 2021a, 421; Raiche 2005). There was no meaningful pattern in the residual contrast plot either. The variance of the first contrast accounted for 10.3% of the total variance, which was much smaller than the variance explained by item difficulties (19.7%). Standardized residual correlations of a pair of items were small, ranging from 0.18 to 0.41, indicating negligible local dependency (Linacre 2021a, 436). These results supported the assumption of unidimensionality and local independence. Model fit statistics indicated an excellent fit of the data to the Rasch model. Infit and outfit mean square residuals were 1.00 and 1.01, with small model standard error 0.13. Standardized Z scores were -0.10 and 0.00 for infit and outfit, respectively.

Person and item reliability and separation

The instrument demonstrated good person and item reliability and separation. The Rasch person reliability (equivalent to Cronbach's alpha) was 0.83, and the item reliability statistic 0.95 indicated

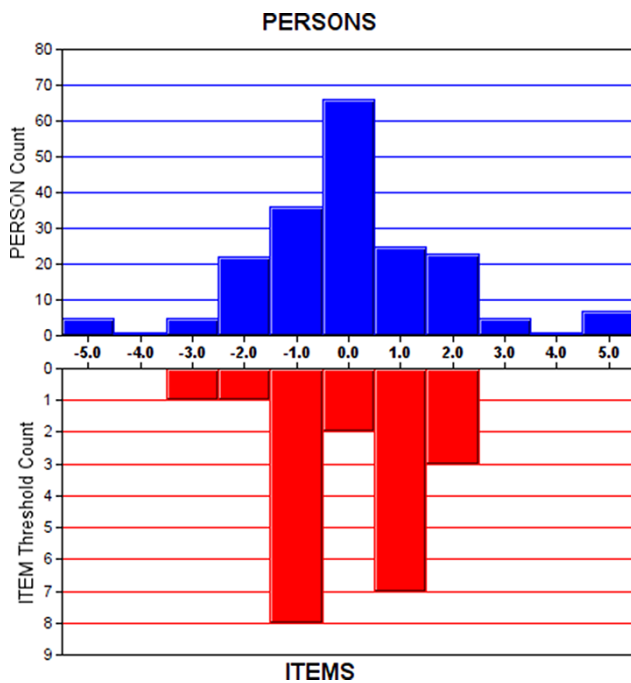


Fig. 3. Person-item threshold map.

Note: The blue bars illustrate the distribution of person locations, and the red bars illustrate the distribution of item Andrich threshold locations.

a wide range of item difficulties (Bond et al. 2021, 41). Person and item separation indices suggested adequate spread of persons and items across the trait continuum. The person separation index was 2.17, and the item separation index was 4.49. The number of statistically significant strata was 2.23 for persons and 6.32 for items, indicating that persons in the sample could be separated into approximately 2 distinct groups, and items in the instrument into approximately 6 distinct groups of trait levels (Wright and Masters 1982, 92).

Targeting

The mean person location was zero logits, indicating a well-targeted measure in terms of mean (Tennant and Conaghan 2007). The person-item threshold map, however, suggested ceiling and floor effects. Figure 3 is the person-item threshold map that illustrates the distribution of person locations and items' Andrich threshold locations. The upper chart in blue shows the distribution of person locations along the construct (x axis). The lower chart in red shows Andrich thresholds of the items along the same x axis. Though person locations extended from -5.0 to 5.0 logits, there were gaps at the top and bottom of the construct hierarchy in terms of item threshold locations. The ceiling effect was more prominent than the floor effect as the gap between the highest person location (the positive end of the x axis) and the highest threshold location was larger than the gap between the lowest person location (the negative end of the x axis) and the lowest threshold location.

Item and person fit

All individual items and most persons showed good fit. Table 3 provides information on item locations in misfit order. Infit and outfit

Table 3. Item location statistics in misfit order of the 3-category responses ($n = 196$)

Item no.	Item location ^a	Model SE	Infit		Outfit	
			MNSQ	ZSTD	MNSQ	ZSTD
3	-1.50	0.14	1.17	1.59	1.36	2.38
7	0.58	0.13	1.32	3.12	1.25	2.25
1	-0.34	0.13	1.18	1.86	1.24	2.28
8	0.48	0.13	1.17	1.72	1.14	1.35
11	-0.08	0.13	1.04	0.48	1.08	0.80
2	0.16	0.13	0.96	-0.43	0.99	-0.07
6	0.94	0.13	0.91	-0.89	0.90	-0.90
4	-0.29	0.13	0.86	-1.53	0.86	-1.43
5	-0.43	0.13	0.84	-1.79	0.86	-1.38
9	0.11	0.13	0.78	-2.52	0.75	-2.69
10	0.36	0.13	0.75	-2.90	0.73	-2.93
Mean	0.00	0.13	1.00	-0.10	1.01	0.00
SD	0.62	0.00	0.18	1.9	0.20	1.9

^aItem location in logits.

MNSQ, mean square residual statistic; ZSTD, standardized Z-score.

mean square residual statistics of all 11 items were within fit criteria ($0.6 < \text{MNSQ} < 1.40$) (Wright and Linacre 1994). For person fit, 5 and 4 out of 98 persons had infit and/or outfit mean square residual statistics greater than 2 at Time 1 and Time 2, respectively (results not shown). Mean square residual statistics values greater than 2 for a particular individual may indicate that the person is not a typical member of the population or that the person answered the questionnaire inaccurately (Anselmi et al. 2015; Wright and Linacre 1994). Summary fit statistics of 184 non-extreme persons were in acceptable ranges (infit MNSQ 1.00 and outfit MNSQ 1.01) (results not shown).

Differential item functioning

One item (managing your own experience of grief) had DIF contrast 0.93 ($p < 0.05$) between genders. It suggested that women tended to endorse 0.93 logit higher on the item compared to men, when their overall palliative care self-efficacy ability levels were taken into account. Three items (communicating with people with different background, recognizing different types of pain, and reducing pain with nonmedical techniques) had DIF contrast 0.84 ($p < 0.05$), 0.54 ($p < 0.05$), and 0.73 ($p < 0.005$), respectively, between tenure groups. It indicated that IDD staff who worked longer in the field tended to endorse higher on the respective items compared to those who worked shorter in the field when their overall palliative care self-efficacy ability levels were considered.

Change over time

There was a significant improvement in palliative care self-efficacy among 98 training participants overall. The mean Rasch measure increased from -0.40 to 0.39 logits between Time 1 and Time 2 ($p < 0.005$ in a t -test). The effect size was between medium and large indicated by Cohen's d 6.15 (Maher et al. 2013). The scatterplot comparing item difficulties between 2 time points indicated

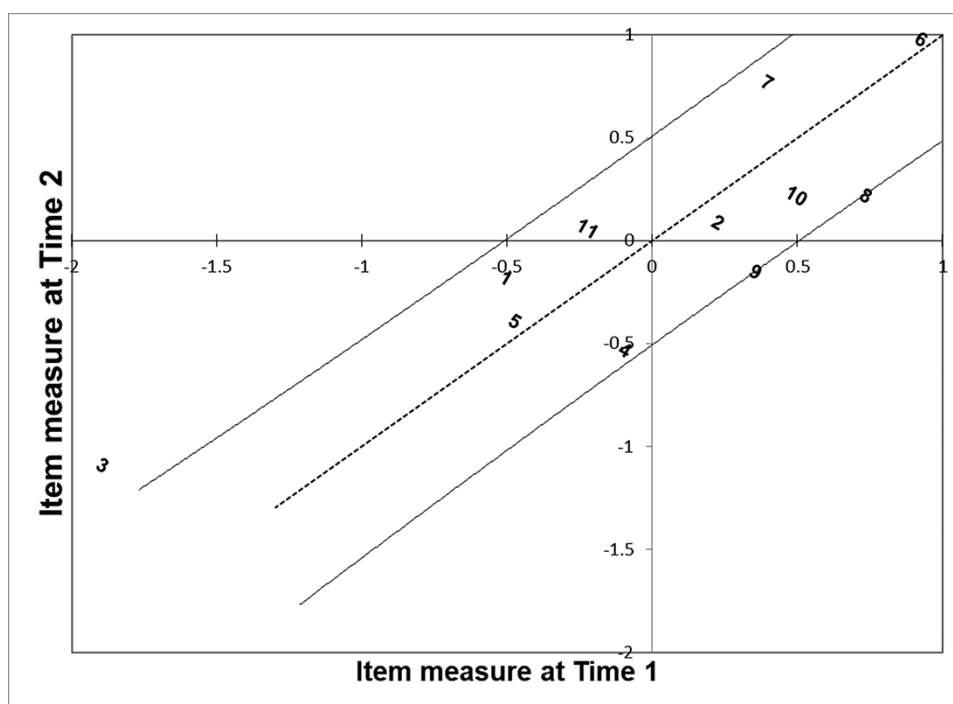


Fig. 4. Item measures at Time 2 (y axis) plotted against those at Time 1 (x axis).

Note: The dotted line is the identity line $x = y$. The solid lines next to the dotted line indicate 95% confidence intervals.

1 item far from the identity line $x = y$ (Figure 4). Item 3 was significantly more difficult at Time 2 (-1.09) than at Time 1 (-1.89), given item 3 is above the solid line signifying the upper limit of a 95% confidence interval. According to the scatterplot comparing person abilities between 2 time points (indicated by id numbers), 12 out of 98 individuals demonstrated a significant improvement in palliative care self-efficacy, while palliative care self-efficacy of 9 individuals decreased (Figure 5). A moderate correlation between the measures at 2 time points ($r = 0.64$) indicated the variation in the intervention's effect on participants. Individuals above the solid curved line over the identity line (the upper limit of a 95% confidence interval) are those who showed a significant improvement, while the individuals below the solid curved line under the identity line (the lower limit of a 95% confidence interval) are those whose palliative care self-efficacy decreased between Time 1 and Time 2.

The solid curved lines next to the dotted line indicate 95% confidence intervals.

Discussion

This study examined the psychometric properties of a palliative care self-efficacy instrument developed for staff who care for people with IDD and the change in palliative care self-efficacy between 2 time points using Rasch analysis. The analysis provided detailed information on how well the instrument items could measure palliative care self-efficacy, the positive and less desirable elements of the instrument, and the change in palliative care self-efficacy in terms of items and persons.

The analysis demonstrated that overall, the instrument performed adequately as a measure of palliative care self-efficacy of IDD staff. Principal components analysis of the Rasch residuals indicated unidimensionality of the instrument, and standardized

residual correlations of item pairs demonstrated negligible local dependency. Overall fit statistics indicated an excellent fit of the data. Person and item reliability was good, and person and item separation indices suggested adequate spread of persons and items across the trait continuum. All individual items and most persons in the sample demonstrated good fit. Fit statistics of all 11 items were within the adequate fit range. Approximately 5% of the sample had person mean square residual statistics greater than 2, indicating that they answered the questionnaire inaccurately or might not be a typical member of the population (Anselmi et al. 2015; Wright and Linacre 1994).

The results, however, indicated issues in rating scale structure, targeting, and DIF. Response patterns to the original 5 response categories were not optimal. Although the observed average measures and thresholds were in linear order and fit statistics were within the acceptable range, some response categories had fewer than 10 observations and the distance between 2 thresholds was less than the acceptable range. Once the response categories were collapsed to 3 categories, the instrument became effective as a rating scale measure.

The ceiling and floor effects indicate the need for improvement. By adding items of more difficulties and less difficulties, the instrument will be able to cover the whole range of person abilities shown in this sample. More specific items regarding "reducing pain with nonmedical techniques," "caring for patients before they die," and "what to do when a client dies" may fill the gap at the top of the construct hierarchy, and more specific items regarding "communicating with people with different background" and "recognizing different kinds of pain" may fill the gap at the bottom of the construct hierarchy.

Gender-related DIF was found in 1 item and work tenure-related DIF was found in 3 items. Women outperforming men with a comparable level of palliative care self-efficacy in managing

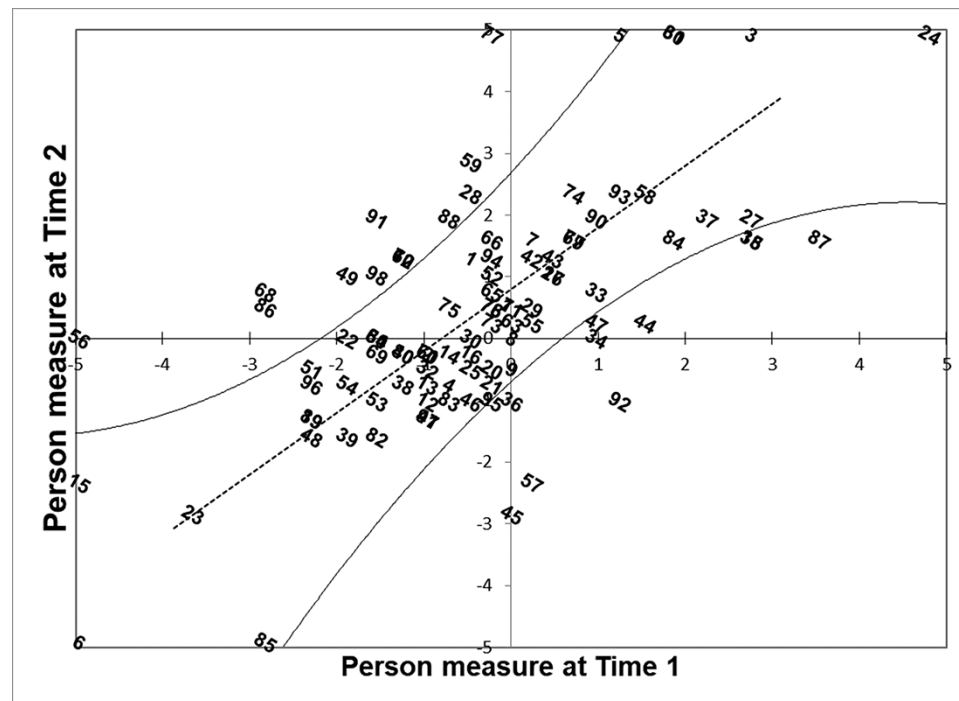


Fig. 5. Person measures at Time 2 (y axis) plotted against those at Time 1 (x axis). Note: The dotted line is the identity line $x = y$.

own experience of grief may be related to gender differences in social support and coping styles, which may affect how men and women approach grief and bereavement (Stroebe et al. 2001). IDD staff who worked longer in the field may have outperformed those with a comparable level of palliative care self-efficacy but with shorter tenure in communicating with people with different background, recognizing different types of pain, and reducing pain with nonmedical techniques, as the former group has gained more work experience regarding these topics. DIF found in this sample, however, informative, should not determine whether the items should be modified or eliminated from the instrument at this point (Du and Yates 1995). Since a significance test with 1 sample cannot differentiate between real DIF and an accident, whether the items continue to exhibit DIF in the same way should be investigated with additional samples (Du and Yates 1995), as is standard statistical analysis protocol.

By estimating a Rasch model on stacked data, persons and items were measured in a common frame of reference encompassing different time points. The stacked analysis provided the information not only on the overall change in palliative care self-efficacy but also on the changes in individual item difficulties and person abilities between the 2 time points. The fact that item 3 was significantly more difficult at Time 2 than at Time 1 may suggest the unexpected impact of the training. The participants may have overestimated their ability to communicate with people from different backgrounds at baseline, then reevaluated it after learning more about communication and cultural diversity in the training. There was variation in the effect of the intervention on different participants, among whom palliative care self-efficacy improved, stayed the same, and worsened through the training.

A few limitations of this study must be considered. Participants were recruited from nonprofit IDD service agencies in a US Midwestern state, which prevents us from generalizing our results

to the entire IDD staff in the US. Our sample, however, was diverse in terms of age, race, length of work experience, and rural-urban/suburban location. While a sample size of 98 participants at both time points was sufficient for Rasch analysis, a larger sample size would have been preferable.

Conclusions

Rasch analysis was a useful tool for measuring the psychometric properties of the PCSE-IDD. The use of Rasch analysis has many advantages over other analysis methods, including generalizability across samples and items, true transformation of ordinal to interval-level variables, distinguishing unusual respondents and items that are not working, ensuring the invariance of the instrument across time points, and enabling examination of dimensionality, rating scale structure, and DIF (Anselmi et al. 2015; Bond et al. 2021, 203; Tennant and Conaghan 2007). This instrument is unique in that it is targeted for IDD staff and it covers self-efficacy about palliative care provision for people with IDD as well as about self-care of IDD staff. The results of this study point to the need to improve the instrument by using 3 response categories (such as, minimally, moderately, and a lot), adding more items of higher- and lower-level difficulties and testing the modified instrument with other samples.

Acknowledgments. The authors thank study participants for their contributions.

Funding. This work was supported by the Broadband Innovation Grant and Tri-County Endowment Award of Northern Illinois University.

Conflicts of interest. The authors declare none.

References

- Adriaansen MJ, Van Achterberg T and Borm G (2005) Effects of a postqualification course in palliative care. *Journal of Advanced Nursing* **49**(1), 96–103. doi:10.1111/j.1365-2648.2004.03268.x
- American Association of Intellectual and Developmental Disabilities (2023) Defining criteria for intellectual disability. <https://www.aaidd.org/intellectual-disability/definition> (accessed 11 January 2023).
- Anselmi P, Vidotto G, Bettinardi O, et al. (2015) Measurement of change in health status with Rasch models. *Health and Quality of Life Outcomes* **13**, 16. doi:10.1186/s12955-014-0197-x
- Bandura A (1986) *Social Foundations of Thought and Action: A Social Cognitive Theory*. Upper Saddle River, NJ: Prentice-Hall.
- Bekkema N, de Veer AJ, Hertogh CM, et al. (2015) “From activating towards caring”: Shifts in care approaches at the end of life of people with intellectual disabilities; a qualitative study of the perspectives of relatives, care-staff and physicians. *BMC Palliative Care* **14**, 33. doi:10.1186/s12904-015-0030-2
- Billings ME, Randall J, Engelberg C, et al. (2009) Medicine residents’ self-perceived competence in end-of-life care. *Academy Medicine* **84**(11), 1533–1539. doi:10.1097/ACM.0b013e3181bbb490
- Bond T, Yan Z and Heene M (2021) *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*, 4th edn. Milton Park, Oxfordshire: Taylor and Francis.
- Buss MK, Alexander C, Switzer GE, et al. (2005) Assessing competence of residents to discuss end-of-life issues. *Journal of Palliative Medicine* **8**(2), 363–371. doi:10.1089/jpm.2005.8.363
- Cartlidge D and Read S (2010) Exploring the needs of hospice staff supporting people with an intellectual disability: A UK perspective. *International Journal of Palliative Nursing* **16**(2), 93–99. doi:10.12968/ijpn.2010.16.2.46755
- Cha E, Lee S, Lee J, et al. (2020) Health personnel’s knowledge, attitudes, and self-efficacy related to providing palliative care in persons with chronic diseases. *The Korean Journal of Hospice and Palliative Care* **23**(4), 198–211. doi:10.14475/kjhpc.2020.23.4.198
- Chang KC, Wang JD, Tang HP, et al. (2014) Psychometric evaluation, using Rasch analysis, of the WHOQOL-BREF in heroin-dependent people undergoing methadone maintenance treatment: Further item validation. *Health and Quality of Life Outcomes* **12**, 148. doi:10.1186/s12955-014-0148-6
- Coppus A. M. W. (2013) People with intellectual disability: What do we know about adulthood and life expectancy?. *Developmental Disabilities Research Reviews* **18**(1), 6–16. doi:10.1002/ddrr.1123
- Desbiens J and Fillon L (2011) Development of the palliative care nursing self-competence scale. *Journal of Hospice & Palliative Nursing* **13**(4), 230–241. doi:10.1097/NJH.0b013e318213d300
- Dobie J, Plumb M and Shepherd S (2016) End of life care education to care home staff: An evaluation. *Nursing and Residential Care* **18**(7), 369–374. doi:10.1186/s12904-016-0123-6
- Dryden H and Addicott R (2009) Evaluation of a pilot study day for health-care assistants and social care officers. *International Journal of Palliative Care Nursing* **15**(1), 6–11. doi:10.12968/ijpn.2009.15.1.37946
- Du Y and Yates F (1995) When to adjust for differential item functioning. *Rasch Measurement Transactions* **9**(1), 414.
- Edwards C, Hardin-Pierce M, Anderson D, et al. (2020) Evaluation of self-efficacy and confidence levels among newly graduated nurses exposed to an end-of-life simulation: A comparison study. *Journal of Hospice & Palliative Nursing* **22**(6), 504–511. doi:10.1097/NJH.0000000000000698
- Fahey-McCarthy E, McCarron M, Connaire K, et al. (2009) Developing an education intervention for staff supporting persons with an intellectual disability and advanced dementia. *Journal of Policy and Practice in Intellectual Disabilities* **6**(4), 267–275. doi:10.1111/j.1741-1130.2009.00231.x
- Fox CR (2007) What health care providers know (and need to know) about palliative care. *Journal of Allied Health* **36**(4), 209–215.
- Frey R, Gott M and Banfield R (2011) What indicators are measured by tools designed to address palliative care competence among “generalist” palliative care providers? A critical literature review. *Progress in Palliative Care* **19**(3), 114–124. doi:10.1179/1743291X11Y.0000000003
- Friedman SL, Choueiri R and Gilmore D (2008) Staff carers’ understanding of end of life care. *Journal of Policy and Practice in Intellectual Disabilities* **5**(1), 56–64. doi:10.1111/j.1741-1130.2007.00142.x
- Friesen L and Andersen E (2019) Outcomes of collaborative and interdisciplinary palliative education for health care assistants: A qualitative metasummary. *Journal of Nursing Management* **27**(3), 461–481. doi:10.1111/jonm.12714
- García-Domínguez L, Navas P, Verdugo M, et al. (2020) Chronic health conditions in aging individuals with intellectual disabilities. *International Journal of Environmental Research and Public Health* **17**(9), 3126. doi:10.3390/ijerph17093126
- Gray J and Kim J (2020) Palliative care needs of direct care workers caring for people with intellectual and developmental disabilities. *British Journal of Learning Disability* **48**(1), 69–77. doi:10.1111/bld.12318
- Grimby G, Tennant A and Tesio L (2012) The use of raw scores from ordinal scales: Time to end malpractice? *Journal of Rehabilitation Medicine* **44**(2), 97–98. doi:10.2340/16501977-0938
- Hahn JE and Cadogan MP (2011) Development and evaluation of a staff training program on palliative care for persons with intellectual and developmental disabilities. *Journal of Policy and Practice in Intellectual Disabilities* **8**(1), 42–52. doi:10.1111/j.1741-1130.2011.00288.x
- Jenkins K, Alberry B, Daniel J, et al. (2010) Beyond communication: The development of a training program for hospital and hospice staff in the detection and management of psychological distress: Preliminary results. *Palliative & Supportive Care* **8**, 27–33. doi:10.1017/S1478951509990678
- Karacsony S, Chang E, Johnson A, et al. (2015) Measuring nursing assistants’ knowledge, skills and attitudes in a palliative approach: A literature review. *Nurse Education Today* **35**(12), 1232–1239. doi:10.1016/j.nedt.2015.05.008
- Kim J and Gray J (2018) Palliative care experiences and needs of direct care workers. *Journal of Palliative Medicine* **21**(8), 1094–1099. doi:10.1089/jpm.2017.0555
- Kim J and Gray J (2021) Effect of online palliative care training on knowledge and self-efficacy of direct care workers. *Intellectual and Developmental Disabilities* **5**(5), 392–404. doi:10.1352/1934-9556-59.5.392
- Kinnear D, Morrison J, Allan L, et al. (2018) Prevalence of physical conditions and multimorbidity in a cohort of adults with intellectual disabilities with and without Down syndrome: Cross-sectional study. *BMJ Open* **8**(2), e018292. doi:10.1136/bmjopen-2017-018292
- Kirkendall AM and Waldrop D (2013) Staff perspectives on the provision of end-of-life care in a community residence for older adults with developmental disabilities. *Journal of Palliative Medicine* **16**(9), 1121–1124. doi:10.1089/jpm.2012.0408
- Lange JW, Mager D, Greiner PA, et al. (2011) The ELDER Project: Educational model and three-year outcomes of a community-based geriatric education initiative. *Gerontology & Geriatrics Education* **32**(2), 164–181. doi:10.1080/02701960.2011.572056
- Linacre JM (2002) Optimizing rating scale category effectiveness. *Journal of Applied Measurement* **3**(1), 85–106.
- Linacre JM (2021a) *Winsteps User Manual (Version 5.1.4)*. Winsteps. <https://www.winsteps.com/manuals.htm> (accessed 1 September 2021).
- Linacre JM (2021b) *Winsteps (Version 4.8.0.0)*. Winsteps. <https://www.winsteps.com/winbuy.htm> (accessed 1 August 2021).
- Lord AJ, Field S and Smith IC (2017) The experiences of staff who support people with intellectual disability on issues about death, dying and bereavement: A metasynthesis. *Journal of Applied Research in Intellectual Disabilities* **30**, 1007–1021. doi:10.1111/jar.12376
- Maher JM, Markey JC and Ebert-May D (2013) The other half of the story: Effect size analysis in quantitative research. *CBE Life Sciences Education* **12**(3), 345–351. doi:10.1187/cbe.13-04-0082
- McCarron M, McCallion P, Fahey-McCarthy E, et al. (2010) Staff perceptions of essential prerequisites underpinning end-of-life care for persons with intellectual disability and advanced dementia. *Journal of Policy and Practice in Intellectual Disabilities* **7**(2), 143–152. doi:10.1111/j.1741-1130.2010.00257.x
- McEvoy J, MacHale R and Tierney E (2012) Concept of death and perceptions of bereavement in adults with intellectual disabilities. *Journal of Intellectual Disability Research* **56**(2), 191–203. doi:10.1111/j.1365-2788.2011.01456.x
- Moura Minosso JS, Martins MM and De Campos Oliveira MA (2017) Cross-cultural adaptation of the Bonn Palliative Care Knowledge Test:

- An instrument to assess knowledge and self-efficacy. *Revista de Enfermagem Referência* 4(13), 31–42. doi:10.12707/RIV16076
- Mulder SF, Bleijenberg G, Verhagen SC, et al.** (2009) Improved competence after a palliative care course for internal medicine residents. *Palliative Medicine* 23(4), 360–368. doi:10.1177/0269216308100250
- Mundfrom DJ, Shaw DG and Ke TL** (2005) Minimum sample size recommendations for conducting factor analyses. *International Journal of Testing* 5(2), 159–168. doi:10.1207/s15327574ijt0502_4
- Nakazawa Y, Miyashita M, Morita T, et al.** (2009) The palliative care knowledge test: Reliability and validity of an instrument to measure palliative care knowledge among health professionals. *Palliative Medicine* 23(8), 754–766. doi:10.1177/0269216309106871
- Ng J and Li S** (2003) A survey exploring the educational needs of care practitioners in learning disability (LD) settings in relation to death, dying and people with learning disabilities. *European Journal of Cancer Care* 12(1), 12–19. doi:10.1046/j.1365-2354.2003.00323.x
- Phillips J, Salamonson Y and Davidson PM** (2011) An instrument to assess nurses' and care assistants' self-efficacy to provide a palliative approach to older people in residential aged care: A validation study. *International Journal of Nursing Studies* 48, 1096–1100. doi:10.1016/j.ijnurstu.2011.02.015
- Raiche G** (2005) Critical eigenvalue sizes in standardized residual principal components analysis. *Rasch Measurement Transactions* 19(1), 1012.
- Resnick B, Galik E, Pretzer-Aboff I, et al.** (2008) Testing the reliability and validity of self-efficacy and outcome expectations of restorative care performed by nursing assistants. *Journal of Nursing Care Quality* 25(2), 162–169. doi: 10.1097/01.NCQ.0000313766.09891.43
- Reynolds S, Guerin S, McEvoy J, et al.** (2008) Evaluation of a bereavement training program for staff in an intellectual disabilities service. *Journal of Applied Research in Intellectual Disabilities* 5(1), 1–5. doi:10.1111/j.1741-1130.2007.00132.x
- Ryan K, Guerin S, Dodd P, et al.** (2011a) Communication contexts about illness, death and dying for people with intellectual disabilities and life-limiting illness. *Palliative & Supportive Care* 9, 201–208. doi:10.1017/S1478951511000137
- Ryan K, Guerin S, Dodd P, et al.** (2011b) End-of-life care for people with intellectual disabilities: Paid carer perspectives. *Journal of Applied Research in Intellectual Disabilities* 24, 199–207. doi:10.1111/j.1468-3148.2010.00605.x
- Ryan K, McEvoy J, Guerin S, et al.** (2010) An exploration of the experience, confidence and attitudes of staff to the provision of palliative care to people with intellectual disabilities. *Palliative Medicine* 24(6), 566–572. doi:10.1177/0269216310371413
- Shipman C, Burt J, Ream E, et al.** (2008) Improving district nurses' confidence and knowledge in the principles and practice of palliative care. *Journal of Advanced Nursing* 63(5), 494–505. doi:10.1111/j.1365-2648.2008.04729.x
- Stein GL** (2008) Providing palliative care to people with intellectual disabilities: Services, staff knowledge, and challenges. *Journal of Palliative Medicine* 11(9), 1241–1248. doi:10.1089/jpm.2008.0130
- Stroebe M, Stroebe W and Schut H** (2001) Gender differences in adjustment to bereavement: An empirical and theoretical review. *Review of General Psychology* 5(1), 62–83. doi:10.1037/1089-2680.5.1.62
- Tennant A and Conaghan P** (2007) The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper. *Arthritis and Rheumatism* 57(8), 1358–1362. doi:10.1002/art.23108
- Tuffrey-Wijne I, Hogg J and Curfs L** (2007) End-of-life and palliative care for people with intellectual disabilities who have cancer or other life-limiting illness: A review of the literature and available resources. *Journal of Applied Research in Intellectual Disabilities* 20, 331–344. doi:10.1111/j.1468-3148.2006.00350.x
- Wark S, Hussain R and Edwards H** (2014) The training needs of staff supporting individuals ageing with intellectual disability. *Journal of Applied Research in Intellectual Disabilities* 27, 273–288. doi:10.1111/jar.12087
- Watters L, McKenzie K and Wright R** (2012) The impact of staff training on the knowledge of support staff in relation to bereavement and people with an intellectual disability. *British Journal of Learning Disabilities* 40(3), 194–200. doi:10.1111/j.1468-3156.2011.00693.x
- Wiese M, Stancliffe RJ, Dew A, et al.** (2014) What is talked about? Community living staff experiences of talking with older people with intellectual disability about dying and death. *Journal of Intellectual Disability Research* 58(7), 679–690. doi:10.1111/jir.12065
- Wilkinson S, Perry R, Blanchard K, et al.** (2008) Effectiveness of a three-day communication skill course in changing nurses' communication skills with cancer/palliative care patients: A randomised controlled trial. *Palliative Medicine* 22, 365–375. doi:10.1177/0269216308090770
- Wittenberg-Lyles E, Goldsmith J, Ferrell B, et al.** (2014) Assessment of an interprofessional online curriculum for palliative care communication training. *Journal of Palliative Medicine* 17(4), 400–407. doi:10.1089/jpm.2013.0270
- World Health Organization** (2015) Palliative care. <http://www.who.int/mediacentre/factsheets/fs402/en/> (accessed 1 September 2021).
- World Health Organization and World Bank** (2011) World report on disability. Geneva, Switzerland: World Health Organization Press.
- Wright BD** (2003) Rack and stack: Time 1 vs. time 2 or pre-test vs. post-test. *Rasch Measurement Transactions* 17, 905–906.
- Wright BD and Linacre JM** (1994) Reasonable mean-square fit values. *Rasch Measurement Transactions* 8(3), 370.
- Wright BD, and Masters GN** (1982) *Rating Scale Analysis*. Chicago, IL: Mesa Press.