Published in final edited form as:

Aust Occup Ther J. 2014 December ; 61(6): 384–393. doi:10.1111/1440-1630.12155.

Development and Validation of the Activity Significance Personal Evaluation Scale

Trudy Mallinson^{1,2}, Stacey L. Schepens Niemiec², Mike Carlson², Natalie Leland^{2,3}, Cheryl Vigen², Jeanine Blanchard², and Florence Clark²

¹Department of Clinical Research and Leadership, School of Medicine and Health Sciences, The George Washington University

²The Division of Occupational Science and Occupational Therapy at the Herman Ostrow School of Dentistry, University of Southern California

³The Davis School of Gerontology, University of Southern California

Introduction

Engagement in meaningful occupation for health promotion is a longstanding tenet of occupational therapy (Kielhofner, 2007; Wilcock, 2006). Beyond the direct physical, cognitive, and social benefits of daily activity (Leland & Elliott, 2012), occupational therapists assert that participation in meaningful activity can result in an improved sense of mastery, life satisfaction, and development of hopeful future life narratives (Berger, McAteer, Schreier, & Kaldenberg, 2013; Fallahpour, Jonsson, Joghataei, Nasrabadi, & Tham, 2013; Foster, Bedekar, & Tickle-Degnen, 2014). Further, Jonsson (2008) has eloquently argued for an experiential view occupation that makes more clear the relationship between engagement in occupation and individual well-being. Consequently, an assessment tool that captures personal perceptions regarding the significance of daily activity for health and wellness could enhance the study of potential benefits of participation in daily activities (Hemmingsson & Jonsson, 2005). Additionally, understanding the mechanisms by which occupational therapy interventions benefit individuals requires the ability to quantify differences in activity significance among individuals, and to capture changes over time. To date, no tool captures adults' perceptions of the significance of daily activity for their personal health and wellness.

In treating community-living older adults, occupational therapy interventions are premised not just on enabling more frequent participation in activities or in more challenging

Correspondence concerning this article should be directed to Florence Clark, PhD, OTR/L, FAOTA, Associate Dean and Professor, Division of Occupational Science and Occupational Therapy at the Herman Ostrow School of Dentistry of USC, University of Southern California, 1540 Alcazar Street, CHP 133, Los Angeles, CA 90033, fclark@usc.edu.

Author Info for Online Submission: Trudy Mallinson, PhD, OTR/L, NZROT, FAOTA, Visiting Associate Professor, trudy@gwu.edu

Stacey L. Schepens Niemiec, PhD, OTR/L, Research Assistant Professor, schepens@osot.usc.edu

Mike Carlson, PhD, Research Professor, mcarlson@osot.usc.edu

Natalie Leland, PhD, OTR/L, BCG, FAOTA, Assistant Professor, nleland@usc.edu

Cheryl Vigen, PhD, Research Assistant Professor, vigen@usc.edu

Jeanine Blanchard, PhD, OTR/L, Project Coordinator, jeanine@osot.usc.edu

Florence Clark, PhD, OTR/L, FAOTA, Professor of Occupational Science and Occupational Therapy, fclark@osot.usc.edu

activities, but on greater engagement in personally meaningful activities (Gustafsson & McKenna, 2010; Pereira & Stagnitti, 2008). Unfortunately, available assessments generally focus on quantifying frequency and/or time devoted to activities. For example, activity checklists solicit the frequency of participating in different activities such as cooking or volunteering and may vary greatly in the number of activities assessed (Everard, Lach, Fisher, & Baum, 2000). Time-use diaries ask participants to record the number of minutes engaged in selected activities within a certain time period (Gerber & Furst, 1992). Participants may additionally be asked to classify or rate dimensions such as satisfaction or performance for activities like self-care or leisure (Archenholtz & Dellhag, 2008; Scanlan & Bundy, 2011). Classifying activity by type may be challenging to interpret, as scholars note that activities can have different meanings depending on social context (Hammell, 2009). For example, cooking could be considered work or leisure depending on the individual's occupational role (Mallinson & Hammel, 2010). Thus, checklists and time-use diaries, which may be cumbersome and time-consuming, may not inform providers about the perceived value and personal meaning of activities in promoting health and wellness.

A brief, easily administered instrument is clearly needed to capture the extent to which older adults consider their everyday activities, in a global sense, to be meaningful and healthy. The aim of this study was to develop such an assessment that captures differences among older adults' perceptions of meaningful activity and that conforms to contemporary guidelines for scale development and testing.

Methods

The Activity Significance Personal Evaluation (ASPEn) scale was fielded during a randomised controlled trial (the University of Southern California Well Elderly 2 Study) of an occupational therapy intervention for community-living older adults (Clark et al., 2011). Data from five assessment waves were used to evaluate the psychometric properties of the ASPEn. Study approval was obtained from the University of Southern California Institutional Review Board. Participants provided informed consent prior to beginning the study.

Participants

Data were collected on 460 older adults aged 60 years and above who participated in the Well Elderly 2 Study (Clark et al., 2011). All participants lived at or used services at residential care settings, community centres, or retirement communities in the Los Angeles metropolitan area. Details of the study are reported elsewhere (Jackson et al., 2009). Briefly, participants were randomly assigned to a six-month activity-based lifestyle intervention or a no-treatment control condition, with individuals in the control condition crossing over to receive the intervention during their second six months of participation (Clark et al., 2011). The sample was predominantly female (66%), was English- (86%) or Spanish-speaking (14%), and had a mean age of 74.9 ± 7.7 years. Demographic data are presented in Table 1.

Phase 1: Item Development

ASPEn items were developed by three content experts in the area of ageing and health in relation to activity. Eighteen items were generated that reflected activity satisfaction and perception of engagement. Development of the item content was not based on a specific conceptual model; rather the items reflect a more global occupational therapy/occupational science perspective of the value of engaging in occupation to promote and maintain health (Christiansen, Baum, & Haugen, 2005). The activity satisfaction items largely focus on the extent to which the respondents' activities contribute to their health, personal satisfaction, and fulfillment, whereas the perception of engagement items focus on benefits commonly associated with valued activities (e.g., ability to connect with others, activity planning). Thirteen of the items used a seven-point Likert-type scale of which four steps were labelled: 1=*Not at all*, 3=*Only a little*, 5=*Somewhat*, and 7=*A great deal*. Steps two, four, and six were unlabelled. Two activity satisfaction questions asked about the past six months and used a seven-point Likert-type scale labelled: 1=*A great deal less*, 2=*Somewhat less*, 3=*A little less*, 4=*The same amount*, 5=*A little more*, 6=*Somewhat more*, and 7=*A great deal*. Two activity satisfaction items and one perception of engagement item were negatively worded and reverse-coded during scoring. Items were translated into Spanish (translation, back translation, and committee review) following recommended guidelines (Acquadro et al., 2008). Prior to the trial, items were pilot tested on ten older adults.

Phase 2: Item Evaluation

This phase involved an evaluation of the construct validity, internal consistency, and measurement precision of the items. The Well Elderly 2 Study data involved up to five measurement waves, occurring at 6-month intervals, for each study participant. Including the same person multiple times in a Rasch analysis violates the assumption of independent data points. Therefore, for the initial calibration effort we followed the method described by Mallinson (2011), randomly selecting participants so that all waves of data were equally represented in the data set but with each participant appearing only once. This data set was used to examine rating scale step structure, item performance, and dimensionality, and to establish the item and rating scale step calibrations. These calibrations were subsequently applied to all participants at all time-points so that all person calibrations were based on the same frame of reference with appropriate standard errors.

Phase 3: Scale Validation

Concurrent validity—When item selection was finalised we examined concurrent validity. We computed Pearson correlation coefficients to test the association between the ASPEn and the Life Satisfaction Index Form Z (LSIZ) and the Medical Outcome Study 36-item Short-Form Health Survey (SF-36) mental (MCS) and physical health (PCS) component scores at baseline. To examine the ability of the ASPEn to distinguish between respondents classified by depressive symptom level, we used independent t-tests to compare participants with Center for Epidemiologic Studies Depression-20 (CESD-20) scores of 12 points or higher versus 11 points or lower (Lewinsohn, Seeley, Roberts, & Allen, 1997). We bootstrapped 10,000 samples with replacement to estimate 95% confidence intervals (CI) for both Pearson correlation coefficients and unpaired t-statistics using Stata SE version 11.c.

Relative validity—We used the method of known-groups validity (Gothwal, Wright, Lamoureux, & Pesudovs, 2010; Hobart et al., 2004; McHorney, Haley, & Ware, 1997) to investigate how well the ASPEn distinguishes between participant groups. When comparing groups known to be different on some clinical indicator, the relative measurement precision is the degree to which an instrument distinguishes between the groups relative to the variation within the groups. The F-statistic from a one-way analysis of variance compares the between group variance (systematic variance) to within group variance (error variance). A relative validity (RV) coefficient is a ratio of the F-statistic for the comparison scale divided by the F-statistic for the standard scale (Kerlinger & Lee, 2000). When the scale with the largest F-statistic is used as the denominator the resulting ratio indicates the difference in relative measurement precision as a percentage (Hobart et al., 2004).

In the current study, F-statistics from one-way analysis of variance were computed by depression symptom grouping and by volunteer status. We expected that participants with high levels of depressive symptoms would report lower levels of activity satisfaction and engagement and those who volunteered would report greater activity significance than those who did not.

Statistical Analysis

Rasch analysis generally proceeds as a logical series of analyses, each time examining overall test performance, rating scale structure, item fit, and dimensionality, ensuring that changes made such as item deletions continue to improve the instrument's psychometric properties. For the first round of analysis, we included all 18 items and three types of rating scales. In the second round, we removed items exhibiting dimensionality and/or misfit and reanalysed the remaining items. In the third round, we evaluated a partial credit model for the remaining 13 items in which each item has its own rating scale structure. In the fourth and final round, we used a single rating scale model for all items, recoding to combine steps two, three, and four to produce a four-point rating scale. Subsequently, we converted logits to a 0-100 scale and rating scale step and created item anchor files. We applied these anchor files to obtain person measures for all study participants at all time points to conduct the validity analyses. This final short version of the scale is referred to as the ASPEn scale. We developed a total raw score-to-Rasch measure conversion, which enables conversion of the total score of the ASPEn into a linear Rasch measure (ranging from 0 to 100).

Rasch analysis was completed using Winsteps® 3.80.1 (Linacre, 2013). Rasch measurement provides logit values (calibrations) as estimates of item challenge and person ability. Negative logits indicate that the item is easier, or the person is less engaged/satisfied, than average. Positive logits indicate that the item is more challenging (harder to endorse), or the person is more able (more engaged/satisfied) than average. Key aspects that we considered in the Rasch analysis are summarised below.

Category threshold order—Rating scale steps should proceed monotonically, that is, there should be an increase in the amount of challenge required to endorse consecutively higher categories.

Item hierarchy—The Rasch model assumes that more challenging items are more challenging for all persons and that more able persons will endorse higher rating scale steps on more challenging items.

Item fit and dimensionality—Mean square fit statistics (MnSqs) are an indication of the extent to which items conform to a hierarchy, which is an underlying assumption of the Rasch model. MnSqs between 0.7 and 1.3 are considered an indication of fit (Wilson, 2005). For completeness we report z-standardised (ZSTD) probabilities. These indicate the likelihood that the MnSq value occurred by chance. When the MnSqs are acceptable the ZSTDs can be disregarded. They are generally inflated in sample sizes over 300 (Linacre, 2012). A principal component analysis (PCA) was undertaken to further examine dimensionality within the item residuals. Unexplained variance of the first factor of less than 10% (Eigenvalue <2.0) indicated that the items were unidimensional (Linacre & Tennant, 2009). In addition, items with loadings greater than 0.3 were considered to load on a contrast.

Precision and reliability—The person separation index (SI) and person separation reliability (PSR) are indicators of measurement precision. Separation ratios greater than 2.0 are considered acceptable, but values closer to 3.0 are preferred for clinical decision-making. The person separation reliability is a correlation coefficient reflecting the ratio of true measure variance to observed measure variance, with values interpreted as per Cronbach's alpha (Linacre, 2012). In addition, internal consistency was evaluated by calculating Cronbach's alpha for each scale. To establish test-retest reliability for the ASPEn Rasch measure, we computed a Pearson product-moment correlation coefficient. This was done for the sub-sample of control group participants who were measured at study onset and prior to crossing over to the intervention. Their engagement was presumed to be stable in that time period.

Differential item functioning—To compare measures across persons, the hierarchical order of the items should remain consistent. Differences among item calibrations across groups of participants are referred to as differential item functioning (DIF). In this study, we examined DIF in English- versus Spanish-speaking participants and by educational level (grade school, high school, college, post-graduate). We considered DIF of < 0.5 to be inconsequential, 0.5-1.0 logits as minimal, and > 1.0 logits to be notable (Lundstrom & Pesudovs, 2009).

Targeting—Comparison of mean person and mean item logit calibrations is an indication of how well the items are targeted to the persons. When the person mean is greater than the item mean, persons are more able than the items are challenging.

Results

Category Thresholds

In the first round of analysis the unlabelled response choices were never the most probable steps, indicating that respondents avoided choosing these categories. A partial credit model in which each item was free to have its own rating scale structure reduced the number of

misfitting items, but did not improve measurement precision (no change in PSR or SI; see Table 2). In the final round of analysis, a single four-step rating scale model (1333557) for all 13 items resulted in a higher PSR and no items misfit, indicating that the unlabelled steps were not adding information about the participants; in fact, they may have been adding measurement noise. Additionally, we recommend that the original category three “somewhat” be replaced with “moderate amount” in order to better differentiate among persons.

Item Fit and Dimensionality

Analysis of all 18 items resulted in three misfitting items (4, 6, 13). In addition, PCA suggested two dimensions in which the items asking about the last six months (items 7 and 8) and three negatively worded items (items 4, 6, 13) loaded on a separate contrast. After removing these five items, the remaining 13 items fit the measurement model, with item discriminations between 0.5 and 1.5 and the PCA eigenvalue below 2.0, suggesting the activity satisfaction and the perception of engagement items cohere to form a single construct.

Item Hierarchy

The final item calibrations after conversion to a 0-100 range are provided in Table 3. It was easiest for participants to acknowledge that daily activities support physical and mental health but harder to endorse items indicating their everyday activities provide a sense of accomplishment and hope for the future. Overall, the items of the ASPEn form a conceptual hierarchy that describes the meaning and significance of older adults' daily activity for health and wellness. A revised version of the ASPEn scale based on the results of this study is presented in Appendix 1, including scoring directions. We also provide a table to enable conversion of raw scores to Rasch-measure (Appendix 2).

Precision and Reliability

In the final model, the SI for the 13 items was 2.7 with a PSR of 0.88. Cronbach's alpha was 0.91. The minor difference is due to Rasch estimates slightly under-estimating and alpha slightly over-estimating the coefficients (Linacre, 2012). These results indicate that the ASPEn is appropriate for individual participant measurement. Test-retest reliability coefficients for the raw score and Rasch measures were 0.69 (p .001) and 0.70 (p .001), respectively.

Differential item function (DIF)

Only one item—how much activities let you connect with other people—showed minimal DIF (exceeding the English-language version by .56 logits) for participants who took the survey in Spanish. Since the DIF was minimal and rating scale changes are recommended, we propose no change to the item at this time.

Targeting

The mean person measure was 64.5 ± 14.9 , suggesting that participants found it relatively easy to endorse these items (by default, items are centred on a mean of 50.0). Twenty

participants had maximum measures (ceiling effect); no participants had minimum measures.

Concurrent Validity

At baseline, ASPEn measures correlated moderately with LSIZ ($r=0.43$, $CI=0.36-0.50$), PCS ($r=0.31$, $CI=0.22-0.40$), and MCS ($r=0.23$, $CI=0.14-0.31$), total scores. Mean scores (SD) for the CESD-20 were 13.73 (10.91); the ASPEn significantly distinguished between participants classified with higher (≥ 12) and lower depressive symptoms (<11), mean difference 8.3, $p<0.001$.

Relative Validity

Table 4 summarises the relative measurement precision estimates for the raw and Rasch versions of the scale. Participants with fewer depressive symptoms (CESD-20 score <12) and persons who volunteered received higher ASPEn raw scores and ASPEn Rasch measures. Rasch measures appear essentially equivalent in precision relative to raw scores for discriminating between participant groups.

Discussion

The purpose of this article was to describe the development and psychometric evaluation of an instrument designed to measure older adults' perceptions of the extent to which their daily activity promotes their overall health and wellness. Following removal of misfitting items and reduction of the rating scale to four steps, the ASPEn scale showed sound construct validity, criterion-related validity, and excellent measurement precision. The final 13 items demonstrated robust internal consistency, the absence of item misfit, and minimal DIF, indicating the ASPEn item hierarchy is stable across a range of older adult groups. Additionally, the ASPEn is related to other health and wellness measures and can effectively distinguish between patients with various levels of depressive symptoms.

Recent literature has demonstrated that traditional categorization of activity into self-care, work/volunteering, and leisure is not necessarily meaningful to older adults (Chilvers, Corr, & Singlehurst, 2010). In arguing for a more client-centred approach to understanding occupations, Hammell (2009) suggests four “experience-based” categories of occupation. Work on the ASPEn precedes Hammell's article but our finding of 13 items that cohered to form a single scale lends empirical support for these categories. Specifically, *restorative occupations* (occupations that contribute to a sense of being in the moment, harmony, peace, reflection, relaxation, and rejuvenation) are captured by five ASPEn items (i.e., Planning for activities to be health promoting; Satisfied with your activities; Activities contribute to your mental health; Activities contribute to your physical health; and Knowledgeable about how activities influence your health). *Occupations fostering belonging, connecting, and contributing* are those that foster perceptions of self-worth, a sense of connection, and making contributions. The ASPEn item, ‘Activities let you connect with other people,’ captures this concept. *Engaging in doing occupations* fosters a sense of capability and accomplishment and includes concepts of interest, reward, routines, and commitment. Five ASPEn items appear to capture this concept (i.e., Activities provide the opportunity for new

experiences or adventure; Needing or having to do activities; Activities provide you with a sense of accomplishment; Confident you can influence your health; and Like the activities you do). Hammells' fourth category, *reflecting life continuity and hope for the future*, includes occupations that preserve biographical continuity, and includes dreams, aspirations, and continuing activities that are rewarding and meaningful. ASPEn items that appear to capture this concept include: Activities provide you with hope for the future, and Activities are personally fulfilling.

The order of the 13 items from easiest to hardest to endorse fits with conceptual and clinical expectations, but also adds to the field's understanding of how older adults perceive the role of daily activity in promoting health and wellness. For example, these older adults found it easiest to agree that their daily activities promote physical and mental health, are enjoyable, and increase their efficacy in influencing their health. It was more challenging for study participants to endorse the idea that their everyday activities provide a sense of accomplishment, fulfillment, and hope for the future. The most challenging items for these adults centred on purposely choosing activities to be healthful, and seeing their activities as connecting them with others and providing new experiences or adventure. These findings concur with recent literature that shows older adults spend most of their time in isolated or solitary activities (McKenna, Broome, & Liddle, 2007). Further, although most activities are routine and not adventuresome, older adults can and do find such opportunities in everyday life as Pereira and Stagnitti (2008) reported in a study of older Italian Australians who compete in Bocce ball.

We removed five items from the analysis due to misfit but additionally because, on reflection, they tapped conceptually different constructs. Asking to what extent activity meaning has changed or level of physical activity has changed in the past six months, while related, do not seem to capture the same information as the significance of occupations for health and wellness. Likewise, items that address the extent to which a person feels bored, would like to change activities, or is fearful of trying new activities may be relevant in a full description of older adults' daily activity choices; however, they do not capture the perception of activity promoting health and wellness.

Three items that were not included in the original scoring (i.e., questions 16-18) were included in the Rasch analysis. The rationale for excluding these items in the original scoring was unclear. Because Rasch analysis demonstrated that these items fit with the overall construct, and there is theoretical rationale for including them (Hammell, 2009), we have maintained them in the instrument and included them in the revised scoring.

The ASPEn is the first scale to attempt to measure older adults' perceptions of the ways daily activity contribute to their health and wellness. This is somewhat surprising because the role of meaningful occupation in promoting health has been a central tenet of occupational therapy since its inception. That the field has not had an instrument to evaluate this concept is a limitation because it hinders researchers' ability to demonstrate how occupational therapy interventions impact health in older adults. In addition, this assessment tool may be useful in clinical settings focused on health promotion to better understand client motivation for engagement (or not) in everyday activities.

This study provides preliminary evidence that the ASPEn is precise, reliable, and has construct and concurrent validity; however, the predictive validity of the tool is yet to be established. Future studies should examine the value of the ASPEn in predicting successful community living, healthcare utilization, and functional abilities. Additional enhancements could include testing the newly recommended four-point scale with the more client-centred labels, further cognitive testing to refine item definitions, and possibly adding more challenging items.

This study examined the ASPEn in the context of community-living, ethnically diverse older adults in the United States. It would be of great value to see if the same hierarchy holds for other groups of older adults, such as those in nursing homes, different cultural and national groups, and varying age groups.

In this vein, it is noteworthy that the scale was able to distinguish between participants with different degrees of depressive symptoms. This suggests that scores on the ASPEn may be able to inform clinical decisions by helping occupational therapists focus on appropriate areas for increasing the meaning of activity in daily life to support clients' mental health not only in old age but possibly in adulthood more generally.

Unlike others (Khan, Chien, & Brauer, 2013; Raczek et al., 1998) we did not find the Rasch version to be significantly more precise than the raw score version. The relationship between raw scores and Rasch measures is, by definition, non-linear (Wright & Stone, 1979), with raw scores being more compressed at the extremes. In addition, participants in this study were quite skewed on their ASPEn scores (20 had maximum scores) and this off-targeting is reflected in the F-ratios. Consequently, the raw scores appear able to more easily detect greater difference in variance between the groups. Similarly, (Fitzpatrick et al., 2004) found that Rasch measures were better at detecting differences in patients at the farther ends of each scale relative to raw scores, but that the approaches were equivalent in the middle of the scale.

Limitations

This study has limitations that influence the generalisability of the results. Participants in this study were community-living and did not represent the full range of older adults treated by occupational therapists. Additionally, data were collected as part of a clinical trial. Although we took steps to minimise the impact of this on study results (e.g., we used baseline data prior to intervention whenever possible), a future prospective study is warranted. Item development was undertaken primarily by content experts who were occupational scientists with experience in geriatric services, and also reflects concepts identified by qualitative studies; however, relevant client-centred concepts may have been omitted. In addition, the items of the scale have a Flesch-Kincaid reading level of 8th to 9th grade, which may be challenging for many older adults. Future work to revise item wording is recommended.

Conclusion

There are few measures capturing older adults' perceptions of how meaningful occupation improves their health and wellness. This scale is unique in that it captures overarching evaluations of personal activity participation and its health-related effects, without relying on combining judgments about particular activity categories. Establishing the scale's sensitivity to change and predictive validity will enable the scale to track changes in occupational meaning for older adults over time and in different settings. In particular, the ASPEn may be useful for predicting onset of health issues or explaining the effectiveness of occupation-based interventions. Information about older adults' activity significance would not only facilitate improved research in occupational therapy interventions, but also expand our understanding of the intricate role of activity in supporting health and wellness.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The research was funded in part by a grant from the National Institute on Aging R01 AG021108 to Dr. Florence Clark. During the time of this study Dr. Leland and Dr. Schepens were funded by awards from the K12 Rehabilitation Research Career Development (RRDC) Program (K12 HD055929). The authors acknowledge the contributions of Abbey Marterella and Ann Kuo on earlier drafts of this manuscript.

References

- Acquadro C, Conway K, Hareendran A, Aaronson N. European Regulatory, Issues, & Quality of Life Assessment, Group. Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value Health*. 2008; 11(3):509–521.10.1111/j.1524-4733.2007.00292.x [PubMed: 18179659]
- Archenholtz B, Dellhag B. Validity and reliability of the instrument Performance and Satisfaction in Activities of Daily Living (PS-ADL) and its clinical applicability to adults with rheumatoid arthritis. *Scandinavian Journal of Occupational Therapy*. 2008; 15(1):13–22.10.1080/11038120701223165 [PubMed: 17852967]
- Berger S, McAteer J, Schreier K, Kaldenberg J. Occupational therapy interventions to improve leisure and social participation for older adults with low vision: a systematic review. *American Journal of Occupational Therapy*. 2013; 67(3):303–311.10.5014/ajot.2013.005447 [PubMed: 23597688]
- Chilvers R, Corr S, Singlehurst H. Investigation into the occupational lives of healthy older people through their use of time. *Australian Occupational Therapy Journal*. 2010; 57(1):24–33.10.1111/j.1440-1630.2009.00845.x [PubMed: 20854562]
- Christiansen, CH.; Baum, CM.; Haugen, JB., editors. *Occupational Therapy: Performance, Participation, and Well-Being*. Vol. 3rd. Thorofare, NJ: Slack; 2005.
- Clark F, Jackson J, Carlson M, Chou CP, Cherry BJ, Jordan-Marsh M, et al. Azen SP. Effectiveness of a lifestyle intervention in promoting the well-being of independently living older people: results of the Well Elderly 2 Randomised Controlled Trial. *Journal of Epidemiology and Community Health*. 2011; 66:782–790.10.1136/jech.2009.099754 [PubMed: 21636614]
- Everard KM, Lach HW, Fisher EB, Baum MC. Relationship of activity and social support to the functional health of older adults. *Journals of Gerontology Series B, Psychological Sciences and Social Sciences*. 2000; 55(4):S208–212.
- Fallahpour M, Jonsson H, Joghataei MT, Nasrabadi AN, Tham K. “I am not living my life”: lived experience of participation in everyday occupations after stroke in Tehran. *Journal of Rehabilitation Medicine*. 2013; 45(6):528–534.10.2340/16501977-1143 [PubMed: 23624528]

- Fitzpatrick R, Norquist JM, Jenkinson C, Reeves BC, Morris RW, Murray DW, Gregg PJ. A comparison of Rasch with Likert scoring to discriminate between patients' evaluations of total hip replacement surgery. *Quality of Life Research*. 2004; 13(2):331–338. [PubMed: 15085905]
- Foster ER, Bedekar M, Tickle-Degnen L. Systematic review of the effectiveness of occupational therapy-related interventions for people with Parkinson's disease. *American Journal of Occupational Therapy*. 2014; 68(1):39–49.10.5014/ajot.2014.008706 [PubMed: 24367954]
- Gerber LH, Furst GP. Validation of the NIH activity record: a quantitative measure of life activities. *Arthritis Care and Research*. 1992; 5(2):81–86. [PubMed: 1390968]
- Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Measuring outcomes of cataract surgery using the Visual Function Index-14. *Journal of Cataract and Refractive Surgery*. 2010; 36(7):1181–1188.10.1016/j.jcrs.2010.01.029 [PubMed: 20610098]
- Gustafsson L, McKenna K. Is there a role for meaningful activity in stroke rehabilitation? *Top Stroke Rehabil*. 2010; 17(2):108–118.10.1310/tsr1702-108 [PubMed: 20542853]
- Hammell KW. Self-care, productivity, and leisure, or dimensions of occupational experience? Rethinking occupational “categories”. *Canadian Journal of Occupational Therapy*. 2009; 76(2): 107–114.
- Hemmingsson H, Jonsson H. An occupational perspective on the concept of participation in the International Classification of Functioning, Disability and Health--some critical remarks. *Am J Occup Ther*. 2005; 59(5):569–576. [PubMed: 16268024]
- Hobart J, Kalkers N, Barkhof F, Uitdehaag B, Polman C, Thompson A. Outcome measures for multiple sclerosis clinical trials: relative measurement precision of the Expanded Disability Status Scale and Multiple Sclerosis Functional Composite. *Multiple Sclerosis*. 2004; 10(1):41–46. [PubMed: 14760951]
- Jackson J, Mandel D, Blanchard J, Carlson M, Cherry B, Azen S, et al. Clark F. Confronting challenges in intervention research with ethnically diverse older adults: the USC Well Elderly II Trial. *Clinical Trials*. 2009; 6(1):90–101.10.1177/1740774508101191 [PubMed: 19254939]
- Jonsson H. A new direction in the conceptualization and categorization of occupation. *Journal of Occupational Science*. 2008; 15(1):3–8.10.1080/14427591.2008.9686601
- Kerlinger, FN.; Lee, HB. *Foundations of Behavioral Research*. 4th. Wadsworth Thompson Learning; 2000.
- Khan A, Chien CW, Brauer SG. Rasch-based scoring offered more precision in differentiating patient groups in measuring upper limb function. *Journal of Clinical Epidemiology*. 2013; 66(6):681–687.10.1016/j.jclinepi.2012.12.014 [PubMed: 23523550]
- Kielhofner, G. *Conceptual Foundations of Occupational Therapy Practice*. 4th. Philadelphia, PA: F.A. Davis; 2007.
- Leland NE, Elliott SJ. Special issue on productive aging: evidence and opportunities for occupational therapy practitioners. *American Journal of Occupational Therapy*. 2012; 66(3):263–265.10.5014/ajot.2010.005165 [PubMed: 22549589]
- Lewinsohn PM, Seeley JR, Roberts RE, Allen NB. Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. *Psychology and Aging*. 1997; 12(2):277–287. [PubMed: 9189988]
- Linacre, JM. A User's Guide to WINSTEPS & MINISTEP Rasch-Model Computer Programs. Program Manual 3.74.0.2012. 2012. from <http://www.winsteps.com/winman>
- Linacre, JM. WINSTEPS®: Rasch Measurement (Version 3.80.1 Date: 09-09-2013). Chicago, IL: MESA Press; 2013. Retrieved from <http://www.winsteps.com/winsteps.htm>
- Linacre JM, Tennant A. More about critical eigenvalue sizes in standardized-residual principal components analysis (PCA). *Rasch Measurement Transactions*. 2009; 23(3):1228.
- Lundstrom M, Pesudovs K. Catquest-9SF patient outcomes questionnaire: nine-item short-form Rasch-scaled revision of the Catquest questionnaire. *Journal of Cataract and Refractive Surgery*. 2009; 35(3):504–513.10.1016/j.jcrs.2008.11.038 [PubMed: 19251145]
- Mallinson T. Rasch Analysis of repeated measures. *Rasch Measurement Transactions*. 2011; 25(1): 1317.

- Mallinson T, Hammel J. Measurement of participation: intersecting person, task, and environment. *Archives of Physical Medicine & Rehabilitation*. 2010; 91(9 Suppl):S29–33.10.1016/j.apmr.2010.04.027 [PubMed: 20801276]
- McHorney CA, Haley SM, Ware JE Jr. Evaluation of the MOS SF-36 Physical Functioning Scale (PF-10): II. Comparison of relative precision using Likert and Rasch scoring methods. *Journal of Clinical Epidemiology*. 1997; 50(4):451–461. [PubMed: 9179104]
- McKenna K, Broome K, Liddle J. What older people do: Time use and exploring the link between role participation and life satisfaction in people aged 65 years and over. *Australian Occupational Therapy Journal*. 2007; 54:273–284.
- Pereira RB, Stagnitti K. The meaning of leisure for well-elderly Italians in an Australian community: implications for occupational therapy. *Australian Occupational Therapy Journal*. 2008; 55(1):39–46.10.1111/j.1440-1630.2006.00653.x [PubMed: 20887432]
- Raczek AE, Ware JE, Bjorner JB, Gandek B, Haley SM, Aaronson NK, et al. Sullivan M. Comparison of Rasch and summated rating scales constructed from SF-36 physical functioning items in seven countries: results from the IQOLA Project. *International Quality of Life Assessment. Journal of Clinical Epidemiology*. 1998; 51(11):1203–1214. [PubMed: 9817138]
- Scanlan JN, Bundy A. Development and validation of the modified occupational questionnaire. *American Journal of Occupational Therapy*. 2011; 65(1):e11.10.5014/ajot.2011.09042
- Wilcock, A. *An Occupational Perspective of Health*. 2nd. Thorofare, NJ: SLACK Incorporated; 2006.
- Wilson, M. *Constructing Measures: An Item Response Modeling Approach*. Mahwah, NJ: Lawrence Erlbaum Associates; 2005.
- Wright, BD.; Stone, M. *Best Test Design*. Chicago, IL: MESA Press; 1979.

Table 1**Sample Characteristics**

Characteristic	Total (n=460)
Demographic	
Age, mean years (SD)	74.9 (7.7)
Gender, No. (%)	
Women	303 (65.9)
Men	157 (34.1)
Race, No. (%)	
White	172 (37.4)
African American	149 (32.4)
Latino	92 (20.0)
Asian	18 (3.9)
Other	29 (6.3)
Primary Language, No. (%)	
English	404 (87.8)
Spanish	56 (12.2)
Education, No. (%)	
Less than high school	136 (29.6)
High school graduate	89 (19.4)
Some college or technical school	158 (34.4)
Four or more years of college	77 (16.7)
Retirement Status, No. (%)	
Retired/Active (student or volunteer)	111 (24.1)
Retired/Inactive	349 (75.9)
Living Situation, No. (%)	
Alone	377 (82.0)
With others	83 (18.0)

Table 2

Rasch Summary Psychometrics

Analysis	Items	Rating Scale Steps	Person Mean (SD) logits	RMSE	Adj. SD	SI	PSR	Number of Misfitting Items	PCA Eigenvalue 1 st contrast (%)
1. All items and rating scales	18	35	0.4 (.4)	0.2	0.4	1.8	.77	3	4.4 (14%)
2. Removal of six-month and reverse coded items	13	14	0.9 (.9)	0.4	0.8	2.2	.82	2	1.8 (7%)
3. Partial credit model	13	90	0.9 (.9)	0.4	0.8	2.2	.82	1	1.8 (7%)
4. Single rating scale, recoding steps 2,4,6 (1333557)	13	4	1.7 (1.6)	0.5	1.5	2.7	.88	0	1.8 (7%)

RMSE=Root mean square error

SI=Separation index

PSR=Person separation reliability

PCA=Principal components analysis

Table 3

Item Statistics in Order of Challenge

Items	Measure (calibration)	Std. Error	Infit MnSq	Infit Zstd	Outfit MnSq	Outfit Zstd
New experiences or adventure?	59.3	.69	1.2	3.0	1.3	4.2
Plan your activities to be health promoting?	51.7	.67	1.0	-.7	1.0	-.1
Need to or have to do your activities?	51.5	.75	1.2	3.5	1.4	4.7
Activities let you connect with other people?	50.9	.75	1.2	3.4	1.3	3.2
Activities provide you with hope for the future?	50.7	.68	1.0	.2	1.0	-.5
Knowledgeable about how your activities influence your health?	49.8	.68	0.8	-3.3	0.8	-3.1
Activities personally fulfilling for you?	49.6	.68	0.9	-2.3	0.9	-1.7
Activities provide you with a sense of accomplishment?	49.3	.69	0.7	-4.8	0.7	-4.1
Satisfied with your activities?	49.2	.69	1.0	-.3	1.0	-.7
Confident that you can influence your health?	45.8	.72	0.8	-3.5	0.8	-2.4
Like the activities that you do?	43.5	.74	0.9	-1.2	0.8	-2.1
Think your activities contribute to your physical health?	41.8	.84	1.0	2.6	1.0	.4
Think your activities contribute to your mental health?	40.2	.89	1.2	3.4	1.2	2.7

Table 4
Relative Precision Estimates of the ASPEn Scale, ASPEn Raw Score, and ASPEn Rasch Measure Using Depressive Symptoms and Volunteer Status as the Grouping Variables

Scale	Depressive Symptoms group means (SD)					Relative precision
	CESD<11 (n=244)	CESD 12 (n=216)	Mean difference	p-value	F-statistic	
ASPEn raw score	76.6 (11.7)	67.8 (14.9)	8.9	<0.001	50.59	1.00
ASPEnRasch measure	69.4 (12.7)	61.1 (12.5)	8.3	<0.001	49.16	0.97
Scale	Volunteer Status group means (SD)					Relative precision
	Not Volunteer (n=378)	Volunteer (n=81)	Mean difference	p-value	F-statistic	
ASPEn raw score	71.5 (14.0)	77.1 (13.4)	-5.6	<0.001	10.88	1.00
ASPEnRasch measure	64.6 (13.1)	69.8 (13.4)	-5.2	<0.001	10.44	0.96

ASPEn=Activity Significance Personal of Evaluation
CESD= Center for Epidemiologic Studies Depression-20 (CESD-20)