Self-Report Measure of Psychological Abuse of Older Adults

Corresponding Author:
Kendon J. Conrad, PhD
Professor of Health Policy and Administration
School of Public Health (MC 923)
University of Illinois at Chicago
1603 West Taylor Street
Chicago IL 60612-4394
Tel: (312) 996-3185 Fax: (312) 996-5356
E-mail: kjconrad@uic.edu

Madelyn Iris, PhD
Director, Leonard Schanfield Research Institute
CJE SeniorLife
3003 W. Touhy Ave.
Chicago, IL 60645
(773) 508-1076
micki.iris@cje.net

John W. Ridings, PhD
Metropolitan Family Services, Chicago
Director of Outcomes and Evaluation
One North Dearborn St.
Suite 1000
Chicago, IL 60614
(312) 986-4037
ridingsj@metrofamily.org

Kate Langley, MPH
School of Public Health
University of Illinois at Chicago
1603 West Taylor Street
Chicago IL 60612-4394
(312) 996-0194
kjlangley@gmail.com

Georgia J. Anetzberger, PhD, ACSW
Lecturer in the Health Care Administration Program
Department of Management and Labor Relations
Cleveland State University
2121 Euclid Avenue, BU 434
Cleveland, Ohio 44115

Tel: (216) 687-3835 Fax: (216) 687-9354
E-mail: g.anetzberger@csuohio.edu

Abstract word count: 214
Text word count: 5,488
Reference word count: 1,406

Running head: PSYCHOLOGICAL ABUSE OF OLDER ADULTS

Abstract

Purpose: This study tested key psychometric properties of the Older Adult Psychological Abuse Measure (OAPAM), one self-report scale of the Older Adult Mistreatment Assessment (OAMA).

Design and Methods: Items and theory were developed in a prior concept mapping study. Subsequently, the measures were administered to 226 substantiated clients by 22 elder abuse staff from seven agencies in a full-scale field test. The resulting database was used to estimate the psychometric properties of the OAPAM using the Rasch item response theory model and traditional validation techniques. Analyses included tests for dimensionality, model fit, and theoretical construct validation. Results from the OAPAM client report were validated against the APS substantiation decision of abuse and the elder abuse staff assessment of psychological abuse.

Results: The client self-report measures met stringent Rasch analysis fit and unidimensionality criteria; had high person (internal consistency) and item reliability. The validity results supported the usefulness of the client measures and led to reconsideration of aspects of the hypothesized theoretical hierarchy. A short form was developed. Cut-points were proposed to distinguish levels of psychological abuse. **Implications:** The measure is now available to aid in the assessment of psychological abuse of older adults by both clinicians and researchers. Theoretical refinements developed using the Rasch item hierarchy may help to improve assessment and intervention.

Keywords: emotional abuse, psychological abuse, elder mistreatment, Rasch measurement, abuse theory

Introduction

The purpose of this study was to develop and test a self-report measure of psychological abuse of older adults. The National Center on Elder Abuse defines emotional or psychological abuse as the infliction of anguish, pain, or distress through verbal or nonverbal acts (NCEA, 2003). Emotional/psychological abuse (terms used synonomously) includes, but is not limited to, verbal assaults, insults, threats, intimidation, humiliation, and harassment. In addition, treating an older person like an infant; isolating an older person from his/her family, friends, or regular activities; giving an older person the "silent treatment" and enforced social isolation are examples of emotional/psychological abuse (NCEA, 2003). Such treatment would typically occur in private and be difficult for third parties to detect.

A range of instruments that assess elder abuse have been developed over the past 20 years (Dyer & Goins, 2000; Canadian Task Force on the Periodic Health Examination, 1994; Mount Sinai, 1988; Fulmer & Cahill, 1984; Fulmer, Paveza, Abraham, & Fairchild, 2000; Reis & Nahmiash, 1998; Bass, Anetzberger, Ejaz, & Nagpaul, 2001). Most have considered multiple abuse forms, sometimes including psychological abuse, but without specific focus on conceptualizing and assessing psychological abuse. Further, most screening instruments rely on clinician assessments rather than self-report by older adults (Marshall et al., 2000), and are designed to evaluate quality of caregiving (e.g., Bravo et al., 1995), identify abusive caregivers of older adults (Reis & Nahmiash, 1995), or help health professionals detect problems (Fulmer, Ramirez, and Fairchild, 1999; Reis & Nahmiash, 1998; Wang, 2005, 2006). An example of a recently developed patient self report is the Elder Abuse Suspicion Index (Yaffe, Wolfson, Lithwick, & Weiss, 2008), a six-item physician to patient interview that includes a psychological abuse item.

In a systematic review of 49 studies of elder abuse (Cooper, Selwood, & Livingston, 2008), 6% of older adults reported significant abuse in the previous month and 5.6% of couples reported physical violence in their relationship in the previous year. These authors reported that nearly a quarter of the older adults reported significant levels of psychological abuse. Sixteen percent of nursing home staff admitted to significant psychological abuse of residents, and a third of family caregivers reported being involved in significant abuse. However, only a small proportion of this abuse was known to protective services. One in 6 professional caregivers reported committing abusive acts but over four-fifths reported observing them. Unfortunately, only 7 of the studies that were reviewed used measures for which any type of reliability and validity had been assessed (Cooper, Selwood, & Livingston, 2008). Cooper et al. concluded that valid, reliable measures and consensus on what constitutes an adequate standard for validity of abuse measures are needed.

The small amount of literature published exclusively on psychological abuse of older adults is understandable given the difficulty in developing a precise definition that would lead to valid and reliable measures. Additionally, any definition of psychological abuse may reflect a cultural perspective (Anetzberger, Korbin, & Tomita, 1996; Moon, Tomita, & Jung-Kamei, 2001). Furthermore, some believe that the meaning of psychological abuse is best represented not through any illustrative act, but rather through the perceived effect of the act on the victim, which then allows for consideration of cultural variation in definition (e.g., Nerenberg, 2008) and reinforces the importance of obtaining client self-reports.

Prevalence

Even though psychological abuse is believed to be under-reported (Cooper et al., 2008; Schofield & Mishra, 2003), the percentages of occurrence reported in extant studies indicate the pervasiveness of the problem. In 1988, Pillemer & Finkelhor conducted one of the few random sample studies of elder abuse, surveying 2020 community-dwelling elderly in the Boston area. Overall, they found a rate of abuse of 3.2%. However, they limited their questions regarding psychological abuse to verbal aggression only, for which they established a rate of 1.1 percent. Most recently, Acierno and colleagues (2010) conducted a national prevalence study, and based on a sample of 5777 older adults (60 and over), found a one-year prevalence rate of 4.6% for emotional abuse, the highest rate for any type of abuse queried. Even higher prevalence rates were found by Beach and associates (2010), in their investigation of financial exploitation and psychological mistreatment among African Americans and non-African Americans, in Allegheny County, PA. They reported significantly higher prevalence rates for psychological abuse of African American elders as compared to non-African Americans: 24.4% vs. 13.2% respectively.

In samples of abused older adults, Brownell, Berman, & Salmone (1999) found that among 402 cases of abuse of older adults, 54% involved psychological abuse; a similar study by Anetzberger (1998) revealed that 41% of incidents of abuse of older adults were psychological. Anetzberger (1998) also found that in cases where there was psychological abuse, additional forms of abuse were present 89.7% of the time, including physical neglect and financial exploitation. Similarly, the National Elder Abuse Incidence Study (1998) found a 35% prevalence rate; Lithwick, Beaulieu et al., (1999) found 87%; Vladescu, Eveliegh, Ploeg & Patterson (1999) and Godkin, Wolf, & Pillemer (1989) also reported high percentages (73% and 72%, respectively), though both studies had small samples. These mixed findings illustrate the difficulties in establishing a consistent prevalence rate for psychological

abuse. Differences in the definition and measurement of psychological abuse used by each study above may account for some discrepancies and variability.

Conceptual Models

The limited research on most forms of elder abuse, including psychological abuse, has lacked an overall conceptual framework to guide data collection efforts and provide effective assessment of the risk factors for and the consequences of different types of abuse. Godkin, Wolf, and Pillemer (1989) developed five conceptual components of abusive relationships. Anetzberger (2000) developed the Exploratory Model for Elder Abuse which examined characteristics of the perpetrator as the primary consideration, and secondarily, characteristics of the victim and the context in a temporal arrangement. The National Research Council's seminal book on elder abuse (2003) presents a structure, process, and outcome model that includes the socio-cultural context and the transactional processes among the parties leading to abuse.

These models have several commonalities; primary among them is that they recognize the importance of including the perpetrator and his/her characteristics as well as the social network. While the models are able to explain the etiology of general abuse, they do not present examples of items that represent psychological abuse, nor do they indicate which components are most important to elder abuse or which are most severe. Understanding these issues is essential to obtaining accurate assessments of types and levels of abuse.

Prior Study: Item Development

In the precursor of the present study (Conrad, Iris, Ridings, Fairman, & Rosen, in press), three-dimensional concept mapping (Trochim, 1989) was used to conceptualize psychological abuse of older adults. Statements were generated from literature review and by local and national panels consisting of 16 experts in the field of psychological abuse. These statements were sorted and

rated on a 1-5 scale for severity, using Concept Systems software which grouped the statements into clusters and depicted them as a map. The clusters represent the distinct conceptual areas of the overall domain of psychological abuse. Based on average ratings for all statements in a particular cluster, the clusters were then ranked in order of severity. These concepts in descending order of severity were: 1) isolation, 2) insensitivity and disrespect, 3) shaming and blaming, 4) threats and intimidation, and 5) trusted other risk factors. This hierarchy formed the basis for a measurement model of the construct of psychological abuse of older adults.

The statements developed for the concept map were subsequently framed as questions, and questionnaires were developed for both third party observation and client self-report. Third party observation included completion of the questionnaire by an elder abuse investigator, based on his/her understanding of the client's report, his/her observations while conducting the investigation, and any information obtained from others, including the alleged abuser. Nine focus groups were convened to review the wording of items and the formats of the questionnaires. Six focus groups were conducted with 44 staff members from elder abuse investigation/treatment provider agencies. Three groups were comprised of 20 consumers. The participants in the staff focus groups consisted of either naturally formed work groups (such as a team of elder abuse staff) or were participants in our earlier study. Groups of clients were formed based on having been served by the same agency. The meetings were held at several local, non-profit, agency and business locations. The focus group process consisted of a review of the psychological abuse items that were compiled. Participants were asked to read each item and evaluate its relevance to psychological abuse, its wording, and its clarity. They were also asked to review the ordering and formatting of the questions and to suggest additional items. The final items are provided in Appendix 1.

Cognitive interviews were conducted with four clients who were substantiated as having experienced elder abuse and who had not participated in the focus groups. Details of these focus groups and other qualitative work may be reviewed in the National Institute of Justice Report from this study (Conrad, Iris, & Ridings, 2009), which resulted in the Older Adult Psychological Abuse Measure (OAPAM), the client self-report measure. The OAPAM is one scale of the Older Adult Mistreatment Assessment (OAMA), which is now being developed as a comprehensive elder abuse assessment procedure (Conrad, Iris, Riley, Mensah, Mazza, 2009). The OAMA, in its current form, consists of third party observations and client self-report measures of financial exploitation and psychological abuse. In addition to demographics, it has draft versions of physical, sexual, and neglect assessments, including short screeners of all of the above types of abuse and descriptive information about alleged abusers.

Objectives

The specific objectives of the present full-scale field test of the OAPAM were:

1. To test the construct dimensionality of the OAPAM, i.e., Did the items form a single overarching psychological abuse construct?

2. To test the fit of the items to the Rasch measurement model, i.e., rating scale model;

3. To assess internal consistency reliability;

4. To develop short forms that would be user-friendly for clinical applications;

5. To examine appropriateness for the target population;

6. To test construct validity by positing a theoretical hierarchy of concept rankings that conforms to expectations developed in a prior research phase and by testing a set of hypothesized relationships using correlation analysis.

7. Propose a reasonable, although speculative given lack of external validation, cut-off to determine psychological abuse.

Design and Methods

Sample

Data collection was supported by a research agreement with the Illinois Department on Aging (IDOA) which advocated the recruitment of the elder abuse providers and clients for the project. With IDOA's support, recruitment was from seven adult protective services agencies in Chicago and its collar counties. Two samples were established: first, twenty-two highly experienced elder abuse staff members were recruited from these agencies. Since interviewing clients with a standardized questionnaire was not previously done as part of their screening procedures, the elder abuse staff members were trained in interviewing for this study by the two lead authors. The staff members also completed the human subjects subcommittee online training program of the University of Illinois at Chicago (UIC). The human subjects research proposal and informed consent forms were approved by the UIC institutional review board via the human subjects subcommittee. All 22 participating elder abuse staff members were volunteers and gave informed consent. Second, the elder abuse staff recruited and screened clients for ability to consent to research participation, and for their ability to serve as reliable reporters of abuse. A key component of the interview was the assessment of cognitive status using the Mini-Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975). To participate in the study, the client had to score at least 17 on the MMSE or in the judgment of the elder abuse investigator demonstrate adequate cognitive capacity to provide self-report. The elder abuse staff was responsible for obtaining clients' consent. In all, 226 clients consented and completed the OAPAM.

The 22 elder abuse staff members administered client self-report measures of psychological abuse via interview in the home to the 226 clients who were substantiated for at least one type of elder mistreatment and spoke English. They also completed a staff observation questionnaire for each of these clients. Recruitment was limited to only substantiated clients to be sure that the population was appropriate for the measures. However, they did not have to be substantiated for psychological abuse. This meant that there would likely be a substantial group in the "floor" to be sure there was representation of a full range of the construct and power for a yes/no cut-point.

Background Characteristics of the Elder Abuse Staff and Clients

The sample of 22 elder abuse staff was predominantly female (86.36%). More than half was Caucasian (59.09%), a quarter was African American (27.27%), and the remainder Hispanic or mixed race. The elder abuse staff's average years of on-the-job experience was 5.46 years.

The sample of 226 clients was also predominantly female (70.4%). The majority of clients was African American (61.3%), more than one third was Caucasian (35.5%), and the remainder was of mixed race or other. Most were non-Hispanic (92.9%). The majority of clients was between 75-90 years of age (58.7%).

Statistical Analysis

Traditional test theory counts the number of items endorsed and uses that as an estimate of the person's level on the construct of interest. The Rasch measurement model (Rasch, 1960) was chosen because of its desirable scaling properties of linear, interval measurement (Embretson & Reise, 2000). It places both persons and items on the same ruler. This is useful in judging which items persons are likely to endorse which is helpful in setting cutoff scores. The model provides an item hierarchy (seen in Figure 1) that is useful to support theory building and test construct validity. Therefore, the Rasch model was needed to test the theoretical hierarchy developed in

prior work. This is a type of construct validation. The Rasch model was also useful in testing unidimensionality, examining usefulness of the rating scale, and testing the fit of items to the model. These are also aspects of construct validation that can be facilitated with the Rasch model.

The Rasch rating scale model (Wright & Masters, 1982) estimates the probability that a respondent will choose a particular response category for an item as:

$$\ln \frac{P_{nij}}{P_{ni(j-1)}} = B_n - D_i - F_j ,$$

where P_{nij} is the probability of respondent n scoring in category j of item i, $P_{ni(j-1)}$ is the probability of respondent n scoring in category j-1 of item i, B_n is the person measure of respondent n, D_i is the difficulty of item I, and F_j is the difficulty of category step j. Rating scale categories are ordered steps on the measurement scale. Completing the j^{th} step can be thought of as choosing the j^{th} alternative over the $(j-1)^{th}$ in the response to the item.

Rasch analysis places persons (B_n) and items (D_i) on the same measurement scale (illustrated in Figure 1) where the unit of measurement is the logit (log odds unit). Person reliability in Rasch is analogous to Cronbach's alpha in traditional testing. It gives an idea of how stably persons are placed on the scale. Since Rasch places both persons and items on the same scale, reliability can be estimated for items as well as for persons. The Winsteps Computer Program was used for these calculations (Linacre, 2009).

Dimensionality. Since the Rasch model requires unidimensionality, principal component analysis of residuals was used to examine whether a substantial factor existed in the residuals after the primary measurement dimension had been estimated (Linacre, 1998a; Smith, E., 2002). Although there are no hard rules for interpreting principal components results, our rule of thumb for unidimensionality was variance explained of >40% by the measurement dimension (Linacre,

2006). For comparison, Reckase (1979) used 20% to define a substantial factor. To be conservative in testing a second dimension, <15% (even lower than Reckase) was set as the criterion for variance explained by the first principal component of the residuals, i.e., the second dimension. Simply put, using 40% and 15% variance as the criteria for the first and second dimensions is a rigorous test in that the measurement dimension must be large at 40%, while the second dimension must be quite small at under 15%. Dimensionality was also tested using Linacre's (1998b) procedure. Two subsets of items were extracted representing the opposite poles of the factor. Each subject was then measured on each subset of items. The subject measures were cross-plotted and correlation coefficients were obtained. Additional criteria for unidimensionality were employed using item fit statistics discussed next.

Quality control with fit statistics. Rasch analysis provides fit statistics to test assumptions of fundamental measurement (Wright & Stone, 1979). Understanding item misfit can lead to improving or dropping items. The following link provides a handy guide to interpreting fit statistics: http://www.rasch.org/rmt/rmt82a.htm. The Rasch model provides two indicators of misfit: infit and outfit. For this analysis, items with values below 1.33 mean square (MNSQ) on both infit and outfit were considered acceptable quality (Wilson, 2005; Smith, 2000)."

Rating scale. The proper functioning of the rating scale was examined using: 1) outfit mean-squares less than 2.0, 2) average measures advance monotonically with each category, and 3) step calibrations increase monotonically (Linacre, 1999; 2002; Zhu, Updike, & Lewandowski, 1997). Based on our focus group work, a "suspected" category was included as intermediate between "yes" and "no" where no=0, suspected=1 and yes=2. Given our prior experience, it was predicted that this would be a little used category that would not conform to the Rasch model, but it was included to be responsive to clinical input that said it was needed for greater sensitivity in

our measures.

For a complete treatment of Rasch analysis, see Bond & Fox (2007) which includes a glossary of Rasch measurement terminology. Terminology may also be accessed online via *Rasch Measurement Transactions* located at http://www.rasch.org/rmt/. The tables below were developed from Winsteps 3.67 (Linacre, 2009) with annotated explanations and interpretations.

Construct validation. In Rasch analysis, the item hierarchy that is created by the item difficulty estimates provides an indication of construct validity (Smith, 2001). The items should form a ladder with low severity symptoms on the bottom to high severity symptoms on the top. In prior work (Conrad, Iris, Ridings, Fairman, & Rosen, in press), 16 experts grouped the items into six groups and rated the severity of the items on a scale from 1-5. These item severities were then averaged within each group. The result was a theoretical hierarchy of five conceptual components of psychological abuse arranged in descending severity (Table 1) as follows (mean expert ranking from 1-5 in parentheses): isolation (1), threats and intimidation (2), insensitivity and disrespect (3), shaming and blaming (4), and trusted other risk factors (5). To test whether this hierarchy was validated by the client respondents in this study, the Rasch calibration on each item was obtained, and these were subsequently averaged within each item grouping to see if the hierarchy would remain the same, i.e., "client groups" rankings were compared to the rankings of the "expert groups."

Multi-trait, multi-method analysis. Construct validation also may be tested by setting up a pattern of theoretical expectations and testing whether those expectations are supported by the data (Campbell & Fiske, 1959). As Campbell and Fiske pointed out, measures of the same construct should be highly correlated and especially so if they use the same method.

The Illinois Department on Aging (IDoA) questionnaire, which is required by IDoA for elder abuse investigations, covers many forms of elder abuse, including psychological abuse. The IDoA form also asks staff members to give a closing status on the case, identifying which types of abuse are substantiated. This closing status substantiation decision on psychological abuse was used to correlate with the OAPAM. The OAPAM was also correlated with OAMA staff data from the 22 elder abuse staff who reported their psychological abuse observations on the 226 substantiated clients. The Older Adult Psychological Abuse Measure (OAPAM), involved these 226 clients providing self-reports on an interview questionnaire. These are described as follows:

1) *Client Gender:* coded male=0, female=1

2) *Psychological Abuse Substantiation Decision of Illinois Dept. on Aging:* Psychological abuse was considered substantiated if it was marked as "verified" or "some indication."

3) *OAMA Staff Psychological Abuse Measure:* The Rasch person reliability was high at .87 which corresponded with the Cronbach's alpha of .92. The Rasch item reliability was very high at .96. The final 53 items of Staff-reported Psychological Abuse met stringent Rasch analysis fit and unidimensionality criteria.

4) OAMA Older Adult Psychological Abuse Measure (OAPAM): Details are described in the results section.

The direction and strength of construct pairs depends on method and theoretical expectations. A pattern of expected correlations roughly corresponding to Cohen's guidelines (1988, 1992) was set up as follows: NS=non-significant, >.1=low, >.3=moderate, and > .5=high. Others have suggested lower values based on reviews of research, e.g., >.2=moderate, and >.3=high (Hemphill, 2003), so there are no absolute guidelines available. This hypothesized pattern and resulting correlations are in the upper right half of Table 2. The diagonal entries are

the person reliabilities. The hypothesized correlations are stated above each correlation coefficient in the table and are bulleted below:

1) *Client Gender:* There was no reason to expect differential exploitation by gender so all gender correlations were expected to be non-significant (NS).

2) *Psychological Abuse Substantiation Decision:*

- Moderate correlation with OAMA Client Psychological Abuse
- High correlation with OAMA Staff Psychological Abuse

3) *OAMA Staff Psychological Abuse:*

- High correlation with OAMA Client Psychological Abuse

In the multi-trait, multi-method analyses, the most complete versions of all OAMA measures were used.

Short form. For the OAPAM to be most useful in both research and clinical settings, a short form would be required. In developing the short form, all 31 items were viewed as valid, and our principal inclusion criterion was representation of the items across the full range of item calibrations. To delete items, more stringent fit criteria were applied, i.e., either (rather than both) infit or outfit greater than 1.33 would qualify the item for possible deletion. However, some items with high outfit (less of a concern than infit) were still included if they were needed to cover the full range or to prevent gaps along the ruler.

Results

In this section each objective is restated in a header with the accompanying findings.

Test the Fit of the Items and Rating Scale

No items were dropped because they all met our criteria for fit. Specifically, both infit and outfit were less than 1.33 on all items. The rating scale performed as expected with the "unsure" category being least used.

Test Construct Dimensionality

The raw variance explained by the measure was 43.1%. This was a large amount, beyond the 40% criterion, and was supportive of a strong principal measurement dimension. The unexplained or residual variance that was explained by the first contrast was a small 10.5%. This, along with the well fitting items, suggested that there was not a substantial rival dimension. This was supportive of unidimensionality. The correlation of the first and second factors using Linacre's procedure (1998b) was .729. This was also supportive of unidimensionality.

In Figure 1, the annotated Rasch ruler, known as a Wright map, is displayed. Persons are arrayed on the left of the dashed line and items on the right (item numbers with item abbreviations are used on the Wright map and in the text). The items form a hierarchy of severity with lower severity items at the bottom and higher severity items at the top. The persons are also displayed according to their measure on the psychological abuse scale. There is a substantial floor of persons at the bottom who are not registering any client-reported psychological abuse. The concept that each item belongs to is indicated in brackets at the end of the item label, i.e., ISO=Isolation, T&I=Threats and Intimidation, I&D=Insensitivity and Disrespect, and S&B=Shaming and Blaming. Only the ISO concept had a coherent cluster of items which was located at the high severity end of the hierarchy. The other concepts were composed of items that were not located together at the same severity level, but were spread throughout the rest of the severity hierarchy. The two items, UncomfortableW/AA and AfraidOfAA, which had formed the Risk Factor cluster, were regrouped with the T&I cluster because of their unexpectedly high severity.

Assess Internal Consistency Reliability Using a Standard of .80

The Rasch person reliability was very high at .86 which corresponds with the Cronbach's alpha of .92. The Rasch item reliability was also very high at .97. The final 31 items of the OAPAM met stringent Rasch analysis fit and unidimensionality criteria. The measure as a whole had high person and item reliability.

Develop Short Forms That Would Be User-friendly for Clinical Applications

To test if a more parsimonious model would also function well, a shorter form was developed containing 18 items. Appendix 1 contains the items by form and factor information. While the short form is most useful, the longer form provides a bank of items that may be used in future development of alternative forms or computerized adaptive tests.

The final 18 items of Client-reported Psychological Abuse met stringent Rasch analysis fit and unidimensionality criteria and maintained the measurement range of the 31 item ruler. The Rasch person reliability for the 18-item form was still reasonably high at .78 which corresponded with the Cronbach's alpha of .87. The Rasch item reliability was very high at .96.

Examine the Appropriateness of the Measure for the Target Population

Although the persons in the floor were included on the Wright map (Figure 1), they were not included in the calculation of the person mean (-.59). This was reasonably well targeted since the person mean was within 1.0 logit and within one SD (.93) of the item mean of zero.

Test Construct Validity with a Hierarchy of Concept Rankings and Hypothesized Relationships

Looking at Table 1, "Original Concept Group," the ordering of the conceptual components of psychological abuse was the same for both experts, averaging their concept map ratings, and clients, averaging their Rasch measurement calibrations. This was supportive of the construct

validity of the measure. The item by item details of the expert concepts and rankings as well as the client item calibrations are located in the appendix.

Multi-trait, multi-method analysis of hypothesized relationships. It was hypothesized that all gender correlations would be NS and the three correlations were (Table 2). The other three correlations, two high and one moderate, were as hypothesized. This was supportive of the criterion validity of the OAPAM.

Identify an Appropriate Cut-off to Determine Psychological Abuse (PA)

Since there is no solely empirical way to determine a cut-point, e.g., using the Wright map (Figure 1), the logic of the cut-point decision is described in the discussion below.

Discussion

A measure consisting of 31 items was validated as a unidimensional measure of client-reported psychological abuse. Subsequently, a shorter form consisting of 18 items was developed. It is notable that only 97 clients (43%) in the sample had some indication of psychological abuse using IDOA criteria, but this IDOA designation lacked specifics about how the decision was arrived at or what it means. However, in Figure 1, the Wright map, there were 189 clients (84%) that endorsed at least one symptom of abuse. The persons are represented by the pound signs (three persons) and dots (one person) to the left of the vertical dashed line. Three persons endorsed all of the symptoms, i.e., in the "ceiling," with a definite "yes." Thirty-seven persons were in the floor, i.e., endorsing 0 symptoms. Above -1.0 on the ruler, the item meanings, i.e., severity of the symptoms going up the scale, and locations indicate that this may be a useful cutoff score for psychological abuse. Above this -1.0 logit level were 126 persons (57%) that were likely to endorse symptoms such as *23Manipulated*, *19SworeOrYelled*, *16HurtEsFeelings*, and *8UncomfortableW/AA*. These persons scored 12 or more of a possible 62 raw score. If 0 on the

ruler is used as the higher criterion for more serious psychological abuse, there were 52 persons (24%) above this level having even more severe symptomatology such as *22MadeFeelSmall*, *25TalkedAsIfNotThere*, *31MadeAshamed*, and *28DelibConfused*. Above 1.0 on the ruler (16 persons, 7%) could be classified as extreme psychological abuse since the four items above 1.0, i.e., *10Confined*, *13DepriveAsstvdevice*, *24ManipW/drugs*, and *11PreventContactOutsd*, all involve serious psychological isolation, deprivation and manipulation that border on or may include physical abuse and/or neglect. Such abuse may have serious, e.g.. depression, long-lasting and even life-threatening sequelae.

Multi-trait, Multi-method Construct Validation

As hypothesized, client gender was not significantly related to any indicators. The OAMA correlations alone were consistent with theoretical expectations. Therefore, based on their concurrence with theoretical expectations, the construct validity of the OAPAM was supported.

Concept Analysis

The concepts of psychological abuse were ranked the same by both the experts and by the client Rasch calibrations (Table 1). This was supportive of construct validity. However, the middle three concepts Threats and Intimidation (T&I), Insensitivity and Disrespect (I&D), and Shaming and Blaming (S&B) were so close in average rank, i.e., within one standard error (SE=.52) that this ranking may not be reliable. Looking at Figure 1, the Wright Map, the Isolation concept clearly had the most severe items (high on the ruler/map). However, the rest of the concepts have their items interspersed throughout the ruler without discernable lines of demarcation. The item *9AfraidOfAA* was fairly high on the severity ruler, i.e., at - .36. This item and *8UncomfortableW/AA* were originally classified as the "Risk Factor" concept. However, such a high calibration was indicative of something more serious than a risk factor so these items were

reclassified into Threats and Intimidation. This was logical since the items, i.e., "uncomfortable with" and "afraid of," can be interpreted as sequelae of threats and intimidation.

The major point that was taken from this conceptual analysis was that Isolation is clearly the most serious type of psychological abuse since it may border on or include physical abuse such as physical and chemical restraints. The other three types, Threats and Intimidation, Insensitivity and Disrespect, and Shaming and Blaming do not form a clear hierarchy as concepts, i.e., each concept is not at a distinct severity level. Rather, the items within each concept vary greatly in severity.

Limitations

While this was the largest sample of substantiated elder abuse clients that was found, it was still limited to 7 agencies in the Chicago area. New measures always require further validation; that includes this one. Ongoing validation of the Rasch-derived theoretical hierarchy, and the cutoff scores proposed here will be needed to understand its most appropriate uses.

Strengths

The OAPAM was developed with expert and client input involving 83 informed stakeholders (Conrad, Iris, Ridings, 2009); data were then collected on 226 substantiated clients and analyzed. The results were supportive of the validity of using the Older Adult Psychological Abuse Measure (OAPAM) in helping to assess the existence and the level of psychological abuse of older adults who are able to self-report using a Mini-Mental State (Folstein, Folstein, & McHugh, 1975) exam score of at least 17 or investigator judgment as the criterion for adequate cognitive capacity.

From a theoretical perspective, this work has classified items into four types of psychological abuse of older adults: Isolation, Threats and Intimidation, Insensitivity and

Disrespect, and Shaming and Blaming. Despite the limitations and need for further development, these items, used as long and short forms, should help to open the neglected area of psychological abuse of older adults for improved services and research. This OAPAM can be widely useful in elder abuse research and practice since there had been no validated client-report measures, and self-report by the alleged victim of his/her internal mental state is an important, some might say essential, indicator of abuse.

The measures provide empirically-derived and theoretically supported gradations along the continuum of psychological abuse severity that can enable better decision-making by clinicians and supervisors. With standardization, decisions will not be so dependent on the staff's training, experience, and idiosyncracies. With further development of validated cutoff scores, cases may be triaged more effectively into appropriate interventions.

Future Directions

This study is part of a program of research that is developing parallel third party measures that may be used by elder abuse staff as well as other reporters such as police, family members, and neighbors. Obtaining information from multiple sources is a good way to cross-validate reports as well as to discover additional information that may be lacking from an individual. This type of triangulation of data is key to accurate assessment, intervention, and adjudication. It should help to improve estimates of prevalence and to study the correlational and causal relationships that will help professionals to understand better and to ameliorate elder abuse.

References

Acierno, R., Hernandez, M.A., Amstadter, A.B., Resnick, H.S., Steve, K., Muzzy, W., and Kilpatrick, D.G. (2010). Prevalence and correlates of emotional, physical, sexual, and financial abuse and potential neglect in the United States: The national elder mistreatment study. *American Journal of Public Health, 100*(2), 292-297.

Anetzberger, G. (1998). *Psychological abuse and neglect: A cross-cultural concern to older Americans. Understanding and Combating Elder Abuse in Minority Communities*, Archstone Foundation, 141-151.

Anetzberger, G. (2000). Caregiving: Primary cause of elder abuse? *Generations, 24*(2), 46-51.

Anetzberger, G. J., Korbin, J. E., & Tomita, S. K. (1996). Defining elder mistreatment in four ethnic groups across two generations. *Journal of Cross-Cultural Gerontology, 11*, 187-212.

Bass, D.M., Anetzberger, G.J., Ejaz, F.K., & Nagpaul, I. (2001). Screening tools and referral protocol for stopping abuse against older Ohioans: A guide for service providers. *Journal of Elder Abuse & Neglect*, *13*(2), 23-38.

Beach, S.R., Schulz, R., Castle, N.G., and Rosen, J. (2010). Financial exploitation and psychological mistreatment among older adults: Differences between African Americans and non-African Americans in a population-based survey. *The Gerontologist,* first published online July 22, 2010, doi:10.1093/geront/gnq053

Bond, T.G. & Fox, C.M. (2007). *Applying the Rasch Model: Fundamental measurement in the human sciences*. (*2nd Ed*.) Mahwah, NJ: Erlbaum Associates.

Bravo, G., Girouard, D., Gosselin, S., Archambualt, C., & Dubois, M. (1995). Further validation of the QUALCARE scale. *Journal of Elder Abuse & Neglect, 7*(4), 29-48.

Brownell, P., Berman, H.J., & Salmone, A. (1999). Mental health and criminal justice issues among perpetrators of elder abuse. *Journal of Elder Abuse & Neglect, 11*(4), 81-94.

Burgess, A.W. (2006). *Elderly victims of sexual abuse and their offenders: Report to National Institute of Justice*, NIJ 216550.

Campbell, D.T & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2): 81-105.

Canadian Task Force on Periodic Health Examination. (1994). Periodic health examination, 1994 update: 4. Secondary prevention of elder abuse and mistreatment. *Canadian Medical Association Journal, 151*(10), 1413-1420.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159.

Conrad, K.J., Iris, M., & Riley, B.B., Mensah, E., Mazza, J. (2009). Developing End-User Criteria and a Prototype for an Elder Abuse Assessment System, National Institute of Justice Project # 2009-IJ-CX-0202.

Conrad, K.J., Iris, M., & Ridings, J.W. (2009). Conceptualizing and measuring financial exploitation and psychological abuse of elderly individuals, NCJ 228632, National Institute of Justice (197 pages). http://www.ncjrs.gov/pdffiles1/nij/grants/228632.pdf

Conrad, K.J., Iris, M., Ridings, J.W., Fairman, K. & Rosen (in press). Conceptual model and map of psychological abuse of older adults. *Journal of Elder Abuse and Neglect.*

Cooper, C., Selwood, A., & Livingston, G. (2008). The prevalence of elder abuse and neglect: A systematic review. *Age and Aging, 37*(2), 151-160.

Dyer, C.B., Connolly, M.T., & McFeeley, P. (2003). The clinical and medical forensics of elder abuse and neglect. In *National Research Council, Elder mistreatment: Abuse, neglect, and exploitation in an aging America* (pp. 339-381). Washington, DC: The National Academies Press.

Dyer, C. B., & Goins, A. M. (2000). The role of interdisciplinary geriatric assessment in addressing self-neglect of the elderly. *Generations, 24*(2), 23-27.

Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologists.* Mahwah, N.J.: Erlbaum Associates.

Folstein MF, Folstein SE, & McHugh PR (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician". *Journal of Psychiatric Research, 12*(3), 189-198.

Fulmer, T., & Cahill, V. M. (1984). Assessing elder abuse: A study. *Journal of Gerontological Nursing, 10* (12), 16-20.

Fulmer, T., Paveza, G., Abraham, I., & Fairchild, S. (2000). Elder neglect assessment in the emergency department. *Journal of Emergency Nursing, 26*(5), 436-443.

Fulmer T., Ramirez M., Fairchild S., Holmes, D., Koren, M., & Teresi, J. (1999). Prevalence of elder mistreatment as reported by social workers in a probability sample of adult day health care clients. *Journal of Elder Abuse and Neglect, 11*(3), 25–36.

Godkin, M., Wolf, R., & Pillemer, K. (1989). A case-comparison analysis of elder abuse and neglect. *International Journal of Aging and Human Development, 28*(3), 207-225.

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation co-efficients. *American Psychologist*, 58, 78-79.

Lachs, M.S. & Pillemer, K. (1995). Abuse and neglect of elderly persons. *The New England Journal of Medicine, 332*(7), 437-443.

Linacre, J.M. (2009). *Winsteps Rasch Measurement* (Version 3.64.2). www.winsteps.com. Author.

Linacre J.M. (2006). Data variance explained by measures. *Rasch Measurement Transactions, 20* (1), 1045.

Linacre, J.M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*(1), 85-106.

Linacre, J.M. (1999). Investigating rating scale category utility. *Journal of Applied Measurement, 3*,103-122.

Linacre, J.M. (1998a). Detecting multidimensionality: Which residual data-type works best? *Journal of Outcome Measurement, 2*(3), 266-283.

Linacre, J.M. (1998b). Structure in Rasch residuals: Why principal components analysis (PCA)? *Rasch Measurement Transactions,1*(2), 636.

Lithwick, M., Beaulieu, M., Gravel, S., & Straka, S. M. (1999). The mistreatment of older adults: Perpetrator-Victim relationships and interventions. *Journal of Elder Abuse & Neglect, 11*(4), 95-112.

Marshall, C. E., Benton, D., & Brazier, J. M. (2000). Elder abuse: Using clinical tools to identify clues of mistreatment. *Geriatrics, 55*(2), 42-53.

Moon, A., Tomita, S.K., & Jung-Kamei, S. (2001). Elder mistreatment among four Asian American groups: An exploratory study on tolerance, victim blaming and attitudes toward third-party intervention. *Journal of Gerontological Social Work, 36*(1-2), 153-169.

Mount Sinai/Victim Services Agency Elder Abuse Project. (1988). *Elder mistreatment guidelines for health care professionals: Direction, assessment and intervention*. New York: Author.

National Center on Elder Abuse. (1998). *The National Elder Abuse Incidence Study: Final report*. Washington, DC: Author.

National Center on Elder Abuse. (2003). *The basics: Major types of abuse.* Retrieved February 18, 2007, from http://www.elderabusecenter.org/default.cfm?p=basics.cfm

National Research Council. (2003). *Elder mistreatment, abuse, neglect, and exploitation in aging America. Panel to review risk and prevalence of elder abuse and neglect.* R. Bonnie & R.Wallace, (Eds.) Washington D.C.: National Academies Press.

Nerenberg, L. (2008). *Elder abuse prevention: Emerging trends and promising strategies*. New York: Springer Publishing Company.

Pillemer, K., & Finkelhor, D. (1988). The prevalence of elder abuse: A random sample survey. *Gerontologist, 28,* 51-57.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests* (Reprint, with Foreword and Afterword by B. D. Wright, Chicago: University of Chicago Press, 1980). Copenhagen, Denmark: Danmarks Paedogogiske Institut.

Reckase, M. (1979). Unifactor latent trait model applied to multifactor tests: Results and implications. *Journal of Educational and Behavioral Statistics, 4*(3), 207-230.

Reis, M. & Nahmiash, D. (1998). Validation of the indicators of abuse (IOA) screen. *The Gerontologist, 28*(4), 471-480.

Reis, M., & Nahmiash, D. (1995). Validation of the caregiver abuse screen (CASE). *Canadian Journal on Aging, 14,* 45-60.

Schofield, M & Mishra, G. (2003). Validity of self-report screening scale for elder abuse: Women's health Australia Study. *The Gerontologist, 43*(1),110-120.

Smith, E.V. (2001). Evidence for the reliability of measures and validity of measure interpretation: A Rasch measurement perspective. *Journal of Applied Measurement, 2*(3), 281-311.

Smith, E.V. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement, 3*(2), 205-231.

Smith, R.M. (2000). *Common oversights in Rasch studies: MESA note 9.* Retrieved March 3, 2009 from: http://www.rasch.org/rn9.htm.

Trochim, W. (1989). An introduction to concept mapping for planning and evaluation. *Evaluation and Program Planning*, *12*(1), 1-16.

Vladescu, D., Eveleigh, K., Ploeg, J . & Patterson, C. (1999). An evaluation of a client-centered management program of elder abuse. *Journal of Elder Abuse & Neglect*, *11*(4), 5-22.

Wang, J.J. (2005). Psychological abuse behavior exhibited by caregivers in the care of the elderly and correlated factors in long-term care facilities in Taiwan. *Journal of Nursing Research, 13*(4), 271-80.

Wang, J.J. (2006). Psychological abuse and its characteristic correlates among elderly Taiwanese. *Archives of Gerontology Geriatrics*, *42*(3), 307–18.

Wilson, M. (2005). *Constructing measures: An item response modeling approach.* Mahwah, NJ: Erlbaum Associates.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement.* Chicago, Illinois: MESA Press.

Wright, B.D., & Stone, M.H. (1979). *Best test design.* Chicago: University of Chicago, MESA Press.

Yaffe MJ, Wolfson C, Lithwick M, Weiss D. (2008). Development and validation of a tool to improve physician identification of elder abuse: The Elder Abuse Suspicion Index (EASI). *Journal of Elder Abuse and Neglect, 20*(3), 276-300.

Zhu, W., Updike, W. F., & Lewandowski, C. (1997). Post-hoc Rasch analysis of optimal categorization of an ordered response scale. *Journal of Outcome Measurement, 1*(4), 286-304.

Table 1.

Expert Item Groups and Rankings Compared with Client Factors and Rankings

Expert Concepts and Ranks with Rasch Measures

Expert Concept Rank	Expert Concept Name	Expert Groups Average Rasch Measure[a]	Client Concept Rank
1	Isolation (ISO)	0.688	1
2	Threats & Intimidation (T&I)	-0.024	2
3	Insensitivity & Disrespect(I&D)	-0.036	3
4	Shaming & Blaming (S&B)	-0.315	4
5	Risk Factors	-0.585	5

Expert and client rankings were the same, but the middle 3 were so close in average rank, i.e., within one standard error (SE=.52), that this ranking may not be reliable.

Two Risk Factor items involving fear of abuser were reclassified as T&I.

[a] Based on the client endorsement of the items but using the items as grouped by the experts. To calculate the average measures, the item calibrations were summed, i.e., where items are located on the ruler in Figure 1, and divided by the number of items in that group, e.g., 7 ISO items. Since most of the ISO items are located high on the ruler, the ISO group/concept has the highest severity.

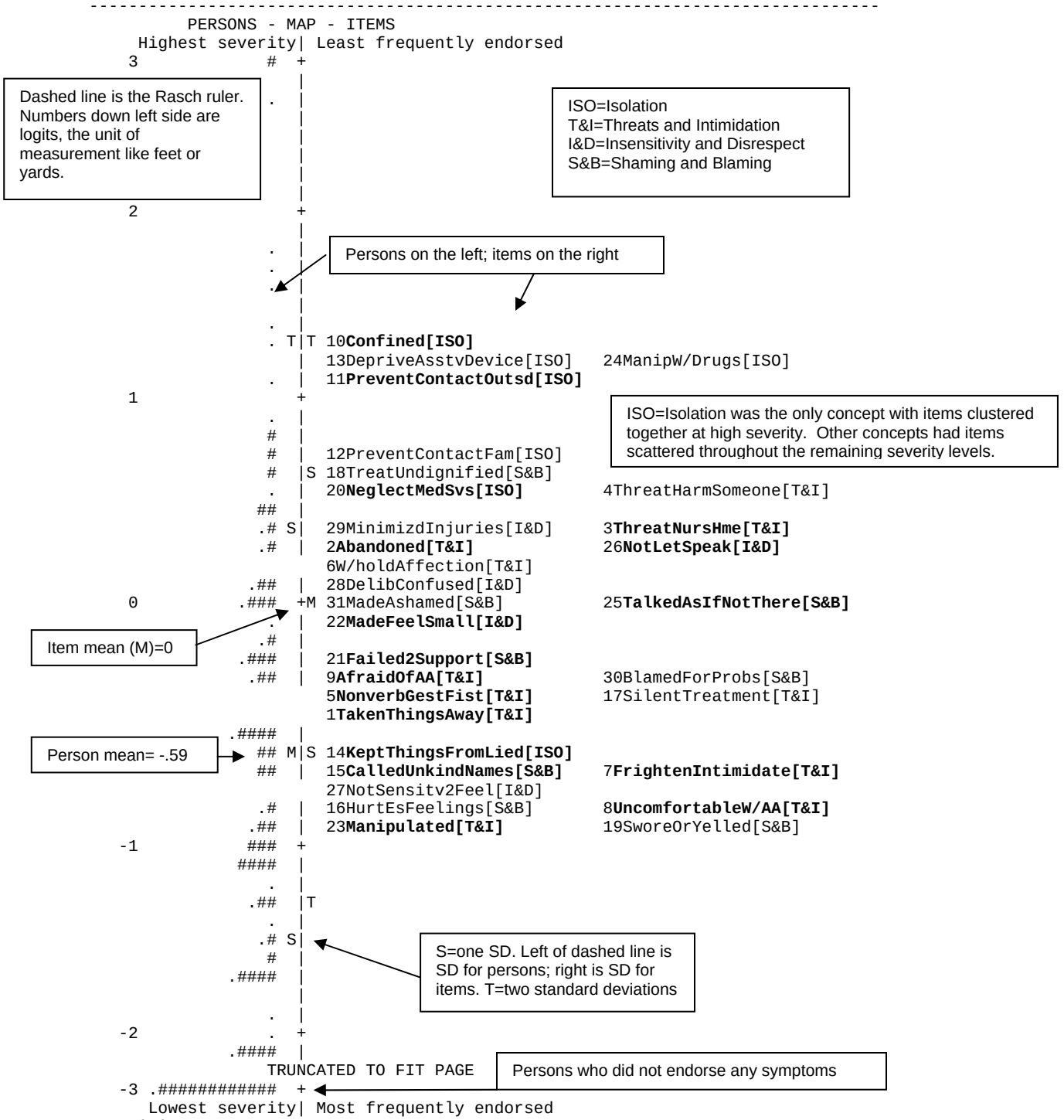
Table 2

Hypothesized and Actual Correlations[a] of OAPAM with Gender, Substantiation Decision, and Staff Psychological Abuse (PA) Assessment

	Client Gender	Psych Abuse Sub. Decision (IDoA)	OAMA Staff PA	OAMA Client PA
Client Gender M=0, F=1	--	NS -.042	NS -.076	NS .026
Em. Abuse Sub. Decision (IDoA)		--	High .478**	Mod .360**
OAMA Staff PA			.87[b]	High .700**
OAMA Client PA				.86

[a] Hypothesized correlations: NS = non-significant, >.1= low, >.3 = moderate, and >.5 = high are listed above the actual correlations.
[b] Person reliabilities of OAMA scales are located on the diagonal.
** Correlation is significant at the 0.01 level (2-tailed)

Figure 1. Wright Map of persons and items on the Rasch ruler of client-reported psychological abuse (item #'s keyed to appendix)



Appendix 1

Final Scale and Item Information for Client Psychological Abuse (item #'s keyed to Figure 1)
Response categories are no=0, suspected=1 and yes=2.

Item Number, Full Item, and Item Abbreviation (number and abbreviation are same as in Figure 1)	**Concept Name (from expert panel)**	**Rasch Measure (from client data)**	**Results of Analyses**
Prompt "In the past 12 months , has (NAME OF ALLEGED ABUSER)":		*Measures based on 31 item analysis*	*Blank means included in 31 item analysis*
1. Taken things away or threatened to take things away from you? **(TakenThingsAway)**	Threats and Intimidation	-.44	**Short Form Item**
2. Abandoned or threatened to abandon you? **(Abandoned)**	Threats and Intimidation	.24	**Short Form Item**
3. Threatened to place you in a nursing home when it was not appropriate? **(ThreatNursHme)**	Threats and Intimidation	.33	**Short Form Item**
4. Harmed or threatened to harm someone or something close to you (kids, pets, etc.)? (ThreatenHarmSomeone)	Threats and Intimidation	.54	
5. Used non-verbal behavior such as shaking a fist, pushing, poking, or slapping, to threaten or scare you? **(NonverbGestFist)**	Threats and Intimidation	-.36	**Short Form Item**
6. Manipulated you by withholding affection and love? (WithholdingAffection)	Threats and Intimidation	.17	
7. Behaved in ways that frighten or intimidate you? **(FrightenIntimidate)**	Threats and Intimidation	-.66	**Short Form Item**
In the past 12 months:			
8. Have you been uncomfortable with _____? **(UncomfortableW/AA)**	Risk Factors reclassified as T&I	-.81	**Short Form Item**
9. Have you been afraid of _____? **(AfraidOfAA)**	Risk Factors reclassified as T&I	-.36	**Short Form Item**
In the past 12 months , has NAME ALLEGED ABUSER:			
10. Confined you against your will? **(Confined)**	Isolation	1.31	**Short Form Item**
11. Prevented you from having contact with the	Isolation	1.08	**Short Form Item**

Item Number, Full Item, and Item Abbreviation (number and abbreviation are same as in Figure 1)	Concept Name (from expert panel)	Rasch Measure (from client data)	Results of Analyses
outside world via telephone, newspapers, television, or radio, etc.?. **(PreventContactOutsd)**			
12. Prevented you from contacting family, friends, or community resources? (PreventedContactFamily)	Isolation	.66	
13. Deprived you of glasses, hearing aids, prosthetics, walker, wheelchair, or any other assistive devices that you needed? (DeprivedOfAssistiveDevices)	Isolation	1.2	
14. Kept things from you or lied about things that you should know about? **(KeptThingsFromEldOrLied)**	Isolation	-.64	**Short Form Item**
15. Called you unkind names or put you down? **(CalledUnkindNames)**	Shaming & Blaming	-.66	**Short Form Item**
16. Deliberately made you feel bad or hurt your feelings? (HurtEldFeelings)	Shaming & Blaming	-.79	
17. Given you the silent treatment? (SilentTreatment)	Threats & Intimidation	-.35	
18. Treated you in an undignified or inappropriate way while assisting you with dressing, eating, bathing and so on? (TreatEldUndignifiedWay)	Shaming & Blaming	.60	
19. Sworn or yelled at you? (SworeOrYelled)	Shaming & Blaming	-.92	
20. Refused or neglected to get medical services that you needed? **(NeglectMedSvs)**	Isolation	.52	**Short Form Item**
In the past 12 months:			
21. Has_____ failed to support you or back you up when you needed it? **(Failed2Support)**	Shaming & Blaming	-.26	**Short Form Item**
In the past 12 months, has the ALLEGED ABUSER:			
22. Made you feel small, for example, treated you like a child? **(MadeFeelSmall)**	Insensitivity & Disrespect	-.07	**Short Form Item**

Item Number, Full Item, and Item Abbreviation (number and abbreviation are same as in Figure 1)	Concept Name (from expert panel)	Rasch Measure (from client data)	Results of Analyses
23. Manipulated or tried to control you in any way? **(Manipulated)**	Threats & Intimidation	-.91	**Short Form Item**
24. Manipulated you with drugs or alcohol? (ManipulatedWithDrugs)	Isolation	1.2	
25. Talked about you as if you were not there? **(TalkedAsIfNotThere)**	Shaming & Blaming	-.04	**Short Form Item**
26. Not let you speak for yourself? **(NotLetSpeak)**	Insensitivity & Disrespect	.21	**Short Form Item**
27. Not been sensitive to your feelings? (NotSensitiveFeelings)	Insensitivity & Disrespect	-.71	
28. Deliberately confused you? (DeliberatelyConfused)	Insensitivity & Disrespect	-.09	**Short Form Item**
29. Minimized your injuries or complaints? (MinimizedInjuries)	Insensitivity & Disrespect	.3	
30. Blamed you for their problems? (BlamedForProblems)	Shaming & Blaming	-.41	
31. Said something about you that made you feel ashamed? (MadeEldFeelAshamed)	Shaming & Blaming	-.04	