Self-report Measure of Psychological Abuse of Older Adults

Kendon J. Conrad, PhD,*1 Madelyn Iris, PhD,2 John W. Ridings, PhD,3 Kate Langley, MPH,1 and Georgia J. Anetzberger, PhD, ACSW4

1School of Public Health, Department of Health Policy and Administration, University of Illinois at Chicago.  
2Leonard Schanfield Research Institute, CJE SeniorLife, Chicago, Illinois.  
3Metropolitan Family Services, Department of Outcomes and Evaluation, Chicago, Illinois.  
4Department of Management and Labor Relations, Cleveland State University, Ohio.

*Address correspondence to Kendon J. Conrad, PhD, School of Public Health, University of Illinois at Chicago, 1603 West Taylor Street, Chicago IL 60612-4394. E-mail: kjconrad@uic.edu

Received August 4, 2010; Accepted October 29, 2010  
Decision Editor: William J. McAuley, PhD

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 7 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 adult protective services substantiation decision of abuse and the elder abuse staff assessment of psychological abuse (PA). Results: The client self-report measures met stringent Rasch analysis fit and unidimensionality criteria and 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 PA. Implications: The measure is now available to aid in the assessment of PA of older adults by both clinicians and researchers. Theoretical refinements developed using the Rasch item hierarchy may help to improve assessment and intervention.

Key Words: Emotional abuse, Psychological abuse, Elder mistreatment, Rasch measurement, Abuse theory

The purpose of this study was to develop and test a self-report measure of psychological abuse (PA) of older adults. The National Center on Elder Abuse (NCEA) defines emotional or PA as the infliction of anguish, pain, or distress through verbal or nonverbal acts (NCEA, 2003). Emotional/PA (terms used synonymously) 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 or her family, friends, or regular activities; and giving an older person the “silent treatment” and enforced social isolation are examples of emotional/PA (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 (Bass, Anetzberger, Ejaz, & Nagpaul, 2001; Canadian Task Force on the Periodic Health Examination, 1994; Dyer & Goins, 2000; Fulmer & Cahill, 1984; Fulmer, Paveza, Abraham, & Fairchild, 2000; Mount Sinai/Victim Services Agency Elder Abuse Project, 1988; Reis & Nahmiash, 1998). Most have considered multiple abuse forms, sometimes including PA, but without specific
focus on conceptualizing and assessing PA. Furthermore, most screening instruments rely on clinician assessments rather than self-report by older adults (Marshall, Benton, & Brazier, 2000) and are designed to evaluate quality of caregiving (e.g., Bravo, Girouard, Gosselin, Archambault, & Dubois, 1995), identify abusive caregivers of older adults (Reis & Nahmiash, 1995), or help health professionals to detect problems (Fulmer et al., 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 PA 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 PA. Sixteen percent of nursing home staff admitted to significant PA 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 six professional caregivers reported committing abusive acts but over four fifths reported observing them. Unfortunately, only seven of the studies that were reviewed used measures for which any type of reliability and validity had been assessed (Cooper et al., 2008). Cooper and colleagues 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 PA 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 PA may reflect a cultural perspective (Anetzberger, Korbin, & Tomita, 1996; Moon, Tomita, & Jung-Kamei, 2001). Furthermore, some believe that the meaning of PA 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 PA is believed to be underreported (Cooper et al., 2008; Schofield & Mishra, 2003), the percentages of occurrence reported in extant studies indicate the pervasiveness of the problem. Pillemer and Finkelhor (1988) 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 PA to verbal aggression only, for which they established a rate of 1.1%. Most recently, Acierno and colleagues (2010) conducted a national prevalence study, and based on a sample of 5,777 older adults (aged 60 years and older), 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, Schulz, Castle, and Rosen (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 PA of African American elders as compared with non-African Americans: 24.4% versus 13.2%, respectively.

In samples of abused older adults, Brownell, Berman, and Salmone (1999) found that among 402 cases of abuse of older adults, 54% involved PA; a similar study by Anetzberger (1998) revealed that 41% of incidents of abuse of older adults were psychological. Anetzberger also found that in cases where there was PA, additional forms of abuse were present 89.7% of the time, including physical neglect and financial exploitation. Similarly, the National Elder Abuse Incidence Study (National Center on Elder Abuse, 1998) found a 35% prevalence rate; Lithwick, Beaulieu, Gravel, and Straka (1999) found 87%; Vladescu, Eveleigh, Ploeg, and Patterson (1999) and Godkin, Wolf, and 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 PA. Differences in the definition and measurement of PA used by each study above may account for some discrepancies and variability.

Conceptual Models

The limited research on most forms of elder abuse, including PA, 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 and colleagues (1989) developed
five conceptual components of abusive relationships. Anetzberger (2000) developed the Exploratory Model for Elder Abuse that 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 (2003) seminal book on elder abuse presents a structure, process, and outcome model that includes the sociocultural 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 or her characteristics as well as the social network. Although the models are able to explain the etiology of general abuse, they do not present examples of items that represent PA 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 PA of older adults. Statements were generated from literature review and by local and national panels consisting of 16 experts in the field of PA. 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 PA. 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 PA 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 or her understanding of the client’s report, his or 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 comprised 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, nonprofit, agency, and business locations. The focus group process consisted of a review of the PA items that were compiled. Participants were asked to read each item and evaluate its relevance to PA, 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 Table A1.

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 PA. 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, that is, Did the items form a single overarching PA construct?
2. To test the fit of the items to the Rasch measurement model, that is, 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, cutoff to determine PA.

**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, 22 highly experienced elder abuse staff members were recruited from these agencies. Because 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 PA 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 PA. 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 were African American (61.3%), more than one third were Caucasian (35.5%), and the remainder were of mixed race or other. Most were non-Hispanic (92.9%). The majority of clients were 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:

\[
\ln \frac{P_{ni}}{P_{ni,j-1}} = B_n - D_i - F_j,
\]

where \( P_{ni} \) 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. Because 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.**—Because 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; E. V. Smith, 2002). Although there are no hard rules for interpreting principal component 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, that is, 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%, whereas 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 less than 1.33 mean square (MNSQ) on both infit and outfit were considered acceptable quality (R. M. Smith, 2000; Wilson, 2005).

**Rating Scale.**—The proper functioning of the rating scale was examined using: (1) outfit MNSQs 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 and 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 (E. V. 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 et al., in press), 16 experts grouped the items into six groups and rated the severity of
the items on a scale from 1 to 5. These item severities were then averaged within each group. The result was a theoretical hierarchy of five conceptual components of PA arranged in descending severity (Table 1) as follows (mean expert ranking from 1 to 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, that is, “client groups” rankings were compared with the rankings of the “expert groups.”

Multitrait Multimethod 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 IDOA questionnaire, which is required by IDOA for elder abuse investigations, covers many forms of elder abuse, including PA. 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 PA 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 PA observations on the 226 substantiated clients. The 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 Department on Aging: PA 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 PA met stringent Rasch analysis fit and unidimensionality criteria.
4. OAMA 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 (1988, 1992) guidelines was setup as follows: NS = nonsignificant, >.1 = low, >.3 = moderate, and >.5 = high. Others have suggested lower values based on reviews of research, for example, >.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 subsequently:

1. Client Gender: There was no reason to expect differential exploitation by gender so all gender correlations were expected to be NS.
2. Psychological Abuse Substantiation Decision:
   • Moderate correlation with OAMA client PA
   • High correlation with OAMA staff PA
3. OAMA Staff Psychological Abuse:
   • High correlation with OAMA client PA

In the multitrait multimethod 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, that is, 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 (1998b) procedure 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 PA scale. There is a substantial floor of persons at the bottom who are not registering any client-reported PA. The concept that each item belongs to is indicated in brackets at the end of the item label, that is, 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

![Wright Map of persons and items on the Rasch ruler of client-reported psychological abuse (item number’s keyed to Table A1)](image-url)
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. Table A1 contains the items by form and factor information. Although 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 PA 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 (−0.59). This was reasonably well targeted because the person mean was within 1.0 logit and within one SD (0.93) of the item mean of 0.

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 PA 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 Table A1.

Multitrait Multimethod Analysis of Hypothesized Relationships

It was hypothesized that all gender correlations would be non-significant 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 Cutoff to Determine PA

Because there is no solely empirical way to determine a cut-point, for example, using the Wright map (Figure 1), the logic of the cut-point decision is described in the Discussion.

Discussion

A measure consisting of 31 items was validated as a unidimensional measure of client-reported PA. 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 PA 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, that is, in the “ceiling,” with a definite “yes.” Thirty-seven persons were in the floor, that is, endorsing 0 symptoms. Above −1.0 on the ruler, the item meanings, that is, severity of the symptoms going up the scale, and locations indicate that this may be a useful cutoff score for PA. 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 PA, there were 52 persons (24%) above this level having even more severe symptomatology, such as 22MadeFeelSmall, 25TalkedAsIfNotThere,
31. **MadeAshamed**, and 28. **DelibConfused**. Above 1.0 on the ruler (16 persons, 7%) could be classified as extreme PA because the four items above 1.0, that is, **10. Confined**, **13. DepriveAsstvdevice**, **24. ManipW/drugs**, and **11. PreventContactOutsd**, all involve serious psychological isolation, deprivation, and manipulation that border on or may include physical abuse and/or neglect. Such abuse may have serious, for example, depression, long-lasting, and even life-threatening sequelae.

### Multitrait Multimethod 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 PA 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, that is, within one standard error ($SE = 0.52$), 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 **9. AfraidOfAA** was fairly high on the severity ruler, that is, at $-0.36$. This item and **8. UncomfortableW/AA** were originally classified as the “Risk Factor” concept. However, such a high calibration was indicative of something more seri-

### Table 1. Expert Item Groups and Rankings Compared With Client Factors and Rankings

<table>
<thead>
<tr>
<th>Expert concept Rank</th>
<th>Expert concept name</th>
<th>Expert groups average Rasch measure&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Client concept rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Isolation (ISO)</td>
<td>0.688</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Threats &amp; Intimidation (T&amp;I)</td>
<td>-0.024</td>
<td>2&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>3</td>
<td>Insensitivity &amp; Disrespect (I&amp;D)</td>
<td>-0.036</td>
<td>3&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>4</td>
<td>Shaming &amp; Blaming (S&amp;B)</td>
<td>-0.315</td>
<td>4&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>5</td>
<td>Risk Factors</td>
<td>-0.585</td>
<td>5&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>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, that is, where items are located on the ruler in Figure 1, and divided by the number of items in that group, for example, seven ISO items. Because most of the ISO items are located high on the ruler, the ISO group/concept has the highest severity.

<sup>b</sup>Expert and client rankings were the same, but the middle three were so close in average rank, that is, within one standard error ($SE = 0.52$), that this ranking may not be reliable.

<sup>c</sup>Two risk factor items involving fear of abuser were reclassified as T&I.

### Table 2. Hypothesized and Actual Correlations<sup>a</sup> of OAPAM With Gender, Substantiation Decision, and Staff Psychological Abuse (PA) Assessment

<table>
<thead>
<tr>
<th>Client gender</th>
<th>Psychological abuse substantiation decision (IDOA)</th>
<th>OAMA staff PA</th>
<th>OAMA client PA</th>
</tr>
</thead>
<tbody>
<tr>
<td>M = 0, F = 1</td>
<td>—</td>
<td>NS, -.042</td>
<td>NS, -.076</td>
</tr>
<tr>
<td>Emotional Abuse Substantiation Decision (IDOA)</td>
<td>—</td>
<td>High .478**</td>
<td>Moderate .360**</td>
</tr>
<tr>
<td>OAMA staff PA</td>
<td>—</td>
<td></td>
<td>.87&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>OAMA client PA</td>
<td></td>
<td>High .700**</td>
<td>.86</td>
</tr>
</tbody>
</table>

<sup>a</sup>Hypothesized correlations: NS = nonsignificant, >.1 = low, >.3 = moderate, and >.5 = high are listed before the actual correlations.

<sup>b</sup>Person reliabilities of OAMA scales are located on the diagonal.

Notes: IDOA = Illinois Department on Aging; OAMA = Older Adult Mistreatment Assessment; OAPAM = Older Adult Psychological Abuse Measure.
ous than a risk factor so these items were reclassified into Threats and Intimidation. This was logical because the items, that is, “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 PA because 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, that is, each concept is not at a distinct severity level. Rather, the items within each concept vary greatly in severity.

Limitations

Although this was the largest sample of substantiated elder abuse clients that was found, it was still limited to seven 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 OAPAM in helping to assess the existence and the level of PA of older adults who are able to self-report using a MMSE (Folstein et al., 1975) 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 PA 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 PA of older adults for improved services and research. This OAPAM can be widely useful in elder abuse research and practice because there had been no validated client-reported measures, and self-report by the alleged victim of his or 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 PA 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 idiosyncrasies. 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 a 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.

Funding

This work was supported by the National Institute of Justice [grant number 2006-MU-MU-0004].

Acknowledgements

Points of view are those of the author(s) and do not necessarily represent the position of the U.S. Department of Justice. We appreciate cooperation for data collection by the Illinois Department on Aging and the participating adult protective services agencies. We are grateful to Jessica Mazza for editorial assistance.

References


**Appendix Table 1.**

<table>
<thead>
<tr>
<th>Item number, full item, and item abbreviation (number and abbreviation are same as in Figure 1)</th>
<th>Concept name (from expert panel)</th>
<th>Rasch measure (from client data)</th>
<th>Results of analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Prompt “In the past 12 months, has (NAME OF ALLEGED ABUSER)”::</td>
<td>Threats and Intimidation</td>
<td>Measures based on 31 item analysis</td>
<td>Blank means included in 31 item analysis</td>
</tr>
<tr>
<td>1. Taken things away or threatened to take things away from you? (TakenThingsAway)</td>
<td>Threats and Intimidation</td>
<td>-0.44</td>
<td>Short form item</td>
</tr>
<tr>
<td>2. Abandoned or threatened to abandon you? (Abandoned)</td>
<td>Threats and Intimidation</td>
<td>0.24</td>
<td>Short form item</td>
</tr>
<tr>
<td>3. Threatened to place you in a nursing home when it was not appropriate? (ThreatNursHme)</td>
<td>Threats and Intimidation</td>
<td>0.33</td>
<td>Short form item</td>
</tr>
<tr>
<td>4. Harmed or threatened to harm someone or something close to you (kids, pets, etc.)? (ThreatenHarmSomeone)</td>
<td>Threats and Intimidation</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>5. Used non-verbal behavior such as shaking a fist, pushing, poking, or slapping, to threaten or scare you? (NonverbGestFist)</td>
<td>Threats and Intimidation</td>
<td>-.36</td>
<td>Short form item</td>
</tr>
<tr>
<td>6. Manipulated you by withholding affection and love? (WithholdingAffection)</td>
<td>Threats and Intimidation</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>7. Behaved in ways that frighten or intimidate you? (FrightenIntimidate)</td>
<td>Threats and Intimidation</td>
<td>-0.66</td>
<td>Short form item</td>
</tr>
<tr>
<td>In the past 12 months:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Have you been uncomfortable with _______? (UncomfortableW/AA)</td>
<td>Risk Factors reclassified as T&amp;I</td>
<td>-0.81</td>
<td>Short form item</td>
</tr>
<tr>
<td>9. Have you been afraid of _______? (AfraidOfAA)</td>
<td>Risk Factors reclassified as T&amp;I</td>
<td>-0.36</td>
<td>Short form item</td>
</tr>
<tr>
<td>In the past 12 months, has NAME ALLEGED ABUSER:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Confined you against your will? (Confined)</td>
<td>Isolation</td>
<td>1.31</td>
<td>Short form item</td>
</tr>
<tr>
<td>11. Prevented you from having contact with the outside world via telephone, newspapers, television, or radio, etc.? (PreventContactOutsd)</td>
<td>Isolation</td>
<td>1.08</td>
<td>Short form item</td>
</tr>
<tr>
<td>12. Prevented you from contacting family, friends, or community resources? (PreventedContactFamily)</td>
<td>Isolation</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>13. Deprived you of glasses, hearing aids, prosthetics, walker, wheelchair, or any other assistive devices that you needed? (DepriveOfAssistiveDevices)</td>
<td>Isolation</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>14. Kept things from you or lied about things that you should know about? (KeptThingsFromEldOrLied)</td>
<td>Isolation</td>
<td>-0.64</td>
<td>Short form item</td>
</tr>
<tr>
<td>15. Called you unkind names or put you down? (CalledUnkindNames)</td>
<td>Shaming &amp; Blaming</td>
<td>-0.66</td>
<td>Short form item</td>
</tr>
<tr>
<td>16. Deliberately made you feel bad or hurt your feelings? (HurtEldFeelings)</td>
<td>Shaming &amp; Blaming</td>
<td>-0.79</td>
<td></td>
</tr>
<tr>
<td>17. Given you the silent treatment? (SilentTreatment)</td>
<td>Threats &amp; Intimidation</td>
<td>-0.35</td>
<td></td>
</tr>
<tr>
<td>18. Treated you in an undignified or inappropriate way while assisting you with dressing, eating, bathing and so on? (TreatEldUndignifiedWay)</td>
<td>Shaming &amp; Blaming</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>19. Sworn or yelled at you? (SworeOrYelled)</td>
<td>Shaming &amp; Blaming</td>
<td>-0.92</td>
<td>Short form item</td>
</tr>
<tr>
<td>20. Refused or neglected to get medical services that you needed? (NeglectMedSvs)</td>
<td>Isolation</td>
<td>0.52</td>
<td>Short form item</td>
</tr>
<tr>
<td>In the past 12 months:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. Has______ failed to support you or back you up when you needed it? (Failed2Support)</td>
<td>Shaming &amp; Blaming</td>
<td>-0.26</td>
<td>Short form item</td>
</tr>
</tbody>
</table>

Appendix continued
## Appendix (continued)

<table>
<thead>
<tr>
<th>Item number, full item, and item abbreviation (number and abbreviation are same as in Figure 1)</th>
<th>Concept name (from expert panel)</th>
<th>Rasch measure (from client data)</th>
<th>Results of analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>In the past 12 months, has the ALLEGED ABUSER:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. Made you feel small, for example, treated you like a child? (MadeFeelSmall)</td>
<td>Insensitivity &amp; Disrespect</td>
<td>-0.07</td>
<td>Short form item</td>
</tr>
<tr>
<td>23. Manipulated or tried to control you in any way? (Manipulated)</td>
<td>Threats &amp; Intimidation</td>
<td>-0.91</td>
<td>Short form item</td>
</tr>
<tr>
<td>24. Manipulated you with drugs or alcohol? (ManipulatedWithDrugs)</td>
<td>Isolation</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>25. Talked about you as if you were not there? (TalkedAsIfNotThere)</td>
<td>Shaming &amp; Blaming</td>
<td>-0.04</td>
<td>Short form item</td>
</tr>
<tr>
<td>26. Not let you speak for yourself? (NotLetSpeak)</td>
<td>Shaming &amp; Blaming</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>27. Not been sensitive to your feelings? (NotSensitiveFeelings)</td>
<td>Insensitivity &amp; Disrespect</td>
<td>-0.71</td>
<td>Short form item</td>
</tr>
<tr>
<td>28. Deliberately confused you? (DeliberatelyConfused)</td>
<td>Insensitivity &amp; Disrespect</td>
<td>-0.09</td>
<td>Short form item</td>
</tr>
<tr>
<td>29. Minimized your injuries or complaints? (MinimizedInjuries)</td>
<td>Insensitivity &amp; Disrespect</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>30. Blamed you for their problems? (BlamedForProblems)</td>
<td>Shaming &amp; Blaming</td>
<td>-0.41</td>
<td></td>
</tr>
<tr>
<td>31. Said something about you that made you feel ashamed? (MadeEldFeelAshamed)</td>
<td>Shaming &amp; Blaming</td>
<td>-0.04</td>
<td></td>
</tr>
</tbody>
</table>