Visual Disability Variables. I: The Importance and Difficulty of Activity Goals for a Sample of Low-Vision Patients

Robert W. Massof, PhD, Cynthia T. Hsu, MD, Frank H. Baker, MD, G. David Barnett, BS, William L. Park, OD, James T. Deremeik, MA, Carol Rainey, BS, Cathy Epstein, BA

ABSTRACT. Massof RW, Hsu CT, Baker FH, Barnett GD, Park WL, Deremeik JT, Rainey C, Epstein C. Visual disability variables. I: The importance and difficulty of activity goals for a sample of low-vision patients. Arch Phys Med Rehabil 2005; 86:946-53.

Objective: To test the validity and reliability of latent trait measures estimated from ratings by low-vision patients of the importance and difficulty of selected activity goals.

Design: Validation of a telephone-administered functional assessment instrument using Rasch analysis of self-assessment ratings.

Setting: Telephone interviews of respondents in their homes.

Participants: Consecutive series of 600 outpatients with low vision.

Interventions: Not applicable.

Main Outcome Measures: Ratings of the importance and difficulty of achieving 41 activity goals. Person and item traits were measured with the Andrich rating scale model. Measurement validity and reliability were tested statistically by comparing response patterns and distributions with measurement model expectations.

Results: Patients could distinguish only 3 categories of importance and 4 categories of difficulty. The distributions of person and item measure fit statistics were consistent with 2 unidimensional constructs: value of independence estimated from importance ratings and visual ability estimated from difficulty ratings. However, 8 of 41 activity goals were poor estimators of value of independence and 7 of 41 activity goals were poor estimators of visual ability. Person measure distributions could be divided into 3 statistically distinct strata for estimates from both importance ratings and difficulty ratings. Item measure distributions could be divided into 21 strata for estimates from importance ratings and 7 strata for estimates from difficulty ratings.

Conclusions: The 2 variables that define visual disability—value of independence and visual ability—are valid constructs that can be estimated accurately and reliably from patient ratings of the importance and difficulty of activity goals.

Key Words: Questionnaires; Rehabilitation; Vision, low.

From the Lions Vision Research and Rehabilitation Center, Wilmer Ophthalmological Institute, The Johns Hopkins University School of Medicine, Baltimore, MD.

Supported by the National Eye Institute, National Institutes of Health (grant no. EY-012045).

No party having a direct interest in the results of the research supporting this article has or will confer a benefit on the author(s) or on any organization with which the author(s) is/are associated.

Reprint requests to Robert W. Massof, PhD, Lions Vision Center, 6th Fl, 550 N Broadway, Baltimore, MD 21205, e-mail: *rmassof@lions.med.jhu.edu*.

0003-9993/05/8605-9118$30.00/0

doi:10.1016/j.apmr.2004.09.016

FROM THE BEGINNING of the 20th century, visual disability in the United States has been defined by measures of visual impairments.[1-4] Impairment measures such as visual acuity, visual fields, and tests of binocular vision have served as surrogates for quantifying the difficulty visually impaired people have working, attending school, driving a car, or performing ordinary activities of independent daily life. More contemporary notions of visual disability follow the World Health Organization's (WHO) original system of impairments, disabilities, and handicaps.[5] In the WHO system, which recently was modified,[6] disability is defined as "any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being."[5(p143)] Although WHO defines visual disability more generally, disability scales that embrace the WHO definitions still equate visual disability with visual impairments.[7]

Within the WHO framework, low vision and visual disability mean the same thing. People with low vision have chronic visual impairments that cause disability. The purpose of low-vision rehabilitation is to overcome visual disability with vision-enhancing devices, adaptive technology, adaptive skills, and environmental modifications.[8] Successful low-vision rehabilitation alters the relation between visual disability and visual impairments. Although the rehabilitated person is still visually impaired, he/she has reduced visual disability. By definition, existing measurement scales of visual disability cannot accommodate successful rehabilitation outcomes. They assume that the relation between disability and impairment is fixed.

The term "visual disability" is often associated with the consequences visual impairments have on employment, education, and social participation. More generally, however, in keeping with the WHO definition, disability means loss of ability.[9] Ability is a personal trait. People with high ability can perform activities easily; people with low ability have greater difficulty performing the same activities. Low-vision rehabilitation can be thought of as the restoration of visual ability that was lost because of chronic visual impairments.

To understand low vision and its rehabilitation and to measure the effects of intervention, we must be able to measure visual disability. Obviously, such measurements must be based on measurements of visual ability. However, ability is a theoretical construct that we infer from observations of the difficulty a person has in performing a sample of activities. We would not say that a person is disabled if the only activities that are abnormally difficult or impossible to do were those he/she would never do anyway. This truism illustrates that any practical definition of visual disability must include a reference to the importance of performing activities without assistance, as well as to a person's visual ability. If a person with low vision reports that a particular activity is not important, it would not factor into the rehabilitation plan and would not be relevant to the assessment of effectiveness of intervention. Thus, activities included in a sample chosen for observation should be of some value to the person being assessed.

The importance of performing activities has been formalized with goal attainment scaling (GAS) in the psychiatric,[10] rehabilitation,[11] and nursing[12] literature, but it has received limited attention in the low-vision literature. Numerous self-assessment rating

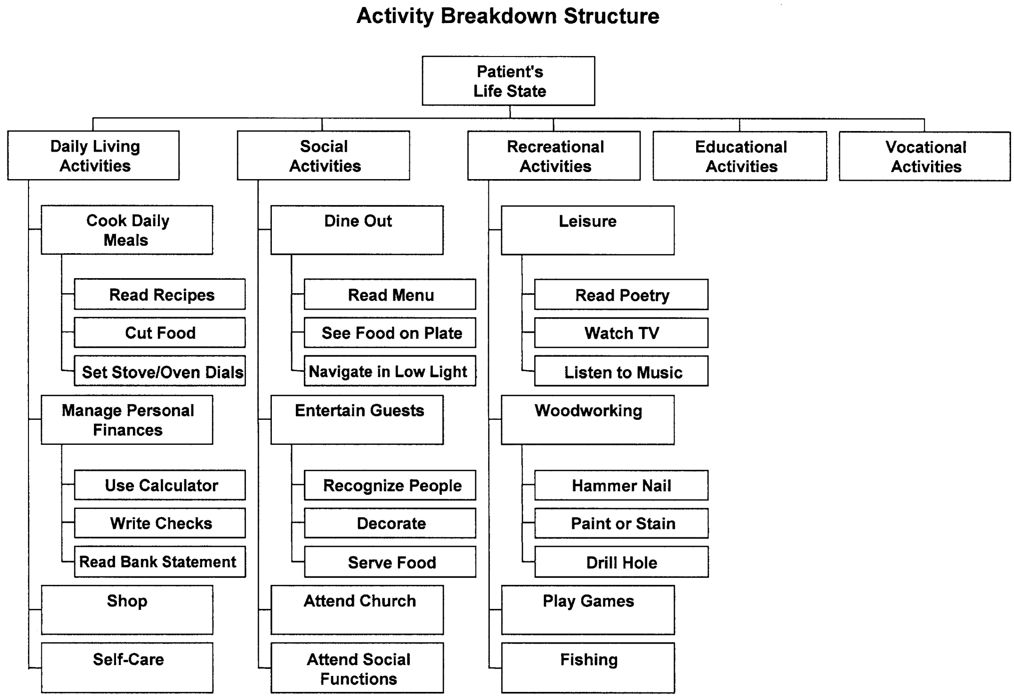

Fig 1. Schematic representation of a hierarchical organization of activities into an ABS. Patients perform everyday activities to meet societal objectives: daily living, social interactions, recreation, education, and work. At the next level under each objective, activities are performed to achieve specific goals. Examples of activity goals listed under the daily living objective are cook daily meals, manage personal finances, shop, and self-care. At the next level under each goal, the specific cognitive and motor activities are listed as tasks. Examples of tasks under the cook daily meals goal are read recipes, cut food, and set stove and oven dials. Tasks can also be grouped across goals and objectives according to visual function. For example, read recipes, read menu, read poetry, and read bank statement can be grouped as reading activities.

scale instruments have been developed over the past 2 decades to evaluate visual ability in visually impaired people.[13] Each instrument asks respondents to rate difficulty, frequency of occurrence, or level of agreement for a fixed set of questions. Many of the instruments permit responses of "do not do for reasons other than vision," "not applicable," or "not interested" for items that are not relevant or important to the respondent.[14-16] However, there has been no systematic analysis of the frequency of such responses of visually impaired people to different items. Rather, such responses are scored as missing data.

In an earlier pilot study, Massof[17] used Rasch analysis to estimate the value of 29 general activities from ratings of importance made by patients at a low-vision clinic. Massof found that daily living activities had the highest value and recreational activities had the lowest. Leat et al,[18] in reporting the frequency that selected tasks were rated as important or very important by low-vision patients, found that reading large-print books was rated as important or very important by the fewest patients and that reading medicine labels was rated of high importance by the largest number of patients.

Activities that are both important and difficult for the low-vision population should be given the greatest emphasis in low-vision rehabilitation programs. Although programs could be designed to emphasize activities that are likely to be considered important and difficult to the average visually impaired patient, low-vision rehabilitation per se is targeted at an individual patient. Ideally, visual disability instruments would be customized individually to measure only those activities important to the person being assessed. However, if we consider all possible general and specific activities, with all conceivable variations for describing those activities, the number of items from which to choose a sample would be unreasonably large.

Activity Breakdown Structure

Several years ago we offered a model for organizing large numbers of activities into a manageable hierarchy called the Activity Breakdown Structure (ABS).[19] As shown in figure 1, at the lowest level of the ABS, activities are described as tasks. Tasks are very specific cognitive and/or motor activities that can be observed by another person. Examples of tasks are: read recipes, write checks, watch television, cross a street, climb stairs, and sew a hem. The importance of a task to a person depends on why it is being performed.

At the second level of the ABS, activities are described as goals. Goals are general activities that represent a group of tasks that are being performed in a coordinated manner for a common purpose. Examples of goals are: prepare a meal, shop, manage finances, dine out, attend church, play games, and leisure entertainment. The tasks are nested under the goals they serve, so goals can be thought of as subsets of the set of tasks.

At the third level of the ABS, activities are described as objectives. Examples of objectives, as shown in figure 1, are daily living activities, social activities, recreational activities, educational activities, and vocational activities. Goals are nested under the objectives they serve, so objectives also are subsets of the set of tasks.

Tasks can also be grouped according to visual function.[20] For example, reading recipes, newspaper articles, price tags, and signs are instances of reading function. Climbing stairs, crossing streets, hiking, and walking to the store are instances of mobility function. Visual functions also are subsets of the set of tasks. If tasks cannot be performed, a person experiences functional limitations. If goals cannot be achieved, a person experiences disabilities (ie, activity limitations[6]). If objectives cannot be met, a person experiences handicaps (ie, participation limitations[6]).

Within the framework of the ABS, visually impaired people experience disability when goals that are important to them are abnormally difficult or impossible to achieve. Thus, for the ABS model, as is assumed by GAS,[10] visual disability is a composite variable: it is a function of both goal importance and goal difficulty. This article is focused on measurements of goal importance and goal difficulty in a sample of patients in a low-vision clinic. To make these measurements, we used the Activity Inventory

(AI), a rating-scale instrument we designed to conform to the ABS and to provide an individualized self-assessment questionnaire.

Activity Inventory

Although most self-assessment visual function questionnaires can be calibrated to produce a common visual ability metric,[21,22] the set of questions in any single instrument represents a small subset of activities that would be important to consider before a rehabilitation plan is developed. Our aim was to develop an instrument to use to obtain both a functional history, from which to develop an individual rehabilitation plan, and a quantitative measure of a patient's functional ability.

Low-vision patients in the Low Vision Rehabilitation Service at the Wilmer Ophthalmological Institute are routinely assessed to determine their functional limitations and rehabilitation goals. As described earlier,[17] we performed retrospective chart reviews of more than 3200 patients and identified 24 frequently cited activities that could be classified as goals under 3 objectives (daily living, social interactions, recreational activities) and more than 200 activities that could be classified as tasks under the goals. We constructed a pilot instrument, the AI, based on these goals and tasks. Through the AI, patients were also asked to identify other goals and tasks of relevance to them that were not included in the original list. The AI subsequently was administered to 445 low-vision patients before they began rehabilitation.[17] They were asked to rate the importance of each goal on a scale of 0 (not important) to 5 (extremely important). If a patient rated a goal to have at least some importance (>0), he/she was asked to rate the difficulty of that goal on a scale of 0 (not difficult) to 5 (extremely difficult or impossible). If a goal presented at least some difficulty (>0), then the patient was asked to rate the difficulty of each subsidiary task. In this way, patients rated the importance of each goal, but they rated the difficulty of only those goals that were of some importance to them. Patients rated the difficulty of only those tasks that served goals that were not easy for them. Consequently, each patient responded to a relevant set of questions that was unique to him/her. That is, each patient received a personalized functional assessment.

We modified the AI for this study based on the results of the pilot study. The new AI has 41 goals under the same 3 objectives (daily living, social interactions, and recreation—the education and vocation objectives have not yet been developed and are not included because the majority of low-vision patients are elderly[23] and do not work or attend school); there are 337 tasks under the goals. Some original goals were modified to make the description more general (eg, "recreational reading" became "leisure entertainment"). New goals and tasks were added to represent the "other" goals and tasks contributed by the patients in the pilot study.

The hierarchical structure of the ABS assumes that goals are subsets of tasks and objectives are subsets of goals. Thus, it is expected that goals inherit difficulty from tasks and inherit importance from the objectives that frame them. Our purpose in this study was to test the theoretical assumptions of the ABS and to test the measurement validity of the AI.

METHODS

Participants

The revised AI was administered to 600 consecutively recruited adult patients who were scheduled for first visit appointments in our Low Vision Rehabilitation Service. None of the patients had participated in the pilot study. Patients ranged in age from 19 to 101 years (median age, 73y). Fifty-eight percent were women. Primary diagnoses were as follows: age-related macular degeneration, 44%; other macular disorders, 2%; glaucoma, 14%; other optic neuropathy, 6%; diabetic retinopathy, 10.5%; other retinal vascular disease, 3%; other retinal disorders, 7%; refractive disorder, 3.5%; cerebrovascular accident or brain injury, 3%; retinitis pigmentosa, 1.5%; other retinal degeneration, 1%; developmental disorder, 1.5%; cataract, 1%; and corneal disorder, 1%. Best corrected visual acuity in the better eye ranged from 20/20 to 20/2400 (median, 20/60; standard deviation [SD], 0.4 log minimum angle of resolution).

Once they made an appointment, patients were informed of the study by mail and asked to return a postcard if they did not want to participate. (Data were collected before the Health Insurance Portability and Accountability Act was implemented.) After 1 week, research assistants contacted by telephone patients who had not returned the card; they described the study, obtained verbal consent from subjects willing to participate, and scheduled a telephone interview. Patients were mailed a large-print medical and functional history questionnaire (routinely mailed to all low-vision patients), a large-print consent form, a written description of the study, and reminders of their appointment dates for the telephone interview and their visit to the low-vision service. Patients returned the history questionnaire and signed consent form in a self-addressed stamped envelope. The participation rate was 95% of 630. The study's protocol and consent procedure followed the tenets of the Declaration of Helsinki and were approved by the Johns Hopkins Institutional Review Board.

Procedures

Participants were administered the AI by telephone by using a computer-assisted interview program we developed. Research assistants called the patients at the appointed time, explained the interview process, and answered questions about it. First, the patient was asked to rate the question, "How important is it for you to [description of goal] without the assistance of another person?" The possible answers were "Not important," "Slightly important," "Moderately important," "Very important," or "Extremely important." The interviewer entered the patient's response, and the computer assigned a rank score on a scale of 0 (not important) to 4 (extremely important). If the patient responded, "Not important," the interviewer moved to the next goal. Otherwise, the interviewer asked the patient, "How difficult is it for you to [description of goal] without the assistance of another person?" The possible answers were "Not difficult," "Slightly difficult," "Moderately difficult," "Very difficult," "Extremely difficult but possible," or "Impossible." The patient's response was entered into the computer, which ranked it on a scale of 0 (not difficult) to 5 (impossible). If the patient responded, "Not difficult," the interviewer moved to the next goal. Otherwise, the patient was asked, "How difficult is it for you to [description of task]?" Possible answers were the same categories as those for goal difficulty, or the patient could respond that the task is "not applicable." The patient's response was entered into the computer and ranked on a scale of 0 to 5 or recorded NA (not applicable). The average interview lasted 25 minutes.

Data Analysis

We used Rasch analysis with the Andrich rating scale model[24] to estimate item and person trait variables on interval scales from patients' importance and difficulty ratings and to test the construct validity and reliability of the psychometric measurements.[25] We used the term *inherent value* to represent the interval-scaled trait of goals that underlies importance ratings (symbolized as ζ_i for goal i), and we used the term *required ability* to represent the interval-scaled trait of goals that underlies difficulty ratings (symbolized as ρ_i for goal i). The corresponding person traits were called *preference for independence* (symbolized as ξ_n for person n) and *visual ability* (symbolized as α_n for person n), respectively. The Andrich model assumes that the probability that person n will respond with rating x to item i is governed by the sum of inherent value and preference for independence, $\xi_n+\zeta_i$ (which is called the *value of independence*), for importance ratings and by the difference between the visual ability of the patient and the visual ability required by the goal, $\alpha_n-\rho_i$ (which is called the *functional reserve*[17,26]), for difficulty ratings.

A third measure estimated by Winsteps[a] for the Andrich model is the value of $\xi_n+\zeta_i$ or $\alpha_n-\rho_i$, at which the probability of responding with importance or difficulty rating x is equal to the probability of responding with rating $x-1$ (symbolized by τ_x). If the ratings are used in the same way within the sample of patients, then the values of τ_x should be in the same order as values of x, and there should be a value of $\xi_n+\zeta_i$ and a value of $\alpha_n-\rho_i$ at which x is the most probable response.

If the value of independence and functional reserve constructs are valid, then the pattern of response ratings by the patients to the goals and tasks will be consistent with the expectations of the Andrich rating scale model, a manifestation of the polytomous Rasch measurement model.[25] We used an unconditional maximum likelihood estimation routine[27] to estimate the values of ξ_n and ζ_i from the matrix of importance ratings of goals, and α_n and ρ_i were estimated from the matrix of difficulty ratings. The infit mean square is the mean square residual (not normalized) divided by the average expected variance of the responses. The distributions of infit mean squares (for patients or items) can be transformed to normal distributions with expected values of zero and unit SDs.[28] For purposes of presentation and interpretation, the infit mean squares for each item and each patient were transformed by Winsteps to normal deviate units (ie, z scores) using a Wilson-Hilferty transformation.[29]

To test the hypothesis that a single construct is responsible for the distribution of patient ratings for the items, we computed correlations between the distributions of residuals for different items (across patients) and between distributions of residuals for different patients (across items). Principal components analyses were made of the covariance matrices to estimate the factors responsible for the variance of the distributions of residuals (Winsteps).[30]

This article focuses on analyses of patient responses to goal-level items; a second article[31] concentrates on analyses of patient responses to task-level items. To test the hypothesis that the different objectives represent independent domains, we estimated sets of person measures for each objective from patient importance and difficulty ratings for the subsidiary goals. Correlations between person measure distributions were computed for each pair of objectives and confirmatory factor analysis was performed on the correlation matrices for goal-based person measures (Systat[b]).[32]

RESULTS

The analysis proceeded in 4 steps. First, we investigated how patients used the importance and difficulty rating categories. Second, we examined estimates of item measures and the validity and reliability of the estimated scales. Third, we evaluated the validity and reliability of estimates of person measures. And fourth, we explored the dimensional properties of person measures across objectives.

Based on the estimated probability of using each importance rating category as a function of the value of independence (ie, sum of person and item measures), the most probable responses were "not important" (0), "very important" (3), and "extremely important" (4). There was no value of $\xi_n+\zeta_i$ at which "slightly important" (1) or "moderately important" (2) was the most probable response, and the estimated values of τ_1 and τ_2 caused disordering of the cross-points of the response probability functions. The peaks of the response probability functions for using rating categories 1 and 2 occurred near the peak probability of using rating category 3. Therefore, the patient responses were recoded as 0 (not important), 1 (slightly, moderately, or very important), and 2 (extremely important), and we repeated the analysis. The resulting response probability distributions for the recoded data formed 3 well-defined and ordered response probability functions.

Based on the estimated probability of using each difficulty rating category as a function of the difference between person ability and the ability required to perform the activity described by the item, the most probable responses were "not difficult" (0), "very difficult" (3), and "impossible" (5). The probability functions for "slightly difficult" (1), "moderately difficult" (2), and "possible but extremely difficult" (4) had low peak probabilities and were disordered. Unlike the case with the probability functions for importance ratings, the probability functions for difficulty rating categories 1 and 2 filled the gap between probability functions for difficulty rating categories 0 and 3. Difficulty rating category 4 covered the same range as category 3. Therefore, the responses were recoded as 0 (not difficult), 1 (slightly or moderately difficult), 2 (very or extremely difficult), and 3 (impossible) and again analyzed. The response probability functions for the recoded data had well-defined peaks, and the values of τ_x were ordered.

Table 1 lists the item measures that were estimated from recoded patients' ratings of the importance and difficulty of accomplishing each goal without the assistance of another person. The item measures estimated from importance ratings corresponded to the inherent social value of each goal (ς) on an interval scale. Personal hygiene had the greatest value for independent performance and leatherwork had the least. All goals that serve the daily living objective occupy the top half of the scale; the goals that serve the social interactions and recreation objectives are in the bottom two thirds. The item measures estimated from difficulty ratings corresponded to required visual ability for independent accomplishment of each goal (ρ). Sewing and needlework required the most visual ability and eating required the least.

The variance in the item and person measures is the sum of the true variance in item or person traits plus variance resulting from estimation errors. Separation reliability is an estimate of the fraction of the observed variance that can be attributed to the true variance. The separation reliabilities for item measures were .999 for ς and .96 for ρ. The separation reliabilities of the person measures were .82 for ξ and .86 for α. These reliability coefficients indicate that the item measure distributions (± 2 SDs) can be divided into 21 statistically distinct strata (3 standard errors [SEs] wide) for ς and 7 strata for ρ.[28] Similarly, the person measure distributions (± 2 SDs) can be divided into 3 statistically distinct strata for both ξ and α.

Table 1 also lists the normal deviates for infit mean square residuals for the inherent social value (ς) and required visual ability (ρ) of each goal. Fifty-four percent of the normalized residuals for ς and 51% for ρ fell within ± 2 SDs of the expected value. The most misfitting goals (infit mean squares significantly different from the expected value after correcting for multiple comparisons; ie, >3 SDs) were: manage personal finances, child care, sewing or needle work, and gardening and lawn care for inherent social value. For required visual ability, they were: use public restroom, personal health care, follow the news, attend church, and attend social functions. For these goals, the rating responses across subjects were inconsistent relative to their responses to the other goals. These misfitting

Table 1: Item Measures and Fit Statistics for Social Value and Required Ability of Goals

Objectives Goals	Social Value Item Measure	Social Value Infit z Score	Required Ability Item Measure	Required Ability Infit z Score
Daily Living				
Toileting	2.37	−1.3	−1.14	4.1
Personal hygiene	3.39	−4.5	−1.75	2.3
Dressing	2.84	−3.8	−1.14	−0.1
Personal health care	3.19	−4.5	−1.08	4.4
Eating	3.13	−4.8	−1.98	0.0
Daily meal preparation	1.30	2.7	−0.65	−2.9
Household tasks	1.67	−1.0	−0.46	−3.4
Personal communication	2.81	−4.3	0.23	1.8
Correspondence	2.97	−5.2	0.56	−4.3
Follow the news	2.67	−2.5	−0.48	6.8
Follow a schedule	1.98	−0.4	−0.49	2.3
Manage finances	1.54	3.8	0.45	−2.7
Shopping	1.86	−1.0	0.69	−2.8
Child care	−1.31	3.2	−0.59	−1.1
Social Interactions				
Social functions	0.84	−0.9	0.11	3.5
Entertain guests	0.70	−0.7	−0.97	−1.2
Prepare food for guests	−1.24	0.0	−0.22	−4.3
Dining out	1.79	−2.2	0.23	2.2
Spectator events	0.79	−1.4	0.51	−0.1
Attend meetings	0.08	0.9	0.23	0.2
Play games	−0.04	2.5	0.28	−3.8
Perform in public	−1.82	2.9	−0.05	1.5
Attend church	1.20	3.0	0.06	4.7
Recreation				
Leisure entertain	1.77	−2.4	−0.61	2.7
Hobbies and crafts	−0.37	−0.3	0.53	−2.7
Sewing or needle	−1.26	3.9	1.47	−0.2
Knit or crochet	−2.20	2.6	1.04	−0.9
Woodworking	−2.68	1.4	0.68	−1.4
Metalwork	−4.09	0.8	0.20	−1.6
Paint or draw	−2.19	1.9	0.42	−1.6
Recreational cooking and baking	−0.98	1.6	−0.23	−3.1
Electrical work	−2.80	2.0	0.31	−1.6
Model building	−3.71	1.2	0.45	0.5
Musical instruments	−2.22	2.3	0.49	1.1
Sightseeing	1.21	−0.7	0.79	2.6
Fishing	−2.08	1.9	0.34	−1.7
Hunt and shoot	−3.56	1.8	0.67	−1.8
Outdoor activities	−0.31	0.5	−0.04	−0.6
Gardening and lawn care	−0.22	4.2	0.18	2.7
Play sports	−1.42	1.5	0.72	−0.2
Leatherwork	−5.61	−0.1	0.19	0.0

goals could be the result of the influence of other factors on subjects' ratings, such as gender roles or living arrangements for inherent social value, and individual adaptations of the activities for required visual ability.

High infit mean squares could result from confounding constructs and/or random error. If confounding constructs are responsible, they will be manifested as independent factors with principal components analysis of the response residuals. Principal components analysis of item residuals (Winsteps) indicated that inherent social value (ς) was the only meaningful construct underlying the importance ratings of goals (accounting for 73% of the variance). Five factors accounted for the other 27% of the variance: factor 1, 8%; factor 2, 6%; factor 3, 5%; factor 4, 4%; and factor 5, 4%. Although representing only 8% of the variance, the first factor for the residuals was not random error; rather, it was highly correlated with inherent social value (r=–.88; $P=2\times10^{-14}$). This correlation means that the principal component of the residual correlated negatively with inherent social value. In other words, there was a covariate that added more variability to patients' importance ratings of less important goals than of more important goals.

This covariate most likely is item-specific "personal preference," which translates to frequency of importance in the population. Socially important goals, such as personal hygiene, are important to everyone because of high social pressure. However, at the other extreme, socially unimportant goals, such as leatherwork, are unimportant to most people but extremely important to a few because of high personal preference for that activity. Another possible covariate could be goal difficulty. The inherent social value of goals correlated negatively with required visual ability (r=–.53, $P<.001$). In other words, there appeared to be a trend for patients to undervalue goals that were more difficult for them. However, when comparing person measures, value of independence did not correlate with visual ability (r=.013, P=.38).

Principal components analysis of goal item residuals for required visual ability (ρ) indicated that goal difficulty ratings for some items were partially confounded by a factor that was independent of ρ (factor 1 accounted for 8% of the variance). The goals that loaded heavily onto the confounding factor were: attend social functions, dine out, attend spectator events, attend church, attend meetings, shopping, sightseeing, and fishing. These goals depend heavily on mobility. In general, however, the difficulty ratings for the goals can be explained almost entirely by the required visual ability construct (73%) and random error (5 factors accounted for 27% of the variance). Unlike the case with importance ratings, for required visual ability there was no evidence of a covariate that produced a systematic relation between the item measures and the residuals.

For value of independence, the person measure (ξ) distribution had a mean of −2.87 logits and an SD of 1.17 logits. These results indicate that the average patient endorsed only those items with the greatest inherent social value. Patients with less-than-average value of independence endorsed fewer items with high inherent social value. Patients with a more-than-average value of independence not only endorsed items with high inherent social value, but they also endorsed items with lower inherent social value. As mentioned earlier, this distribution of person measures can be divided into 3 statistically distinct strata.

For visual ability, the person measure (α) distribution had a mean of .78 logits and an SD of 1.3 logits. These results indicate that the average patient would be positioned between playing sports and sightseeing. That is, the average patient would have insufficient visual ability to knit or sew, but would have adequate ability to achieve all goals below playing sports (albeit with greater-than-usual difficulty). Thus, half of the low-vision patients in this study were fairly high functioning, which is consistent with the 20/60 median visual acuity. Like value of independence, the visual ability person measure distribution can be divided into 3 statistically distinct strata.

Approximately 11% of the patients had infit mean squares for estimates of ξ that exceeded model expectations by more than 2 SDs, and 2.6% had infit mean squares that were more than 2 SDs below the model expectations. (The expected value was 2.5% of patients in each tail of the normal distribution.) The most misfitting patients (ie, the 6.5% whose infit and outfit mean square z

scores were greater than expected values by 2.5 SDs or more) had a wide range of ξ values and most likely had anomalous response patterns because of idiosyncratic personal preferences. Principal components analysis on the correlation matrix of patients for distributions of residuals across items indicated that there was mostly random error with respect to the person measures. The first 5 factors accounted for 28% of the variability.

Approximately 13% of the patients had infit mean squares for estimates of α that exceeded model expectations by more than 2 SDs, and 5% had infit mean squares that were more than 2 SDs below the model expectations. Approximately 4% of the patients had infit mean squares that were greater than the expected values by 2.5 SDs or more. These most misfitting patients had a wide range of α values. Although analyses of the effects of comorbidities are beyond the scope of this article, we might hypothesize that the misfitting patients had response patterns governed by other types of impairments (eg, cognitive, physical, neurologic, psychologic). However, principal components analysis on the correlation matrix of patients for distributions of residuals across items indicated there was mostly random error. The first 5 factors account for 27% of the variance.

To answer the question of whether value of independence is the same construct for the different objectives, we did separate Rasch analyses on goal importance ratings for each objective. Consequently, we obtained 3 estimates of ξ for each patient: 1 for daily living, 1 for social interactions, and 1 for recreation. The interobjective product-moment correlations were .50 for daily living versus social interactions, .44 for daily living versus recreation, and .55 for social interactions versus recreation (all correlations were significant, $P<.001$). Confirmatory factor analysis on the correlation matrix (Systat) led to the conclusion that 1 factor was sufficient—that is, the value of independence construct was the same for all 3 objectives (eigenvalues: factor 1=1.521, factor 2=0.018, factor 3=0.000). The relatively low interobjective correlations can be attributed to random error.

We did the same analyses of goal difficulty ratings for each of the 3 objectives. Separate Rasch analyses were done on difficulty ratings for each objective and 3 estimates of α were obtained for each patient: 1 for daily living, 1 for social interactions, and 1 for recreation. The interobjective product moment correlations were .67 for daily living versus social interactions, .62 for daily living versus recreation, and .60 for social interactions versus recreation (all correlations were significant, $P<.001$). Confirmatory factor analysis of the correlation matrix led us to conclude that 1 factor was sufficient—that is, the visual ability construct was the same for all 3 objectives (eigenvalues: factor 1=1.898, factor 2=0.005, factor 3=0.000). As for value of independence, the low interobjective correlations can be attributed to random error.

DISCUSSION

We defined visual disability as abnormally high difficulty or inability because of impaired vision to achieve the goals of activities that are important in serving personal and societal objectives. With this definition, estimates of visual disability require measurements of 2 variables: a person's ability to achieve activity goals and the importance of those goals to the person and/or society. If we applied the reasoning behind GAS,[10] we would simply use the product of these 2 variables. However, numerous arbitrary and untested assumptions are built into GASs.[33,34] Therefore, deeper discussions of how these 2 measurements might be combined to produce a single visual disability variable are required, but they would take us beyond the scope of this article. In this study, we described and evaluated the AI, a hierarchical individualized self-assessment questionnaire that provides estimates of functional ability and value of independence for each person and each activity goal.

Functional ability (α) and the value placed on independence (ξ) are attributes of individual people. For visual disability, these theoretically constructed variables are an emergent property of the low-vision patient population. We hypothesized that these attributes are scalable and that low-vision patients differ from one another in attribute magnitudes. Our study supported these hypotheses by showing that the AI can be used to estimate valid and reliable measures of ξ and α on an interval scale and that there were significant differences in attribute magnitudes between low-vision patients in our sample. Given the precision of the estimated magnitudes relative to the distribution of attributes, we can stratify our sample of 600 low-vision patients into 3 statistically distinct categories for both ξ and α.

Value of Independence

The various goals of activities have different levels of importance in achieving a larger objective. The importance of any specific goal is likely to vary across subjects, but the importance to society is estimated by the average for the sample. Not unexpectedly, the goals with greatest inherent social value are those that serve the daily living objective. This apparent inheritance of goal importance from the objective is especially evident for cooking. Preparing daily meals, which serves the daily living objective, has much higher inherent social value than does preparing food for guests, which serves the social interaction objective. The goals with the least inherent social value are specific hobbies, but all of the specific hobbies and crafts have much lower inherent social value than does the general hobbies and crafts goal. The most likely explanation is that our subjects agreed that hobbies and crafts have value, but they disagreed on which hobby or craft is most important. It is interesting that for this mainly geriatric sample of patients, the general hobbies and crafts goal had higher inherent social value than the child care goal.

The infit mean square residual for items corresponds to the variance of relative goal importance between subjects.[22] As corroborated by principal components analyses of the residuals, our study shows that as the average goal importance decreases, the variance increases. This strong covariance most likely is a consequence of the process used to select goals for the AI. That is, the goals that are included in the instrument are those that some subset of patients identified as important to them. If patients agree that the goal is important, the mean value will be high and the variance will be low. As disagreement about goal importance increases, the mean value decreases (ie, more patients rate the goal as unimportant) and the variance increases. At the other extreme, the variance of goal importance again decreases as consensus about goal unimportance increases (improved fit statistics in table 1 with lower inherent social value).

This study shows that only 3 importance rating categories were used consistently by our low-vision patients. Three of the original 5 response categories (slightly, moderately, very important) appear to be used interchangeably within the sample. When these categories are combined, the response probability functions for importance ratings have 3 clearly defined peaks and the values of the cross points, τ_x, are properly ordered. These results suggest that patients make coarse discriminations in the importance of activities (ie, unimportant, important, very important) and that resolution of the scale cannot be improved by increasing the number of response alternatives.

Results of the Rasch analysis led us to conclude that the importance ratings of goals can be used to estimate reliable measures of a *value of independence* construct on an interval scale. The validity of the construct measurement is confirmed by the person and item fit statistics. When the sample was limited to patients with infit mean squares less than 3 SDs from the expected value, 3% had mean squares between 2 and 3 SDs from the expected value (vs 2.1% predicted by the normal distribution). Eight percent of all patients had infit mean squares that exceeded the expected value by 3 SDs or more. The patterns of importance ratings to the items by these outlying patients were highly idiosyncratic compared with the rest of the sample. Some of these patients may have misunderstood the rating scale, some may have had cognitive disorders, some may have had unusual living arrangements, and some may have had an unusually high preference for activities with low inherent social value and low preference for activities with high inherent social value. Although the responses of these patients add error to the estimates of the item measures, the number of patients with outlying mean square fit statistics was too small to influence the instrument calibration.

Four of the 41 goals (10%) had infit mean squares that exceeded the expected values by 3 SDs or more. These items contributed excessive variability to the person measure estimates. That is, patient importance ratings of these items were inconsistent with the pattern of ratings to the other items. Five of the 41 goals (12%) had infit mean squares less than the expected value by 3 SDs or more. These items exhibited less response variability across patients than expected—that is, the responses were too predictable. If responses to the items with the largest infit mean squares were excluded from the analysis, the estimates of the person measures would be more accurate and precise, and the 5 items with unexpectedly low infit mean squares would appear more consistent with the revised estimate of the distribution of response residuals. However, in addition to making measurements, our aim is to use the AI as a rehabilitation planning tool. Therefore, at this stage of development, we do not want to exclude items that will be helpful in setting rehabilitation priorities, although responses to these items could be excluded in future analyses.

One might expect differences among the importances of the 3 objectives. Indeed, goals serving the daily living objective have higher inherent social value than do goals that serve the social interactions or recreation objectives. However, the results of the separate analyses of responses to goals under each of the 3 objectives led us to conclude that there is only 1 value of independence construct. That is, the correlations among person measures for value of daily living activities, social interaction activities, and recreation activities indicated that only 1 identifiable factor and random variation were responsible for the variance in the person measures.

Visual Ability

If a person possesses far more ability than is required to achieve some goal, we would expect that person to rate that goal as "easy." At the other extreme, if a person has less ability than the goal requires, then we expect that person to rate that goal as "impossible." The difference between a person's ability (α) and the ability required to perform an activity (ρ) is called functional reserve.[17,26] The premise of our visual ability measurement model is that difficulty ratings represent ordered categories of functional reserve. In terms of the model, functional reserve is the difference between the person measure and the item measure (ie, $\alpha_n - \rho_i$).

In this study, patients appeared to use either the "slightly difficult" or "moderately difficult" rating categories. They also appeared to use either the "very difficult" or "extremely difficult" rating categories. Our results indicate that the average patient can discriminate only 4 categories of difficulty. This apparently coarse resolution of differences in functional reserve cannot be improved by increasing the number of response categories.

The goals that require the least visual ability are fundamental activities of daily living (ADLs). Only the most severely impaired low-vision patients reported difficulty performing ADLs. The goals that require the most visual ability are visually demanding hobbies and some important instrumental activities of daily living (IADLs). Even patients with relatively mild visual impairments reported difficulty performing these IADLs.

Many of the goals require the same level of ability (within estimation error). From a measurement perspective, goals that have the same difficulty provide redundant information, so redundant items would be culled. However, because no difficulty ratings are solicited for goals that are rated not important, item measure redundancy in the AI is useful. Redundancy does not necessarily refer to repetition of an item. For example, in cooking, there is no difference in required ability when done for social or recreational objectives, so these items would be redundant. However, significantly less ability is required (4 SEs) for cooking when done for the daily living objective as opposed to when done for social or recreational objectives. In this case, the repeated cooking item is not redundant because the required ability to achieve the goal is determined in part by the parent objective.

The fit statistics support the conclusion that the AI provides a valid measurement of a unidimensional visual ability variable. Based on the infit mean square, the most misfitting items to the measurement model were: follow the news, attend church, personal health care, use public restroom, and attend social functions. Large mean square values indicate high variability in the population in the functional requirements of the activity. For example, some people may follow the news mainly through television, some with newspapers and magazines, and some by radio. These 3 methods have very different visual demands; therefore, this item will contribute error to the estimate of visual ability. Normally, future iterations of development would cull the most misfitting items from the instrument to improve measurement accuracy and precision. However, these items make important contributions when using the AI to plan an individual patient's rehabilitation. Therefore, we will not remove any goals from the AI at this stage of its development, although in practice the patient responses to the misfitting items should be edited out to avoid perverting the measurements of patient ability. We currently are building our database with the AI, so we will reserve specific recommendations on editing items until we have a stronger instrument calibration.

Several self-assessment visual function instrument developers claim there are different domains assessed by their instruments.[16,35,36] Typically, developers recommend that the items be divided into subscales and that a separate score be computed for each domain. For example, the Impact of Visual Impairment Profile suggested 5 domains: leisure, household, social, mobility, and emotional.[36] To the extent that the leisure domain is the same as the recreation objective, the household domain is the same as the daily living objective, and the social domain matches the social interactions objective, one might expect there to be 3 different visual ability constructs measured by the AI. Contrary to this expectation, the present study concludes that there is a single visual ability construct. That is, the person measures are the same whether estimated from difficulty ratings for daily living items, recreation items, or social interaction items—these 3 objectives do not represent independent domains.

CONCLUSIONS

This study provides evidence that patient ratings of the importance and difficulty of specific activity goals can be used to estimate valid and reliable measures of the value the patient places on independence, his/her visual ability, the value to the patient's cohort of independently achieving each goal, and the threshold level of visual ability required to achieve each goal. Although different activity goals serve different objectives, there were only 2 constructs: value of independence and visual ability. Our sample of patients, mainly visually impaired geriatric outpatients, could distinguish 3 levels of importance and 4 levels of difficulty in their use of the ordinal rating categories. The precision of the person measure estimates from AI ratings could separate 3 strata of value of independence and 3 strata of visual ability within the patient sample. The precision of the item measure estimates could separate 21 strata of inherent social value and 7 strata of visual ability thresholds among the 41 activity goals in the AI.

References

1. Black NM, Snell AC, Patton J, Gradle HS. Report of Committee on Compensation for Eye Injuries. JAMA 1925;85:113-5.
2. GUIDES to the evaluation of permanent impairment; the visual system. J Am Med Assoc 1958;168:475-88.
3. American Medical Association. Committee on Rating of Mental and Physical Impairment. Guides to the evaluation of permanent impairment. Chicago: AMA Pr; 1971.
4. Houston TP, Doege TC, editors. Guides to the evaluation of permanent impairment. 4th ed. Chicago: AMA Pr; 1993.
5. World Health Organization. International classification of impairments, disabilities, and handicaps: a manual of classification relating to the consequences of disease. Geneva: WHO; 1980.
6. World Health Organization. International classification of functioning, disability, and health. Geneva: WHO; 2001.
7. Colenbrander A. The visual system. In: Cocchiarella L, Andersson BJ, editors. Guides to the evaluation of permanent impairment. 5th ed. Chicago: AMA Pr; 2000. p 277-304.
8. Massof RW, Dagnelie G, Deremeik J, DeRose JL, Alibhai SS, Glasner NM. Low vision rehabilitation in the U.S. health care system. J Vis Rehabil 1995;9:3-31.
9. Colenbrander A. Dimensions of visual performance. Trans Am Acad Ophthalmol Otolaryngol 1977;83:332-7.
10. Kiresuk T, Sherman R. Goal attainment scaling: a general method of evaluating comprehensive mental health programs. Community Ment Health J 1968;4:443-53.
11. Ottenbacher KJ, Cusick A. Goal attainment scaling as a method of clinical service evaluation. Am J Occup Ther 1990;44:519-25.
12. Stanley B. Evaluation of treatment goals: the use of goal attainment scaling. J Adv Nurs 1984;9:351-6.
13. Massof RW, Rubin GS. Visual function assessment questionnaires. Surv Ophthalmol 2001;45:531-48.
14. Steinberg EP, Tielsch JM, Schein OD, et al. The VF-14. An index of functional impairment in patients with cataract. Arch Ophthalmol 1994;112:630-8.
15. Mangione CM, Phillips RS, Seddon JM, et al. Development of the "Activities of Daily Vision Scale": a measure of visual functional status. Med Care 1992;30:1111-26.
16. Mangione CM, Lee PP, Pitts J, Gutierrez P, Berry S, Hays RD. Psychometric properties of the National Eye Institute Visual Function Questionnaire (NEI-VFQ). Arch Ophthalmol 1998;116:1496-504.
17. Massof RW. A systems model for low vision rehabilitation. II. Measurement of vision disabilities. Optom Vis Sci 1998;75:349-73.
18. Leat SJ, Legge GE, Bullimore MA. What is low vision? A re-evaluation of definitions. Optom Vis Sci 1999;76:198-211.
19. Massof RW. A systems model for low vision rehabilitation. I. Basic concepts. Optom Vis Sci 1995;72:725-36.
20. Committee on Disability Determination for Individuals with Visual Impairments; Lennie P, Van Hemel SB, editors; National Research Council. Visual impairments: determining eligibility for Social Security benefits. Washington (DC): Natl Acad Pr; 2002.
21. Fisher WP Jr, Harvey RF, Taylor P, Kilgore KM, Kelly CK. Rehabits: a common language of functional assessment. Arch Phys Med Rehabil 1995;76:113-22.
22. Massof RW. The measurement of vision disability. Optom Vis Sci 2002;79:516-52.
23. Massof RW. A model of the prevalence and incidence of low vision and blindness among adults in the U.S. Optom Vis Sci 2002;79:31-8.
24. Andrich D. A rating formulation for ordered response categories. Psychometrika 1978;43:561-73.
25. Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human sciences. London: Erlbaum; 2001.
26. Kirby RL. The nature of disability and handicap. In: Basmajian JV, Kirby RL, editors. Medical rehabilitation. Baltimore: Williams & Wilkins; 1984. p 14-8.
27. Linacre JM, Wright B. A user's guide to Winsteps Rasch model computer program. Chicago: Mesa Pr; 2001.
28. Wright B, Masters GN. Rating scale analysis: Rasch measurement. Chicago: Mesa Pr; 1982.
29. Smith RM. The distributional properties of Rasch item fit statistics. Educ Psychol Meas 1991;51:541-65.
30. Linacre JM. Detecting multidimensionality: which residual datatype works best? J Outcome Meas 1998;2:266-83.
31. Massof RW, Hsu CT, Baker FH, et al. Visual disability variables. II: The difficulty of tasks for a sample of low-vision patients. Arch Phys Med Rehabil 2005;86:954-67.
32. Stensen H, Wilkinson L. Factor analysis. SYSTAT 10: statistics I. Chicago: SPSS; 2000. p 327-64.
33. Seaberg JR, Gillespie DF. Goal attainment scaling: a critique. Soc Work Res Abstr 1977;13:4-9.
34. Cytrynbaum S, Ginath Y, Birdwell J, Brandt L. Goal attainment scaling: a critical review. Eval Q 1979;3:5-40.
35. Sloane ME, Ball K, Owsley C, Bruni JR, Roenker DL. The visual activities questionnaire: developing an instrument for assessing problems in everyday visual tasks. Technical digest, noninvasive assessment of the visual system. Washington (DC): Optical Society of America; 1992. p 26-9.
36. Weih LM, Hassel JB, Keeffe J. Assessment of the impact of vision impairment. Invest Ophthalmol Vis Sci 2002;43:927-35.

Suppliers

a. Winsteps, PO Box 811322, Chicago, IL 60681.
b. Version 10; Systat Software, 501 Canal Blvd, Ste E, Point Richmond, CA 94804.