

Measuring Low-Vision Rehabilitation Outcomes with the NEI VFQ-25

Joan A. Stelmack,^{1,2,3} Thomas R. Stelmack,^{2,3,4} and Robert W. Massof⁵

PURPOSE. To evaluate the sensitivity of the National Eye Institute Visual Functioning Questionnaire-25 (NEI VFQ-25) to change in visual abilities after low-vision rehabilitation in two different Veterans Administration (VA) low-vision programs

METHODS. Seventy-seven legally blind veterans from the Blind Rehabilitation Center (BRC) at Hines VA Hospital and 51 partially sighted veterans from the Visual Impairment Center to Optimize Remaining Sight (VICTORS) program at the Chicago Health Care Network, West Side Division, were administered the NEI VFQ-25 plus supplement in interview format at admission and discharge. Instructions for administration were modified to have study participants answer all the questions as if they were wearing glasses or contact lenses or were using low-vision devices. Interval measures of person ability and item difficulty were estimated from the patients' responses to 34 of the 39 items on the VFQ-25 plus supplement before and after rehabilitation, by the polytomous rating scale measurement model of Wright and Masters.

RESULTS. In VICTORS patients, item order by difficulty before rehabilitation agreed with item order for BRC patients. Visual ability scales are used similarly by different patients with different degrees of low vision. Based on prerehabilitation person measure distributions, VICTORS patients were less disabled, as would be predicted by visual acuity, than were BRC patients. After rehabilitation, estimated item difficulty for 4 of the 34 items decreased significantly in both BRC and VICTORS patients.

CONCLUSIONS. The present study demonstrates that the NEI VFQ-25 plus supplement can be used to measure the effects of low-vision rehabilitation; however, only 7 of the 34 items tested are sensitive to change after rehabilitation. Targeted activities, such as reading ordinary print, small print, and street signs are easier to perform for graduates of both programs after rehabilitation. The patients' visual ability also shows improvement in both BRC and VICTORS. Improvement in visual ability is independent of change in difficulty of targeted items. Although this was not a controlled clinical trial, the decrease in

difficulty of targeted items may reflect the use of low-vision aids and training to make tasks easier. The change in visual ability may reflect positive outcomes of rehabilitation or may be the consequence of patients' overestimates of their functional ability at the time of discharge. (*Invest Ophthalmol Vis Sci.* 2002;43:2859-2868)

There are numerous anecdotal reports of the value of low-vision rehabilitation, but there have been no clinical trials and only a few observational studies of the effectiveness of intervention. Unlike other interventions based on established protocols, such as treatment of diabetic retinopathy, low-vision rehabilitation consists of a wide range of devices and training that can be offered to patients. Clinical programs differ from one another in the range and intensity of the services they provide. At one extreme are the comprehensive 6-week inpatient blind rehabilitation centers in the VA system.¹ At the other extreme are primary care outpatient low-vision practices that are limited to dispensing low-vision aids and providing patients with minimal instruction in their use.² Despite the wide variation in the intensity and range of services offered, different programs may be able to demonstrate successful outcomes, but with success limited to the specific services that they provide. Rehabilitation outcomes would not be expected to be equivalent for different programs when there are differences in program goals. For example, dispensing a magnifier would be expected to help the patient read, but it would not be expected to help the patient with mobility.

Studies³ have reported success with low-vision rehabilitation that ranges from 23% to 100%. However, there has been little consensus on what constitutes an outcome measure. Measurement strategies frequently used include questionnaires on patient satisfaction, frequency or duration of use of low-vision devices, health-related quality of life, whether the patient perceives that the device is worth the expense, and tasks for which the patient uses the device and evaluation of performance on reading and writing tasks or measurements of visual acuity with the device.³

Although it is incontrovertible that many patients can improve their performance with low-vision devices, not all patients are successful users. In the evaluation of outcomes, some studies have recognized this difference between efficacy (the ability to perform tasks in the clinic) and effectiveness (the use of devices and skills in everyday life, once the patient returns home). Many of the early studies measured outcomes by using tests administered during evaluation and training, such as tests of reading speed or change in visual acuity with devices,³ whereas more recent studies, such as those by Watson et al.^{4,5} and Del'Aune et al.⁶ measured outcomes in terms of patient satisfaction with services provided and the ability to perform activities after rehabilitation. They assessed use of low-vision devices among military veterans who participated in either the VA Visual Impairment Center to Optimize Remaining Sight (VICTORS) program or blind rehabilitation center programs. Most (85.4%) of the prescribed devices were still in use and considered by the veterans to be of substantial benefit. Del'Aune et al.⁶ observed on patient self-assessments a high level of satisfaction with VA blind rehabilitation services and

From the ¹Blind Rehabilitation Center, Edward Hines VA Hospital, Hines, Illinois; the ²Illinois College of Optometry, Chicago, Illinois; the ³Department of Ophthalmology and Visual Science, University of Illinois at Chicago, Chicago, Illinois; the ⁴Chicago Health Care Network, West Side Division, Chicago, Illinois; and the ⁵Wilmer Eye Institute, Johns Hopkins University, Baltimore, Maryland.

Supported by the Office of Research and Development, Rehabilitation Research and Development Service of the Department of Veterans Affairs (JAS, TRS); Illinois Society for Prevention of Blindness, Chicago (JAS, TRS); and Grant EY-12045 of the National Eye Institute, National Institutes of Health, Bethesda, Maryland (RWM).

Submitted for publication October 30, 2001; revised April 10, 2002; accepted April 26, 2002.

Commercial relationships policy: N.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: Joan A. Stelmack, Blind Rehabilitation Center, Edward Hines Hospital, Hines, IL 60141; joan.stelmack@med.va.gov.

large gains in ability to read mail or cross a street with a traffic light after rehabilitation.

Rather than focus on satisfaction and the performance of specific activities, some have argued that the outcomes of low-vision services should be measured by improvement in quality of life.^{3,7-9} Although there are many different definitions of quality of life, current thinking is that quality of life is multidimensional. Included are a physical dimension (disease symptoms and their treatment), a functional dimension (self-care, mobility, activity level, and activities of daily living), a social dimension (social contact, interpersonal relationships), and a psychological dimension (cognitive function, emotional status, well-being, life satisfaction, and happiness).¹⁰ Because the patient's own perception of his or her quality of life may be one of the most meaningful outcomes of a treatment, and because most health-related quality of life self-assessment instruments presume that the patient has systemic disease, the National Eye Institute (NEI) contracted with the Rand Corporation (Santa Monica, CA) to develop the NEI Visual Functioning Questionnaire (NEI VFQ), an instrument designed to measure vision-specific health-related quality of life.¹¹⁻¹⁵

Recently, there has been interest in using the NEI VFQ to measure outcomes of low-vision care. Both Scott et al.⁸ and Stelmack et al.^{7,9} have reported that the NEI VFQ-25 raw scores before rehabilitation change after the training. Such changes indicate a trend in the right direction, but raw scores are not measures.¹⁶ Furthermore, the recommended method of scoring the NEI VFQ, based on sums of ratings for each of 12 arbitrarily defined domains,¹⁴ invites multiple statistical comparisons and confuses the selection of a primary outcome variable. Finally, as worded, the NEI VFQ-25 instructions to the patient do not allow for use of low-vision aids. Stelmack et al.⁹ pointed out the need to modify directions on the instrument to include use of low-vision aids, particularly for measuring the outcome of low-vision rehabilitation. Low-vision rehabilitation enhances but does not restore remaining vision. The NEI VFQ-25 includes questions that ask patients to rate their difficulty in performing activities of daily living when using glasses or contact lenses. Low-vision devices enable patients to perform tasks such as reading newsprint. If patients rate their difficulty in reading newsprint according to the NEI VFQ directions, while using only glasses or contact lenses, the outcome or change in patients' performance attributed to the use of a low-vision device will not be measured.¹⁷

Before a meaningful measure of the outcome of low-vision rehabilitation can be made, there must be a clear understanding of the construct that intervention is attempting to modify. Chronic visual impairments reduce the patient's ability to perform everyday activities with customary ease. In other words, chronic visual impairments cause vision disabilities. Each activity requires some threshold level of visual ability. Each patient possesses a level of visual ability that is modulated by the type and extent of visual impairment. If a patient's visual ability far exceeds that required to perform an activity, the patient will find that activity easy to perform. If the patient's visual ability is only slightly greater than the visual ability required to perform the activity, the patient will be able to perform it, but it will be difficult. If the patient's visual ability is less than that required to perform the activity, the patient will not be able to perform the activity at all. The difficulty the patient experiences with an activity depends on the difference between the patient's visual ability and the visual ability necessary to perform the activity. This difference is called functional reserve.¹⁸⁻¹⁹

The goal of low-vision rehabilitation is to make it easier for visually impaired persons to perform everyday activities. Regarding the visual ability construct just discussed, the goal of low-vision rehabilitation is to increase functional reserve. Func-

tional reserve for a given activity can be increased either by increasing the visual ability of the patient or by decreasing the visual ability needed to perform the activity.¹⁸⁻¹⁹ For example, refractive error correction increases the patient's visual ability by correcting the impaired vision. Therefore, refractive error correction would increase functional reserve for many activities, depending on the magnitude of the refractive error. In contrast, low-vision rehabilitation consists of providing the patient with assistive devices and training. Unlike refractive error correction, low-vision aids do not modify the patient's vision, they modify the visual activity by reducing the visual ability needed to perform the activity, thus increasing functional reserve for that activity. Training—for example, eccentric viewing training in which the patient is taught to use a part of the retina that has higher visual acuity—can improve a patient's visual ability or it can have an effect that is limited to a specific activity—for example, organizing clothes in closets and drawers. In either case, functional reserve for the activity is increased, thereby making the activity easier to perform.

With instruments such as the NEI VFQ, in which the patient is asked to rate the difficulty of performing a specific activity, we presume the patient is judging his or her functional reserve for that activity.¹⁹ Less-specific questions, such as rating eyesight quality or rating the frequency of obtaining help from others, are most likely to ascertain the patient's judgment of a weighted average functional reserve across several activities. Thus, despite different response scales and differences in reasons for asking the questions (i.e., different domains), all questions are likely to ascertain a single visual ability variable, albeit with different levels of precision. This expectation was confirmed for the NEI VFQ by Rasch analysis of the responses of a large group of patients with low vision.¹⁶

The questions on the NEI VFQ that require difficulty ratings for specific activities are likely to be selectively sensitive to the effects of interventions targeted to easing the performance of those activities. In contrast, the more general questions are likely to be sensitive only to changes in the patient's visual ability or to targeted reductions in difficulty in a broad range of activities (which, in practice, may be indistinguishable from a change in the patient's visual ability). Responses to some questions, such as rating general health, most likely would not be affected at all by low-vision rehabilitation. Because in most cases low-vision rehabilitation is targeted to ease the performance of specific activities and does not alter the patient's vision, we expected that changes in functional reserve as a consequence of rehabilitation would be unequal for different items on the NEI VFQ.

The purpose of this study was to evaluate the sensitivity of the NEI VFQ-25 to change in visual ability after low-vision rehabilitation. The effects of rehabilitation on the visual ability of patients with low vision and on the minimum visual ability needed for each item in the NEI VFQ were measured. Rasch analysis was used to estimate visual ability separately in each patient and required visual ability for each item, by using the patient's responses to the NEI VFQ before and after rehabilitation. In the study, we compared the observed outcomes with the NEI VFQ of two low-vision rehabilitation programs of differing intensity in the VA.

METHODS

Programs

The sensitivity of the NEI VFQ-25 to change after rehabilitation was compared in two low-vision programs, the Visual Impairment Center to Optimize Remaining Sight (VICTORS) at the VA Chicago Health Care System, West Side Division (Chicago, IL) and the Blind Rehabilitation Center (BRC) at Edward Hines VA Hospital (Hines, IL). These programs

represent the continuum of care provided by the Department of Veterans Affairs for patients who have mild to severe visual loss. The VICTORS, a less-intense program, is an interdisciplinary low-vision rehabilitation service that includes optometry, ophthalmic examination, psychological evaluation, assessment by a social worker, audiologic examination, and sessions with a low-vision training specialist. VICTORS patients are housed as lodgers in the hospital for 3 to 4 days or receive services on an outpatient basis.¹ The VICTORS program at West Side Division is targeted to serving visually impaired veterans who are not legally blind.

The Hines VA BRC, a more intense program, is a comprehensive rehabilitation service usually offered inpatients at the residential center. An interdisciplinary team of optometrists, psychologists, nurses, physicians, social workers, and blind rehabilitation specialists guide the individual through a program with the broad goals of "maximum adjustment to the disability, reorganization of the person's life, and return to a contributing place in the family and community."²⁰ Veterans participate in a variety of skill courses (including low-vision evaluation and training, orientation and mobility training, daily-living skills, and manual skills to help achieve a realistic level of independence) and receive counseling to develop a healthy attitude toward themselves, vision loss, and the future.²⁰ The duration of the BRC program averages 42 days. Patients must be legally blind, with best corrected visual acuity of 20/200 or less in the better-seeing eye or visual field diameter of 20° or less measured with Goldmann perimetry to qualify for services. Eligibility for services is based on visual acuity and or visual fields reported by the referring VA Eye Clinic. Because these tests are performed by students, residents, attending physicians, nurses, and technicians, the accuracy of the findings varies. Visual acuity is carefully measured and manifest refraction performed during admission. It is the policy of the rehabilitation center to allow veterans with best corrected visual acuity of 20/100 or less to complete the comprehensive blind rehabilitation program. These veterans receive all services and prosthetic equipment provided by the BRC. Veterans with visual acuity better than 20/100 and who are not legally blind according to visual field testing are discharged after a brief low-vision rehabilitation program and were not included in the study sample.

Subjects

Subjects included 51 consecutive patients with low vision from the VICTORS program and 77 consecutive legally blind patients from the BRC. The study followed the tenets of the Declaration of Helsinki for research in human subjects. Written informed consent was waived by the Institutional Review Board based on Section 46.116D#1 Code of Federal Regulations General Requirement for Informed Consent, because the research involved no more than minimal risk to the patients.²¹ Excluded were patients with severe cognitive or hearing deficits who could not participate in the interview. Five veterans at the BRC were withdrawn from the study. These veterans could not complete the rehabilitation program because of deterioration in health status or a visual condition that resulted in inability to use the prescribed low-vision devices. These veterans were excluded from the study, because it is the policy of the rehabilitation center to offer veterans who do not complete the rehabilitation program the opportunity to reapply for training at the BRC after their visual condition or health status has stabilized. The time frame of the study did not permit these patients to be observed through readmission and program completion. Two veterans in the VICTORS program died.

Mean best corrected visual acuity was 0.54 log minimum angle of resolution (logMAR; 20/63) in the VICTORS patients and 1.00 logMAR (20/200) in the BRC patients. Thirty-six patients had best corrected visual acuity measured at the BRC greater than 1.0 logMAR, and eight patients had visual acuity less than 0.7 logMAR (20/100 Snellen). VICTORS patients included 40 men and 11 women. Mean age of VICTORS patients was 68 years (range, 44–87). BRC patients included 72 men and 5 women, with a mean age of 72 years (range, 38–88). Diagnoses most frequently reported in the VICTORS patients were

macular degeneration (39.2%), glaucoma (21.6%), and diabetic retinopathy (17.6%). Diagnoses most frequently reported in the BRC patients were macular degeneration (66.2%), diabetic retinopathy (15.6%), and glaucoma (11.7%).

Instruments

The NEI VFQ-25 plus supplement was administered to all subjects by personal interview.¹⁵ The interviews were conducted by two interviewers at Hines and one interviewer at West Side Division at baseline, before intervention, and at the conclusion of the rehabilitation program, before discharge. Because the intent of this study was to evaluate the NEI VFQ-25, not to conduct a clinical trial of low-vision outcomes, masking was not included in the study design. The NEI VFQ-25 consists of 25 items and a supplement of 14 additional items, all of which were taken from the original 52-item NEI VFQ.¹⁴ Among the 39 items of the NEI VFQ-25 plus supplement, 6 ask patients to grade their general health and vision, 20 rate difficulty with activities, and 13 ask for the level of agreement with statements describing the severity of problems associated with vision loss. The questions on difficulty with activities were rated on a 1-to-6 scale, with response choices including no difficulty, a little difficulty, moderate difficulty, extreme difficulty, stopped doing this because of your eyesight, and stopped doing this for other reasons/not interested. A rating response of 6 was scored as missing data. The questions on level of agreement with statements describing role limitations due to vision loss were rated on a 5-point scale ranging from agree all of the time to agree none of the time for five of the items and ranging from definitely true to definitely false for the remaining eight items. Two items in the supplement rated overall health and vision on a 0 (worst) to 10 (best) scale.

The general directions (copyrighted by the Rand Corporation) are, "I am going to read you some statements about problems which involve your vision or feelings that you have about your vision conditions. After each question I will read you a list of possible answers. Please choose the response that best describes your situation. Please answer all the questions as if you were wearing your glasses or contact lenses (if any)."¹⁵ These directions were modified for both the pre- and postrehabilitation interviews to add consideration of use of low-vision devices and to clarify the information requested in three ways: (1) The directions for part 2: difficulty with activities question 5 to 14 were restated to "The next questions are about how much difficulty, if any, you have doing certain activities wearing your glasses, contact lenses or using low-vision devices if you have them for that activity." (2) The administration protocol for the NEI VFQ-25 does not address or consider the need for repetition of the directions. The directions were repeated as often as necessary during administration of the questionnaire, because veterans frequently forgot the instructions provided by the interviewer for answering the questions.⁹

Analysis

The recommended NEI VFQ scoring algorithm that consists of linear transformations of raw scores for each of 12 domains was not used in this study. To achieve the goals of the present study, we estimated the visual ability of each patient and the minimum visual ability required for a threshold response to each item. The visual ability of individual patients may be related to the raw score, but such a relationship would be nonlinear because of the ceiling and floor in the rating scale and the implicit assumption in Likert scoring that every item has the same value.¹⁶ Also, the domains were arbitrarily defined by the NEI VFQ developers and are highly intercorrelated¹²; thus, individual domain scores are redundant estimates of the same variable.¹⁶

We used the rating scale model of Wright and Masters,²² which is based on the partial credit model of Masters,²³ the polytomous response scale model of Andrich,²⁴ and the logistic version of the probabilistic measurement model of Rasch.²⁵ The Rasch model assumes that the probability patient n will give response x to item i depends only on functional reserve. That is, the probability of obtaining response x depends on the difference between the patient's visual ability

(α_n) and the visual ability required by the item (ρ_i). The only other parameter in the model is the threshold value of functional reserve required for response x (τ_x).^{16,19}

The Rasch analysis involves the estimation of α_n (person ability) for each patient p_i (item difficulty) for each item, and τ_x for each response category from the matrix of patients' ratings of the items on the questionnaire. It includes parametric tests of construct validity and estimation reliability. We used an unconditional maximum-likelihood estimation routine (Winsteps, see Ref. 26) to perform Rasch analyses on the pre- and postrehabilitation NEI VFQ-25 data for both the BRC and VICTORS patient samples. Winsteps provides estimates of the three model parameters and tables of estimation errors, reliability coefficients, and fit statistics.

Although all 39 items were administered during the study, the three items on driving (questions 15, 16, and A10) were eliminated from the final analysis because very few of the patients were allowed to drive, and the data were insufficient for analysis. The two questions in the supplement that rate vision and health (A1 and A2) were also eliminated. These are the only two items with other than a 5-point rating scale (they have a 10-point scale), and therefore the responses cannot be analyzed with the other data without collapsing categories. Because these items are nearly identical with questions 1 and 2 in the main part of the NEI VFQ-25, the information was not lost by excluding them from the analysis.

Response polarity had to be reversed for items 17 through 25 and A11a through A13. For these 13 items, a response of 5 is most positive and 1 is most negative. For the other 25 items, a response of 1 is most positive and 5 is most negative. The response scale was reversed by subtracting the response rating from 6.

RESULTS

Prerehabilitation Results

Rasch analyses were performed on the prerehabilitation NEI VFQ-25 ratings of BRC patients and separately on the ratings of VICTORS patients.

Figure 1 shows item measures (ρ_i) estimated from the VICTORS prerehabilitation data versus item measures estimated from BRC prerehabilitation data. Each data point represents a different item on the NEI VFQ-25; the error bars are ± 1 SE. If the item measures for the two patient samples were in perfect agreement, they would fall along the identity (solid) line.

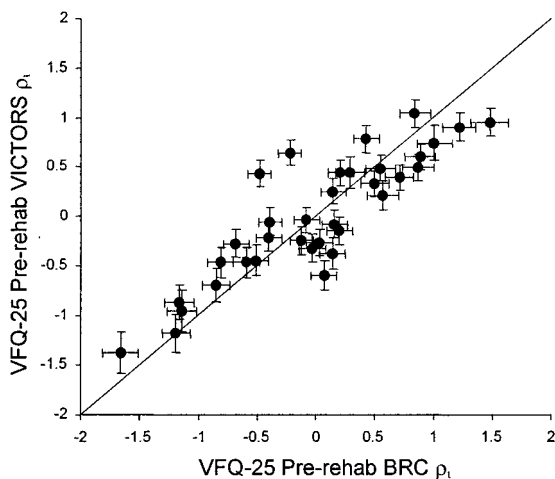


FIGURE 1. Item measures (ρ_i) estimated from the VICTORS prerehabilitation data versus those estimated from BRC prerehabilitation data. Each data point represents a different item on the NEI VFQ-25. Error bars: ± 1 SE.

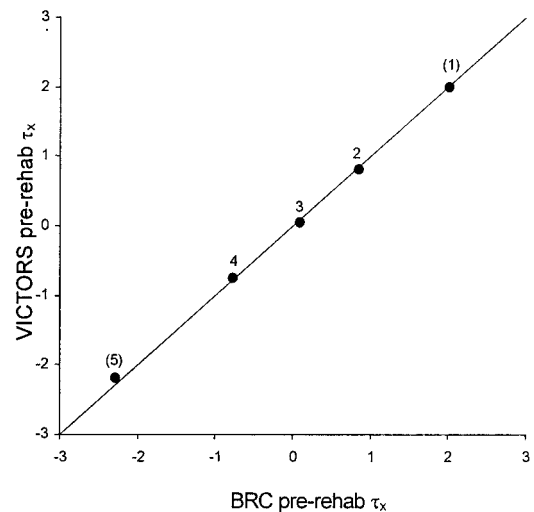


FIGURE 2. Comparison of the VICTORS and BRC prerehabilitation rating category thresholds, τ_x (in functional reserve units of $\alpha - \rho$). Numbers beside data points: corresponding rating categories for each pair of τ values. The extreme categories 1 and 5 are in parentheses, because those values were estimated from the average functional reserve for the patient-item encounters that had that rating. Solid line: identity line.

The Pearson product moment correlation between the two sets of item measures is 0.87. The bivariate linear regression has a slope of 0.85 and an intercept of 0.0017, not significantly different from 1 and 0, respectively ($P = 0.57$). These results indicate that visual ability scales estimated from the BRC and VICTORS prerehabilitation patients' responses to the NEI VFQ-25 are the same.

Figure 2 compares the VICTORS and BRC prerehabilitation rating category thresholds τ_x (in functional reserve units of $\alpha - \rho$). The numbers beside the data points indicate the corresponding rating categories for each pair of τ values (τ represents functional reserve corresponding to the maximum response probability for that rating category—not “step measures”). The extreme categories, 1 and 5, are in parentheses, because those values were estimated from the average functional reserve for the patient-item encounters that had that rating, rather than the peak of the response probability function. The solid line through the data is the identity line. The bivariate linear regression has a slope of 0.97 and an intercept of -0.01 , values that are not significantly different from 1 and 0, respectively ($P = 0.2$). These results indicate that the prerehabilitation BRC and VICTORS patients used the five rating categories on the NEI VFQ-25 items in the same way with the same average response criteria.

As might be expected from the programs' eligibility criteria, VICTORS patients had greater visual ability than did BRC patients.

Figure 3 compares the distribution of visual ability measures (α) for VICTORS patients (gray bars), which has a mean of 0.43 logits (median = 0.48) and an SD of 0.63 logits, to the distribution of visual ability measures in BRC patients (black bars), which has a mean of -0.08 logits (median, -0.19) and an SD of 0.60 logits. Although this difference does not achieve statistical significance (t -test: $P = 0.09$), based on a linear relationship between the person measure and logMAR visual acuity observed earlier,¹⁶ the 0.51-logit difference between the means is equivalent to a 0.425-logMAR difference in visual acuity.¹⁶ This agrees with the 0.45-logMAR difference in the present study between the average visual acuities of the VICTORS and BRC patients.

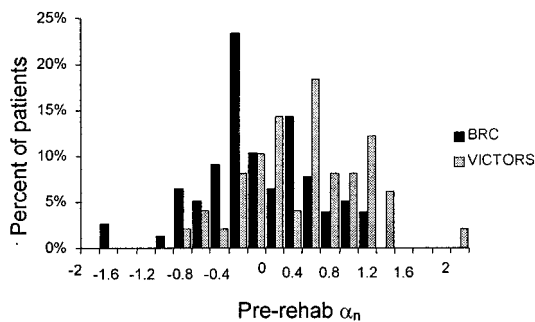


FIGURE 3. Comparison of the distribution of vision ability measures (α) for VICTORS patients (*shaded bars*) with the distribution of vision ability values for BRC patients (*filled bars*).

Comparison of Fit Statistics for Pre- to Postrehabilitation Estimates of Person Ability and Item Difficulty

Rasch analyses were performed separately on pre- and postrehabilitation responses of VICTORS and BRC patients to the NEI

VFQ-25. Figure 4 compares the normalized information weighted fit statistics (*infit*)²² for pre- versus postrehabilitation item difficulty estimates (Figs. 4a, 4b) and person ability estimates (Figs. 4c, 4d), for both the BRC (Figs. 4a, 4c) and VICTORS (Figs. 4b, 4d) data sets. The fit statistics are consistent with the expectations of the model (expected value is 0 with a tolerance of ± 2 SD, which is represented by the box in each figure) for both BRC and VICTORS pre- and postrehabilitation data. There is no item that consistently misfits model expectations across data sets (points above or to the right of the box in Figs. 4a, 4b). These results confirm the construct validity of the measurements.

Comparison of Pre- with Postrehabilitation Estimates of Response Category Thresholds

The logit scales for person ability (α) and item difficulty (ρ) are dimensionless. Because of variability in the interpretation of the item or in personal circumstances that apply to the item (e.g., illumination, working distance, print size), ρ_i has some distribution in the patients. The variances of the item difficulty distributions may not be the same for different items (these differences are captured in the *infit* mean square). The item

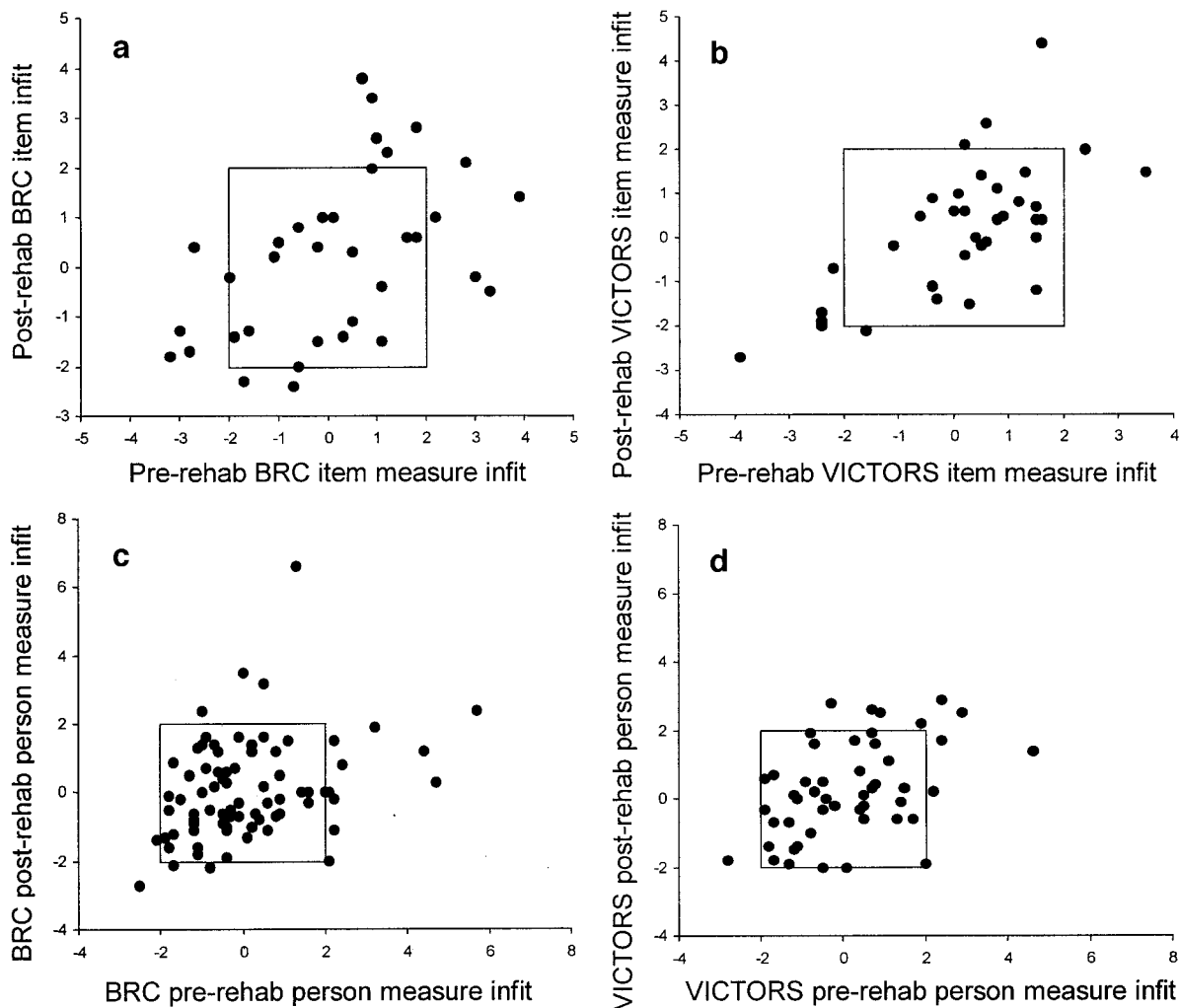


FIGURE 4. Normalized information weighted fit statistics (*infit*) for pre- versus postrehabilitation. Fit statistics are consistent with the expectations of the model (expected value is 0 with a tolerance of ± 2 SD, which is represented by the *box*). (a) Pre- versus postrehabilitation *infit* for item measures estimated from BRC data. Each point corresponds to an item. (b) Pre- versus postrehabilitation *infit* for item measures estimated from VICTORS data. (c) Pre- versus postrehabilitation *infit* for person measures estimated from BRC data. Each point corresponds to a patient. (d) Pre- versus postrehabilitation *infit* for person measures estimated from VICTORS data.

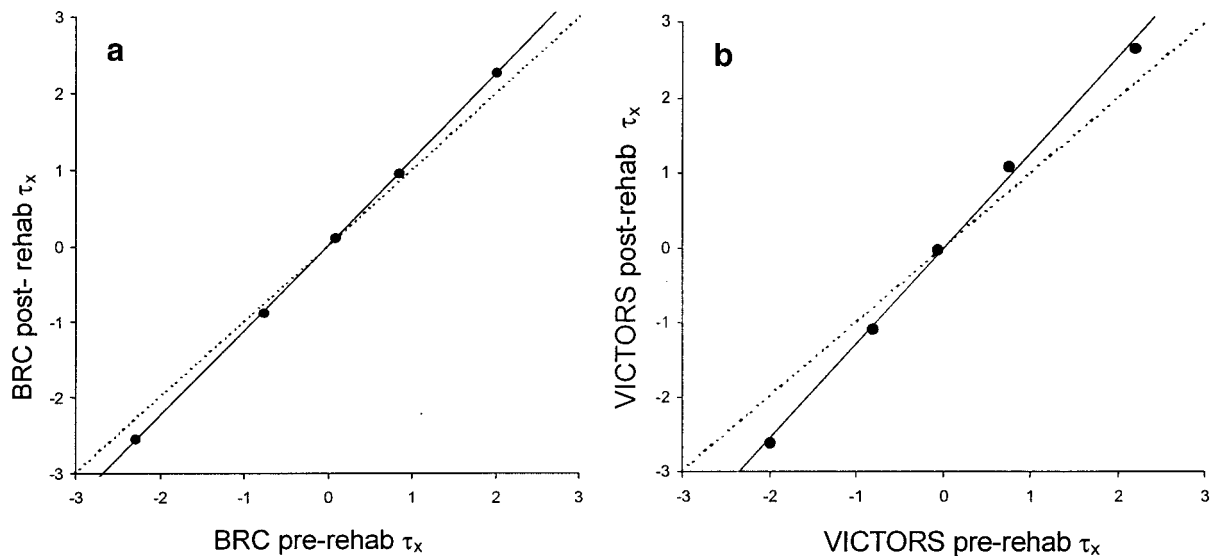


FIGURE 5. Post- versus prerehabilitation estimates of τ_x for each of the five rating categories for the (a) BRC and (b) VICTORS patients' responses. Dashed line: identity line; solid line: bivariate regression.

difficulty estimated by Winsteps is equivalent to the mean of the item difficulty distribution normalized to the average of the item difficulty distribution SDs. Using default parameters, Winsteps sets the origin of the scale to the average value of ρ for the items included in the instrument. Consequently, if rehabilitation changes the difficulty of some items but not of others and changes the average item difficulty variability, then pre- and postrehabilitation logit scales would be normalized differently and have different origins. In other words, pre- and postrehabilitation logit scales might be linearly related, but with a slope that reflects the ratio of the average item difficulty variances and an intercept that reflects the difference between the average item difficulty means.

Based on these properties of the logit scale, we hypothesize that the postrehabilitation visual ability scale (α') is a linear transformation of the prerehabilitation visual ability scale (α), where the slope (m) corresponds to the ratio of the average item difficulty SDs and the intercept (b) corresponds to the difference between the average item difficulty means. This hypothesis predicts that postrehabilitation functional reserve at the step between response categories (τ_x) is proportional to the prerehabilitation value, with a proportionality constant m . In other words, because $\tau = \alpha - \rho$ and

$$\alpha' - \rho' = m\alpha + b - m\rho - b = m(\alpha - \rho)$$

we conclude that

$$\tau_x' = m\tau_x$$

which means that we can estimate m from the linear relationship between pre- and postrehabilitation response category thresholds. The linear relationship will have an intercept at the origin.

Figure 5 shows scatterplots of the post- versus prerehabilitation estimates of τ_x for each of the five rating categories of the BRC (Fig. 5a) and VICTORS (Fig. 5b) patients' responses. The dashed line in each figure is the identity line, and the solid line is the bivariate regression. The bivariate regression for the BRC data has a slope of 1.12, which is significantly different from 1 ($F_{4,4} = 130$; $P < 0.001$) and an intercept of 0.004, which is not significantly different from 0 ($P = 0.39$). The bivariate regression for the VICTORS data has a slope of 1.27,

which is significantly different from 1 ($F_{4,4} = 17.5$; $P < 0.01$) and an intercept of -0.01 , which is not significantly different from 0 ($P = 0.49$). These results agree with our hypothesis and indicate that the average variance of the item difficulty distributions is greater for post- than for pre-rehabilitation data from both BRC and VICTORS patients.

Comparison of Pre- to Postrehabilitation Estimates of Item Difficulty

To determine whether rehabilitation of BRC patients altered the value of ρ for specific items, postrehabilitation measures of item difficulty were subtracted from rescaled prerehabilitation measures of item difficulty (i.e., prerehabilitation item difficulty multiplied by 1.12, the slope of the bivariate regression in Fig. 5a).

Figure 6 illustrates the difference in ρ for each item included in the analysis (points ± 1 SE of the estimate). Seven of the items (open circles) had ρ estimated from postrehabilitation patient responses that were significantly less than ρ estimated from prerehabilitation responses ($P < 0.01$ with Bonferroni

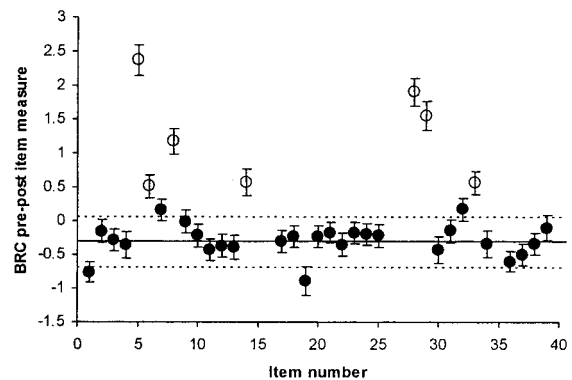


FIGURE 6. Difference in ρ for each item included in the analysis (points ± 1 SE of the estimate). Seven of the items (○) had ρ estimated from postrehabilitation patient responses that was significantly less than ρ estimated from prerehabilitation responses. For the other items (●), there was no significant difference in ρ , or the prerehabilitation ρ was less than the postrehabilitation value.

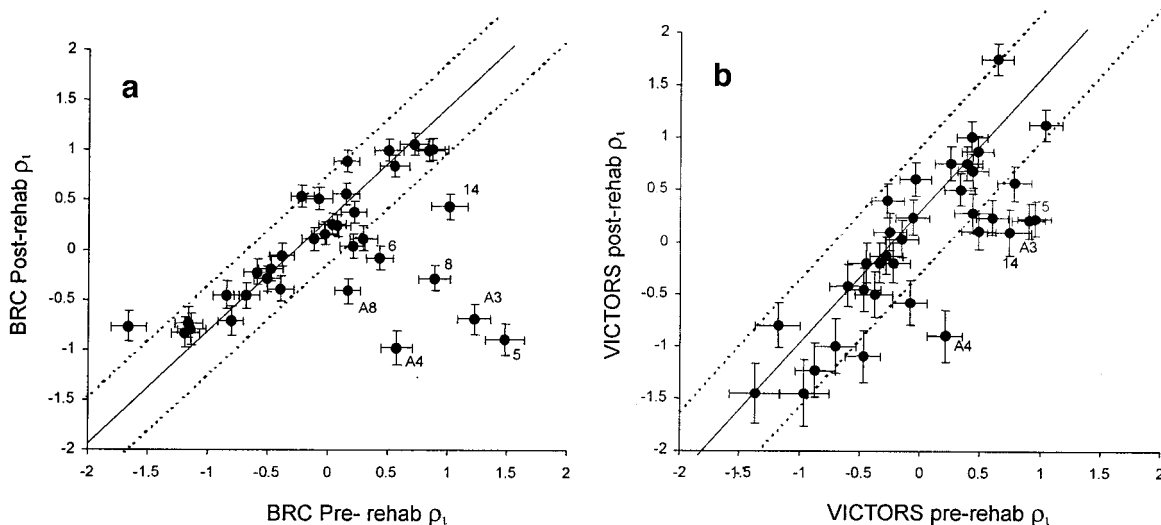


FIGURE 7. (a) Comparison of the pre- and postrehabilitation item measures for the (a) BRC and (b) VICTORS patients (points ± 1 SE of the estimate). *Solid line*: hypothesized identity line.

correction). For the other items (filled circles), there was no significant difference in ρ , or the prerehabilitation ρ was less than the postrehabilitation ρ (i.e., the items were more difficult after rehabilitation).

Excluding the seven items that showed significant improvement after rehabilitation, the average item difficulty change was -0.3 logit (Fig. 6, solid horizontal line). The seven items that had significantly lower ρ after rehabilitation would be expected to change the origin of the postrehabilitation logit scale. Thus, the -0.3 logit average item difficulty change for the other 27 items reflects the change in origin and corresponds to the intercept (b) for the hypothesized linear relationship between pre- and postrehabilitation scales. The dashed lines are ± 3.6 SE, the 99% confidence limits around the average difference with Bonferroni correction for multiple comparisons.

Figure 7a compares the measures of pre- and postrehabilitation item difficulty for the BRC patients (points ± 1 SE of the estimate). The solid line through the points is the hypothesized identity line with a slope of 1.12 and an intercept of 0.3. The dashed lines are the 99% confidence limits around the identity line. The seven points below and to the right of the identity line are the seven items that had significantly lower ρ after rehabilitation. These activities, which are labeled with the item number on the NEI VFQ, were easier, in comparison with other items, for BRC patients to perform after rehabilitation than before rehabilitation.

The ρ for item 5 (difficulty reading ordinary print in the newspaper) changed the most after rehabilitation. That item went from being the most difficult item before rehabilitation to being the easiest item after rehabilitation. All the items that exhibited significant reductions in ρ after rehabilitation elicited difficulty ratings from the patient for specific daily activities (5, difficulty reading ordinary print in newspapers; 6, difficulty doing work or hobbies that require you to see well up close; 8, difficulty reading street signs or the names of stores; 14, difficulty going out to see movies, plays, or sports events; A3, difficulty reading small print in a telephone book, on a medicine bottle, or on a legal form; A4, difficulty figuring out whether bills you receive are accurate; and A8, difficulty seeing and enjoying programs on television).

The same analysis of item difficulty was applied to estimates of ρ from pre- and postrehabilitation responses of VICTORS patients. When the slope value of 1.27 from the bivariate

regression illustrated in Figure 5b was used to rescale the prerehabilitation measures of item difficulty, the average difference for those items that did not significantly improve was also -0.3 logit.

Figure 7b compares postrehabilitation estimates of ρ to prerehabilitation estimates for the VICTORS patients (solid points ± 1 SE of the estimate). The solid line is the hypothesized identity line with a slope of 1.27 and an intercept of 0.3. The dashed lines bound the 99% confidence interval (with correction for multiple comparisons). For four items, which are identified in Figure 8, ρ after rehabilitation was significantly less than ρ before rehabilitation: 5, difficulty reading ordinary print in newspapers; 14, difficulty going out to see movies, plays, or sports events; A3, difficulty reading small print; and A4, difficulty figuring out whether bills you receive are accurate.

The four items that were significantly less difficult for the VICTORS patients after rehabilitation were also significantly less difficult for the BRC patients after rehabilitation. Because the prerehabilitation item difficulty scales for BRC and VICTORS patients are the same (Fig. 1), the postrehabilitation measures of item difficulty were back-transformed to the prerehabilitation item difficulty scale to compare item difficulty

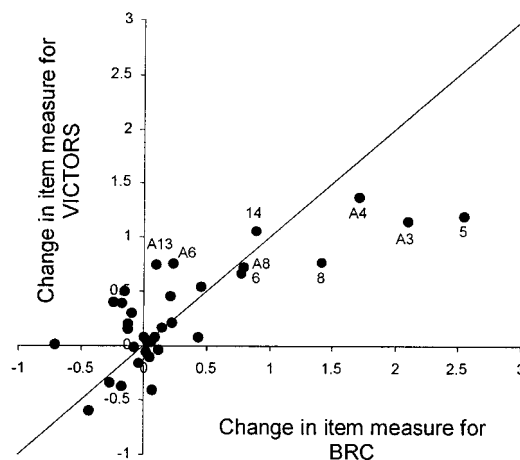


FIGURE 8. Change in item measure from prerehabilitation to postrehabilitation in the VICTORS compared with that in BRC patients.

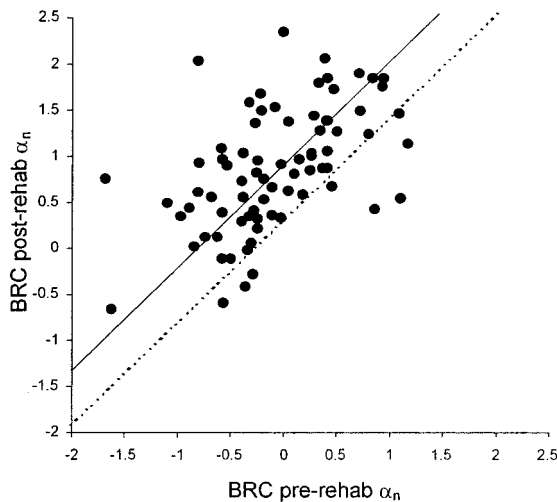


FIGURE 9. BRC post- versus prerehabilitation person measures. *Dashed line:* hypothesized identity line. Most of the data points fall *above* the identity line, indicating that postrehabilitation α was, on average, greater than the prerehabilitation value. *Solid line:* bivariate regression.

changes for BRC patients with those for VICTORS patients in the same units.

Figure 8 illustrates the change in item difficulty from before to after rehabilitation for the VICTORS compared to BRC patients. If the change in item difficulty after rehabilitation were the same for the two groups of patients, the points would fall along the identity (solid) line. For three of the items (8, difficulty reading street signs or the names of stores; A3, difficulty reading small print in a telephone book, on a medicine bottle, or on a legal form; and 5, difficulty reading small print in the newspaper), the change for BRC patients was significantly larger than the change for VICTORS patients. For two of the items, (A6, difficulty you have recognizing people you know across a room; and A13, agree that I do not go out of my home alone because of my eyesight), the change for VICTORS patients was larger (but not significantly so) than the change for BRC patients.

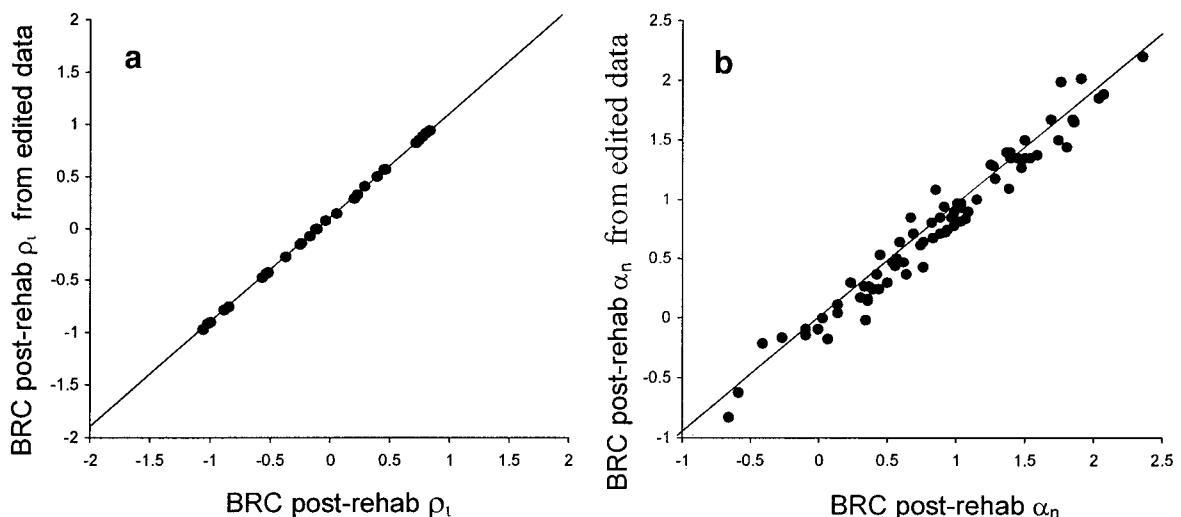


FIGURE 10. (a) Item measures estimated from the edited data set are the same as the postrehabilitation item measures estimated earlier, but with a different origin (due to a 0.1-log-unit difference in the average postrehabilitation item measure with the seven items deleted). (b) Postrehabilitation person measures are the same when responses to the seven items are deleted as they are when those responses are included. *Solid line:* identity line.

Comparison of Pre- with Postrehabilitation Estimates of Person Abilities

If low-vision rehabilitation has only the effect of making targeted activities easier to perform, then postrehabilitation person abilities (α) should be the same as prerehabilitation person abilities, once corrections are made for linear changes in the scale (i.e., change in origin and item difficulty variability with postrehabilitation changes in item difficulty).

Figure 9 shows BRC post- versus prerehabilitation person abilities. The dashed line is the identity line (slope of 1.12 and intercept of 0.3 logit). Most of the data points fall above the identity line, indicating that postrehabilitation α was, on average, greater than the prerehabilitation α . The solid line through the data is the bivariate regression with the slope constrained to 1.12. The intercept of this best-fit line is 0.87 logit. In prerehabilitation person ability units, this intercept represents an average increase in α after rehabilitation of 0.51 logit.

For the rating-scale version of the Rasch model used here, the measures of person ability and item difficulty are separable.^{19,22} Therefore, the increase in person visual ability after rehabilitation should be independent of the observed decrease in item difficulty after rehabilitation for seven of the items. To test this prediction, the postrehabilitation VFQ-25 data were reanalyzed, after deletion of data from the seven items for which BRC patients' responses had changed significantly.

As shown in Figure 10a, the measures of item difficulty estimated from the edited data set are the same as the postrehabilitation measures of item difficulty estimated earlier, but with a different origin (due to a 0.1 logit difference in the average postrehabilitation item difficulty with the seven items deleted). Figure 10b shows that postrehabilitation person abilities were the same when responses to the seven items were deleted as they were when those responses were included. The solid line through the data is the identity line. These results confirm the expectation based on measures of person ability and item difficulty separability in the model that the increase in α after rehabilitation is independent of the postrehabilitation decrease in item difficulty for the seven items. Person abilities estimated from the responses of VICTORS patients also exhibited an increase in visual ability after rehabilitation.

Figure 11 illustrates postrehabilitation α in comparison with prerehabilitation α in the VICTORS patients. The dashed line is

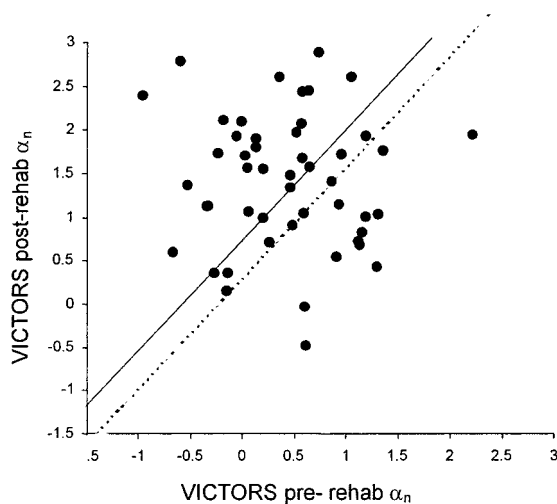


FIGURE 11. Postrehabilitation versus prerehabilitation α in the VICTORS patients. *Dashed line:* identity line; *solid line:* bivariate regression.

the identity line with a slope of 1.27 and an intercept of 0.3 logit. The solid line is the best-fit bivariate regression with the slope constrained to 1.27. The intercept for the regression line is 0.74 logit. This intercept means that the average increase in α after rehabilitation is 0.35 logit (in prerehabilitation units).

The improvement in visual ability of BRC patients after rehabilitation is significant ($P < 0.001$ for paired-comparison *t*-test) and equivalent to improved ability associated with a 0.425-logMAR improvement in visual acuity.¹⁶ In VICTORS patients, the postrehabilitation improvement in visual ability is also significant ($P < 0.001$) and is equivalent to improved ability associated with a 0.3-logMAR improvement in visual acuity.

DISCUSSION

Seven items on the NEI VFQ-25 that ask patients to rate the difficulty of performing activities of daily living were responsive to the effects of rehabilitation. These seven activities were significantly less difficult after rehabilitation in the BRC program, and four of these activities were also significantly less difficult after rehabilitation in the VICTORS program. The activities that were significantly less difficult after rehabilitation in both programs were reading ordinary print in newspapers (item 5); going out to see movies, plays, or sports events (item 14); reading the small print in a telephone book, on a medicine bottle, or on a legal form (item A3); and figuring out whether bills you receive are accurate (item A4). The ability to perform these activities would be expected to be responsive to near magnification (items 5, A3, A4) or telescopic distance magnification (item 14). Dispensing such magnifiers and training patients to use them are major components of both the VICTORS and BRC programs. The other three activities that became significantly less difficult for BRC patients after rehabilitation (and also less difficult for VICTORS patients after rehabilitation, but not significantly so) involved abilities that also would be expected to be responsive to magnification: see well up close (item 6), reading street signs or the names of stores (item 8), and seeing and enjoying programs on television (item A8).

Patients from both the VICTORS and BRC programs exhibited a significant increase in visual ability after completing the rehabilitation programs. This increase in visual ability is independent of the decrease in difficulty observed for the seven activities (see Figures 10a and 10b). The average increase in

visual ability was 0.5 logit for the BRC and 0.35 logit for the VICTORS. Improvements of these magnitudes would correspond to functional improvements that accompany four-line and three-line improvements in visual acuity, respectively.¹⁶ BRC patients appeared to gain more than VICTORS patients in person visual ability and some item scores. Patients' visual impairments after rehabilitation were the same as they were before. Therefore, improvements in visual ability would have to be attributed to education, counseling, training, and general improvement in visual skills. However, we must also consider the possibility that the observed improvement in visual ability could be an error. That is, the observed increase in visual ability might be the result of a halo effect. Postrehabilitation ratings were obtained from the patients immediately before discharge, before they had any experience with daily activities outside the supporting environment of the institution. It is possible that, in the use of the rating scale, patients consistently overestimated their functional abilities at the conclusion of the rehabilitation program.²⁷ This possibility should be taken into consideration in the design of future clinical studies that use self-assessment rating scales.

One explanation for differences between the BRC and VICTORS in item scores is eligibility for prosthetic equipment. At the time this study was conducted, VA policy restricted eligibility for closed circuit television systems (CCTVs) to legally blind veterans. Another explanation is that the legally blind veterans had greater opportunity for improvement after rehabilitation, because they entered the program reporting more difficulty performing activities. Thus, these persons had more room for improvement than those initially reporting less difficulty.

The seven items that were responsive to the effects of low-vision rehabilitation require patients to rate the difficulty of performing daily activities. An earlier study of item responses on the 25-item NEI VFQ demonstrated that items in part 2, which requires the patient to respond with difficulty ratings of very specific activities, work together to produce a valid interval functional ability scale for patients with low vision.¹⁶ The more general items that require frequency or level-of-agreement ratings also can be mapped onto the functional ability scale, but with far less reliability than the items in part 2. In the present study, these more general items did not change in difficulty after rehabilitation, but did contribute to estimates of patient functional ability, which increased after rehabilitation.

Although the present study demonstrates that the NEI VFQ-25 can be used to measure outcomes of low-vision rehabilitation, it is questionable whether the items in the instrument adequately represent the rehabilitation needs of patients with low vision over the range of vision loss in this population.⁹ For example, only three items are relevant to mobility: going down steps, stairs, or curbs in dim light or at night (item 9); noticing objects off to the side while you are walking along (item 10); and taking part in active sports or other outdoor activities that you enjoy (like golf, bowling, jogging, or walking; item A7). Unlike the seven activities that most likely were responsive to magnification, none of these mobility-related activities changed in difficulty after rehabilitation for either BRC or VICTORS patients. This result is surprising, because the BRC program includes intensive mobility training. It is possible that the wrong questions were asked.

The present study was observational, designed to test the sensitivity of the NEI VFQ-25 to change, it is not a controlled clinical trial. Therefore, we do not know why there was a significant change in item difficulty and person ability after rehabilitation. However, it is reasonable to hypothesize that the decrease in difficulty of targeted items reflects the benefit of low-vision devices, which are designed to make tasks easier.

Refractive error correction, eccentric viewing training, and counseling may be responsible for the improvement in visual ability, or the improvement may be a halo effect (i.e., overestimates of visual ability because the patients had not yet had experience performing activities outside the rehabilitation center).²⁷ Before conclusions can be drawn, however, a control group is needed, to compare person ability and item difficulty of those undergoing rehabilitation with that of those who are not receiving the same treatment.

Acknowledgments

The authors thank Melinda Dunlap, Charlotte Joslin, OD, and Beth Ann Swetland, OD, for assistance with data collection.

References

- Maino JH. Low vision and blindness rehabilitation in the VA: inpatient rehabilitation. In: Massof RW, Lidoff L, eds. *Issues in Low Vision Rehabilitation Service Delivery, Policy, and Funding*. New York: AFB Press; 2001:187-202.
- Massof RW, Dagnelie G, Deremeik JT, DeRose JL, Suleiman SA, Glasner NM. Low vision rehabilitation in the U. S. health care system. *J Vis Rehabil*. 1995;9:3-31, reprinted in: Massof RW, Lidoff L, eds. *Issues in Low Vision Rehabilitation Service Delivery, Policy, and Funding*. New York: AFB Press; 2001:267-299.
- Raasch TW, Leat SJ, Kleinstein RN, Bullimore MA, Cutter GR. Outcomes of low-vision services. *J Am Optom Assoc*. 1997;68:287-293.
- Watson G, DeL'Aune W, Stelmack J, Maino J, Long S. National survey of the impact of low vision device use among veterans. *Optom Vis Sci*. 1997;74:249-259.
- Watson GR, De l'Aune W, Long S, Maino, Stelmack J. Veterans' use of low vision devices for reading. *Optom Vis Sci*. 1997;74:260-265.
- DeL'Aune W, Welsh R, Williams M. A national outcomes assessment of the rehabilitation of adults with visual impairments. *J Vis Impair Blind*. 2000;94:281-291.
- Stelmack J. Quality of life of patients with low vision and outcomes of low-vision rehabilitation. *Optom Vis Sci*. 2001;78:335-342.
- Scott IU, Smiddy WE, Schiffman J, Feuer WJ, Pappas CJ. Quality of life of patients with low vision and the impact of low-vision services. *Am J Ophthalmol*. 1999;128:54-62.
- Stelmack J, Szlyk J, Joslin C, Swetland B, Myers L. Pilot study: use of the NEI VFQ-25 to measure outcomes of low vision rehabilitation services in the Department of Veterans Affairs. In: Stuen C, Arditi A, Horowitz A, et al., eds. *Vision Rehabilitation: Assessment, Intervention and Outcomes*. Lisse, The Netherlands: Swets & Zeitlinger; 2000:774-776.
- Aaronson NK. Quality of life: what is it? How should it be measured? *Oncology (Huntingt)*. 1988;2:69-76.
- Mangione CM, Berry S, Spritzer K, et al. Identifying the content area for the 51-item National Eye Institute Visual Function Questionnaire: results from focus groups with visually impaired persons. *Arch Ophthalmol*. 1998;116:227-233.
- Mangione CM, Lee PP, Pitts J, Gutierrez P, Berry S, Hays RD. Psychometric properties of the National Eye Institute Visual Function Questionnaire. (NEI-VFQ): NEI-VFQ Field Test Investigators. *Arch Ophthalmol*. 1998;116:1496-1504.
- Kupfer C. Outcomes research: translating efficacy into effectiveness. *Optom Vis Sci*. 1998;75:235-236.
- Mangione CM, Lee PP, Gutierrez PR, Spritzer K, Berry S, Hays RD. Development of the 25-item National Eye Institute Visual Function Questionnaire. *Arch Ophthalmol*. 2001;119:1050-1059.
- NEI VFQ-25, Interviewer Administered Format. Rand Corporation, Santa Monica, CA, 1996.
- Massof RW, Fletcher DC. Evaluation of the NEI visual functioning questionnaire as an interval measure of visual ability in low vision. *Vis Res*. 2001;41:397-413.
- Massof RW, Babcock JL, Head DN, Stelmack J. Sensitivity of the NEI VFQ-25 to the effects of low vision rehabilitation depends on patient instructions [ARVO Abstract]. *Invest Ophthalmol Vis Sci*. 2001;42(4):S859. Abstract nr 4613.
- Massof RW. A systems model for low vision rehabilitation. I: basic concepts. *Optom Vis Sci*. 1995;72:725-736.
- Massof RW. A systems model for low vision rehabilitation. II: measurement of vision disabilities. *Optom Vis Sci*. 1998;76:349-373.
- Department of Veterans Affairs, Veterans Health Administration, Blind Rehabilitation Service. Coordinated Services for Blinded Veterans IB 11-59 (revised) P87250. Washington DC: Department of Veterans Affairs; 1996:4.
- Dunn CM, Chadwick G. Code of Federal Regulations. In: Allen W, ed. *Protecting Study Volunteers in Research*. Boston: MA: 1999: 203.
- Wright BD, Masters GN. *Rating Scale Analysis*. Chicago, IL: MESA Press; 1982.
- Masters GN. A Rasch model for partial credit scoring. *Psychometrika*. 1982;47:149-174.
- Andrich D. A rating formulation for ordered response categories. *Psychometrika*. 1978;43:561-573.
- Rasch G. Probabilistic models for some intelligence and attainment tests. Copenhagen: Danmarks Paedagogiske Institut; 1960.
- Linacre JM, Wright B. A user's guide to Winsteps Rasch-model computer program. Chicago, IL: MESA Press; 2001.
- Wolfe G, Fakhourny N, Wu S, Stelmack J, Massof R. A follow-up of changes in HRQL three years after low vision rehabilitation (Abstract). *Optom Vis Sci*. 2001;78(suppl):297.