Guttman Scale Analysis of the Distance Vision Scale

Vijaya K. Gothwal,1,2 Thomas A. Wright,1 Ecosse L. Lamoureux,3,4,5 and Konrad Pesudovs1

PURPOSE. The Distance Vision Scale (DVS) is a self-assessment of visual acuity (VA). Like VA testing in which letter reading becomes progressively difficult through the test, DVS questions have a hierarchy of difficulty (Guttman scale). The aims were to determine whether the DVS fulfills Guttman scaling criteria and to test the relationship between DVS score and VA in a cataract population.

METHODS. Responses of 113 participants to the DVS were subjected to Guttman analysis. Standard criteria of scalability were evaluated that included the coefficient of reproducibility (CR), minimal marginal reproducibility (MMR), and coefficient of scalability (CS). The relationship between total item score and binocular visual acuity was determined.

RESULTS. Five participants were excluded because of missing data. Regularity in the banding pattern of the scalogram of the 108 participants was suggestive of a deterministic Guttman scale. Analyses showed that DVS satisfies the criteria for classification as a valid unidimensional and cumulative scale, as CS (0.93), CR (0.99), and MMR (0.85) values fell within the desired range. The statistically significant correlation between the total item score and binocular VA was 0.24.

CONCLUSIONS. The DVS fits the Guttman scale, supporting the deterministic model underlying the scale. It correlates poorly with VA, suggesting it taps aspects of visual performance and other issues beyond high-contrast VA. The DVS could be used as an outcome measure to evaluate change over time and could be used to set achievable treatment objectives because of its hierarchical properties. (Invest Ophthalmol Vis Sci. 2009;50: 4496–4501) DOI:10.1167/iovs.08-3330

It has been hypothesized that as cataract forms in the eye, vision becomes blurred and performance of a range of activities deteriorates. Although clinically apparent, this hypothesis has not been empirically demonstrated. In particular, little is known about the order in which visual activities are affected. If progression occurs in an ordered, predictable, stepwise manner (i.e., hierarchically), then successive tasks will be more difficult; we would expect distance vision activities that are relatively more difficult, such as recognizing people across the street, to be affected initially and to be followed by easier activities, including those at closer distances such as watching TV.

The application of patient-reported outcomes (or questionnaires) to assess visual disability is increasing in popularity. To date most visual disability researchers have applied questionnaires with “sumiated” scales or Likert scales. Likert-scaled questionnaires conveniently provide a total score that is obtained by summing responses to individual items in the questionnaire. This total score is assumed to reflect the individual’s level of a given trait (i.e., visual disability). However, Likert scales are limited by the assumption that all items in a questionnaire have an equal level of difficulty. Similarly limited is the assumption that response categories (assigned ordinal values 1, 2, 3...) are equally spaced along an interval level measure. These assumptions are invariably invalid and are problematic because summing of ordinal scores may assign inappropriate abilities to people.

An alternative to Likert scaling is Guttman or “cumulative” scaling. In a Guttman scale, responses to items are contingent on the amount of the underlying construct (i.e., visual disability) an individual has. Different items in a questionnaire are not assumed to be of the same inherent difficulty; a subject’s response depends on his or her ability and on the difficulty of the item, the latter of which is referred to as the conjoint structure. Establishing a hierarchy with a Guttman scale helps to legitimize the use of a summed score because the rank ordering of scale items is confirmed. In practice, the Guttman scale is used far less frequently than Likert scales, partly because Guttman scales are comparatively more difficult to construct than Likert scales and have reduced reliability and validity because of their brevity.

Consequently, most researchers prefer to construct Likert scales and use Likert scoring or summary scoring. Despite these limitations, Guttman scales are advantageous because a single response can be used to predict responses to all items on the scale; therefore, the Guttman scale is deterministic. Guttman scales are characterized by the “implicational” or “scalable” nature of their items. That is, tasks that can be successfully completed only when component subtasks or preconditions are completed in a certain order are considered implicational or scalable in nature. The final score obtained from Guttman scaling is equivalent to the highest item the participant has agreed with or answered correctly. From this final score, one can surmise all other items that the participant has agreed with or answered correctly. Under these conditions the scale is said to be fully scalable or implicational. The Guttman scale is not statistical because it leaves no room for error estimation. To function in this way, the items in a Guttman scale must be ordered from the most to the least difficult; this is usually determined by sorting the items in descending order according to the proportion of people passing or failing the items. According to this approach, the most able person will respond positively to difficult and easy items whereas the least able person will respond positively to easy items only. The Guttman scale is most commonly used when there is a need to develop short questionnaires with good discriminant...
The DVS consists of five questions ordered by increasing difficulty (Table 1). For each question, the participants are asked to respond positively or negatively. For the first 4 items a positive response is scored as a 0 (yes) and a negative response as 1 (no). On the last item, the scoring system is reversed. The response options of “some” or “none” are also different for this item. This should conform to a Guttman scale because the respondent can only theoretically affirm the next question after having affirmed the previous question. This is analogous to a visual acuity chart whereby the reader should only be able to read the next line after having read the previous line.

Thus responses to each item in a Guttman scale should be capable of being divided into two (“yes” or “no”) categories, which then differentiates whether a participant “passed” or “failed” an item. An item hierarchy is obtained by sorting items from most to least difficult based on the proportion of participants who passed or failed a question.

**Participants**

Participants were patients awaiting cataract extraction at the Flinders Eye Centre, Adelaide, South Australia. Patients were 18 years of age or older, spoke English, and had no severe cognitive impairment. Patients were mailed the DVS for self-completion while they were on the cataract surgery waiting list. One hundred thirteen patients completed the DVS. Mean age of the patients was 74.9 years (SD, 9.1). Patients had coexisting ocular and systemic comorbidities that appeared to be representative of the elderly cataract population in Australia. Participant characteristics are detailed Table 2.

**Clinical Assessment**

Routine clinical assessments were performed before cataract extraction. Visual acuity assessments were performed using computerized testing based on LogMAR principles with screen illumination of 150 cd/m². All assessments were performed binocularly because binocular acuity was considered representative of real-world ability. Therefore, binocular visual acuity was used in all analyses.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y), mean ± SD</td>
<td>74.9 ± 9.1</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td>Male 48 (42.5), Female 65 (57.5)</td>
</tr>
<tr>
<td>Binocular visual acuity</td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>LogMAR 0.26 ± 0.21, Snellen 6/12 (H11002)</td>
</tr>
<tr>
<td>Range</td>
<td>LogMAR −0.26 to 0.92, Snellen 6/3 to 6/48 (H11002)</td>
</tr>
<tr>
<td>Awaiting second-eye surgery, n (%)</td>
<td>49 (45.4)</td>
</tr>
<tr>
<td>Ocular comorbidity, n (%)</td>
<td>Present 59 (52.2), Absent 54 (47.8)</td>
</tr>
<tr>
<td>Median</td>
<td>1</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>3</td>
</tr>
<tr>
<td>Systemic comorbidity,† n (%)</td>
<td>Present 102 (90.3), Absent 11 (9.7)</td>
</tr>
</tbody>
</table>

n = 113 participants.
* Such as glaucoma, diabetic retinopathy, age-related macular degeneration.
† Such as diabetes, hypertension, angina.
Statistical Analysis

Data were analyzed according to the following steps: (1) ordering of items defined by the total number of endorsements per item, to form a rank ordering of item difficulty; (2) prediction of endorsement order per subject from their total number of endorsements; (3) calculation of total number of errors from mismatch of actual order and predicted order of endorsements; and (4) calculation of statistical values with standard formulas.

Guttman scales are unidimensional, which implies that component items measure a single underlying dimension. The unidimensionality of a set of questions is assessed by the extent to which a score of 1 to any question is associated with a score of 1 on all other items ranked as less difficult. Second, Guttman scales are cumulative. The cumulative nature of the DVS can be examined by ordering items according to the total score and ordering participants such that a participant who responds positively to a difficult question will respond affirmatively to another less difficult question.

Three criteria determine whether DVS items conform to a Guttman scale.25–27 First, we must assess how often responses fit the ideal pattern. This is indicated by the coefficient of reproducibility (CR), which varies from 0 to 1. This value is calculated as 1 minus the result of the total number of errors divided by the total number of responses. A CR value of more than 0.90 is considered acceptable and suggests that it is a valid cumulative and unidimensional Guttman scale. However, if many participants pass or fail all items of the DVS (i.e., “extreme” items), the CR will be spuriously high. This situation can also occur if all participants pass an item or if an item is too difficult for all participants.

The second criterion is that the minimal marginal reproducibility (MMR) value, which is represented by the average overall frequency of each response, be close to the CR value. MMR corrects for the chance appearance of a hierarchy.28 The difference between CR and MMR represents the percentage improvement (PI), which is an indication of the extent to which CR reflects the response patterns rather than the inherent cumulative interrelation of the variables used. The third and the most important criterion is the coefficient of scalability (CS). This value indicates the proportion of responses that can be correctly predicted from the total summed score, thereby allowing for the relative frequencies with which different items are passed. Essentially, the CS tests the degree to which data fit the model. The CS is obtained by dividing PI by the difference between 1 and MMR. The CS varies between 0 and 1. A CS value of >0.60 is accepted to confirm the validity of the Guttman scale.

The high CR (0.99) indicates that both patterns of items are well enough to recognize a friend if you get close to his face.

The second criterion is that the minimal marginal reproducibility (MMR) value, which is represented by the average overall frequency of each response, be close to the CR value. MMR corrects for the chance appearance of a hierarchy.28 The difference between CR and MMR represents the percentage improvement (PI), which is an indication of the extent to which CR reflects the response patterns rather than the inherent cumulative interrelation of the variables used. The third and the most important criterion is the coefficient of scalability (CS). This value indicates the proportion of responses that can be correctly predicted from the total summed score, thereby allowing for the relative frequencies with which different items are passed. Essentially, the CS tests the degree to which data fit the model. The CS is obtained by dividing PI by the difference between 1 and MMR. The CS varies between 0 and 1. A CS value of >0.60 is accepted to confirm the validity of the Guttman scale.

To two items, they included only items 1 and 2. As evident in Table 5, the DVS satisfies the criteria for classification as a valid Guttman scale because CS, CR, and MMR values fall within the desired range.

Participants with poorer visual acuity reported difficulty with a greater number of items (Fig. 2). There was a monotonic, though not linear, relationship between visual acuity and DVS total score ($r = 0.24; P = 0.002$).

RESULTS

One hundred thirteen participants self-administered the DVS. Five patients were excluded from analyses because of incomplete data. Hence, complete data from 108 participants were available for analysis.

The responses given to the DVS items are presented in Table 4. Regularity in the banding pattern of the Guttman scalogram (Fig. 1) suggests that the responses on the DVS do follow a determinist Guttman scale (i.e., they are essentially ordered from the most difficult to the least difficult). In Figure 1, green represents a positive response, red represents a negative response, and yellow is applied to responses that are inconsistent with responses to preceding questions. Two participants had inconsistent responses. For all participants who responded affirmatively to only one item, this included item 1 and no others. For all participants who responded affirmatively to two items, this included only items 1 and 2. As evident in Table 5, the DVS satisfies the criteria for classification as a valid Guttman scale because CS, CR, and MMR values fall within the desired range.

Participants with poorer visual acuity reported difficulty with a greater number of items (Fig. 2). There was a monotonic, though not linear, relationship between visual acuity and DVS total score ($r = 0.24; P = 0.002$).

DISCUSSION

The Guttman scale described in the present study provides a model for analysis of assessment of patient-reported distance visual acuity through the DVS. The items in the DVS meet all the requirements of a Guttman scale, and logically all the items relate to blurred vision, typical of vision loss in patients with cataract.

The high CR (0.99) indicates that both patterns of items are cumulative and that the DVS is reliable. It further suggests that a subject’s scores can be legitimately summed.31 A hierarchy of scale item difficulty can be established from the rank order; therefore, it is appropriate to use a cumulative total score. Because the CR exceeds 0.9, we can predict a subject’s response to an “easier” or more frequently passed item when a subject’s most “difficult” item is also passed. For example, if a participant answered affirmatively to “Can you see well enough to recognize a friend who is an arm’s length away?,” we can also predict responses to other questions such as “Can you see well enough to recognize a friend if you get close to his face?.” The coefficient of scalability is further evidence of an ideal Guttman scale. The CS is 0.93, which is significantly above 0.6, the generally accepted minimum level of scalability.

### Table 4. Responses Given to DVS

<table>
<thead>
<tr>
<th>Question No.</th>
<th>Question Description</th>
<th>Response (n = 108)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>See well enough to recognize a friend if you get close to his face</td>
<td>107 1</td>
</tr>
<tr>
<td>2</td>
<td>See well enough to recognize a friend who is an arm’s length away</td>
<td>107 1</td>
</tr>
<tr>
<td>3</td>
<td>See well enough to recognize a friend across a room</td>
<td>91 17</td>
</tr>
<tr>
<td>4</td>
<td>See well enough to recognize a friend across a street</td>
<td>63 45</td>
</tr>
<tr>
<td>5*</td>
<td>Problems seeing distant objects</td>
<td>94 14</td>
</tr>
</tbody>
</table>

* For Question 5, the response options are Some or None, respectively.

### Table 3. Ideal Response Pattern for a Perfect Guttman Scale on the DVS

<table>
<thead>
<tr>
<th>Scale Score</th>
<th>Q5</th>
<th>Q4</th>
<th>Q3</th>
<th>Q2</th>
<th>Q1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Q. question.
Our findings of acceptable Guttman fit for the DVS are in contrast to findings reported for the ADVS, NEI-VFQ, VAQ, and VF-14 visual function questionnaires.\textsuperscript{22} None of these questionnaires functioned as deterministic measurement instruments.\textsuperscript{22} These visual disability instruments did show a hierarchy of items within a Rasch model but not a Guttman model. However, the DVS contains items related to distance vision only and therefore is not a true “visual disability instrument.” The wording of the DVS represents a surrogate measure of distance visual acuity. The deterministic Guttman principle underlying the DVS is comparable to the logarithmic progression of the rows on a distance visual acuity chart.\textsuperscript{34}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Guttman scalogram for the five items on the Distance Vision Scale. Participants (\textit{rows}) are ordered by their total score across all items, and items (\textit{columns}) are ordered by their total score across all 108 participants. Response categories are color coded (\textit{green}, affirmed; \textit{red}, rejected), with the lowest score (person with best functional vision or easiest item to endorse) in \textit{green} and the highest score (person with worst functional vision or hardest item to endorse) in \textit{red}. \textit{Yellow} represents an inconsistent response (i.e., error).}
\end{figure}

\begin{table}[h]
\centering
\caption{Evaluation of Guttman Properties of the DVS}
\begin{tabular}{|c|c|}
\hline
\textbf{Evaluation Criteria} & \textbf{Value} \\
\hline
Coefficient of reproducibility & 0.99 \\
Minimal marginal reproducibility & 0.85 \\
Percent improvement & 0.14 \\
Coefficient of scalability & 0.93 \\
\hline
\end{tabular}
\end{table}
Furthermore, because the items can be ordered by difficulty, it is possible for clinicians to identify visual acuity when the response to the first and easiest item (“Can you see well enough to recognize a friend if you get close to their face?”) is negative. The clinician can use this response as a basis for checking the visual acuity data. This might help avoid recording errors not uncommon in busy clinical settings. After observing this response pattern, clinicians might use this as an opportunity to begin a conversation with the patient to confirm difficulties with distance vision activities. In this way, the DVS might enhance the productiveness of patient-clinician interactions and increase the probability of cataract surgery referrals at the appropriate time.

In conclusion, the DVS can be used as a patient-reported measure to assist the early identification of visual loss in those with ocular conditions such as cataract. The DVS is cost-effective and convenient and may allow for optimal use of health resources.

References


