Modifying Health Outcome Measures for People With Aphasia

Frances M. Tucker, Dorothy F. Edwards, Leslie Kirchner Mathews, Carolyn M. Baum, Lisa Tabor Connor

KEY WORDS
- aphasia
- communication
- comprehension
- outcome assessment (health care)
- self report

OBJECTIVE. The goal for this study was to determine methods to modify outcome measures for people with aphasia and to provide beginning support for the efficacy of these suggested modifications.

METHOD. Twenty-nine community-dwelling people with aphasia participated. Modified outcome measures included the Stroke Impact Scale, the 36-item short form Medical Outcomes Study, Reintegration to Normal Living Scale, and Activity Card Sort. Participants were supported in their responses by systematically applying a hierarchy of support. An Independence Scale score was determined for each measure for each participant.

RESULTS. Data from an examiner-rated Independence Scale, internal consistency of participants’ responses, correlations among subscale scores across measures, and correlations between aphasia severity and reported outcome provide evidence that people with aphasia are able to understand and respond with these supports.

CONCLUSION. We provide key recommendations for making self-report measures accessible for people with aphasia.

Outcomes research has resulted in the development of self-reported quality-of-life measures of the consequence of a disease or condition (e.g., Williams, Weinberger, Harris, Clark, & Biller, 1999). The measures depend on the respondent’s ability to process language both receptively (i.e., listening, reading) and expressively (i.e., speaking, writing, pointing, marking a written choice). Any linguistic deficit present in aphasia potentially interferes with participation in self-report studies (Bose, McHugh, Schollenberger, & Buchanan, 2009; Cruice, Worrall, Hickson, & Murison, 2005; Hilari & Byng, 2001).

Indeed, aphasia is often an exclusion criterion for research using self-report measures (e.g., Edwards, Hahn, Baum, & Dromerick, 2006; Townend, Brady, & McLaughlan, 2007). Determining the extent of exclusion is problematic because of vaguely worded criteria, such as “too confused or dysphasic to undergo these assessments themselves” (Patel et al., 2006, p. 274) or “unable to understand a two-step command” (Lai, Studenski, Duncan, & Perera, 2002, p. 1840). Considering that people with aphasia make up at least one-fourth of the population of people with stroke (National Aphasia Association, n.d.), excluding them limits the scope of stroke outcome studies. Including people with aphasia is important because proxy judgments by family members or close associates can be inaccurate (e.g., Cruice et al., 2005; Dorman, Waddell, Slattery, Dennis, & Sandercok, 1997; Knapp & Hewison, 1999).

A body of literature has demonstrated that people with moderate to severe aphasia are capable of interacting effectively when communication supports are available. Analogous to wheelchair ramps, “communication ramps” (Kagan, 1995, p. 20) can support the participation of a person with aphasia. For example, providing a written statement or question to view or providing written...
choices to select can support auditory comprehension (Garrett & Huth, 2002; Lasker, Hux, Garrett, Moncrief, & Eischeid, 1997). Reading comprehension can be enhanced by changing font style and size as well as letter and line spacing (Brennan, Worral, & McKenna, 2005; Rose, Worral, & McKenna, 2003). The literature on supporting communication has described a variety of techniques to facilitate responses in people with severe aphasia (e.g., Kagan, Black, Duchan, Simmons-Mackie, & Square, 2001).

As a result of these principles, Cruice et al. (2005) modified measures, including the 36-item Short-Form Medical Outcomes Study (SF–36; Ware & Sherbourne, 1992), for administration to people with aphasia using a three-stage cueing procedure: (1) slower repetition, (2) rephrased item, and (3) personalizing items (explaining how the item could relate to the person) until the person understood (Cruice et al., 2005, p. 119). In addition, response format was altered by substituting yes–no questions for declarative sentences and providing a modified 5-point scale of agreement. Others used existing measures as a springboard for creating new measures. Helm-Estabrooks, Haley, and Womack (2007) modified the Activity Card Sort (ACS; Baum & Edwards, 2001) by substituting line drawings for photographs and simplifying sorting procedures. Hilari and Byng (2001) modified the Stroke Specific Quality of Life Scale (Williams et al., 1999) to create the Stroke and Aphasia Quality of Life Scale (Hilari, Byng, Lampung, & Smith, 2003). Their modifications included increasing font size, administering by interview, reducing linguistic complexity (e.g., shortening questions to reduce reading demands, giving examples orally, providing lead-in information to each section of questions), and modifying response format.

A systematic means to adapt a wide variety of measures for people with aphasia has yet to be developed. Although aphasia quality-of-life measures had common features (e.g., all used an interview format and closed-choice response formats), different types and levels of support were also reported. In this study, we provided practitioners with methods to modify outcome measures for people with aphasia, and we show beginning support for the efficacy of these suggested modifications.

Method

Research Design

We used a prospective cross-sectional research design. The protocol was approved by the Human Research Protection Office. All participants provided written consent.

Participant Selection

Inclusion criteria were a diagnosis of aphasia secondary to stroke of ≥3 mo duration, living in the community, and hearing and vision (aided or unaided) sufficient to perform the task as determined by self-report. Potential participants were excluded if they were unable to tolerate a several-hour testing session per self-report or were unable to travel to the testing site by taxicab or car. Participants were recruited from local clinics, from a community aphasia group, and through referrals from physicians and rehabilitation therapists.

Instruments

We included four self-report outcome measures and the Boston Diagnostic Aphasia Examination–3rd Edition, Short Form (BDAE–3; Goodglass, Kaplan, & Barresi, 2001) in this study. We selected these outcome measures because they have been used in many other stroke outcome studies; however, published studies using these measures have typically reported that people with communication difficulties were excluded from participation. Additionally, these measures are included in a larger, 4-hr assessment protocol for stroke outcome research at our institution. The BDAE–3 provided a measure of severity of language impairment to examine the relationship between performance on modified self-report measures and severity of aphasia.

The SF–36 was designed as a health-related quality-of-life instrument. The maximum score is 100. The SF–36 has satisfactory internal consistency, with Cronbach’s αs > .70.

The Stroke Impact Scale (SIS; Duncan et al., 1999) assesses changes in impairments, disabilities, and handicaps after a stroke. The maximum score is 100. The eight domains of the SIS have high reliability; Cronbach’s αs range from .83 to .90. Interclass correlation coefficients (ICCs) ranged from .70 to .92, except for the Emotion domain (ICC = .57). Moreover, validity has been established and found to be good by correlating SIS domain scores with other measures of that function.

The Reintegration to Normal Living Index (RNL; Wood-Dauphiné, Opzoomer, Williams, Marchand, & Spitzer, 1988) assesses satisfaction with activities of daily living (ADLs), home and community mobility, social roles and responsibilities, and productive pursuits as indicators of quality of life. The maximum total score is 55. The RNL has high content validity and internal consistency, with Cronbach’s αs of ≥.90. Construct validity was demonstrated by a correlation of .68 with the Quality of Life Index.

The ACS assesses participation in 80 instrumental, social, and high- and low-physical-demand leisure
activities. The maximum score is 100% retained. The ACS has high internal consistency. Everard, Lach, Fisher, and Baum (2000) reported \( \alpha \) coefficients \( \geq .71 \) for the four domains measured, and Carpenter et al. (2007) reported \( \alpha \) coefficients \( \geq .83 \). Test–retest reliability has been reported as high, with intraclass correlations ranging from .71 (Carpenter et al., 2007) to .98 (Chan, Chung, & Packer, 2006). ACS scores have content, construct, and predictive validity (Baum & Edwards, 2008).

We used the BDAE–3 Short Form (Goodglass et al., 2001) to determine aphasia severity and obtained three measures of performance: (1) The Expressive component indicates the participant’s ability to produce language, (2) the Auditory Comprehension component indicates the participant’s ability to comprehend language, and (3) the Language Competency Index (LCI) provides a combined score indicating aphasia severity. In addition, all participants were given sentence and paragraph reading comprehension subtests. All scales on the BDAE–3 range in value from 0 to 100; higher scores indicate better language performance. The BDAE–3 has high internal consistency (average \( \alpha = .81 \)). We did not modify BDAE–3 subtests; they were used to measure aphasia severity and were thus administered as originally developed.

Procedure for Modifying the Outcome Measures

We followed a multistep procedure for modifying outcome measures: reviewing current literature for modifications reported to be successful; trying out preliminary modifications, including a hierarchy of support with 5 participants; and administering the entire protocol to 24 additional participants.

To determine whether modifications were successful, we examined correlations among scales with similar constructs and correlations of reported outcome with aphasia severity measures. Consistency in reporting across measures of similar constructs (e.g., the SF–36 Motor and the SIS Physical Function scores) would indicate that participants understood the questions and were able to respond. Likewise, we expected that more severe aphasia would result in lower perceptions of stroke outcome; therefore, we anticipated high correlations between scores on aphasia severity measures and outcome scales.

Step 1: Reviewing the Literature. On the basis of the literature, we considered two main areas for modification: (1) adapting the measures’ vocabulary and syntax, test presentation formats, and response presentation formats and (2) developing a protocol for examiner support for participants. We retained the original vocabulary and syntax in the initial presentation of all test items. We made modifications to presentation format and provided a method for testers to incrementally support the person with aphasia’s response to questions.

Test administration modifications. To enhance visual presentation for the questionnaires, each question was placed on a single horizontal sheet of paper. Modifications included increasing white space on the page, simplifying the font and increasing the font size, and increasing character spacing (e.g., Brennan et al., 2005; Rose et al., 2003). Questions and response choices were printed in Arial 24-point font. Spacing between characters was 3 points. The question was placed at the top of the page, and response choices were placed along the bottom. The middle of the page was print free.

Response format modifications. For the SIS and RNL, response formats included both a written choice and its representative number (e.g., \( 4 = \text{quite a bit of strength} \)). Written choices were paired with numbered choices across the bottom of the page from left to right. For the SIS and RNL, presentation order of numbers and written choices were reversed from 5–1 to 1–5. This modification was made to ensure that all scales for all measures were consistent, with 1 representing the lowest score and 5 representing the highest score. In the case of the SF–36, numbers were assigned to questions on the basis of the number of options. After the question and response choices were read aloud by the examiner, the participant could respond by pointing to written or number choices.

Standard administration of the ACS requires participants to sort photographs of 80 activities into piles using five printed response options (never done, still doing, given up due to stroke, do less due to stroke, and started after stroke). In the standard administration, sorts into the five response piles are done simultaneously. In this study, we increased the number of sorts from one to three separate sorts of no more than three choices to minimize confusion among response options. At each step, we removed photographs eliminated in the sort. For instance, in the first sort the participant placed cards into two piles, activities that were never done and those that had been done in the past or were currently done. The never-done pile was then removed. In the second sort, participants divided remaining cards into three piles: still doing, given up, or do less. Finally, participants were asked whether any activities had been started after stroke.

Examiner support. For the SIS, RNL, and SF–36, we made administration procedures for each item more explicit than the original protocols. Printed questions and responses were read aloud, and the participant selected a response. For example, the examiner read aloud SIS Question 1a, “In the past week, how would you rate the
strength of your arm that was most affected by your stroke?” (Duncan et al., 1999). The examiner verified each response (Kagan et al., 2001) after the participant provided it by pointing to the number and stating the meaning corresponding to the written choice (e.g., “You rate the strength in your arm most affected by the stroke as ‘4,’ quite a bit of strength”).

If a participant did not respond or the response was ambiguous, we used a 5-step hierarchy of examiner support:

1. Repeat the question and choices.
2. Simplify and restate the question (e.g., “How is the strength in your right arm?” [the examiner provides the name of affected limb]), and review choice scale.
3. Re-explain the entire choice scale (e.g., “This [point to no strength at all] would mean you have no strength in your right arm. This [point to a lot of strength] would mean you have a lot of strength in your right arm.” Repeat the restated question.
4. Combine a yes–no question with scale (e.g., “Do you have a lot of strength in your right arm? Would you say that you have no strength at all in your right arm?”)
5. Present the next question.

By the fourth level of support, the participant had received maximal cueing involving successive exposure to the question. Thus, with no response, the examiner moved to the next question. All verified responses, regardless of support level, were accepted.

**Step 2: Testing Modifications on a Preliminary Group of Participants.** To test modifications that we had gleaned from the literature and to determine whether additional changes in wording or sentence structure needed to be made, we administered the measures with original wording to the initial 5 participants. The examiner took note of their behaviors if participants failed to respond, such as requests to repeat questions or need for clarification of response choices. At that point, the examiner supported the participant’s request for clarification (e.g., repeating or restating the question). On the basis of the initial 5 participants’ feedback, we verified the need for simultaneous auditory and visual presentation of questions and responses and further simplified the response choice format. Participants indicated that selecting a single choice among five or more written choices and their corresponding numbers was confusing. They preferred a format that presented numbers with a written anchor only at each end of the choice continuum.

**Step 3: Administration of the Battery to the Other Participants.** On the basis of the initial 5 participants’ feedback, we confirmed that modifying wording of items was unnecessary to elicit a response. Adaptations were made, therefore, only to test administration and response presentation formats as we have outlined. The support hierarchy used informally for the first 5 participants was used for the remaining 24 participants. All participants were included in the analyses reported, because the test methods for all were similar.

**Independence Scale**

To indicate how much examiner support was required to complete each measure, we created a 7-point Independence Scale. This 7-point scale was based on independence ratings of the FIM™ (Smith, Illig, Fielder, Hamilton, & Ottenbacher, 1996) in which 1 = does not produce response with maximal support and 7 = responds with no need for additional support. The examiner rated the participant’s overall independence on each measure. In addition, an occupational therapy student naïve to aphasia severity was trained to make criteria-based judgments of independence to assess interrater reliability.

**Data Collection**

Participants traveled by car or taxicab to our laboratory. They completed a 2- to 3-hr test session using adapted materials, with breaks given as requested. Two participants completed testing over two sessions on consecutive days. A speech–language pathologist administered all measures. Audio recordings were made of the first 2 participants’ testing sessions; other sessions were videotaped.

**Data Analysis**

We analyzed the data with SPSS Version 15.0 (SPSS, Inc., Chicago). Internal consistency, a measure of reliability, was evaluated with Cronbach’s α coefficient (Portney & Watkins, 2009). The included measures all have data supporting their construct validity in people without aphasia. Our goal in this investigation was to provide converging evidence that the adapted instruments similarly measure the constructs measured by the original instruments. This goal was accomplished by providing (1) Pearson correlation coefficients among scales evaluating the same constructs, (2) Pearson correlation coefficients among measures of aphasia severity and self-reported stroke outcome, and (3) Spearman rank order correlations among independence ratings and measures of aphasia severity.

**Results**

The characteristics of the 29 study participants are presented in Table 1, including average independence ratings. Mean age was 60.8 yr (standard deviation [SD] = 12.5, range = 41–78). Participants were well educated...
and had a mean of 14.8 yr of education (SD = 2.5, range = 12–20). Average interval from stroke onset to testing for this study was 4 yr (SD = 3), although the range was 0.5–15 yr.

The mean Expressive component score from the BDAE–3 was 49.14 (SD = 27.74; range = 5–100). The mean Auditory Comprehension component score was 65.97 (SD = 28.51, range = 13–100). The mean LCI score was 57.72 (SD = 26.87, range = 18–100). The mean score on sentence and paragraph reading comprehension was 61.03 (SD = 36.58, range = 0–100). Thus, this group represented a wide range of language ability from severe to mild aphasia, with average auditory comprehension being greater than expression.

**Independence Ratings**

Interrater reliability was .90 for the independence rating of the RNL, .88 for the SIS and SF–36, and .75 for the ACS. Mean independence scores (5 for the SIS and SF–36 and 6 for the ACS and RNL) indicated that participants required minimal to moderate assistance to respond, although we found the full range of assistance levels in our sample. Only 2 respondents had a >1-point difference in independence rating across the four measures; these 2 respondents had a 2-point difference. Independence ratings for each measure were highly correlated with aphasia severity scores, and Spearman correlations ranged from .57 to .80. Thus, the greater the aphasia severity was, the more support participants needed to respond.

To further examine the degree of support participants required to complete the measures, we also obtained frequency counts for each independence rating score for all measures. For 21% of all the assessments administered, ratings indicated fully independent performance. For 35% of the ratings, participants needed only minimal support. Minimal–moderate (14%), moderate (15%), and moderate–

### Table 1. Participant Characteristics

<table>
<thead>
<tr>
<th>Age</th>
<th>Race</th>
<th>Sex</th>
<th>Education, yr</th>
<th>Time Since Stroke, yr</th>
<th>LCI</th>
<th>EC</th>
<th>AC</th>
<th>Average Independence Ratinga</th>
</tr>
</thead>
<tbody>
<tr>
<td>70b</td>
<td>C</td>
<td>M</td>
<td>16</td>
<td>1.7</td>
<td>56</td>
<td>35</td>
<td>77</td>
<td>Not obtained</td>
</tr>
<tr>
<td>51b</td>
<td>C</td>
<td>F</td>
<td>18</td>
<td>1.7</td>
<td>64</td>
<td>55</td>
<td>73</td>
<td>Not obtained</td>
</tr>
<tr>
<td>75b</td>
<td>C</td>
<td>F</td>
<td>12</td>
<td>6.2</td>
<td>18</td>
<td>15</td>
<td>20</td>
<td>3.00</td>
</tr>
<tr>
<td>57b</td>
<td>C</td>
<td>F</td>
<td>12</td>
<td>5.4</td>
<td>55</td>
<td>50</td>
<td>60</td>
<td>6.25</td>
</tr>
<tr>
<td>43b</td>
<td>C</td>
<td>M</td>
<td>16</td>
<td>8.7</td>
<td>53</td>
<td>45</td>
<td>60</td>
<td>6.25</td>
</tr>
<tr>
<td>77</td>
<td>C</td>
<td>M</td>
<td>18</td>
<td>2.6</td>
<td>53</td>
<td>45</td>
<td>60</td>
<td>4.00</td>
</tr>
<tr>
<td>69</td>
<td>AA</td>
<td>F</td>
<td>16</td>
<td>15.0</td>
<td>36</td>
<td>25</td>
<td>47</td>
<td>4.25</td>
</tr>
<tr>
<td>74</td>
<td>C</td>
<td>M</td>
<td>13</td>
<td>4.3</td>
<td>36</td>
<td>15</td>
<td>57</td>
<td>6.00</td>
</tr>
<tr>
<td>74</td>
<td>C</td>
<td>M</td>
<td>12</td>
<td>2.1</td>
<td>34</td>
<td>30</td>
<td>37</td>
<td>5.75</td>
</tr>
<tr>
<td>78</td>
<td>C</td>
<td>F</td>
<td>12</td>
<td>3.2</td>
<td>41</td>
<td>35</td>
<td>47</td>
<td>4.50</td>
</tr>
<tr>
<td>65</td>
<td>C</td>
<td>F</td>
<td>14</td>
<td>3.1</td>
<td>19</td>
<td>10</td>
<td>27</td>
<td>4.75</td>
</tr>
<tr>
<td>71</td>
<td>C</td>
<td>F</td>
<td>12</td>
<td>5.4</td>
<td>75</td>
<td>50</td>
<td>100</td>
<td>6.00</td>
</tr>
<tr>
<td>63</td>
<td>C</td>
<td>M</td>
<td>20</td>
<td>5.5</td>
<td>93</td>
<td>85</td>
<td>100</td>
<td>7.00</td>
</tr>
<tr>
<td>53</td>
<td>AA</td>
<td>F</td>
<td>18</td>
<td>1.5</td>
<td>83</td>
<td>85</td>
<td>80</td>
<td>6.00</td>
</tr>
<tr>
<td>53</td>
<td>C</td>
<td>M</td>
<td>14</td>
<td>1.3</td>
<td>26</td>
<td>15</td>
<td>37</td>
<td>4.00</td>
</tr>
<tr>
<td>70</td>
<td>C</td>
<td>M</td>
<td>18</td>
<td>0.5</td>
<td>95</td>
<td>90</td>
<td>100</td>
<td>7.00</td>
</tr>
<tr>
<td>53</td>
<td>C</td>
<td>F</td>
<td>16</td>
<td>1.4</td>
<td>80</td>
<td>70</td>
<td>90</td>
<td>7.00</td>
</tr>
<tr>
<td>44</td>
<td>AA</td>
<td>M</td>
<td>12</td>
<td>4.6</td>
<td>83</td>
<td>85</td>
<td>80</td>
<td>6.25</td>
</tr>
<tr>
<td>78</td>
<td>C</td>
<td>M</td>
<td>16</td>
<td>2.2</td>
<td>75</td>
<td>50</td>
<td>100</td>
<td>5.75</td>
</tr>
<tr>
<td>74</td>
<td>C</td>
<td>M</td>
<td>19</td>
<td>1.3</td>
<td>71</td>
<td>65</td>
<td>77</td>
<td>6.00</td>
</tr>
<tr>
<td>61</td>
<td>C</td>
<td>F</td>
<td>12</td>
<td>0.7</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>6.75</td>
</tr>
<tr>
<td>44</td>
<td>C</td>
<td>M</td>
<td>14</td>
<td>5.9</td>
<td>26</td>
<td>35</td>
<td>17</td>
<td>2.25</td>
</tr>
<tr>
<td>54</td>
<td>C</td>
<td>F</td>
<td>14</td>
<td>1.5</td>
<td>80</td>
<td>70</td>
<td>90</td>
<td>6.00</td>
</tr>
<tr>
<td>48</td>
<td>C</td>
<td>F</td>
<td>16</td>
<td>2.9</td>
<td>31</td>
<td>15</td>
<td>47</td>
<td>3.50</td>
</tr>
<tr>
<td>44</td>
<td>C</td>
<td>M</td>
<td>16</td>
<td>9.0</td>
<td>73</td>
<td>65</td>
<td>80</td>
<td>5.50</td>
</tr>
<tr>
<td>69</td>
<td>C</td>
<td>M</td>
<td>12</td>
<td>6.2</td>
<td>21</td>
<td>5</td>
<td>37</td>
<td>2.75</td>
</tr>
<tr>
<td>41</td>
<td>C</td>
<td>F</td>
<td>16</td>
<td>1.4</td>
<td>85</td>
<td>70</td>
<td>100</td>
<td>6.50</td>
</tr>
<tr>
<td>65</td>
<td>C</td>
<td>F</td>
<td>14</td>
<td>2.1</td>
<td>93</td>
<td>85</td>
<td>100</td>
<td>5.75</td>
</tr>
<tr>
<td>44</td>
<td>C</td>
<td>F</td>
<td>12</td>
<td>1.2</td>
<td>19</td>
<td>25</td>
<td>13</td>
<td>4.00</td>
</tr>
</tbody>
</table>

Note. AA = African-American; AC = Auditory Comprehension component of the LCI; BDAE–3 = Boston Diagnostic Aphasia Examination, 3rd ed.; C = Caucasian; EC = Expressive component of the LCI; F = female; M = male; LCI = Language Competency Index of the Boston Diagnostic Aphasia Examination, 3rd ed.

bParticipants who gave feedback regarding adequacy of modifications.

aThe average independence rating is for the four measures administered.
maximum (12%) support categories accounted for 41% of the scores. Only 1% of ratings fell within the maximum support category, and <2% fell within the did-not-respond category. With few exceptions, given sufficient support, participants could clearly respond to the measures.

Reliability of Modified Measures
The second step in evaluating adaptations was to examine internal consistency of the adapted scales. We did not include the ACS because it does not have multiple items that evaluate the same construct. Cronbach's α coefficients for SIS subscales ranged from .86 to .90, with an overall coefficient of .92, indicating high levels of internal consistency. For SF–36 domain scores, α coefficients were somewhat lower, although all were within acceptable levels (ranging from .66 to .90). The α coefficient for the RNL was .84, also indicative of a high level of internal consistency. The α coefficients obtained for modified measures were similar to or higher than those reported in the literature for the standard administration format.

Construct Validity
As further indication that supports were successful, we examined consistency of responses across measures with similar constructs. If participants were responding consistently, we would expect moderate to high correlations among these scales. For instance, the correlation between the SIS Physical Domain and SF–36 Physical Function scales was high (r = .85). We examined several of these relationships; correlations for selected subscales are reported in Table 2. ACS percentage retained for High Demand Leisure correlated with SIS Physical Domain (r = .62) and SF–36 Physical Function scores (r = .58) because high-demand leisure activities require motor function for participation. Domains associated with social participation (i.e., SIS Communication and SF–36 Social Function; ACS Social Domain and SF–36 Social Function; ACS total percentage retained and SIS Participation) showed moderate to high correlations, as did the RNL with SIS Mobility and SIS participation (r’s > .52).

Finally, we expected that aphasia severity would be associated with poorer scores on quality-of-life assessments. Overall language ability (BDAE LCI) was moderately to highly correlated with SIS Communication, SIS Participation, and SIS Recovery (see Table 3). Moreover, language ability was correlated with total percentage of activities retained and percentage retained for Instrumental and Low Demand Leisure activities on the ACS. Contrary to our predictions, we found no significant correlations among language abilities and scales from the SF–36 or RNL. Specifically, we expected the SF–36 Social Function scale and the RNL scores to correlate with language abilities. The pattern of correlation of language abilities with outcome measures was quite similar for both expressive language and comprehension. Although these data did not conform precisely to our predictions, the correlations with the language measures combined with the pattern of correlations among the scales themselves provide preliminary support for the construct validity of the adapted outcome measures.

Discussion
People with aphasia are often excluded from self-report quality-of-life studies under the assumption that language impairment interferes with ability to respond independently to self-report assessments. Recent aphasia research has demonstrated, however, that modifying the communication environment supports and enhances the ability of a person with aphasia to provide relevant information. A variety of supports are described in the literature, but a systematic approach applicable across a variety of measures has yet to be developed. In this study, we sought to incorporate and clarify supports used in previous studies that could be applicable to other stroke outcomes and measures in order to provide practitioners with those methods of designing modifications for people with aphasia.

Our results demonstrated that participants with a range of aphasia severity were able to participate in surveys about their subjective states given sufficient external support. External supports developed for this study included modifying administration format and providing systematic examiner support when needed to facilitate responding. Degree of support needed was captured by independence ratings. We obtained strong correlations between aphasia severity scores and independence ratings, and participants were overwhelmingly successful in completing questions for the four instruments that we examined. Three other types of information bolster our conclusions: Participants demonstrated internal consistency of responses within scales, correlations among scales measuring similar constructs were moderate to high, and correlations among aphasia severity measures and outcome scales were moderate to high.

We recognize several limitations to our study. Data were not collected on participants’ performance without modifications. Independence ratings for adapted measures, however, demonstrate that many participants...
with LCI scores of 100 (the maximum value) needed some support. Other limitations stem from the fact that we recruited a convenience sample. Our findings, therefore, may not generalize to the population of people with aphasia. Specifically, our sample was not representative of our own stroke population in terms of ethnicity or educational attainment. Only 3 of the participants were non-White, all of whom were African-American; this proportion is uncharacteristic of the stroke population in our geographic area. All study participants had completed high school or higher education. Although education levels were unrepresentative of the general stroke population, the high education level of this sample enabled us to rule out prestroke literacy as a confound with aphasia.

An additional limitation was that our participants generally had better auditory comprehension than verbal expression (see Table 1). We recognize that expecting that all people with severe aphasia can participate in outcome testing is unrealistic. Future work will have to establish the limits of comprehension impairment that can be successfully supported with these modifications. Moreover, the broad range of time since stroke (mean = 4 yr, SD = 3, range = 0.5–15 yr) may limit comparison with other stroke outcome data for nonaphasic stroke survivors.

Despite these limitations, several findings of this study make it an important contribution to the field. First, the range of language abilities was wide, with mean LCI falling in the moderate severity range. This wide range of language abilities enabled us to test the levels of support people needed to participate using these outcome measures. Second, most participants were able to complete measures with minimal modifications. Initially, we considered modifying the vocabulary and syntactic structure of the questions to facilitate comprehension. Original questions were, however, left unmodified. If the person demonstrated a need for support, questions were simplified in terms of vocabulary and syntax through an interaction between participant and examiner. Third, given the support structure, participants completed several hours of testing.

### Table 2. Correlations of Related Subtests Between Measures

<table>
<thead>
<tr>
<th>Scales</th>
<th>ACS</th>
<th>SF–36</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Instrumental</td>
<td>LDL</td>
</tr>
<tr>
<td>RNL</td>
<td>ns</td>
<td>.54**</td>
</tr>
<tr>
<td>SIS Recovery</td>
<td>ns</td>
<td>.42*</td>
</tr>
<tr>
<td>SIS Communication</td>
<td>.35</td>
<td>.48*</td>
</tr>
<tr>
<td>SIS Participation</td>
<td>ns</td>
<td>.50*</td>
</tr>
<tr>
<td>SIS Physical Domain</td>
<td>.58**</td>
<td>.51**</td>
</tr>
<tr>
<td>SF–36 Phys Fun</td>
<td>.52**</td>
<td>.47*</td>
</tr>
<tr>
<td>SF–36 Social</td>
<td>ns</td>
<td>ns</td>
</tr>
</tbody>
</table>

*Note. ACS = Activity Card Sort; HDL = High Demand Leisure; LDL = Low Demand Leisure; ns = not significant; Phys Fun = Physical Function; RNL = Reintegration to Normal Living Scale; SF–36 = 36-item Short-Form Medical Outcomes Study; SIS = Stroke Impact Scale.

\*p = .05 level (two-tailed). **p = .01 level (two-tailed).
Implications for Occupational Therapy Practice
On the basis of our findings in this study, we offer the following recommendations for making self-report quality-of-life measures accessible for people with aphasia.

- The interview format should be face-to-face. People with aphasia have difficulty comprehending telephone interviews or completing written evaluations independently.
- Closed-ended questions should be used with response options to which the participant can point (e.g., yes–no, Likert scale).
- Examiners should read aloud each question with the written question available simultaneously. Some people with aphasia will find the auditory modality more supportive, and some will find the reading modality more supportive.
- Written questions should be provided in a simplified format (e.g., one question-and-response set per page, simple and large font, increased character spacing, white space in the middle of the page).
- Response options should be presented in a manner that is clear to the participants (e.g., written anchors corresponding to number choices).
- Examiner support should be hierarchical and based on the participant’s need. Provide the minimal level of support necessary per item to elicit a response; move to the next level of support only if the participant does not respond or the response is ambiguous.

Recommendations for Future Research
Future studies will be needed to address the limitations of this research, particularly those introduced by recruiting a convenience sample. A more representative sample in terms of race, ethnicity, educational attainment, and time since stroke will need to be recruited to determine whether this study’s results generalize to the stroke population at large. In addition, a study designed to recruit people with more severe comprehension impairments is warranted to determine the limits of the simple modifications made in this study. Finally, a study to specifically compare performance in people with aphasia on self-report quality-of-life assessments both with and without modification would further clarify the extent of modifications needed to support participation in these types of measures.

Conclusions
Format modifications enable people with a wide range of aphasia severity to participate in research using quality-of-life and participation assessments. Results of this study should encourage researchers and practitioners to include people with aphasia in studies that assess self-reported quality of life. These modifications and supports allow people with aphasia to be included in self-reported quality-of-life studies, thus enriching the understanding of consequences of stroke. ▲

Acknowledgment
This work was supported by Grant 220020087 from the James S. McDonnell Foundation. We acknowledge the assistance of our student Emilia Epperley and thank the people with aphasia who participated in this study. We also thank two anonymous reviewers for suggestions on an earlier version of this article.

References


