Reliability of standardized assessment for adults who are deafblind

Dawn M. Guthrie, PhD;1* Robyn Pitman, MA;1 Paul Stolee, PhD;2 Graham Strong, OD, MSc;3 Jeff Poss, PhD;2 Erin Y. Tjam, PhD;4 Lindsay Bowman, MPH;1 Melody Ashworth, MSc;1 John P. Hirdes, PhD2
1Department of Kinesiology and Physical Education, Wilfrid Laurier University, Waterloo, Ontario, Canada; 2Department of Health Studies and Gerontology and 3Centre for Sight Enhancement, School of Optometry, University of Waterloo, Waterloo, Ontario, Canada; 4St. Mary’s General Hospital, Kitchener, Ontario, Canada

Abstract—This study assessed the reliability of the interRAI Community Health Assessment (interRAI CHA) and Deafblind Supplement (DbS). The interRAI CHA and DbS represents a multidimensional, standardized assessment instrument for use with adults (18 and older) who are deafblind. The interrater reliability of the instrument was tested through the completion of dual assessments with 44 individuals who were deafblind in the province of Ontario, Canada. Overall, nearly 50% of items had a kappa value of at least 0.60, indicating fair to substantial agreement for these items. Several items related to psychosocial well-being, mood, and sense of involvement had kappa scores of less than 0.40. However, among these items with low kappa values, most (78%) showed at least 70% agreement between the two assessors. The internal consistency of several health subscales, embedded within the assessment, was also very good and ranged from 0.63 to 0.93. The interRAI CHA and DbS represents a reliable instrument for assessing adults with deafblindness to better understand their needs, abilities, and preferences.

Key words: blindness, deafblind disorders, deafness, dual sensory loss, needs assessment, psychometrics, reliability, reproducibility, sensation disorders, standardized assessment.

INTRODUCTION

Deafblindness also affects an individual’s ability to function and carry out daily activities. Individuals with deafblindness are typically classified into two groups: congenital and acquired deafblindness. Individuals with congenital deafblindness experienced vision and hearing loss before age 2 [2]. Congenital deafblindness is often caused by a prenatal virus (e.g., congenital rubella) or chromosomal abnormalities (e.g., CHARGE syndrome; CHARGE stands for coloboma of the eye, heart defects, atresia of the nasal choanae, retardation of growth and/or development, genital and/or urinary abnormalities, and ear abnormalities and deafness). On the other hand, the

Abbreviations: ADL = activity of daily living, ADL SHS = Activities of Daily Living Self-Performance Hierarchy Scale, CAP = clinical assessment protocol, CDBA = Canadian Deafblind Association, CHESS = Changes in Health End-Stage Disease Signs and Symptoms, CNIB = Canadian National Institute for the Blind, DbS = Deafblind Supplement, DbSI = Deafblind Severity Index, DRS = Depression Rating Scale, IADL = instrumental ADL, interRAI CHA = interRAI Community Health Assessment, interRAI HC = interRAI Home Care, O&M = orientation and mobility, RAI = Resident Assessment Instrument.

*Address all correspondence to Dawn M. Guthrie, PhD; Department of Kinesiology and Physical Education, BA 516, Wilfrid Laurier University, 75 University Ave W, Waterloo, ON, N2L 3C 5 Canada; 519-884-0710, ext 2590; fax: 519-747-4594. Email: dguthrie@wlu.ca

DOI:10.1682/JRRD.2010.09.0175
onset of acquired deafblindness occurs later in life but can also be caused by genetically inherited disorders (e.g., Usher Syndrome) [3], postnatal or early childhood infections, and acquired brain injury [2]. In Canada, approximately 55 to 70 percent of individuals who are deafblind become deafblind after age 2 [2].

In the province of Ontario, the provincial government, through the Ministry of Community and Social Services, provides funding for intervenor services for individuals who are deafblind. Intervenors are trained professionals who communicate with and facilitate the person’s interaction with people, places, and things by providing information about the person’s environment. Ongoing concern exists regarding the equity of service provision across the groups of deafblind clients in Ontario. To address this concern, Dalby et al. developed a standardized assessment tool to better understand the needs of this group [4]. The assessment tool was developed to be compatible with existing instruments used in Ontario in the home-care and complex continuing care sectors. For example, in Ontario, the interRAI Home Care (interRAI HC) has been mandated for use with all long-stay (>60 days) home-care clients since 2002. Similarly, the Resident Assessment Instrument (RAI) 2.0 [5] has been mandatory for all complex continuing-care hospitals/units since 1996, and it is currently being implemented on a staged basis in long-term-care facilities. The standardized instrument for individuals with deafblindness is the interRAI Community Health Assessment (interRAI CHA) [6] and a Deafblind Supplement (DbS) to the CHA [7]. The interRAI CHA includes roughly 150 items, all of which are identical to those found in the interRAI HC and many of which are also found in the RAI 2.0. These instruments were designed by interRAI (http://www.interrai.org/), a not-for-profit research network of more than 30 countries that focuses on the development and testing of assessment systems to improve the quality of life and delivery of services for vulnerable populations, including older persons and persons with disabilities.

The interRAI CHA captures basic background information about the person (e.g., date of birth, marital status, living arrangement) and detailed information across 13 domains, including abilities in activities of daily living (ADLs) (e.g., bathing, dressing, grooming), instrumental ADLs (IADLs) (e.g., using the telephone, managing finances), social functioning, mental health, pain, hearing, and vision. The DbS includes approximately 150 items and gathers further data across 11 domains considered important to assess for individuals with deafblindness, including vision and hearing (e.g., age of onset of vision/hearing loss, diagnoses, visual acuity and field diameter, alerting to sounds, devices used), communication (e.g., communication modes used, ability to communicate with family members), mood and behavior, level and type of informal support from friends/family, psychosocial well-being, and orientation and mobility (O&M) (e.g., ability to move about in both familiar and unfamiliar environments). In both the interRAI CHA and the DbS, the items are closed-ended with response options varying from item to item. Typically, response options are dichotomous (i.e., yes/no) or they are provided on an ordinal-type scale ranging from 0 to 5, or sometimes as high as 8. The time frame for assessment is typically the previous 3 days, although a few items ask about the past 90 days.

The instrument was developed for use with adult clients (18 and older) who reside in the community and are typically not receiving formal home-care services. The purpose of the assessment is to help an assessor identify individuals who are functionally deafblind and gain insight into their needs and abilities. The assessment also provides the assessor with detailed guidelines, known as clinical assessment protocols (CAPs). The CAPs are written documents that include suggestions for further assessment and considerations for development of a service plan. Items in the assessment map onto the CAPs through computer-generated algorithms.

The instrument was based on significant input from content experts and service providers in the field; initial psychometric testing provided evidence of the convergent validity and internal consistency of the instrument [4]. In the current project, we conducted further testing of the instrument, including an assessment of its interrater reliability through dual assessments in a group of individuals with deafblindness.

**METHODS**

**Participants**

A total of 44 participants were recruited and two assessments were completed with each individual. Clients were eligible if they were 18 or older and were receiving services from either of the two participating organizations.
Eleven assessors from the Canadian Deafblind Association (CDBA) \((n = 3)\) and the Canadian National Institute for the Blind (CNIB) \((n = 8)\) voluntarily participated in the project. On average, they had 12 years of experience in working with the deafblind (5 of the 11 had at least 20 years) and all had completed the interRAI CHA and DbS as part of a previous project. All assessors attended a 1-day training session. This session focused on the project protocols and a discussion of techniques for assessing and coding items to ensure consistency between assessors.

**Procedures**

The CDBA and CNIB were responsible for the recruitment of study participants. They approached eligible clients and obtained informed consent from the client or a proxy (i.e., legal guardian or caregiver). All clients were assessed using the interRAI CHA and DbS during an in-person oral interview with the client. If the client required an intervenor or an interpreter to complete the assessment, one was provided.

Each participant was assessed twice, by two different assessors, within 7 days of each other. The 7-day time frame was chosen because clients within this population were not expected to show significant clinical or functional changes within 1 week. The assessors did not share information from the assessment and did not discuss the client. Where possible, assessors were randomly assigned to go in as the “first” assessor. This was done to minimize any potential bias caused by having a given assessor always completing the first assessment with the client.

Following data collection, all assessors were invited to participate in a one-on-one interview with the study research assistant. The purpose of the interview was to gain their feedback on the assessment itself (e.g., length of time to complete it, items included in the assessment, areas where they had the most difficulty) and their suggestions for future training sessions. All assessors agreed to an interview and signed a consent form to allow taping and transcription of the interviews.

**Measures**

The interRAI CHA and DbS constitutes a comprehensive assessment of an individual’s health and functional abilities as well as more specific assessment and documentation of domains of importance when determining needs and abilities for persons with dual sensory loss and deafblindness.

Once completed, a series of 10 health subscales can be automatically generated from the assessment. The Depression Rating Scale (DRS) assesses signs and symptoms of depression. A score of 3 or higher is typically used as a cut-off predictive of psychiatrist ratings and clinical thresholds used with the Hamilton and Cornell scales [8]. The Activities of Daily Living ADL Self-Performance Hierarchy Scale (ADL SHS) uses four ADL items and rates individuals from 0 to 6, with higher scores indicating lower levels of independence in ADLs. It has been shown to be a valid and reliable measure [9]. The IADL Involvement Scale indicates the level of independence persons have in performing routine activities (e.g., managing their finances, preparing meals) around the home or in the community. The score ranges from 0 to 9, with higher values indicating lower levels of independence [10]. The Changes in Health End-Stage Disease Signs and Symptoms (CHESS) Scale was developed to detect health instability. Higher scores on the CHESS are associated with reduced survival over time [11–12]. The Pain Scale uses two items to create a score that ranges from 0 (no pain) to 3 (severe daily pain); it has been shown to be highly predictive of pain on the visual analog scale in nursing home residents [13].

Two new subscales have recently been developed. The O&M Scale contains three items based on the person’s capacity for O&M in familiar indoor and outdoor environments and unfamiliar indoor environments. Scale scores range from 0 (independent) to 18 (full dependence on others). The Deafblind Severity Index (DbSI) utilizes two variables on functional vision and hearing to create a severity score ranging from 0 (no impairment in either sense) to 5 (severe impairment in both senses).

**Analysis**

The kappa statistic was used to assess interrater reliability for each interRAI CHA and DbS item. The unweighted kappa was used for dichotomous variables, and the weighted kappa, using Fleiss-Cohen (quadratic) weights [14], was used for ordinal data. The kappa statistic is commonly used to measure agreement between two raters, with values ranging from –1.00 to 1.00. A value of zero reflects no agreement and negative values suggest higher levels of disagreement than would be expected by chance alone. Landis and Koch have proposed the following as standards for evaluating the strength of agreement for the kappa statistic: \(<0.00 = \text{poor agreement}, 0.00 \text{ to } 0.20 = \text{slight agreement}, 0.21 \text{ to } 0.40 = \text{fair agreement}, \)
0.41 to 0.60 = moderate agreement, 0.61 to 0.80 = substantial agreement, and 0.81 to 1.00 = almost perfect agreement [15]. Internal consistency of the health subscales was measured using the first assessment completed for each of the 44 participants. Cronbach alpha (standardized) was used as the measure of internal consistency. An alpha value between 0.70 and 0.90 was considered evidence of good reliability [16]. All analyses were completed using SAS software version 9.1 (SAS Institute; Cary, North Carolina) [14].

RESULTS

Among the 44 participants (aged 18–71 years), 28 (63.6%) had congenital deafblindness and the remainder had acquired deafblindness. The acquired group was significantly older than the congenital group (mean age: 52.1 vs 33.3, \( p < 0.001 \)). In addition, they were significantly less likely to be male (18.8% vs 53.6%, \( p = 0.02 \)) and to have never been married (56.3% vs 96.4, \( p = 0.002 \)). The groups did not significantly differ in their level of education and the severity of their dual sensory loss, as measured by the DbSI (Table 1). The DbSI is based on two items within the interRAI CHA, one related to vision and one to hearing. It can range from 0 (no vision/hearing impairment) to 5 (severe impairment in both senses).

Kappa Results for interRAI Community Health Assessment

Table 2 contains the mean kappa value and ranges for multiple domain areas of the interRAI CHA. Overall,
at least fair agreement (0.21–0.40) [15] existed between
the assessors for the majority of the items under each
domain. However, the items with the lowest level of
agreement were related to psychosocial well-being and
mood. For the psychosocial well-being domain, two
items, “fearful of a family member” (κ = 0.09) and “other
interaction with long-standing social or family member”
(κ = 0.08), were considered to have slight agreement.
The other items in this domain were considered to have
slight to substantial agreement. For the mood domain,
three items, “repetitive health complaints” (κ = 0.06),
“made negative statements” (κ = 0.04), and “repetitive
anxious complaints/concerns” (κ = 0.16), were consid-
ered to have slight agreement. All others in this domain
were considered to have at least fair agreement.

Kappa Results for IntraRAI Deafblind Supplement

Overall, at least fair agreement existed between the
assessors for the majority of the items. Several areas had
only slight agreement, namely, sense of involvement and
the home environment. In the sense of involvement
domain, two items had kappa values less than 0.20, namely
“reacts positively to interactions initiated by others” (κ = 0.17) and “adjust easily to changes in routines in last 30 days” (κ = 0.16). For the home environment domain, one item, “limited access to home or rooms in the house” (κ = −0.03) was considered to have poor agreement (Table 3).

The Figure shows the cumulative percentage of kappa values across the entire assessment tool. Although some specific items had a lower level of reliability, these results show that, overall, nearly 50 percent of items had a kappa value of 0.60 or higher, which is considered to be at least a moderate level of agreement [15]. Among those items with kappa scores of less than 0.40, the vast majority (78%) showed at least 70 percent agreement between the two raters.

The Figure shows the cumulative percentage of kappa values across the entire assessment tool. Although some specific items had a lower level of reliability, these results show that, overall, nearly 50 percent of items had a kappa value of 0.60 or higher, which is considered to be at least a moderate level of agreement [15]. Among those items with kappa scores of less than 0.40, the vast majority (78%) showed at least 70 percent agreement between the two raters.

The Figure shows the cumulative percentage of kappa values across the entire assessment tool. Although some specific items had a lower level of reliability, these results show that, overall, nearly 50 percent of items had a kappa value of 0.60 or higher, which is considered to be at least a moderate level of agreement [15]. Among those items with kappa scores of less than 0.40, the vast majority (78%) showed at least 70 percent agreement between the two raters.

The Figure shows the cumulative percentage of kappa values across the entire assessment tool. Although some specific items had a lower level of reliability, these results show that, overall, nearly 50 percent of items had a kappa value of 0.60 or higher, which is considered to be at least a moderate level of agreement [15]. Among those items with kappa scores of less than 0.40, the vast majority (78%) showed at least 70 percent agreement between the two raters.

The Figure shows the cumulative percentage of kappa values across the entire assessment tool. Although some specific items had a lower level of reliability, these results show that, overall, nearly 50 percent of items had a kappa value of 0.60 or higher, which is considered to be at least a moderate level of agreement [15]. Among those items with kappa scores of less than 0.40, the vast majority (78%) showed at least 70 percent agreement between the two raters.

The Figure shows the cumulative percentage of kappa values across the entire assessment tool. Although some specific items had a lower level of reliability, these results show that, overall, nearly 50 percent of items had a kappa value of 0.60 or higher, which is considered to be at least a moderate level of agreement [15]. Among those items with kappa scores of less than 0.40, the vast majority (78%) showed at least 70 percent agreement between the two raters.

The Figure shows the cumulative percentage of kappa values across the entire assessment tool. Although some specific items had a lower level of reliability, these results show that, overall, nearly 50 percent of items had a kappa value of 0.60 or higher, which is considered to be at least a moderate level of agreement [15]. Among those items with kappa scores of less than 0.40, the vast majority (78%) showed at least 70 percent agreement between the two raters.

The Figure shows the cumulative percentage of kappa values across the entire assessment tool. Although some specific items had a lower level of reliability, these results show that, overall, nearly 50 percent of items had a kappa value of 0.60 or higher, which is considered to be at least a moderate level of agreement [15]. Among those items with kappa scores of less than 0.40, the vast majority (78%) showed at least 70 percent agreement between the two raters.

The Figure shows the cumulative percentage of kappa values across the entire assessment tool. Although some specific items had a lower level of reliability, these results show that, overall, nearly 50 percent of items had a kappa value of 0.60 or higher, which is considered to be at least a moderate level of agreement [15]. Among those items with kappa scores of less than 0.40, the vast majority (78%) showed at least 70 percent agreement between the two raters.
Internal Consistency of Health Subscales

The Cronbach alpha value ranged from 0.63 (DRS) to 0.93 (ADL SHS) for the health subscales (Table 4). This indicates that for all but the one scale, internal consistency was considered to be good.

Assessor Feedback

In-person or telephone interviews were conducted with all 11 assessors involved in the project. They reported that the assessment, which lasted close to 2 hours on average, can be physically and emotionally draining for clients. As a result, they suggested that completing the assessment over the course of several meetings may be helpful. General agreement existed that the items within the DbS will be useful in the determination of service provision and client needs, while some voiced concern that the interRAI CHA captures medical information that is not useful for service providers.

The items most frequently identified by the assessors as being difficult to complete were related to mood/mental health, continence, pain, medication use, sense of involvement, and classification of vision loss. Although consensus existed that items related to mood and psychological well-being were important and should be kept in the assessment, the assessors felt that they often did not have the skills to assess these areas. They recommended that future training sessions include a more detailed review of the mood items and also include some fictitious case studies, role playing, and observations of another experienced assessor.

DISCUSSION

This project provides evidence of a moderate level of interrater reliability of the items in the interRAI CHA and DbS and good internal consistency for three of the four embedded scales. For the majority of items, the kappa values were at least 0.50, indicating a moderate level of agreement [15]. However, some items, particularly in the areas of mood and psychological well-being, did not reach an acceptable level of agreement. One possible explanation for the low reliability was the reports from the assessors that they had difficulty completing these items because they did not have specific training or background in the area of mental health. They suggested that future training on using the assessment tool should include a more thorough discussion of these items and include strategies that would assist them in reliably assessing them. However, it is acknowledged that domains that involve more abstract concepts or require judgment about internal states, such as mood or quality or level of comfort with social interaction, are likely more difficult to assess reliably, especially among individuals with cognitive impairment. Some individuals with congenital deafblindness also have cognitive impairments and experience difficulties in receptive and expressive communication to varying degrees, making it difficult for them to articulate or self-report their feelings and social experiences. As a consequence, assessors may need to rely more on direct observations of the client and possibly behavioral indicators of internal states.

The interRAI CHA showed excellent reliability, consistent with assessments of the interRAI HC instrument. Across multiple items, direct comparisons can be made, because the interRAI CHA is a subset of those items contained in the interRAI HC. Across the items related to ADLs and IADLs, the current kappa values were nearly identical to those from a recent six-country study of interrater reliability across multiple interRAI instruments that included 220 interRAI HC assessments [17]. In the case of communication, pain, and preventive health, the current kappa values were slightly lower than those in the multicountry study (differences ranging from 0.09 to 0.25) [17].

The current kappa values met or exceeded the scores for most countries, in another multicountry study among long-term-care residents, for items related to ADLs and were at least as high as the lowest kappa value reported for both communication and behavior symptoms [18]. Although higher than in the current study, the kappa values reported by Sgadari et al. for mood and behavior symptoms also tended to be much lower than for other domain areas, which is consistent with our findings.

The single items related to vision and hearing were comparable with, or at times higher, than data collected in

Table 4.
Internal consistency of select health subscales (n = 44 participants).

<table>
<thead>
<tr>
<th>Subscale</th>
<th>No. of Items</th>
<th>Cronbach Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADL Self-Performance Hierarchy</td>
<td>4</td>
<td>0.93</td>
</tr>
<tr>
<td>IADL Involvement</td>
<td>3</td>
<td>0.74</td>
</tr>
<tr>
<td>Depression Rating</td>
<td>7</td>
<td>0.63</td>
</tr>
<tr>
<td>Orientation and Mobility Capacity</td>
<td>3</td>
<td>0.80</td>
</tr>
</tbody>
</table>

ADL = activities of daily living; IADL = instrumental ADL.
home care and nursing homes in the United States [19] and in a seven-country study of long-term-care residents [18]. The reliability for items related to pain and IADLs was also similar to that reported for inpatient psychiatry [20].

Although this study has several strengths, several limitations of the research should also be noted. Although random assignment was used for the completion of the initial assessment, this was not always possible because of scheduling limitations on the part of the service providers. Therefore, a bias may have been introduced if the initial assessor consistently achieved different scores that reflected his or her rapport with the individual who is deafblind. Based on our interviews with the assessors, we have no reason to believe that this situation created a serious flaw in the study.

In terms of the study sample, the participants do not represent a random sample of individuals with deafblindness in the province. Most participants had been previously assessed as part of an earlier pilot study [4]. This may have lead to an overestimate of the reliability results, given that individuals, and their caregivers, were familiar with the assessment tool and the process. However, in virtually all cases, there was nearly a year between the previous and current assessment, which should have minimized the influence of practice effects.

To date, a limited number of studies have published results related to internal consistency of the health subscales. Two such articles measured internal consistency in a group of clients with intellectual disability [21] and also in a population of inpatient mental health clients [20]. Where comparisons were possible, the current alpha values were nearly identical for the ADL SHS and slightly lower for both the DRS (lower by 0.15) and the IADL Involvement Scale (lower by 0.18) [20–21]. The higher DRS score likely reflects the enhanced training and background of the assessors in the area of mental health. In addition, the alpha values in the current project are very similar to those previously obtained from a group of 182 adults with deafblindness [4].

Overall, the interRAI CHA and DbS represent a reliable assessment system for measuring the needs, abilities, and strengths of individuals with deafblindness. The assessment, and accompanying service-planning protocols and health subscales, provides a wealth of information to assessors. The feedback from participating assessors has been very positive, and they continue to highlight the value of having standardized information available for each client. The Ministry of Community and Social Services plans to implement this assessment in the spring of 2011 for anyone interested in intervenor support provided by the government. By using this standardized assessment, assessors will be able to accurately identify the needs of clients and ensure that service delivery is optimized through the development of client-specific service plans. Furthermore, the assessments will populate an electronic database that will enable the government to better track client demographics and needs and how the population they serve is changing over time. This ultimately will benefit clients and families, because it will create a greater level of consistency in assessment across service providers and better equip policy makers to make evidence-based decisions as they plan for the future allocation of resources.

CONCLUSIONS

To our knowledge, this is one of the first studies of the psychometric properties of an assessment instrument for use with adults who are deafblind. Coupled with our previous research [4], this study provides some preliminary evidence of the reliability of the items within the interRAI CHA and DbS. The interrater reliability was at least moderate for nearly half the items, and the instrument appears to have reliability levels comparable with those of other interRAI instruments. However, items related to mood and psychological well-being did not reach an acceptable level of agreement. In order to foster further improvements in the instrument’s reliability, more in-depth training of assessors appears to be needed in these particular areas. Given that this is the first study to test the interrater reliability of this tool, further research is needed to substantiate these findings and also to determine whether enhanced training of assessors can lead to improved levels of reliability.

ACKNOWLEDGMENTS

Author Contributions:
Conceived original idea: D. M. Guthrie, J. P. Hirdes.
Primary lead on developing manuscript: D. M. Guthrie.
Data analysis: D. M. Guthrie, R. Pitman, M. Ashworth.
Coordinated daily activities of project: L. Bowman, M. Ashworth.
Primary data collection and analysis: L. Bowman, M. Ashworth.  
Financial Disclosures: The authors have declared that no competing interests exist.  
Funding/Support: This material was based on work supported by the CNIB through the EA Baker Applied Research Grant program (no grant number was assigned). We are also grateful for the ongoing assistance provided by the Ontario Ministry of Community and Social Services.  
Additional Contributions: This project could not have been completed without the support and participation of a very committed group of assessors. We appreciate the time and effort they dedicated to this project. Additionally, we would like to acknowledge the clients and family members who have participated in this project. Melody Ashworth is now with the Ontario Institute for Studies in Education, and family members who have participated in this project. Additionally, we would like to acknowledge the clients and family members who have participated in this project. Melody Ashworth is now with the Ontario Institute for Studies in Education, and family members who have participated in this project.  
Institutional Review: All study participants, or a substitute decision maker, signed an informed consent statement, and the research protocol was reviewed and approved by the Wilfrid Laurier University Research Ethics Board.  
Participant Follow-Up: The authors do not plan to inform participants of the publication of this study because contact information is unavailable.

REFERENCES


Submitted for publication September 7, 2010. Accepted in revised form December 13, 2010.

This article and any supplementary material should be cited as follows: