



Published in final edited form as:

Qual Life Res. 2016 October ; 25(10): 2417–2427. doi:10.1007/s11136-016-1273-y.

HDQLIFE: The development of two new computer adaptive tests for use in Huntington disease, Speech Difficulties and Swallowing Difficulties

N. E. Carlozzi¹, S.G. Schilling^{1,2}, J.-S. Lai^{3,4}, J.S. Perlmutter^{5,6}, M.A. Nance^{7,8}, J.F. Waljee⁹, J.A. Miner¹, S.K. Barton⁵, S.M. Goodnight¹, and P. Dayalu¹⁰

¹Department of Physical Medicine and Rehabilitation, University of Michigan, Ann Arbor, MI, USA

²Institute for Social Research, University of Michigan, Ann Arbor, MI, USA

³Center on Outcomes, Research and Education, Evanston northwestern Healthcare, Northwestern University, Evanston, IL, USA

⁴Institute for Health Services Research & Policy Studies, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

⁵Departments of Neurology, Radiology, and Anatomy & Neurobiology, Washington University School of Medicine, St. Louis, MO, USA

⁶Program in Occupational Therapy and Program in Physical Therapy, Washington University School of Medicine, St. Louis, MO, USA

⁷Hennepin County Medical Center, Minneapolis, MN, USA

Address reprint requests to: Noelle E. Carlozzi, Ph.D., University of Michigan, Department of Physical Medicine & Rehabilitation, North Campus Research Complex, 2800 Plymouth Road, Building NCRC B14, Room D037, Ann Arbor, MI 48109-2800, Phone: (734) 763 – 8917, Fax: (734) 763-7186, carlozzi@med.umich.edu.

COMPLIANCE WITH ETHICAL STANDARDS

CONFLICT OF INTEREST:

Carlozzi, N.E. currently has research grants from the NIH; she is also supported by grant funding from the NIH, NIDILRR, and CHDI; she declares no conflicts of interest.

Schilling, S.G. has a research grant from NSF. He also is supported by grant funding from NIH. He declares no conflicts of interest.

Lai J.-S. currently has research grants from the NIH; she declares no conflicts of interest.

Perlmutter, J.S. currently has funding from the NIH, HDSA, CHDI, and APDA. He has received honoraria from the University of Rochester, American Academy of Neurology, Movement Disorders Society, Toronto Western Hospital, St Lukes Hospital in St Louis, Emory U, Penn State, Alberta innovates, Indiana Neurological Society, Parkinson Disease Foundation, Columbia University, St. Louis University, Harvard University and the University of Michigan.

Nance, M.A. declares no conflicts of interest.

Waljee, J.F. currently has research grants from the Agency for Healthcare Research and Quality, the American College of Surgeons, and the American Foundation for Surgery of the Hand; she declares no conflicts of interest.

Miner, J.A. is supported by research grants from the NIH; she declares no conflict of interest.

Barton, S.K. is supported by grant funding from the Huntington's Disease Society of America, CHDI Foundation and the NIH. She declares no conflicts of interest.

Goodnight, S.M. is supported by grant funding from the NIH and the Craig H. Neilsen Foundation; she declares no conflicts of interest.

Dayalu, P. currently has research grants from the NIH, Astra-Zeneca, and Vaccinex. He declares no conflicts of interest.

Ethical approval:

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent:

Informed consent was obtained from all individual participants included in the study.

⁸Department of Medical and Molecular Genetics, Indiana University, Indianapolis, IN, USA

⁹Department of Surgery, University of Michigan, Ann Arbor, MI, USA

¹⁰Department of Neurology, University of Michigan, Ann Arbor, MI, USA

Abstract

Purpose—Huntington disease (HD) is an autosomal dominant neurodegenerative disease which results in several progressive symptoms, including bulbar dysfunction (i.e., speech and swallowing difficulties). Although difficulties in speech and swallowing in HD have a negative impact on health-related quality of life, no patient-reported outcome measure exists to capture these difficulties that are specific to HD. Thus, we developed a new patient reported outcome measure for use in the Huntington Disease Health-Related Quality of Life (HDQLIFE) Measurement System that focused on the impact that difficulties with speech and swallowing have on HRQOL in HD.

Methods—Five hundred seven individuals with prodromal and/or manifest HD completed 47 newly developed items examining speech and swallowing difficulties. Unidimensional item pools were identified using exploratory and confirmatory factor analyses (EFA and CFA, respectively). Item response theory (IRT) was used to calibrate the final measures.

Results—EFA and CFA identified two separate unidimensional sets of items: Speech Difficulties (27 items) and Swallowing Difficulties (16 items). Items were calibrated separately for these two measures and resulted in item banks that can be administered as computer adaptive tests (CATs) and/or 6-item, static short forms. Reliability of both of these measures was supported through high correlations between the simulated CAT scores and the full item bank.

Conclusions—CATs and 6-item calibrated short forms were developed for HDQLIFE Speech Difficulties and HDQLIFE Swallowing Difficulties. These measures both demonstrate excellent psychometric properties, and may have clinical utility in other populations where speech and swallowing difficulties are prevalent.

Keywords

Health-related quality of life; Neuro-QoL; PROMIS; HDQLIFE; Huntington disease; Speech difficulties; Swallowing difficulties; Patient reported outcome (PRO)

Huntington disease (HD) is an autosomal dominant neurodegenerative disorder which causes profound cognitive, behavioral, and motor decline, leading to death in 15-20 years after diagnosis [1]. Although HD produces multiple motor deficits, most research has focused on chorea [2-5], and to a lesser extent dystonia, bradykinesia, rigidity and gait impairment. While dysarthria and dysphagia are also common (resulting in difficulties with speech and swallowing), much less work has been conducted examining these deficits. Key features of HD dysarthria include a slower and irregular rate of speech, prolonged pauses, abnormal pitch and prosody, and imprecise articulation of consonants [6-8]. The communication difficulties characteristic of dysarthria are related to impaired social functioning, loss of dignity and independence, and sometime an increase in irritability, frustration, and depression. Dysphagia in HD is also complex [9]. Complications of

dysphagia include weight loss and aspiration, which is a major cause of death in individuals with HD [10; 11]. Thus, both dysarthria and dysphagia may cause substantial disability and impair health-related quality of life (HRQOL) in individuals with HD.

Despite the morbidity and mortality associated with dysarthria and dysphagia among individuals with HD, their assessment has been relatively neglected in clinical research; few studies have analyzed the quality and trajectory of speech and swallowing dysfunction in HD. For example, the most widely used outcomes instrument in HD clinical trials, the Unified Huntington Disease Rating Scale (UHDRS) [12], has only a single item for dysarthria, and none for dysphagia. This is especially problematic since a range of motor disorders (chorea, incoordination, hypokinesia, and dystonia) can affect a number of different oropharyngeal structures (lips, jaw, tongue, palate, and pharynx) that are associated with different speech (respiration, phonation, and articulation) or swallowing (preparatory, oral, and pharyngeal) difficulties. Cognitive and behavioral challenges, such as frontal-executive dysfunction, impulsivity, and loss of insight, can also exacerbate these problems.

To date, the influence of bulbar dysfunction on health-related quality of life (HRQOL) among patients with HD is unclear. In modern medical research, patient reported HRQOL provides an essential measure of efficacy for new treatments [13]. In HD, there is an urgent need for more meaningful and sensitive patient reported outcome (PRO) measures that capture HRQOL experienced by these individuals. New measurement systems have been developed to sensitively assess HRQOL in individuals with other chronic diseases (e.g., the Quality of Life for Neurological Disorders [Neuro-QoL] measurement system [14; 15] and Patient Reported Outcomes Measurement Information System [PROMIS] [16]). These new systems permit cross-disease comparison, are customizable, and can be administered more efficiently due to their computer adaptive testing administration format. Yet, these new measurement systems do not capture bulbar dysfunction. Given the negative effects of HD, a measurement system which allows for cross disease comparison, customizable domain assessment/selection, and computer adaptive test administration format are critical to accurately and efficiently capture the outcomes that matter most to these individuals.

Existing measures of bulbar dysfunction do not assess the impact of bulbar dysfunction on HRQOL. For example, a recent 11-item scale developed to assess dysphagia in HD [17] evaluates the presence or absence of dysphagia-related problems such as drooling, nasal regurgitation, coughing, and choking. However, this scale does not affect that dysphagia has on quality of life. Other existing measures designed to capture how bulbar dysfunction affects HRQOL, such as the SWAL-QOL [18; 19] (developed to assess quality of life in dysphagia more generally), do not have validation data in in HD. Furthermore, the SWAL-QOL's 44 items may be impractical for use in HD where testing burden is especially problematic. We know of no dysarthria scale specifically developed for HD.

To fill this urgent need, we aimed to develop a battery of quality of life measures, specifically focused on Huntington disease, the Huntington Disease Health-Related Quality of Life (HDQLIFE) measurement system [20]. HDQLIFE is designed to capture HRQOL across several domains including speech and swallowing dysfunction. Other domain development included chorea and end of life concerns which are highlighted elsewhere [21;

22]. In addition, we capitalized on the existing measurement system capabilities inherent in the existing Neuro-QoL/PROMIS measures of HRQOL; therefore, measurement development followed established PROMIS guidelines [23]. In this report, we describe the development of new measurement tools to evaluate how speech and swallowing difficulties affect HRQOL in HD.

Methods

Participants

We invited individuals with either prodromal or manifest HD to participate in this study. Participants were included with a clinical diagnosis of manifest or prodromal HD (gene-positive status for the HD CAG expansion and no clinical diagnosis), and were at least 18 years of age, able to read and understand English, and have the ability to provide informed consent. Recruitment occurred at several specialized HD treatment centers across the United States (University of Michigan, University of Iowa, University of California-Los Angeles, Indiana University, Johns Hopkins University, Rutgers University, Struthers Parkinson's Center, and Washington University). Participants were also recruited through established registries, such as the HD Roster, online medical record data capture systems [24], and established research cohorts, such as the Predict-HD study [25].

Measures

HDQLIFE Difficulties with Speech and Swallowing Item Pool—To capture all possible concerns regarding speech and swallowing difficulties, focus groups were held [26] and literature reviews were conducted, yielding 102 questions in the initial item pool. An iterative process comprised of expert review, translatability review, literacy review, and cognitive interviews with individuals with HD, resulted in a final item pool of $n = 47$ items. See Carlozzi et al. [20], for a detailed description of the development of these items and this iterative process.

The Unified Huntington's Disease Rating Scales (UHDRS) [12]—Several measures from the UHDRS were administered to study participants to evaluate overall functioning in HD. This included the Total Functional Capacity (TFC) scale [27], a 5-item clinician-rated measure designed to evaluate total functional capacity. Clinical ratings of functional capacity are provided for occupation, finances, domestic chores, ADLs, and care level; scores for each item range from 0-3 or 0-4 depending on the number of response options for each item (e.g., for clinician-rated financial functional capacity, response options include: unable [0 points], major assistance [1 point], slight assistance [2 points] or normal functioning [3 points]). The TFC score ranges from 0 to 13 with higher scores indicating better functioning. TFC scores were used to classify participants with manifest HD as either: early-stage (sum scores of 7-13) or late-stage (sum scores of 0-6) HD.

Statistical Analyses

Analyses were conducted in accordance with established PROMIS measurement development guidelines [23].

The unidimensionality of the difficulties with speech and swallowing item pool was examined using both exploratory (EFA) and confirmatory factor analyses (CFA). The sample was randomly assigned to two separate datasets: one for the EFA with PROMAX rotation and the other for the CFA; analyses were conducted using MPLUS 6.11 [28]. For EFA, the numbers of factors were identified based on eigenvalue (criterion: >1), scree plot (criterion: the number of factors before the break in the scree plot), the amount of variance accounted, and item factor loadings (criterion > 0.4). A promax rotation was then used to examine the association among factors by examining their loadings (criterion > 0.4) and inter-factor correlations. CFA was conducted on the second sample to confirm the factor structure from EFA. The following criteria were used to assess model fit: a comparative fit index (CFI) greater than 0.90; Tucker-Lewis index (TLI) greater than 0.90; Root Mean Square Error of Approximation (RMSEA) less than 0.1; R-square greater than 0.3; and residual correlations less than 0.2 [29-31].

The Graded Response Model (GRM) [32] that was implemented in IRTPRO [33] was used to estimate item parameters. IRT-based ordinal logistic regression [34] as implemented in LORDIF freeware [35] was used to evaluate differential item functioning (DIF). Specifically, DIF was examined for age (< 50 years vs. >50 years and < 40 years vs. >40 years) to ensure that potential bias was due to age and not to the manifestation of clinically significant symptoms around age 40 in HD), gender, and education (high school graduate or lower vs. at least some college) to ensure that each item does not unfairly discriminate against individuals in any one group and to ensure consistency in measurement properties across groups.

Post hoc CAT simulation was conducted using Firestar CAT simulation software [36]. In this simulation, Firestar generated respondents with predefined scores on the newly developed measures. These 10,000 “virtual” respondents were used to estimate scores. The number of items needed for administration prior to meeting the stopping rules (SE of measurement less than 3.0, or number of items exceeds 12) was generated. These simulated scores were compared to the CAT scores based on the administration of all items.

Results

Five hundred seven (507) individuals with prodromal or manifest HD participated in this study. Of these individuals, 196 individuals had prodromal HD (CAG ≤ 36 , but did not yet have a clinical diagnosis of HD), 193 had early-stage HD (sum scores of 7-13 on the TFC) and 117 had later-stage HD (sum scores of 0-6 on the TFC). Participants ranged from 18-81 years of age ($M = 49.01$, $SD = 13.21$) and 59.2% of participants were female. Significant group differences were seen for age (as symptoms are progressive with age), $F(2, 503) = 47.360$, $p < .0001$, with individuals who were prodromal ($M = 42.60$, $SD = 12.04$) being significantly younger than the early-HD group ($M = 51.91$, $SD = 12.41$) and the late-HD group ($M = 55.07$, $SD = 11.89$). Additionally, the early-HD group was younger than the late-HD group. Furthermore, groups did not differ on gender, $X^2(2, N = 506) = 3.193$, $p = .20$. The majority of participants were Caucasian (96.4%); 2.0% were African American, 1.4% were classified as “other,” and 0.2% were unknown. Participants’ education ranged from 4 to 26 years ($M = 15.06$, $SD = 2.88$). While there were group differences in education, $F(2,$

501) = 14.781, $p < .0001$, these differences were small; early- ($M = 14.72$, $SD = 2.78$) and late-HD ($M = 14.22$, $SD = 2.62$) had 1 to 1.5 years less education relative to the prodromal HD group ($M = 15.88$ years, $SD = 2.94$).

EFA and CFA Findings

Exploratory Factor Analyses—Four factors had eigenvalues greater than 1; Factor 1 included 23 items that examined speech difficulties, the second factor included 17 items that reflected swallowing difficulties, the third factor included two items about pureed food and two items about actions taken with food to avoid choking, and the fourth factor included 4 items about the impact of speech difficulties (see Table 1). For the remaining analyses, Factors 1 and 4 were combined in a single set to represent Speech Difficulties (27 items). In addition, all 17 items from Factor 2, Swallowing Difficulties were examined further. Factor 3 was regarded as a spurious factor with limited clinical utility and was comprised of too few items to warrant further development; thus, this factor was excluded from further examination.

Confirmatory Factor Analyses—For Speech Difficulties (Factor 1 in EFA results), acceptable fit indices were found in CFA results: CFI = 0.98, TLI = 0.98, RMSEA = 0.09, all $R^2 > 0.3$. All residual correlations were < 0.14 . Cronbach's alpha for this scale was 0.98 and all item-total correlations were > 0.4 . For Swallowing Difficulties (factor 2 in EFA results), acceptable fit indices were also found in CFA results: CFI = 0.98, TLI = 0.98, RMSEA = 0.11, all $R^2 > 0.3$. A high residual correlation of 0.180 was identified between items “During the past 7 days I needed help eating.” and “In the past 7 days how often did you choke?” The study team decided to remove “During the past 7 days I needed help eating.” from the measure. For the 16 remaining items the fit statistics were good (CFI = 0.99, TLI = 0.98, RMSEA = 0.11, all $R^2 > 0.3$) and all residual correlations were < 0.14 . Cronbach's alpha for this scale was 0.97 and all item-total correlations were > 0.4 .

Item response Theory (IRT) Analyses

Speech Difficulties—For *Speech Difficulties*, the 27 items had slopes ranging from 2.91 to 6.32 and thresholds ranging from -0.41 to 2.42 (Table 2). Information was good between scaled scores of -1.1 and 2.7 (see Figure 1 for the scale information function), and the marginal reliability was 0.92. S-X² model fit statistics were examined; all but 9 of the 27 items had adequate or better model fit statistics ($p > 0.05$). No items demonstrated DIF on age (<50 vs ≥ 50 and <40 vs ≥ 40), gender (male vs female), or education (some college and lower vs college degree and higher).

CAT Simulation: The correlation between the CAT scores and the full item-bank was 0.99, indicating that CAT scores based on the Speech Difficulties item bank can produce results that are very similar to those obtained with administration of the entire 27 item set. In addition, the CAT simulation always used all 27 items in the item bank at -1 SD units; from -0.1 to $+2$ SD units, the CAT always used the minimum number of 4 items in the item bank; and at $+3$ SD units the CAT used all 27 items in the item bank (Figure 2). Thus, the CAT simulation indicates that fewer items were needed to estimate scores for individuals with greater speech difficulties than for individuals with less speech difficulties.

Using this calibrated measure, a 6-item short form was created based on item calibration statistics (e.g., slope, item characteristic curves, item information, average item difficulty), as well as input on clinical characteristics (e.g., items were selected that represent different clinical components of speech difficulties). Specifically, we balanced the psychometric considerations with clinical content to ensure representativeness of the items that were selected for the short form (see bolded items in Table 2).

Scoring of short forms: The IRT-scaled scores (theta) were converted into a standardized score utilizing a t metric (mean = 50, SD = 10); see Table 3 for a summed score scale conversion table. Higher scores indicate greater swallowing difficulty; thus, scores 60 indicate that the individual is functioning more poorly than 68.27% of people with HD and scores 70 indicate that the individual is functioning more poorly than 95.45% of individuals with HD. The short forms scores had a marginal reliability of 0.88.

Swallowing Difficulties—For *Swallowing Difficulties*, parameter estimates of the 16 items from GRM analysis indicated slopes ranging from 3.06 to 7.11 and thresholds ranging from 0.11 to 3.00. S-X2 model fit statistics were examined using the IRTFIT macro program; all but 5 of the 16 items had adequate or better model fit statistics ($p > 0.05$). We decided to delete item “*I ate slowly to avoid choking.*” due to the poor fit statistics and the substantial content overlap with another item within the bank (“During the past 7 days I had to eat slowly to avoid choking.”), resulting in a final set of 15 items. GRM for the remaining 15 items indicated IRT parameter estimates indicated slopes ranging from 3.06 to 7.11 and thresholds ranging from 0.11 to 3.00 (Table 4). Information was good between scaled scores of -0.5 and 3.0 (see Figure 3 for the scale information function) and the marginal reliability was 0.82. No items demonstrated DIF on age, gender, and education.

CAT Simulation: The average number of items administered to 10,000 virtual respondents by the IRTPRO CAT simulation software was 9.11 (Figure 4). The correlation between the CAT scores and the full item-bank was 0.99, indicating that CAT based on the Swallowing Difficulties item bank can produce results that are very similar to those obtained with administration of the entire 15 item set. In addition, the CAT simulation always used all 15 items in the item bank at -1 SD units; from $+0.4$ to $+2.2$ SD units, the CAT always used the minimum number of 4 items in the item bank; and at $+3$ SD units the CAT used all 15 items in the item bank. Thus, the CAT simulation indicates that fewer items were needed to estimate scores for individuals with greater swallowing difficulties than for individuals with less swallowing difficulties.

Using the same criteria as “Speech Difficulties,” a 6 item short form was created as indicated by bolding in Table 4.

Scoring of short forms: The IRT-scaled scores (theta) were converted into a standardized score utilizing a t metric (mean = 50, SD = 10); see Table 5 for a summed score scale conversion table. Higher scores indicate greater speech difficulty; thus, scores 60 indicate that the individual is functioning more poorly than 68.27% of people with HD and scores 70 indicate that the individual is functioning more poorly than 95.45% of individuals with HD.

Discussion

The purpose of this study was to evaluate the psychometric properties of newly developed measures of Speech Difficulties and Swallowing Difficulties for patients with HD. Results supported the development of two new item banks that could be administered as CATs, as well as corresponding 6-item short forms that evaluate speech and swallowing difficulties: “Speech Difficulties” and “Swallowing Difficulties.” These measures were developed for inclusion in the HDQLIFE measurement system, a system designed to capture both the generic and HD-specific aspects of HRQOL for individuals with HD [20]. The HDQLIFE includes the first computer adaptive tests that have been developed for use specifically in HD. The computer adaptive test format for both Speech Difficulties and Swallowing Difficulties offer many advantages to more standard HRQOL assessment. Test administration requires that only the most relevant items be administered to an individual, thus shortening administration time without sacrificing sensitivity/specificity, decreasing overall test administration burden (which is especially problematic in this population). Specifically, the CAT performs well (i.e., less than 10 items are administered) for individuals with Speech Difficulties scale scores between -0.6 and $+2.5$, and for individuals with Swallowing Difficulties scale scores between 0.0 and $+2.7$. As it is not uncommon for a CAT to require the administration of more items for extreme responders, it is not especially surprising that the CAT performs well for people that are experiencing difficulties, but requires substantially more items for individuals that are either not experiencing difficulties or are experiencing significant difficulties. Specifically, individuals that are functioning in the top 1% or the bottom 24% for Speech Difficulties would need to take between 11 and 26 items in order to achieve the same precision as those individuals with scale scores between -0.6 and $+2.5$. Similarly, individuals that are functioning in the top 1% or the bottom 50% for Swallowing Difficulties would need to take between 11 and 26 items in order to achieve the same precision as those individuals with scale scores between 0.0 and $+2.7$. As these measures are designed to capture clinical dysfunction, we would recommend setting the CAT cutoff criterion to the administration of no more than 10 items to ensure an adequate balance between precision and test burden.

The HDQLIFE Speech Difficulties and Swallowing Difficulties item banks were developed according to established methodology [23], and meet the established psychometric standards including being comprised of homogenous item sets (i.e., they are unidimensional), having excellent reliability, and being devoid of items that are biased against age, gender, or education. Scores on these new measures also conform to a standard t metric, with a mean of 50 and standard deviation of 10; higher scores indicate greater difficulty. This type of scoring metric increases the clinical utility of the measure, since scores immediately provide an estimation of an individual’s functioning relevant to a reference group (in this case, other individuals with HD). For example, an individual with a score of 60 or greater would suggest that he/she is functioning more poorly than 68.27% of people with HD, whereas an individual with scores of 70 or above are functioning more poorly than 95.45% of individuals with HD. Such scores provide potential cut-points for establishing when additional clinical follow-up might be warranted. Similarly, such scores will help to

determine the effectiveness of clinical interventions designed to evaluate speech and/or swallowing function.

This study has several limitations. Both newly developed measures are more sensitive for individuals that are more symptomatic (i.e., are experiencing more speech and/or swallowing difficulties). Thus, more items are needed to estimate scores in individuals where impairments are subtle or nonexistent. While this characteristic is advantageous for clinical trials focused on improving dysfunction, it may limit the ability to detect subtle changes over time for prodromal individuals. Ultimately, longitudinal assessment is needed to confirm or refute this hypothesis. Furthermore, while simulation data support the sensitivity of CAT administration, prospective data is needed to confirm the validity of these new measures. Future work in other HD samples is needed to fully understand both the utility and the strengths and weaknesses of these measures. In particular, future work is needed to examine the relationship of these measures to objective speech and swallowing assessments, as well as other PRO measures of HRQOL and bulbar dysfunction.

Regardless, these new HDQLIFE measures provide an exciting tool to assessing self-reported bulbar dysfunction in individuals with HD. Furthermore, the relationship of measurements with more established, but generic measures of HRQOL (i.e., Neuro-QoL and PROMIS) provide a complementary arsenal of tools that can aid both sensitive, disease-specific assessment, as well as support cross-disease examination of HRQOL. Together, these new measures and the larger HDQLIFE system are designed to provide researchers and clinicians with a comprehensive tool to assess all relevant aspects of HRQOL in individuals with HD. Furthermore, these measures provide the first comprehensive assessment of self-reported bulbar dysfunction and their relation to everyday life for use in individuals with HD. Finally, although these measures have been developed in a heterogeneous group of individuals with prodromal and manifest HD, these measures may also have clinical utility in other populations where speech and swallowing difficulties are prevalent (e.g., stroke, spinal cord injury, traumatic brain injury).

Acknowledgements

Work on this manuscript was supported by the National Institutes of Health (NIH), National Institute of Neurological Disorders and Stroke (R01NS077946) and the National Center for Advancing Translational Sciences (UL1TR000433). In addition, a portion of this study sample was collected in conjunction with the Predict-HD study. The Predict-HD was supported by the NIH, National Institute of Neurological Disorders and Stroke (R01NS040068), the NIH, Center for Inherited Disease Research (provided supported for sample phenotyping), and the CHDI Foundation (award to the University of Iowa). Jennifer Waljee's effort was supported in part by a Mentored Clinical Investigator Award through the Agency for Healthcare Research and Quality (K08HS023313). We thank the University of Iowa, the Investigators and Coordinators of this study, the study participants, the National Research Roster for Huntington Disease Patients and Families, the Huntington Study Group, and the Huntington's Disease Society of America. We acknowledge the assistance of Jeffrey D. Long, Hans J. Johnson, Jeremy H. Bockholt, Roland Zschiegner, and Jane S. Paulsen. We also acknowledge Roger Albin, Kelvin Chou, and Henry Paulsen for the assistance with participant recruitment. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

HDQLIFE Site Investigators and Coordinators: Noelle Carlozzi, Praveen Dayalu, Stephen Schilling, Amy Austin, Matthew Canter, Siera Goodnight, Jennifer Miner, Nicholas Migliore (University of Michigan, Ann Arbor, MI); Jane Paulsen, Nancy Downing, Isabella DeSoriano, Courtney Shadrick, Amanda Miller (University of Iowa, Iowa City, IA); Kimberly Quaid, Melissa Wesson (Indiana University, Indianapolis, IN); Christopher Ross, Gregory Churchill, Mary Jane Ong (Johns Hopkins University, Baltimore, MD); Susan Perlman, Brian Clemente, Aaron Fisher, Gloria, Obialisi, Michael Rosco (University of California Los Angeles, Los Angeles, CA); Michael McCormack, Humberto Marin, Allison Dicke (Rutgers University, Piscataway, NJ); Joel Perlmutter, Stacey Barton,

Shineeka Smith (Washington University, St. Louis, MO); Martha Nance, Pat Ede (Struthers Parkinson's Center); Stephen Rao, Anwar Ahmed, Michael Lengen, Lyla Mourany, Christine Reece, (Cleveland Clinic Foundation, Cleveland, OH); Michael Geschwind, Joseph Winer (University of California – San Francisco, San Francisco, CA), David Cella, Richard Gershon, Elizabeth Hahn, Jin-Shei Lai (Northwestern University, Chicago, IL).

References

1. Roos RAC, Hermans J, Vegtervandervlis M, Vanommen GJB, Bruyn GW. Duration of Illness in Huntingtons-Disease Is Not Related to Age at Onset. *Journal of Neurology Neurosurgery and Psychiatry*. 1993; 56(1):98–100.
2. Albin RL, Reiner A, Anderson KD, Penney JB, Young AB. Striatal and Nigral Neuron Subpopulations in Rigid Huntingtons-Disease - Implications for the Functional-Anatomy of Chorea and Rigidity-Akinesia. *Annals of Neurology*. 1990; 27(4):357–365. [PubMed: 1972318]
3. Coyle JT, Schwarcz R. Lesion of Striatal Neurons with Kainic Acid Provides a Model for Huntingtons-Chorea. *Nature*. 1976; 263(5574):244–246. [PubMed: 8731]
4. Marshall FJ, Walker F, Frank S, Oakes D, Plumb S, Factor SA, Fahn S, Hunt VP, Jankovic J, Shinaman A, Shoulson I, Goldstein J, Corey-Bloom J, Belden J, Wojcieszek J, Nickerson C, Higgins D, McCall M, Sanchez-Ramos J, Stepanov N, Kostyk S, Harrison J, Testa C, Shannon B, Feigin A, Radtke D, Nance M, Stanton P, Ashizawa T, Hunter C, Ondo W, Rose K, Geschwind M, Jaglin J, Shannon K, Matthews M, Colcher A, Riendl P, Blindauer K, Dustin K, Anderson K, Fowler S, Mehta D, Hanna P, Lindsay P, Eberly S, Zhao HW, MacDonald M, Young A, Grp HS. Tetrabenazine as antichorea therapy in Huntington disease - A randomized controlled trial. *Neurology*. 2006; 66(3):366–372. [PubMed: 16476934]
5. Metman LV, Morris MJ, Farmer C, Gillespie M, Mosby K, Wu J, Chase TN. Huntington's disease - A randomized, controlled trial using the NMDA-antagonist amantadine. *Neurology*. 2002; 59(5): 694–699. [PubMed: 12221159]
6. Hartelius L, Carlstedt A, Ytterberg M, Lillvik M, Laakso K. Speech disorders in mild and moderate Huntington disease: Results of dysarthria assessments of 19 individuals. *Journal of Medical Speech-Language Pathology*. 2003; 11(1):1–14.
7. Skodda S, Schlegel U, Hoffmann R, Saft C. Impaired motor speech performance in premotor stages of Huntington's disease (HD) over time - A longitudinal investigation. *Movement Disorders*. 2014; 29:S217–S217.
8. Skodda S, Schlegel U, Hoffmann R, Saft C. Impaired motor speech performance in premotor stages of Huntington's disease (HD). *Movement Disorders*. 2014; 29:S217–S217.
9. Heemsker AW, Roos RAC. Dysphagia in Huntington's Disease: A Review. *Dysphagia*. 2011; 26(1): 62–66. [PubMed: 20838817]
10. Lanska DJ, Lavine L, Lanska MJ, Schoenberg BS. Huntingtons-Disease Mortality in the United-States. *Neurology*. 1988; 38(5):769–772. [PubMed: 2966305]
11. Sorensen SA, Fenger K. Causes of Death in Patients with Huntingtons-Disease and in Unaffected 1st Degree Relatives. *Journal of Medical Genetics*. 1992; 29(12):911–914. [PubMed: 1479606]
12. Huntington Study Group. Unified Huntington's Disease Rating Scale: reliability and consistency. *Mov Disord*. 1996; 11(2):136–142. [PubMed: 8684382]
13. Basch E. The Missing Voice of Patients in Drug-Safety Reporting. *New England Journal of Medicine*. 2010; 362(10):865–869. [PubMed: 20220181]
14. Cella D, Nowinski C, Peterman A, Victorson D, Miller D, Lai J-S, Moy C. The Neurology Quality of Life Measurement (Neuro-QOL) Initiative. *Archives of Physical Medicine and Rehabilitation, Supplement*. 2011; 92(Suppl 1):S28–S36.
15. Gershon RC, Lai JS, Bode R, Choi S, Moy C, Bleck T, Miller D, Peterman A, Cella D. Neuro-QOL: quality of life item banks for adults with neurological disorders: item development and calibrations based upon clinical and general population testing. *Quality of Life Research*. 2011
16. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, Amtmann D, Bode R, Buysse D, Choi S, Cook K, DeVellis R, DeWalt D, Fries J, Gershon R, Hahn E, Lai J-S, Pilkonis P, Revicki D, Rose M, Weinfurt K, Hays R. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested in its first wave of adult self-reported health outcome item banks: 2005-2008. *Journal of Clinical Epidemiology*. 2010; 63:1179–1194. [PubMed: 20685078]

17. Heemskerk AW, Verbist BM, Marinus J, Heijnen B, Sjogren EV, Roos RAC. The Huntington's Disease Dysphagia Scale. *Movement Disorders*. 2014; 29(10):1312–1316. [PubMed: 24862624]
18. McHorney CA, Martin-Harris B, Robbins J, Rosenbek J. Clinical validity of the SWAL-QOL and SWAL-CARE outcome tools with respect to bolus flow measures. *Dysphagia*. 2006; 21(3):141–148. [PubMed: 16715210]
19. McHorney CA, Robbins J, Lomax K, Rosenbek JC, Chignell K, Kramer AE, Bricker DE. The SWAL-QOL and SWAL-CARE outcomes tool for oropharyngeal dysphagia in adults: III. Documentation of reliability and validity. *Dysphagia*. 2002; 17(2):97–114. [PubMed: 11956835]
20. Carlozzi NE, Schilling SG, J.-S. L, Paulsen JS, Hahn EA, Perlmutter JS, Ross CA, Downing NR, Kratz AL, McCormack MK, Nance MA, Quaid KA, Stout J, Gershon RC, Ready R, Miner JA, Barton SK, Perlman SL, Rao SM, Frank S, Shoulson I, Marin H, Geschwind MD, Dayalu P, Foroud T, Goodnight SM, Cella D. HDQLIFE: Development and assessment of health-related quality of life in Huntington disease (HD). *Quality of Life Research*. (Under Review).
21. Carlozzi NE, Downing NR, McCormack MK, Schilling SG, Perlmutter JS, Hahn EA, Lai J-S, Frank S, Quaid KA, Paulsen JS, Cella D, Goodnight SM, Miner JA, Nance MA. New measures to capture end of life concerns in Huntington disease: Meaning and Purpose and Concern with Death and Dying from HDQLIFE (a patient reported outcomes measurement system). *Quality of Life Research*. (Under Review).
22. Carlozzi NE, Downing NR, Schilling SG, J.-S. L, Goodnight SM, Miner JA, Frank S. The development of a new computer adaptive test to evaluate chorea in Huntington Disease: HDQLIFE Chorea. *Quality of Life Research*. (Under Review).
23. PROMIS® Instrument Development and Psychometric Evaluation Scientific Standards. http://www.nihpromis.org/Documents/PROMIS_Standards_050212.pdf
24. Hanauer DA, Mei Q, Law J, Khanna R, Zheng K. Supporting information retrieval from electronic health records: A report of University of Michigan's nine-year experience in developing and using the Electronic Medical Record Search Engine (EMERSE). *J Biomed Inform*. 2015; 55:290–300. [PubMed: 25979153]
25. Paulsen JS, Hayden M, Stout JC, Langbehn DR, Aylward E, Ross CA, Guttman M, Nance M, Kiebertz K, Oakes D, Shoulson I, Kayson E, Johnson S, Penziner E, Investigators, P.-H. Preparing for preventive clinical trials - The Predict-HD study. *Archives of Neurology*. 2006; 63(6):883–890. [PubMed: 16769871]
26. Carlozzi NE, Tulsy DS. Identification of health-related quality of life (HRQOL) issues relevant to individuals with Huntington disease. *J Health Psychol*. 2013; 18(2):212–225. [PubMed: 22427174]
27. Shoulson I, Fahn S. Huntington Disease - Clinical Care and Evaluation. *Neurology*. 1979; 29(1):1–3. [PubMed: 154626]
28. Muthén, LK.; Muthén, BO. *Mplus User's Guide*. Muthén & Muthén; Los Angeles, CA: 2011.
29. Cook KF, Kallen MA, Amtmann D. Having a fit: impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption. *Qual Life Res*. 2009; 18(4):447–460. [PubMed: 19294529]
30. Reise SP, Morizot J, Hays RD. The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Qual Life Res*. 2007; 16(Suppl 1):19–31. [PubMed: 17479357]
31. McDonald, RP. *Test theory: a unified treatment*. Lawrence Erlbaum Associates, Inc; Mahwah, NJ: 1999.
32. Samejima, F.; van der Liden, WJ.; Hambleton, R. The graded response model. In: van der Liden, WJ., editor. *Handbook of modern item response theory*. Springer; NY, NY: 1996. p. 85-100.
33. Cai, L.; Thissen, D.; du Toit, SHC. *IRTPRO for Windows* [Computer software]. Scientific Software International; Lincolnwood, IL: 2011.
34. Crane PK, Gibbons LE, Jolley L, van Belle G. Differential item functioning analysis with ordinal logistic regression techniques. *DIFdetect and difwithpar*. *Med Care*. 2006; 44(11 Suppl 3):S115–123. [PubMed: 17060818]
35. Choi SW, Gibbons LE, Crane PK. *lordif: An R Package for Detecting Differential Item Functioning Using Iterative Hybrid Ordinal Logistic Regression/Item Response Theory and Monte Carlo Simulations*. *J Stat Softw*. 2011; 39(8):1–30. [PubMed: 21572908]

36. Choi SW. Firestar: Computerized Adaptive Testing Simulation Program for Polytomous Item Response Theory Models. *Applied Psychological Measurement*. 2009; 33(8):644–645.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

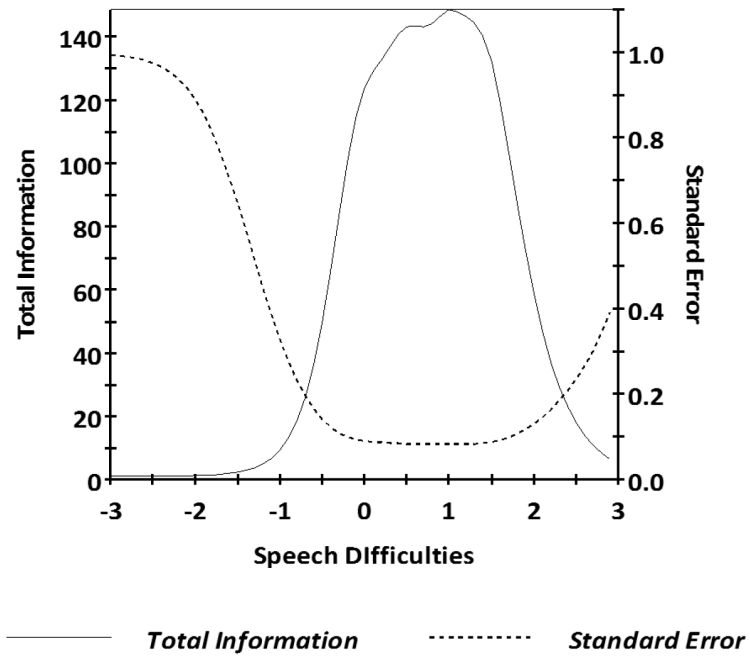


Figure 1.
 Speech Difficulties Test Information Plot
Caption: In general, we want total information to be > 9.0 and standard error to be < 0.33 (this provides a reliability of 0.9). This figure shows excellent total information and standard error for Speech Difficulties scale scores between -1.1 and $+2.7$.

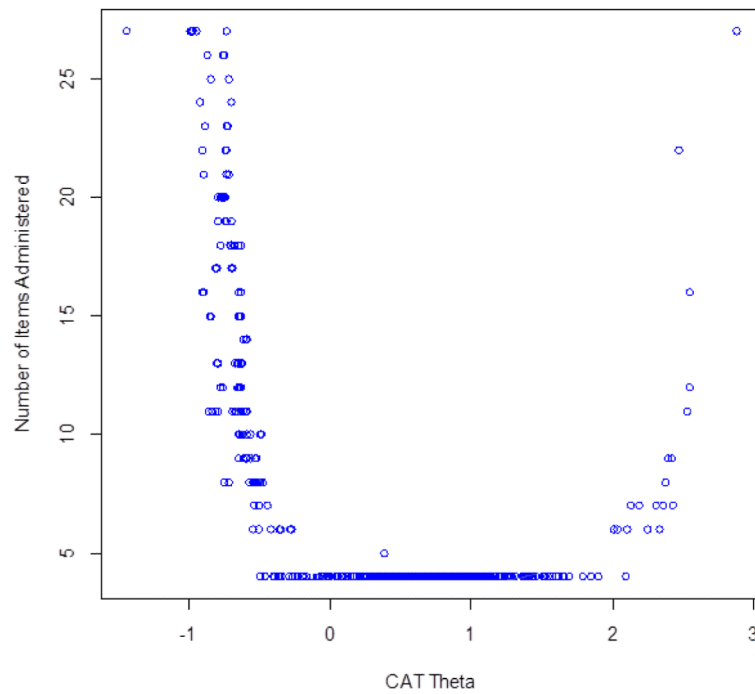


Figure 2.

Speech Difficulties Number of CAT Items by CAT Theta

Caption: This figure shows the number of CAT items used for different scale scores in standard deviation units (circles represent individual simulated observations): at -1 SD units, the CAT always used all 27 items in the item bank; from -0.1 to $+2.0$, the CAT always used the minimum number of 4 items in the item bank; and at 3 SD units the CAT used all 27 items in the item bank.

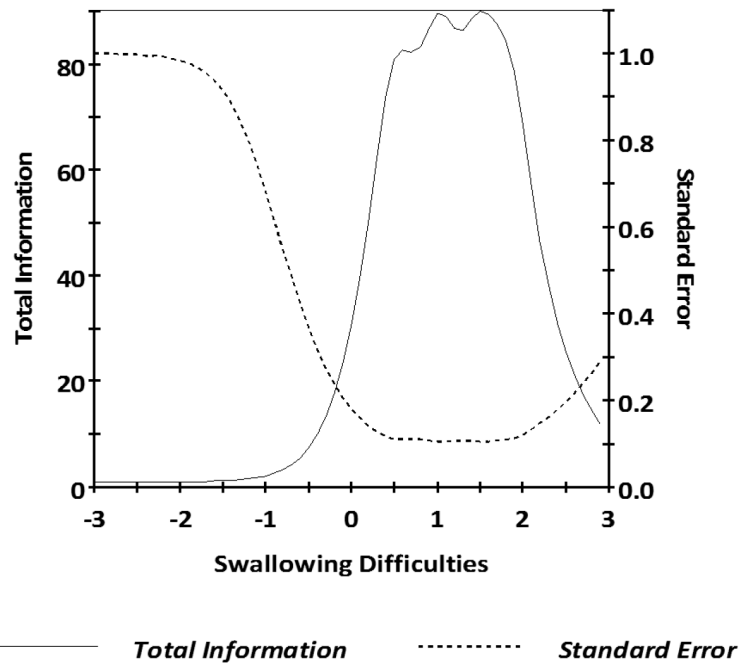


Figure 3.

Swallowing Difficulties Test Information Plot

Caption: In general, we want total information to be > 9.0 and standard error to be < 0.33 (this provides a reliability of 0.9). This figure shows excellent total information and standard error for Swallowing Difficulties scale scores between -0.3 and $+3.0$.

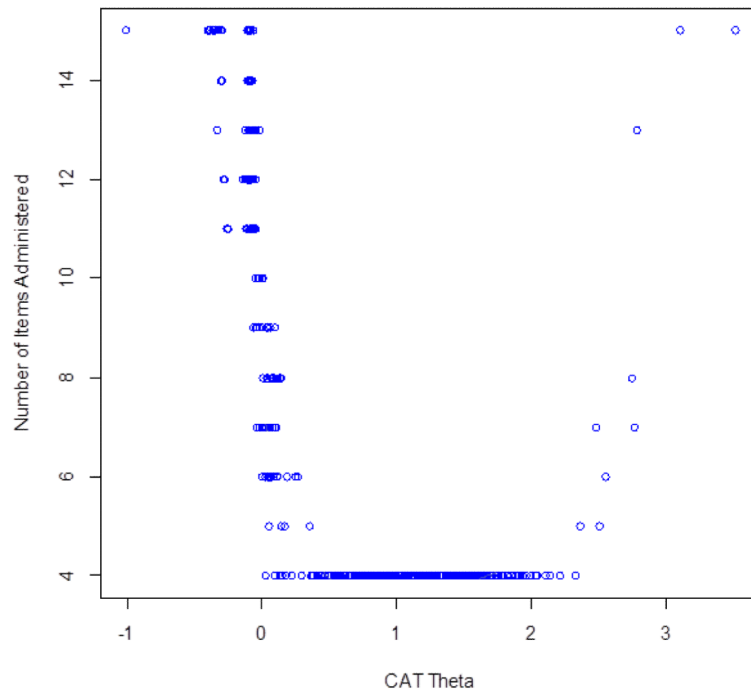


Figure 4.
 Swallowing Difficulties Number of CAT Items by CAT Theta
 Caption: This figure shows the number of CAT items used for different scale scores in standard deviation units (circles represent individual simulated observations): at -1 SD units, the CAT always used all 15 items in the item bank; from $+0.4$ to $+2.2$ SD units, the CAT always used the minimum number of 4 items in the item bank; and at 3 SD units the CAT used all 15 items in the item bank.

Table 1

Exploratory Factor Analysis Results For Speech and Swallowing Item Pool

Item Description	Factor 1	Factor 2	Factor 3	Factor 4
*It was difficult for other people to understand me.	0.89	0.12	-0.03	-0.03
*How often did you slur your words when you spoke?	0.94	0.06	-0.10	0.01
*It was difficult to speak clearly.	0.91	0.16	-0.05	-0.07
How much difficulty do you currently have...saying what you want to say?	0.82	0.15	-0.21	0.06
How much difficulty do you currently have...speaking clearly?	0.97	0.04	-0.11	-0.02
How much difficulty do you currently have...speaking?	0.82	0.16	-0.09	0.03
*How often did you have to speak slowly for other people to understand you?	0.88	0.08	-0.05	-0.04
*How often did you limit your social activities because you had difficulty speaking?	0.51	0.32	0.02	0.20
*How often did you mumble?	0.82	0.12	-0.07	-0.03
*How often did you worry about slurring your words when you spoke?	0.76	0.06	-0.08	0.12
*How often did your speech difficulties interfere with your ability to socialize with your family?	0.33	0.15	0.04	0.54
*How often did you speech difficulties impact your ability to enjoy life?	0.49	0.11	0.02	0.39
*How often did your speech difficulties impact your ability to socialize with your family?	0.35	0.09	-0.06	0.69
*How often did your speech difficulties impact your ability to socialize with your friends?	0.32	0.14	0.06	0.62
*How often did other people have difficulty understanding you?	0.79	0.05	-0.03	0.18
*How often were you bothered by slurring your words when you spoke?	0.63	0.01	0.07	0.34
*How often were you unable to maintain a conversation?	0.49	0.14	0.05	0.20
*How often did you have to make an effort to carry on a conversation?	0.62	0.10	-0.01	0.24
*How often were you bothered by your speech difficulties?	0.66	0.06	0.07	0.27
*My speech difficulties interfered with my ability to work (include work at home).	0.46	0.12	0.03	0.43
^1 had to limit my social activity because of my speech difficulties.	0.35	0.15	0.17	0.46
^1 had to speak slowly for other people to understand me.	0.72	0.03	0.10	0.15
^1 had trouble speaking.	0.73	0.04	0.11	0.17
^1 had trouble speaking clearly.	0.84	-0.12	0.25	0.06

Item Description	Factor 1	Factor 2	Factor 3	Factor 4
^I slurred my words when I spoke.	0.86	-0.11	0.23	0.03
^I was frustrated by my speech difficulties.	0.65	-0.15	0.20	0.35
^My speech difficulties made me feel self-conscious.	0.52	-0.11	0.19	0.45
*How often did you choke?	-0.01	0.99	-0.21	0.03
*I ate slowly to avoid choking.	0.19	0.50	0.41	-0.16
How much difficulty do you currently have...swallowing?	0.27	0.68	0.12	-0.17
How much difficulty do you currently have...chewing?	0.29	0.53	0.22	-0.09
How much difficulty do you currently have...eating?	0.37	0.56	0.15	-0.08
*I limited how much I ate because it was difficult to swallow.	0.11	0.65	0.18	0.11
*How often did you have trouble finishing your meal because of your difficulty swallowing?	-0.01	0.78	0.14	0.12
*How often did you worry about choking?	0.03	0.70	0.19	0.08
*How often did choking interfere with your ability to eat?	-0.01	0.95	-0.01	0.02
*How often did difficulty chewing interfere with your ability to eat?	0.15	0.57	0.20	0.18
*How often were you bothered by your choking?	0.02	0.97	-0.13	0.09
*How often were you unable to swallow?	0.08	0.60	0.16	0.14
*How often did you choke on average?	-0.13	0.97	-0.13	0.20
^I had to cut up my food into small pieces to avoid choking.	0.27	0.33	0.58	-0.20
^I had to eat pureed food to avoid choking.	-0.16	0.08	0.93	0.25
^I had to eat pureed food because of difficulty swallowing.	-0.09	0.10	0.85	0.25
^I had trouble chewing.	0.19	0.58	0.19	0.09
^I had trouble eating because I choked.	0.01	0.88	0.03	0.08
^I had to eat slowly to avoid choking.	0.21	0.44	0.57	-0.19
^I needed help eating.	0.17	0.46	0.34	0.14

* Note. = In the past 7 days...;

^ = During the past 7 days...

Table 2

Difficulty With Speech Item Parameters

Difficulty With Speech Item Description	Slope	T1	T2	T3	T4
It was difficult for other people to understand me.	5.48	-0.05	0.39	0.97	1.39
How often did you slur your words when you spoke?	5.05	-0.35	0.32	0.85	1.36
It was difficult to speak clearly.	5.69	-0.16	0.35	0.95	1.27
How much difficulty do you have saying what you want to say?	2.95	-0.41	0.48	1.12	2.32
How much difficulty do you have speaking clearly?	4.59	-0.23	0.53	1.03	1.93
How much difficulty do you have...speaking?	3.93	-0.05	0.66	1.22	2.42
How often did you have to speak slowly for people to understand you?	3.91	-0.12	0.35	1.00	1.49
How often did you limit your social activities because of difficulty speaking?	3.65	0.44	0.83	1.42	1.87
How often did you mumble?	2.91	-0.23	0.41	1.15	1.66
How often did you worry about slurring your words when you spoke?	3.37	-0.08	0.43	0.98	1.51
How often did speech difficulties interfere with your ability to socialize with family?	3.78	0.32	0.73	1.29	1.76
How often did speech difficulties impact your ability to enjoy life?	3.85	0.28	0.72	1.25	1.84
How often did speech difficulties impact your ability to socialize with family?	4.41	0.31	0.73	1.27	1.77
How often did speech difficulties impact your ability to socialize with friends?	4.77	0.25	0.61	1.23	1.69
How often did other people have difficulty understanding you?	5.01	-0.22	0.35	0.95	1.48
How often were you bothered by slurring your words when you spoke?	4.80	0.00	0.45	1.05	1.57
How often were you unable to maintain a conversation?	2.91	0.16	0.64	1.19	1.72
How often did you have to make an effort to carry on a conversation?	2.93	-0.10	0.40	1.01	1.51
How often were you bothered by your speech difficulties?	4.61	-0.08	0.41	0.98	1.45
My speech difficulties interfered with my ability to work (include work at home).	4.07	0.30	0.66	1.20	1.53
I had to limit my social activity because of my speech difficulties.	4.07	0.66	1.02	1.42	1.81
I had to speak slowly for other people to understand me.	4.40	-0.02	0.55	1.03	1.68
I had trouble speaking.	6.32	0.02	0.59	0.99	1.45
I had trouble speaking clearly.	5.39	-0.02	0.55	1.03	1.41
I slurred my words when I spoke.	5.40	-0.02	0.61	1.03	1.39
I was frustrated by my speech difficulties.	4.01	0.04	0.58	1.05	1.43
My speech difficulties made me feel self-conscious.	3.44	0.10	0.66	1.09	1.53

Note. Items that are bolded were included on the 6-item static short form.

Table 3Difficulty With Speech Short-Form Summed Score to *t* Score Conversion Table

Difficulty With Speech	
Summed Score	<i>t</i> Score
6	38
7	44
8	47
9	49
10	50
11	51
12	52
13	54
14	55
15	55
16	56
17	57
18	58
19	59
20	60
21	61
22	61
23	62
24	63
25	64
26	65
27	66
28	68
29	70
30	74

Note. This table provides the conversion for HDQLIFE Difficulty with Speech Short Form total raw scores (i.e., summed score) to standardized *t* scores (with a mean of 50 and SD 10; i.e., *t* score). Higher scores indicate greater speech difficulty. Scores ≥ 60 indicate that the individual is functioning more poorly than 68.27% of people with HD and scores ≥ 70 indicate that the individual is functioning more poorly than 95.45% of individuals with HD.

Table 4

Difficulty With Swallowing Item Parameters

Difficulty With Swallowing Item Description	Slope	T1	T2	T3	T4
How often did you choke?	3.43	0.11	0.77	1.64	2.52
How much difficulty do you currently have...swallowing?	3.74	0.13	0.97	1.69	
How much difficulty do you currently have...chewing?	3.19	0.63	1.19	1.87	3.00
How much difficulty do you currently have...eating?	3.86	0.56	1.12	1.62	2.60
I limited how much I ate because it was difficult to swallow.	4.26	0.79	1.06	1.57	2.11
How often did you have trouble finishing your meal?	4.23	0.76	1.16	1.75	1.97
How often did you worry about choking?	3.72	0.36	0.81	1.34	1.67
How often did choking interfere with your ability to eat?	7.11	0.50	1.02	1.52	1.91
How often did difficulty chewing interfere with your ability to eat?	4.52	0.56	1.02	1.52	1.83
How often were you bothered by your choking?	5.11	0.36	0.87	1.39	1.84
How often were you unable to swallow?	3.52	0.60	1.14	1.75	2.16
How often did you choke on average?	4.23	0.23	0.90	1.61	2.24
I had trouble chewing.	4.26	0.64	1.13	1.51	1.89
I had trouble eating because I choked.	6.85	0.50	1.04	1.44	1.77
I had to eat slowly to avoid choking.	3.71	0.37	0.84	1.16	1.51

Note. Items that are bolded were included on the 6-item static short form.

Table 5Difficulty With Swallowing Short Form Summed Score to *t* Score Conversion Table

Difficulty With Swallowing	
Summed Score	<i>t</i> Score
6	41
7	49
8	51
9	53
10	54
11	55
12	56
13	57
14	58
15	59
16	60
17	61
18	62
19	63
20	63
21	64
22	65
23	66
24	67
25	68
26	69
27	71
28	73
29	77

Note. This table provides the conversion for HDQLIFE Difficulty with Speech Short Form total raw scores (i.e., summed score) to standardized *t* scores (with a mean of 50 and SD 10; i.e., *t* score). Higher scores indicate greater swallowing difficulty. Scores ≥ 60 indicate that the individual is functioning more poorly than 68.27% of people with HD and scores ≥ 70 indicate that the individual is functioning more poorly than 95.45% of individuals with HD.