Citation: Swinburn, K., Best, W., Beeke, S., Cruice, M., Smith, L., Willis, E. P., Ledingham, K., Sweeney, J. & McVicker, S. J. (2018). A concise patient reported outcome measure for people with aphasia: the aphasia impact questionnaire 21. Aphasiology, 33(9), pp. 1035-1060. doi: 10.1080/02687038.2018.1517406

A concise Patient Reported Outcome Measure for people with aphasia: The Aphasia Impact Questionnaire 21

Swinburn K. [1], [2] Best W. [2], Beeke S. [2], Cruice M.[3], Smith L.[4], Pearce Willis E. [1], Ledingham K. [2], Sweeney J.[2], & McVicker S. J.[1, 3]

Acknowledgments

We would like to thank all those who were involved in data collection for the AIQ study; people with aphasia across the south of the UK, and all the administrators who worked so diligently to gather data.

[1] Formerly Connect, Kings Bench Street, Southwark, London.

[2] Language & Cognition, Division of Psychology & Language Sciences, University College London.

[3] Division of Language & Communication Science, City, University of London, Northampton Square, London

[4] Department of Psychology, Social Work & PH needs to be in full, Faculty of Health & Life Sciences, Oxford Brookes University.

Correspondence to Kate Swinburn, kateswinburnaiq@gmail.com

Abstract

Background: There are many validated and widely used assessments within aphasiology. Few, however, describe language and life with aphasia from the perspective of the person with aphasia. Across healthcare, patient experience and user involvement are increasingly acknowledged as fundamental to person-centred care. As part of this movement, Patient Reported Outcome Measures (PROMs) are being used in service evaluation and planning.

Aims: This paper reports the quantitative aspects of a mixed methods study that developed and validated a concise PROM, the Aphasia Impact Questionnaire (AIQ), co-produced with People with Aphasia (PWA).

Methods & Procedures: The AIQ was developed within the social model of disability and all stages of the development of the AIQ were performed in partnership with PWA. It was adapted from a pre-existing and lengthier PROM for PWA, the Communication Disability Profile. The first iterations of the AIQ focused on domains of communication, participation and well-being/emotional state. Subsequently the AIQ was extended to include additional items relating to reading and writing (AIQ-21). The research design was iterative. Initially concurrent validity, internal consistency, and sensitivity of the AIQ-prototype were obtained. The AIQ-prototype was modified to become the AIQ-21. Statistical testing with a new group of PWA was performed, investigating internal consistency and concurrent validity of the AIQ-21.

Outcomes & Results: Results for both the AIQ-prototype and AIQ-21 showed statistically significant concurrent validity and good internal consistency. Repeated measurement using the AIQ-prototype demonstrated statistically significant change after PWA accessed a community intervention.

Conclusions: The AIQ-21 is a PROM that has great potential to be one of the core set of aphasia tests for clinical and research use. Results can be used alongside language assessment to enable person-centred goal setting and partnership working for people with aphasia.

Key words

Aphasia; AIQ; reliability; assessment, co-production; patient-reported-outcome-measures;

A concise Patient Reported Outcome Measure for people with aphasia: The Aphasia Impact Questionnaire 21

Background and Aims

For decades, measuring change in aphasia has required an assessment that is “reliable enough to give consistent measures; sensitive enough to measure the improvement that the particular therapy involved is intended to produce; and valid so that it measures changes that are of real consequence in the patients’ lives” (Howard & Hatfield, 1987, p.113). There are many validated and widely used aphasia assessments. They vary in relation to the conceptual framework that underpins them, their intended clinical setting, and the manner in which information is gathered (Simmons-Mackie, Threats, & Kagan, 2005). Assessments of language impairment allow clinicians to explore language processing in People With Aphasia (PWA), upon which therapy plans can be built (Whitworth, Webster, & Howard, 2005). However, crucially, these assessments do not provide information about the consequence of a language impairment or its significance in a person’s life.

User involvement has been pertinent within aphasiology for many years. Whilst nomenclature describing this approach varies (client/person/patient-centred care, patient engagement, co-production), the thread remains the same: putting the user/client/patient/person first and foremost in whatever activity is being undertaken, be it research, goal setting, service evaluation, service planning or indeed the development of an assessment tool. Patient Reported Outcome Measurement has come out of this user involvement movement. Patient Reported Outcome Measures (PROMs) were introduced into the National Health Service (UK Government-funded medical and health care services) in 2009 (Devlin & Appleby, 2010), and “contribute to helping patients feel cared for, and the information provides a structured basis for patients' discussions with clinicians (p49)”.

The need for assessment and outcome measurement to be relevant and meaningful to the user in aphasiology is equally understood. Kagan (1995, p.182) identifies communication as the means through which others judge our social, intellectual and emotional competence “thus, when communication is impaired others might perceive the individual as generally incompetent. The effect on identity can be devastating”. These threats to deep-seated feelings of identity and emotional response to the world, are intrinsic to the experience of acquiring aphasia. Increasingly they are the focus of intervention, indeed Wallace et al., (2016) demonstrate that PWA prioritise regaining pre-morbid identity, confidence, not being defined by aphasia, and accepting their changed circumstances. The Royal College of Speech & Language Therapists’ clinical guidelines state “the process of assessment should encompass the perception of the individual … with regard to the impact of the communication disability on their lives” (RCSLT, 2005, p.104), and the Australian Aphasia Pathway calls for "outcome measures for PWA to be relevant, meaningful, and important to stakeholders" (Best Practice Statement 3.6, Power et al., 2015).

The disparity between therapeutic intervention and the outcomes that measure its effectiveness has been well documented. Many authors comment that though in recent years language-based interventions have broadened to include participation in life roles and relationship, and psychological well-being, there is often a mismatch between the treatment approaches and the outcome measures used to gauge their effectiveness (Wallace, Worrall, Rose & Le Dorze, 2017; Guo, Togher, & Power, 2014; Johansson, Carlsson, & Sonnander, 2011; Klippi, Sellman, Heikkinen, & Laine, 2012; Verna, Davidson & Rose, 2009). In addition, surveys report that clinicians more frequently use impairment-based measures than PROMs in clinical practice (Rose, Ferguson, Power, Togher, & Worrall, 2014; Simmons-Mackie, Threats, & Kagan, 2005; Verna et al, 2009), and thus, potentially under-assess aspects of aphasia function, activity, participation and quality of life (QOL). While there has been some shift of late towards more frequent use of QOL measures, an international survey of 400 speech and language therapists (SLTs) by Hilari et al (2015) reported that usage is still low for guiding or evaluating intervention (10% and 19% of SLTs used QOL measures for these purposes, respectively). Without a wider perspective on assessment that targets how people experience aphasia, there is a gap in evaluation. This threatens the practice of person-centred care, specifically negotiation of meaningful intervention, and support for SMARTER goal setting (Hersh, Worrall, Howe, Sherratt, & Davidson, 2012).

Crucially, it is also beneficial if PWA can express how they feel about acquiring and living with aphasia. Kleinman (1988) suggests that it is possible to talk to patients, even those who are most distressed, about the actual experience of illness, and that witnessing and helping them to order that experience, can be therapeutic. He further suggests that within chronic illness "neither the interpretation of illness meanings nor the handling of deeply felt emotions within intimate personal relationships can be dismissed as peripheral tasks. They constitute rather the point of medicine" (p253).

Within aphasiology, Cardol, de Haan, van den Bos, deJong & de Groot (1999) noted how a well-designed PROM (here, the VASES), tailored to the needs of PWA can encourage a therapeutic dialogue. They suggest that an accessible, acceptable tool "provides the basis for a conversation about feelings and expression of distress which would otherwise have been difficult to start" (Cardol et al., 1999, p.119). Swinburn (2003) proposes that the way in which this can be achieved relates to numerous factors including the domains of measurement and the way the tool establishes the context for measuring these domains. She suggests the constructs chosen should relate to the consequences and significance of aphasia for PWA; where possible the scene for a joint discussion or exploration should be set. Value must be given to the expert views of the PWA.

Aphasiologists have recognised this gap in aphasia assessment and addressed it in different ways via tools such as the Assessment for Living with Aphasia (ALA), (Kagan et al., 2013), the Communication Disability Profile (CDP) (Swinburn with Byng, 2006), and the Stroke and Aphasia Quality of Life Scale-39 (SAQOL), (Hilari, Byng, Lamping, & Smith, 2003). These tools enable accessible, subjective measurement of the consequence and significance that aphasia has on life participation for individual PWA. Though categorisation varies, the ALA and the SAQOL-39 are described by their authors as QOL tools, while the ALA and CDP overlap most in relation to content; with the majority of scored items focusing on communication, participation and emotional state. The SAQOL-39, in line with traditional QOL measures, includes additional items on physical function and energy, which may not necessarily be expected to be affected by functional communication change. However, all three are, necessarily, detailed and lengthy. Both the ALA and the SAOQL have been reduced in size to now include 37 and 39 items respectively, with the CDP being the longest at 48 rated items (7 unrated). Both the ALA and the CDP aim to be presented as a conversation which further adds to the length of administration. A perceived lack of clinical time has been reported as *the* major barrier to administering QOL assessments (Cruice, 2016), with Simmons-Mackie et al. (2005) finding that 53% of reported barriers to aphasia outcome assessment related to time constraints. In terms of services for PWA in general, Code & Petheram (2011) show many PWA are unlikely to receive more than 9 hours of clinical contact with an SLT in total. Clearly, this presents SLTs with a dilemma. There are outcome measures available that give a detailed, person-centred view of life with aphasia beyond impairment measures. But the very detail that makes these tools relevant to PWAs' lives can pose a significant time-related challenge to regular use in clinical practice. Maintaining the breadth of assessment must be balanced by the practicalities of what is actually possible (rather than what is ideal).

Additionally, it has been suggested that traditional speech and language assessment can jeopardize the full partnership of therapy (Jordan & Kaiser, 1996). According to these authors, if the process of assessment and measurement itself leaves the person with aphasia feeling less confident, competent, or powerful, then the aim of equal partnership is undermined from the start, and crucially, risks the person beginning their life with aphasia experiencing incompetence. Therapeutic practice should be established from the outset. The assessment process itself therefore must be as accessible and acceptable as possible to PWA, to maximise feelings of competence, confidence and strength.

In summary, there is a gap in aphasia assessment, and a space exists for a subjective outcome measure that addresses how PWA experience life with aphasia, is psychometrically robust, can be conducted within a therapeutic process and yet is concise enough to be practical within a time-pressured clinical environment. The Aphasia Impact Questionnaire (the AIQ) has been developed to fill this gap. The objective is to concisely and robustly explore communication, participation and well-being/emotional state from the perspective of PWA, who were integral to its development, being fully involved at every stage. The study described in this paper set out to test the psychometric properties of this new tool.

The aims of this paper are: to introduce the AIQ-21; to describe the involvement of PWA in this co-production; to describe the iterative development of the AIQ from prototype through first published version (AIQ-prototype) to final version (AIQ-21); and to report the psychometric testing of the AIQ and resultant quantitative data.

Development of the AIQ

Prior to this study, a prototype AIQ had been developed. This prototype has a long history, beginning as the Disability Questionnaire (DQ) within the Comprehensive Aphasia Test (CAT - Swinburn, Porter & Howard, 2004). The DQ is a 34-item tool designed to gain the perspective of the person with aphasia, to be used alongside the impairment sections of the CAT (language and cognitive screening). It was, however, developed without direct input from PWA. Therefore, an advisory group of PWA was recruited to rework the DQ into the Communication Disability Profile (the CDP, Swinburn with Byng, 2006). The CDP was derived from 26 in-depth qualitative interviews with PWA, group interviews with PWA and SLTs, and a co-production process with PWA over 18 months (for full details of the development process, see Swinburn with Byng, 2006). As a 56-item tool, the resulting PROM is considerably longer than the DQ with a broader conceptual base, underpinned by the social model of disability. It explores and rates communication, participation, barriers and facilitators, and emotional consequences of life with aphasia. See Chue, Rose, & Swinburn (2010) for pilot validation. The CDP has been used as an outcome measure in clinical research (Best, Greenwood, Grassly, Hickin, 2008) and by clinicians to measure quality of life of PWA (Hilari et al., 2015).

Although the CDP is in regular use in both research and clinical practice, in 2011 for the reasons mentioned in the preceding text, the CDP underwent an iterative process of modifications to shorten it. The resultant tool was the AIQ-prototype. The AIQ-prototype then underwent psychometric testing. After the first stage of testing, 2 of the 21 items were deleted due to redundancy and lack of responsiveness, and the AIQ-19 was made available to SLT users. Later, 2 items (reading and writing) were added to the AIQ-19 to create the AIQ-21, the current available version, which then underwent further psychometric testing. This paper reports the quantitative results of this 2-stage process[1].

---

[1] As part of the study, additional qualitative research addressed the views of PWA on AIQ acceptability, accessibility and comprehensiveness. These findings will be reported in a companion paper (in preparation).

The AIQ is a subjective, pictorial, self-report questionnaire. It is divided into 3 sections, each containing questions exploring domains of living with aphasia: communication; participation; and well-being/emotional state. Depending on the authors orientation, some commentators may describe it as a QOL measure. An administrator supports the person with aphasia to answer each item using a pictorial rating scale.

From its inception (as with its predecessors the DQ and the CDP), the focus of the AIQ has been on *measuring* the subjective experience of living with aphasia. The choice of a rating scale was, therefore, especially important. Rating scales are commonly used to investigate feelings, attitudes and behaviours (Bowling, 1997), with evidence to suggest that people with cognitive or communication disability benefit from the simplicity of a 5-point Likert-type scale (Cummins, 1997). Thus a 5-point scale was chosen. But there is choice in what the scales might compromise; numbers, words, symbols, pictures? The issue of 'self' within self-report tools was crucial to the design of the scale. One function of the AIQ is to support PWA to begin to explore their new identity of 'self-with-aphasia', and the representation of 'self' in the scale is therefore critical. Pictorial rating scales have a precedent within aphasiology e.g. Kagan uses them in the ALA as anchors, and in 2005, when co-producing the CDP, PWA chose a pictorial scale to support the concept of 'self'. This was, therefore, repeated in the AIQ scale, giving a 0-4 pictorial rating scale with written and numerical descriptors. Eight versions of the scale, that vary in relation to gender and race, are available (see figure 1). Before the AIQ items are presented, to make the interaction as personally relevant as possible for the person with aphasia, the respondent chooses which of the 8 scales they most closely identify with in relation to gender and ethnicity.

(FIGURE 1. Racial and gender options for AIQ scale ABOUT HERE)

Each item is presented on its own page. Each page has an identical header; a text box containing the words 'This week…', an identical footer; and the chosen scale. Between the header and the footer is a text box containing each item and a picture depicting/supporting the construct being explored e.g. 'How easy was it for you to talk to a stranger?' (see figure 2). The person with aphasia identifies the picture on the scale that best fits his/her response to the question.

(FIGURE 2. An example page of the AIQ ABOUT HERE)

The administrator has a conversational script which can be read to support the person with aphasia to understand each item, and to record the choice made. For each domain of the AIQ, the range of scores for each question is 0 to 4. The anchor point descriptors vary between domains and items. Communication items are 0 anchored on 'no problem' and 4 anchored on 'impossible'. Participation items include anchors of 'no problem' and 'impossible', but also 'nothing positive to do' to 'lots of positive things to do', and finally 'very good' and 'very bad'. In the well-being/emotional state domain, each anchor reflects the emotion being rated e.g. 'very frustrated' to 'not at all frustrated'. Pointing can be used as the response mode negating the need for a verbal response.

The numerical scores from each scale are entered onto a summary score sheet and summed to give domain totals and a AIQ total. All sections are rated, however there are additional (non-rated) questions at the end of the AIQ that give opportunity to address positive aspects of life with aphasia ("things you enjoy?").

Throughout the measure, the format of the questions and the design (including pictorial response options) is deliberately repetitive to reduce the comprehension and cognitive load of the assessment. This is one of many design features supported by the literature on accessibility for PWA e.g. use of short, high frequency, concrete words, within simple linguistic structures (both spoken and written), supported with pictures (Nickels & Howard, 1995). Large font size, san serif font, numbers as numbers not letters, blank space, simplified format, presenting one item per page, and using a large font have been shown to increase understanding for people with mild-moderate aphasia (Rose, Worrall, & McKenna, (2003); Rose, Worrall, Hickson, Hoffman (2012)). In addition, relevant line drawn pictures are used to support each construct. Graphics are considered helpful, however their use, especially line drawings, is not ubiquitously appreciated by PWA (Rose et al. 2012).
All the AIQ design features aim to maximise face validity and acceptability and thus make the assessment process as person centred as possible to optimise the chance of assessment being a positive 'therapeutic' encounter.

Co-production methodology
Typically, research contact with PWA is carried out during individual consultations and confined to ascertaining the views of PWA about their own condition. This traditional approach therefore engages PWA as research *participants.* Co-production methodology however, engages PWA as research *partners;* contact with PWA as research partners is typically protracted and requires them to reflect beyond their own immediate views and consider the perspective of others. This project engaged with PWA in both ways, at different stages across the lifetime of the research.

The involvement of PWA in the development of the AIQ
When developing the AIQ-prototype, six PWA based at Connect[2] were recruited to an AIQ Development Group, to oversee the refinement of the CDP. Explanation was given to the group about the need for a shorter tool, and some key psychometric characteristics that needed to be maintained to protect the validity of the emergent

---

[2] A community-based UK charity that supported PWA

tool (e.g. questions must be clear and time bounded, responses must be unambiguous and quantifiable). The advantageous nature of repetitive, pictorial presentation to aid comprehension was also discussed. The AIQ Development Group was presented with all 56 CDP items. A series of meetings took place to choose the most pertinent items, and further refine the content and style of the emergent AIQ. This was then field tested with PWA beyond the AIQ Development Group, to elicit feedback. Suggested modifications were incorporated to produce the AIQ-prototype, which was then validated.

(Table 1. Type of involvement of PWA in development of the AIQ-21 ABOUT HERE)

All items for the AIQ were chosen by PWA. Efforts were made to ensure that constructs identified as significant for PWA were present in AIQ items. Unsurprisingly, many of the constructs concur with previous exploration into the impact of aphasia on people's lives (Parr, Byng, Gilpin, & Ireland, (1997); Le Dorze & Brassard 1995), and overlap with many items contained in other aphasia specific tools such as the BOSS (Doyle, Mcneil, Hula, & Mikolic, (2003), the COAST (Long, Hesketh, Paszek, Booth & Bowen (2008), and the ALA (Kagan et al., (2013). The overlap is not exact, however, and significant discussion took place when selecting items. For example, there was considerable thought and attention given to which emotions were selected for the AIQ, and the balance between positive and negative emotions. Whilst aware that the AIQ risked being a catalogue of negativity, the AIQ development group selecting the items (from the original CDP items) felt that if there was a limit on the number of items, those items that most frequently and best described their lives with aphasia should be included. For this reason, negative emotions were chosen over positive ones by this group. The group stated that if PWA were asked questions implying a positive emotional state, that this could imply that the administrator did not understand the negative consequences of aphasia. The inclusion of 'things you enjoy' was intended to give respondents the opportunity to reflect on areas of participation that are positive. It is however, acknowledged, it does not address issues relating to positive self-image and positive self-esteem. This view was not unanimous for PWA who were consulted. Though some felt the AIQ was negative, finding the items patronising, many felt the pictorial presentation was "friendly" and "positive", and suggested it was cathartic to be able to express themselves in such depth despite having very limited language output. This topic will be explored in more detail in a forthcoming paper that describes the qualitative aspects of the AIQ-21 project.

Though PWA were the guiding influence for the development of the tool, the AIQ needed to be clinically relevant. SLT administrators involved in this project gave feedback in relation to the item selection suggested by the AIQ Development Group. There were times when the preferences of PWA and SLTs diverged. One notable example was the inclusion of the item 'stupid'. Administrators (across a range of sites) were uncomfortable asking this question. This divergence was addressed multiple times with different groups of PWA. On each occasion PWA unanimously felt it was important that this item remained, as for some PWA feeling stupid affected their confidence in interactions with others and their feelings of self-worth. They felt giving respondent PWA the opportunity to express this emotion, outweighed the administrators discomfort at asking the question. Another divergence was that of the domains of ‘talking’ versus ‘expressing yourself.’ SLTs felt it was important to include items that represented non-verbal modes of expression, alongside ‘talking’. However, PWA felt there were more important priorities, notably emotions, so the domain ‘expressing yourself’ was excluded, despite its potential clinical relevance.

The resultant AIQ is a concise, condition specific PROM for PWA, for use once PWA have left the hospital setting and have experienced life with aphasia in their own environment. It enables a person with aphasia to share their perspective, revealing their priorities and views on their current situation, and thereby providing information that enables subsequent person-centred assessment, goal setting and treatment planning. It is designed to complement the impairment-focused perspective of assessment and thus intervention. Its aim is to ensure that intervention is always based on a discussion about an individual’s perception of life with aphasia, not their abilities or disabilities. The AIQ enables documenting and exploring communication, participation and well-being/emotional state when living with aphasia. It enables an individual to personally rate the relative impact of these areas, as a basis and foundation for the clinical discussions mentioned by Cruice (2008), on which to plan interventions that address broader goals in parallel with impairment focused goals. Equally, the AIQ supports beginning to explore a new sense of self for the person with aphasia.

Methods, procedures and results

We will now outline the methods, procedures and results for evaluating the psychometric properties of the AIQ, which took place in two stages; stage 1 - testing the AIQ-prototype and then stage 2 - testing the further refined AIQ-21.

Stage 1: Testing the AIQ: AIQ-prototype validation

The AIQ-prototype underwent validation to establish:

A) concurrent validity,
B) internal consistency and
C) responsiveness.

A) AIQ-prototype Concurrent validity

<u>Procedure </u>*(concurrent validity)*

The AIQ-prototype was administered alongside selected comparable domains of the Burden of Stroke Scale (the BOSS, Doyle et al. 2003). The BOSS domains selected covered self-report on communication (BOSS Communication Difficulty BOSS-CD), psychological distress (Communication-Associated Psychological Distress CAPD), social relations, positive and negative emotional states and resultant functional limitations. The correlation between mean scores of the BOSS and the AIQ-prototype were calculated.

Recruitment and participants (concurrent validity)

Participants for this part of the study were recruited from the community of people who attended Connect's London centre, over 3 months in 2011 (see table 2). Eligible participants were at least eighteen years old and had: a diagnosis of aphasia; hearing, vision and general health sufficient to participate; and sufficient English to complete the assessments. Participants recruited to explore concurrent validity had to have stable aphasia (more than 2 years post onset). Purposive sampling ensured a range of demographic characteristics were represented: gender, age, ethnicity, and aphasia severity. Aphasia severity was rated using the Boston Diagnostic Aphasia Examination (BDAE) Severity Rating Scale (Goodglass, Kaplan & Brand, 1983). Informed consent was obtained from all participants through use of an accessible information sheet and consent form.

(Table 2. AIQ-prototype Participant Characteristics ABOUT HERE)

Data collection (concurrent validity)

The AIQ-prototype and selected domains of the BOSS were administered individually with 31 PWA at Connect's London offices. The same administrator (a Connect staff member author EPW) carried out all assessments and used a conversational script. Communication support was provided to all participants throughout. Sessions lasted approximately 30-40 minutes and included a break when necessary. To control for order effects, the presentation of the AIQ-prototype and the BOSS was alternated between participants.

Data analysis (concurrent validity)

Analysis of AIQ-prototype data was carried out by an independent academic statistician (author LS) using SPSS (Version 21, IBM, 2012). Concurrent validity of each domain of the AIQ-prototype was compared with equivalent domains of the BOSS using Spearman's rank correlation coefficient. Items from the well-being/emotional state domain were separated into positive and negative emotions. This enabled comparison with the 2 emotional domains of the BOSS. Items for the AIQ-prototype, (and the AIQ-21, EQ-5D and the BOSS) are rated so that a high score indicates a negative health state (e.g. AIQ-prototype 4 = major problem, 0 = no problem). To complete the analysis, data were initially normalised by log-transformation.

Results (concurrent validity)

There were significant positive correlations between scores for the AIQ-prototype and BOSS in all domains (See figure 3). A positive correlation indicates higher AIQ scores (worse health state) association with higher scores (worse state) on BOSS. Moderate association was found for *communication* (rho 0.583, p<.001, figure 3a) and for *participation* (rho 0.551, p<.001, figure 3b). Very close association was found for *well-being/emotional state* (*negative* emotions) (rho 0.903, p<.001, figure 3c).

Moderate association was found for *well-being/emotional state* (*positive* emotions) (rho 0.725, p<.001, figure 3d).

(FIGURE 3 a, b, c, d ABOUT HERE)

B) AIQ-prototype Internal Consistency

Procedure (internal consistency)

The AIQ-prototype was administered as part of PWA's first 2 sessions with Connect (T1). The internal consistency was obtained by establishing Cronbach's alpha for each domain of the AIQ-prototype.

Recruitment and participants (internal consistency)

Recruitment took place in a variety of rural and urban locations across London, Cornwall, Gloucestershire and Sussex (See table 3). All PWA referred to Connect over an 18-month period, between 2010-2012, were recruited if they consented, and met the inclusion criteria above (n=137). Presence of stable aphasia was not a requirement (other than for concurrent validity) so time post stroke varied (see table 2). A further 76 PWA were referred, but either did not complete the AIQ-prototype (so were not recruited) or were lost to follow up.

(Table 3. Participant numbers across locations. ABOUT HERE)

Data collection (internal consistency)

Participants were seen in a variety of locations including their own homes, group venues (often community halls), care homes or residential homes, and at Connect's London offices. Communication support was provided to all participants throughout. Administrators were either Connect staff, SLTs or student SLTs, all of whom were trained to maximise communication access. London administrators and the Cornish administrator were trained in AIQ administration by the lead author (KS) and had conversational scripts to follow. Administrators in other locations were given identical conversational scripts, supplemented by written guidance on how to administer the AIQ-prototype and telephone follow up with KS. It is notable that, of over 200 PWA assessed, only 4 were excluded due to the test being inaccessible to them given the severity of their aphasia. (see table 4).

(Table 4. Reasons for dropping out of the responsiveness study ABOUT HERE)

Data analysis (internal consistency)

Cronbach's alpha, was used in accordance with the recommendations of Helms, Henze, Sass, & Mifsud, (2006). According to Kline (1999) and DeVellis (1991), an acceptable range of values are 0.7 to 0.9. A Cronbach's alpha was obtained for each of the 3 AIQ domains: communication; participation; and well-being/emotional state (combined positively and negatively worded emotions).

Results (internal consistency)

Cronbach's alpha fell within the accepted range of 0.7 – 0.9 for the *communication* (α = 0.79) and *participation* domains of the AIQ-prototype (α = 0.77). For *well-being/emotional state*, alpha fell slightly outside of this range (α = 0.92), suggesting that one or more items may be redundant. Recalculation of this domain with each item deleted identified Question 19 'determined' as redundant and so it was deleted (see table 5). For full details see tables 6, 7, 8 in Appendix 1.

(Table 5. AIQ-prototype domain 3 with each question removed in turn ABOUT HERE)

C) AIQ-prototype (and EQ-5D) responsiveness

Procedure (responsiveness)

Sensitivity of the AIQ-prototype to change was tested using repeated measures data, taking a subset of PWA who performed the AIQ-prototype at the beginning of a community support programme (T1) and 6 months after this intervention (T2). The community support programmes varied widely. All sessions were group based, each session lasting approximately 2 hours. PWA accessed between 1 and 5 sessions per week. The groups were predominately conversation groups, often facilitated by PWA trained as peer group facilitators. These data were tested by comparing changes in AIQ-prototype domain scores at T1 and T2 using paired samples t-tests (comparing the mean scores for each participant on each of the 3 domains) which was then used to establish the statistical significance of these differences for each domain.

This group of participants also completed the EQ-5D (EuroQol group 1990), at both T1 and T2, as a comparator tool to compare sensitivity to change. The EQ-5D was chosen because it is the most widely used PROM in the UK and Europe (Brooks, 1996), and has previously shown change for PWA (van der Gaag, Smith, Davies, Moss & Mowles, 2005). The EQ-5D was modified to make it maximally accessible to PWA during this project; simple pictures were added, key words emboldened. No changes were made to the EQ-5D wording or the administration process.

Recruitment and participants (responsiveness)

All PWA who had been recruited at TI for Internal Consistency and were still in contact with Connect 6 months later in any location, were reassessed at T2 on both the AIQ-prototype and the EQ-5D (n=90).

Data collection (responsiveness)

Of the total recruited at T1 (n=137), a significant number (n=47) were recruited too late to undergo a reassessment appointment so completed the AIQ on one occasion only, and EQ-5D data was missing for 6 participants. Therefore, AIQ-prototype data from 90 PWA, and EQ-5D for 84 PWA still in contact with Connect at T2 were analysed. In the majority of cases, reassessment was completed by the same administrator as at T1.

Data analysis (responsiveness)

Responsiveness of the *AIQ-prototype* was calculated using scores obtained from all 3 domains of the AIQ-prototype (combined positive and negative emotions) and the EQ-5D at both T1 and T2. Changes in AIQ-prototype and the EQ-5D scores were compared using a paired samples t-test.

Results (responsiveness)

The changes in domain score on the AIQ-prototype between T1 and T2 are shown shown in table 9. All domains showed statistically significant change between the 2 measurement points, with the T2 scores significantly lower than at T1, indicating improved communication, participation and well-being/emotional state.

(Table 9. AIQ-prototype sensitivity to change ABOUT HERE)

There was no significant difference between the overall EQ-5D score between T1 and T2 (p values ranging from 0.50-0.86). Details can be seen in table 10. These results indicate that the EQ-5D did not demonstrate any statistically significant change in quality of life after the community intervention.

(Table 10. EQ5D sensitivity to change ABOUT HERE)

During this validation process, comments of the administrators and PWA were shared with the AIQ Development Group for further debate and refinement of the tool. This refined tool was named the AIQ-19. Because it had been developed as an outcome measure specifically to assess the impact of Connect support (which did not target reading and writing), reading and writing concepts were deliberately excluded from the AIQ-19. However, when the AIQ-19 was launched and used beyond Connect, professionals who *did* target reading and writing advocated for the addition of reading and writing items. Therefore, a group of 5 Connect service users formed an Item Selection Panel, selecting 2 additional items (one each for reading and writing) to add to the AIQ-19. This panel was shown options, all taken from the reading and writing sections of the CDP (Swinburn with Byng, 2006). They chose between 4 pictographic options for both reading and writing e.g. choose between 'read one word only', 'read a headline', 'read a story in a newspaper', 'read an official letter'. They discussed each option, with the final choice being agreed collectively and incorporated into the AIQ-19 to become the AIQ-21.

Stage 2: Testing the AIQ: AIQ-21 validation

The AIQ-21 underwent validation to establish:

A) concurrent validity, and
B) internal consistency.

The same participants were recruited for both tests, so procedure, recruitment, and data collection will be described jointly, with data analysis and results described separately for concurrent validity and internal consistency.

AIQ-21 procedure

Each participant completed the AIQ-21 alongside the same selected domains of the BOSS used for validity testing of the AIQ-prototype i.e. communication, psychological distress, and social activity. Evaluation of the concurrent validity and internal consistency of the AIQ-21 then took place.

AIQ-21 recruitment and participants

Twenty PWA were recruited to complete psychometric testing of the AIQ-21 (none had been involved in AIQ-prototype testing). Recruitment criteria and informed consent were identical to concurrent validity testing of the AIQ-prototype, though all participants were from London. Participant characteristics are shown in table 11.

(Table 11. AIQ-21 Participant characteristics ABOUT HERE)

AIQ-21 data collection

All AIQ-21 data collection took place at Connect’s London offices between August 2014 and February 2015. AIQ-21 assessments were conducted by 2 student SLTs as part of their master’s dissertations (authors KL, JS), undertaken whilst completing the final year of pre-registration training at University College London (UCL) under the supervision of authors SB, WB, KS.   Participants completed the AIQ-21 and selected domains of the BOSS within a single session in a private room. Presentation of the measures was alternated to control for order effects. Communication support was provided throughout.

AIQ-21 data analysis

A) AIQ-21 Concurrent Validity

AIQ-21 data was analysed by authors KL, JS. Concurrent validity of each domain of the AIQ- 21 was compared with equivalent domains of the BOSS using Spearman's rank correlation coefficient. Items from the well-being/emotional state domain were separated into positive and negative emotions. This enabled comparison with the 2 emotional domains of the BOSS. AIQ-21 and BOSS scores for all participants were analysed using SPSS (Version 22, IBM, 2013).

B) AIQ-21 Internal Consistency

As with the AIQ-prototype, internal consistency was analysed for each domain of the AIQ-21 using Cronbach's alpha, in accordance with the recommendations of Helms et al., (2006). Cronbach's alpha was obtained for each of the 3 AIQ domains: communication; participation; and well-being/emotional state (combined positively and negatively worded emotions).

AIQ-21 results

A) AIQ-21 Concurrent Validity

The *communication* and *well-being/emotional state* domains of the AIQ-21 showed a strong positive relationship with equivalent domains of the BOSS (communication, positive emotions and negative emotions. The *participation* domain showed a weak positive correlation with the social relationship domain of the BOSS. These relationships are shown through scatter graphs in figure 4 (figures 4 a, b, c, d) with detail of the strength of correlations appearing below.

(FIGURE 4. Scatter graphs showing relationships between selected domains of BOSS and AIQ-21 ABOUT HERE)

Communication domain

Figure 4a shows significant positive correlation (rho = 0.81, p <0.001) between AIQ communication domain score and BOSS communication domain score, indicating higher scores (worse state) on AIQ related to higher scores (worse state) on the BOSS.

Participation domain

Figure 4b shows a positive correlation (rho = 0.36, p = 0.12, not significant) between the AIQ participation domain score (transformed) and BOSS social relationships domain score, indicating higher scores on AIQ (worse state) agreed with higher scores (worse state) on BOSS. Two slight outliers were identified from graph 6b. Qualitative data for these participants was evaluated and no reason for the identified outliers was found.

Negative emotions domain

Figure 4c shows a positive correlation (rho = 0.69, p = 0.001) between the AIQ negative emotional state domain score and BOSS negative emotions domain score, indicating higher scores on AIQ (worse state) were associated with higher scores (worse state) on the BOSS.

Positive emotions domain

Figure 4d shows a significant positive relationship (rho = 0.78, p = <0.001) between the AIQ positive emotional state domain score and the BOSS positive emotions domain score also.

B) AIQ-21 Internal Consistency

Internal consistency (Cronbach's alpha) was acceptable for all domains of the AIQ-21 except *participation*, where alpha fell slightly below the accepted range at 0.65. For further details see tables 12, 13, 14 in Appendix 1

Communication Domain

Cronbach's alpha reliability coefficient of this sample for the communication domain was 0.787. Participants' scores for most items in this domain correlated with total communication scores at r = 0.483 or above, including the new reading item, Question 6 ("… read a whole story in the newspaper?"). Scores for the new writing item, Question 5 ("… write a letter?") had a weaker correlation with total domain scores at r = 0.282. The figure for alpha, if this item was deleted, is very marginally higher than the alpha obtained here (α = 0.789), suggesting that internal consistency of this domain would not be substantially increased by removal of Question 5 and thus it was retained (see table 12 in Appendix 1).

Participation Domain

The obtained alpha reliability coefficient of this sample for the participation domain was 0.647. This falls slightly below the accepted range (see table 13 in Appendix 1). Scores for all items in this domain showed weaker correlations with total participation scores than in other domains, except for Question 9 ("How were things with friends?"). This correlated strongly (rho = 0.707) with participation domain-total scores. The weakest correlation with total participation scores was found in Question 10 "How were things with your family?" (rho = 0.223). Subsequent Cronbach's alpha (α = 0.686) shows that internal consistency for this sample would increase towards the accepted range of 0.7 to 0.9 if this question were deleted. However, the content of this item (especially in comparison with "How are things with friends") was felt to be too important to be deleted despite the potential gains in the psychometric properties.

Well-being/emotional state domain

The obtained alpha reliability coefficient of this sample for the well-being/emotional state domain was 0.894, falling within the accepted range (see table 14 in Appendix 1). Stronger correlations were found in this domain between participants' scores for each item and domain-total scores than in other domains. Question 20 "Have you felt confident?" showed the weakest correlation with total domain scores (r = 0.370) but the alpha coefficient if this item was deleted was not larger than the coefficient currently obtained.

Discussion

This project developed, modified and tested two iterations of a new PROM for PWA. Significant time and attention was given to the design of the AIQ through co-production methodology. The psychometric qualities of both the AIQ-prototype and the AIQ-21 were acceptable; the AIQ-prototype showed moderate or very close association with the gold standard BOSS (Doyle et al 2003) acceptable internal consistency and was sensitive to change when the EQ-5D (EuroQol Group (1990) was not. The AIQ-21 similarly had acceptable concurrent validity and internal consistency except for the participation domain.

Psychometric qualities

The AIQ's psychometric results compare favourably with other similar tools. Reliability (as measured by internal consistency) is good (Cronbach's α of 0.79 Communication domain, 0.65 Participation domain & 0.89 Well-being/emotional state), and comparable to both the ALA (Cronbach's α ranging from 0.59 to 0.89 – Language 0.59, Participation 0.86, Environment 0.72, Personal 0.89, see Kagan et al 2013), and the SAQOL-39 (Cronbach's α for subdomains ranging from 0.74 to 0.94, (Hilari, Byng, Lamping, & Smith, 2003). Validity (as measured by comparison with the BOSS) was statistically significant across all but one domain (rho = 0.81, rho = 0.78, rho = 0.69 and rho = 0.36). Again, these scores are comparable to similar tools (ALA r = 0.69 and SAQOL-39 rho = 0.38 to 0.58). In addition, the AIQ-prototype, when administered over time with people who have aphasia following community intervention, has been shown to be sensitive to change (with statistically significant change obtained in all 3 domains; (communication t = 3.38 p = 0.01, participation t = 2.74, p = 0.07, Well-being/emotional state t = 5.66, p < 0.001). The mean change across the three domains was therefore 8.6 across 21 items (each rated 0-4). Though the reason for change cannot be attributed, these results do demonstrate the AIQ's sensitivity to change.

It is worth noting that the AIQ-prototype was more responsive to change than the EQ-5D. This could indicate that there is no change following Connect's support, or that the EQ-5D is not sensitive enough to demonstrate change in people with aphasia. Given the statistically significant change identified in all domains by the AIQ, the first explanation seems unlikely. Equally the EQ-5D *had* demonstrated statistically significant change previously with PWA in a similar context (van der Gaag et al., 2005). It is not entirely clear why the EQ-5D was not sensitive to change in this sample, but much more of the content of the AIQ was related to the intervention provided (in contrast to the generic QOL, the EQ-5D), and thus the AIQ has clearly demonstrated sensitivity to change in this sample.

Design features

Test development is not straightforward, and the development of the AIQ has been protracted and iterative. But at the heart of this process, throughout, was a central tenet: to produce a robust tool (and methodology) that enables a sensitive exploration of the *lived experience* of aphasia, upon which to build a positive therapeutic relationship. The AIQ-21 is accessible and, because of its design and the attention paid to this, people with very little access to language can report and rate these experiences. Only four PWA were excluded from this study because their language disability was too severe.

Co-production

Values-based methodology was integral to the development and design of all iterations of the AIQ. This co-production methodology drew strongly on the social model of disability described by Byng & Duchan (2005), Sarno (2004) and Simmons-Mackie (2000, 2001). It has long been known that the views of PWA and of service providers can differ. Oxenham, Sheard, & Adams (1995) showed that SLTs and spouses differed in their perceptions of disability and handicap. And in 2017, Wallace, Worrall, Rose & Le Dorze showed that PWA, aphasia clinicians and managers have differing views of important outcomes, as do PWA and family members. It is crucial to capitalise on the perspective of PWA to ensure that outcome measurements reflect their view and priorities. Tools such as the AIQ-21 highlight for the aphasia clinician what intervention and outcomes they should focus on.

Limitations of the study and suggestions for further investigation

The most significant limitation of the study is the different samples used for testing across different elements of the psychometric evaluation. This was inevitable, given data gathering across two distinct time periods, but it represents complexity when analysing and reporting the results.

There is a discrepancy between the two concurrent validity results. All domains for both AIQ-prototype and AIQ-21 showed statistically significant correlations with the BOSS, apart from the *participation* domain of the AIQ-21. This is curious as there is no difference between the participation sections of the AIQ-prototype and AIQ-21, and there *was* concurrent validity demonstrated between the participation section of the BOSS and the AIQ-prototype. It could be that the difference is accounted for by the larger sample size used when testing the AIQ-prototype (n=31), which did show concurrent validity, than when testing the AIQ-21 (n=20), which did not. Another factor complicating this analysis, could be variation in the manner of enquiry. The BOSS participation explores ‘interaction with people you’re meeting for the first time’, ‘enjoying leisure activities with friends or relatives’, ‘keep old friendships going’, and ‘maintaining your role as a friend or family member’. The AIQ, in contrast, explores how easy it is to ‘do things you have to do’, whether ‘you have enough positive things you do’, ‘how things are with friends’, and ‘how things are with family’. Although there is overlap, these are not the same constructs. Participation is a broad and difficult construct to capture, especially when there are constraints on the time taken to do so. But, as Chapey et al. (2000) suggests, “this does not mean that treatment comprises only life resumption processes, but rather that enhanced participation in life “governs” management from its inception” (Core value #4, in Life Participation Approach to Aphasia). And Wallace et al. (2016; 2017) demonstrate that for all parties (PWA, families, clinicians and managers) participation, i.e. ‘the person with aphasia being able to participate in different roles and contexts’, is one of the top priorities for treatment. It is encouraging that the correlation seen between the AIQ-prototype and BOSS is statistically significant, but it is thought this area is one for further investigation, and certainly will require greater exploration when developing therapy goals.

Finally, there is insufficient information on the nature of the community support given, so no concrete conclusions can be made relating to the nature of change. However, it is still noteworthy that the AIQ could demonstrate significant change following this support, where the more widely used EQ-5D could not.

Clinical implications

The AIQ-21 has demonstrated promising psychometric properties. It is responsive to areas PWA identify as key to the quality of their lives. It is suggested that the AIQ-21 can be used for information gathering, goal setting and outcome measurement. The rationale for reducing the CDP to the AIQ was to save clinical time whilst maintaining clinical utility. The AIQ-21 consists of 21 rated items compared with 37 and 39 respectively for the CDP and the ALA. Though administration time frames are not always reported in descriptions of assessments, experience suggests that the AIQ can be administered in approximately 25-30 minutes, which compares favourably to the CDP (approximately 1 hour) and the ALA (45 minutes). The AIQ-21 offers clinicians a concise means of appraising language disability from the perspective of the person with aphasia, validating their subjective lived, experience, though with caution, at this stage, around participation. The AIQ-21 is a PROM that has great potential to be one of the core set of aphasia tests for clinical and research use. Results can be used alongside language assessment to enable person-centred goal setting and partnership working for people with aphasia. As such it represents a useful addition when assessing how successfully someone is living with aphasia and contributes to the empowering of the person with aphasia, through acknowledgement, ordering, and elucidation of the impact of aphasia on that person’s life.

7,911 words including abstract

References

Australian Aphasia Pathway http://www.aphasiapathway.com.au/

Best W., Greenwood A., Grassly J., Hickin J. (2008) *International Journal of Language and Communication Disorders* 43, 4, 390-407

Bowling A. (Research Methods in Health: Investigating Health and Health Services. *Open University Press.* Buckingham. Philadelphia.

Brooks R. (1996) EuroQol: the current state of play. *Health Policy.* Jul; 37(1) p53-72.

Byng, S., & Duchan, J. (2005). Social model philosophies and principles: Their applications to therapies for aphasia. *Aphasiology*, 19, 906–922.

Cardol M., de Haan R. J., van den Bos P., deJong B. A. & de Groot I. J. M. (1999) The development of a handicap questionnaire: the impact of participation and autonomy. *Clinical Rehabilitation* vol. 13, pp411-419

Chapey R., Duchan J. F., Eliman R.J., Garcia L.J., Kagan A., Lyon J.G., and Simmons Mackie N. (2000) Life Participation Approach to Aphasia: A Statement of Values for the Future *The ASHA Leader,* February 2000, Vol. 5, 4-6. doi:10.1044/leader. FTR.05032000.4

Chue, W. L., Rose, M. L., & Swinburn, K. (2010). The reliability of the Communication Disability Profile: A patient-reported outcome measure for aphasia. *Aphasiology*, *24*(6–8), 940–956. doi.org/10.1080/02687030903490541

Code & Petheram (2011) Delivering for aphasia *International Journal of Speech-Language Pathology* 13, (1), 3-10 Published online: 15 Feb doi.org/10.3109/17549507.2010.520090

Cruice, M. (2008). The contribution and impact of the International Classification of Functioning, Disability and Health on quality of life in communication disorders. *International Journal of Speech---Language Pathology*, *10,* (1-2), 38–49. doi:10.1080/17549500701790520

Cruice, M. (2016). *Making a meaningful difference in aphasia rehabilitation*. Paper presented at the 17th International Aphasia Rehabilitation Conference, 14-16 December 2016, London.

Cummins R. A. (1997) Self-rated quality of life scales for people with an intellectual disability: a review *Journal of Applied Research in Intellectual Disability* vol. 10, no 3 pp199-216

DeVellis, R. F. (1991). *Scale development: Theory and applications.* Newbury Park, CA: Sage.

Devlin J., Appleby J. (2010) Getting the most out of proms: Putting health outcomes at the heart of NHS decision-making. *The King’s Fund UK*

Doyle, P., Mcneil, M., Hula, W., & Mikolic, J. (2003). The Burden of Stroke Scale (BOSS): Validating patient-reported communication difficulty and associated psychological distress in stroke survivors. *Aphasiology, 17* (3), 291–304. doi.org/10.1080/729255459

EuroQol Group (1990) EuroQol – a new facility for the measurement of health-related quality of life. *Health Policy,* 16, 199-208

Goodglass, H., Kaplan, E., & Brand, S. (1983). *Boston Diagnostic Aphasia Examination.* Philadelphia: Lea & Febiger.

Guo, Y. E., Togher, L., & Power, E. (2014). Speech pathology services for people with aphasia: What is the current practice in Singapore? *Disability and Rehabilitation*, 36, 691–704. doi:10.3109/09638288.2013.804597

Helms, J. E., Henze, K. T., Sass, T. L., & Mifsud, V. A. (2006). Treating Cronbach’s Alpha Reliability Coefficients as Data in Counseling Research. *The Counseling Psychologist*, *34* (5), 630–660. doi.org/10.1177/0011000006288308

Hersh,D., Worrall, L., Howe, T., Sherratt,S., & Davidson, B. (2012). SMARTER goal setting in aphasia rehabilitation. *Aphasiology*, 26 (2), 220-233. doi: 10.1080/02687038.2011.640392

Howard D. & Hatfield F.M., (1987) *Aphasia Therapy: Historical and Contemporary Issues.* Lawrence Erlbaum Associates

Hilari K., Byng S., Lamping D.L., & Smith S.C. (2003) Stroke and aphasia quality of life scale-39 (SAQOL-39): evaluation of acceptability, reliability and validity. *Stroke, 34 (8),* 1944-1950

Hilari, K., Klippi, A., Constantinidou, F., Horton, S., Penn, C., Raymer, A., Wallace, S., Zemva, N., & Worrall., (2015). An international perspective on quality of life in aphasia: A survey of clinician views and practices from sixteen countries. *Folia Phoniatrica et Logopedica, 67* (3), 119-130.

Johansson, M. B., Carlsson, M., & Sonnander, K. (2011). Working with families of persons with aphasia: A survey of Swedish speech and language pathologists. *Disability and Rehabilitation*, 33, 51–62. doi:10.3109/09638288.2010.486465

Jordan L. & Kaiser W. (1996) *Aphasia – a social approach.* London Chapman & Hall

Kagan A. (1995) Revealing the competence of aphasic adults through conversation: a challenge to health professionals. *Topics in Stroke Rehabilitation* vol. 2 no. 1 pp15-28

Kagan, A., Simmons-Mackie, N., Rowland, A., Huijbregts, M., Shumway, E., McEwen, S., Threats, T. & Sharp, S. (2008). Counting what counts: A framework for capturing real-life outcomes of aphasia intervention. *Aphasiology,* Vol. 22(3), 258-280.

Kagan, A., Simmons-Mackie, N., Victor, J. C., Carling-Rowland, A., Hoch, J., Huijbregts, M., & Streiner D.L. (2013) *Assessment for Living with Aphasia (ALA)* (2nd ed.) Toronto, ON: Aphasia Institute.

Klippi, A., Sellman, J., Heikkinen, P., & Laine, M. (2012). Current clinical practices in aphasia therapy in finland: Challenges in moving towards national best practice. *Folia Phoniatrica Et Logopaedica : Official Organ of the International Association of Logopedics and Phoniatrics (IALP)*, 64, 169–178. doi:10.1159/000341106

Kleinman A. (1988) *The Illness narratives: suffering, healing and the human condition.* Basic books USA

Kline P. (1994) *An easy guide to factor analysis.* Routledge, New York

Le Dorze G. & Brassard C. (1995) A description of the consequences of aphasia on aphasic persons and their relatives and friends, based on the WHO model of chronic diseases *Aphasiology* vol.9 no.3, pp239-255

Long A. F., HeskethA., Paszek G., Booth M.& Bowen A. (2008); Development of a reliable self-report outcome measure for pragmatic trials of communication therapy following stroke: The Communication Outcome after Stroke (COAST) scale. *Clinical Rehabilitation* 22: 1083–1094

Nickels L.A. & Howard D. (1995) Aphasic naming: what matters? *Neuropsychologia*, vol. 33, pp1281-1303.

Oxenham, D., Sheard, C., & Adams, R. (1995). Comparison of clinicians and spouse perceptions of the handicap of aphasia: Everybody understands `understanding'. *Aphasiology*, 9, 47-493.

Parr, S., Byng, S., Gilpin, S., & Ireland, C. (1997). *Talking about aphasia: living with loss of language after stroke*. Philadelphia: Open University Press.

Rose, T., Worrall, L., & McKenna, K. (2003). The effectiveness of aphasia-friendly principles for printed health education materials for people with aphasia following stroke. *Aphasiology, 17* ,947 – 963.

Rose T., Worrall L., Hickson L., Hoffman T. (2012), Guiding principles for printed education materials: Design preferences of people with aphasia. *International Journal of Speech-Language Pathology*, 2012; 14(1): 11–23

Rose, M., Ferguson, A., Power, E., Togher, L., & Worrall, L. (2014). Aphasia rehabilitation in Australia: Current practices, challenges and future directions. *International Journal of Speech-Language Pathology, 16*(2), 169-180.

Royal College of Speech & Language Therapists' clinical guidelines. (2005) Bicester (UK) Speechmark Publishing Ltd

Sarno, M. T. (2004). Aphasia therapies: Historical perspectives and moral imperatives. In F. J. Duchan & S. Byng (Eds.), *Challenging aphasia therapies* (pp. 17–31). Hove, UK: Psychology Press.

Simmons-Mackie, N. (2000). Social approaches to the management of aphasia. In L. Worrall & C. Frattali (Eds.), *Neurogenic communication disorders: A functional approach*. New York: Thieme.

Simmons-Mackie, N. (2001). Social approaches to aphasia intervention. In R. Chapey (Ed.), L*anguage intervention strategies in adult aphasia* (pp. 246–268). Baltimore, MD: Williams & Wilkins.

Simmons-Mackie, N., Threats, T. T., & Kagan, A. (2005). Outcome assessment in aphasia: A survey. *Journal of Communication Disorders*. doi.org/10.1016/j.jcomdis.2004.03.007

Streiner D. L.& Norman G. R. (1995) *Health Measurement Scales: a practical guide to their development and use.* Oxford Medical Publications. Oxford

Swinburn K. with Byng S. (2006) *The Communication Disability Profile.* London: Connect Press

Swinburn, K., Porter, G., & Howard, D. (2004). *Comprehensive Aphasia Test*. Psychology Press.

Swinburn C.A. (2003) Unpublished PhD *Evolution of the concepts and methods associated with exploring and measuring the impact of aphasia.* City University. London

van der Gaag, A., Smith, L., Davies, S., Moss, B. and Mowles, C., 2005, Connect therapy and support services for people with long term stroke and aphasia and their relatives; a six month follow up study. *Clinical Rehabilitation,* 19, 372–381.

Verna, A., Davidson, B., & Rose, T. (2009). Speech pathology services for people with aphasia: A survey of current practice in Australia. *International Journal of Speech-Language Pathology,* 11, 191–205

Wallace, S. J., Worrall, L., Rose, T., & Le Dorze, G. (2014a). A good outcome for aphasia. *Aphasiology*, 28, 1400–1404. doi:10.1080/02687038.2014.935119

Wallace, S. J., Worrall, L., Rose, T., & Le Dorze, G. (2014b). Measuring outcomes in aphasia research: A review of current practice and an agenda for standardisation. *Aphasiology*, 28, 1364–1384. doi:10.1080/02687038.2014.930262

Wallace (2015). Development and validation of Australian aphasia rehabilitation best practice statements using the RAND/ UCLA appropriateness method. *BMJ Open, 5,* e007641.

Wallace S.J., Worrall L., Rose T., Le Dorze G., Cruice, M., Isaksen, J., Pak Hin Kong, A., Simmons-Mackie, N., Scarinci N., & Alary Gauvreau (2016): Which outcomes are most important to people with aphasia and their families? an international nominal group technique study framed within the ICF, *Disability and Rehabilitation*, doi:10.1080/09638288.2016.1194899

Wallace S.J., Worrall L., Rose T.& Le Dorze G. (2017) Which treatment outcomes are most important to aphasia clinicians and managers? An international e-Delphi consensus study, *Aphasiology*, 31:6, 643-673, doi:10.1080/02687038.2016.1186265

Whitworth A., Webster, J., & Howard, D. (2005). *A Cognitive Neuropsychological Approach to Assessment and Intervention in Aphasia*. Psychology Press.

FIGURES

(for figure captions see page 45)

1. Racial and gender options for AIQ scale

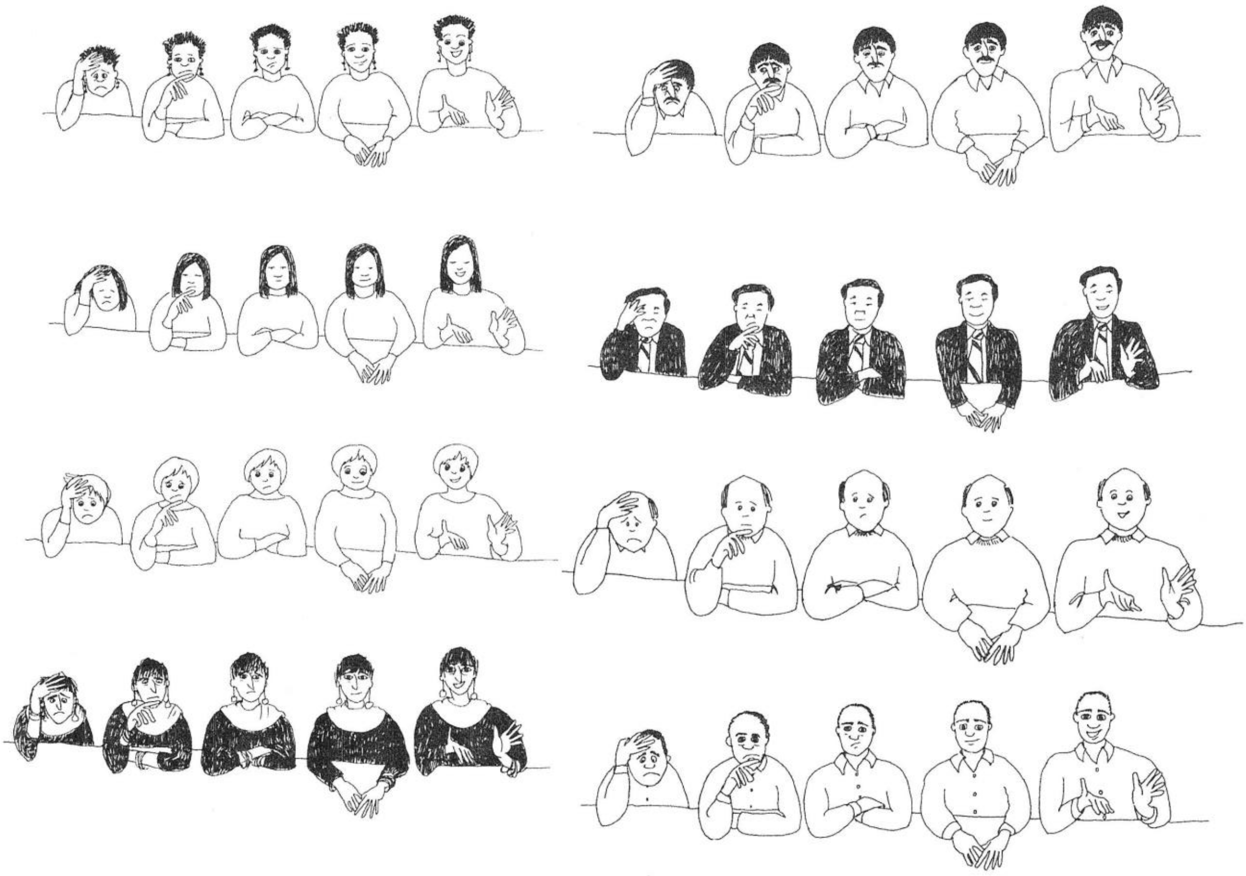
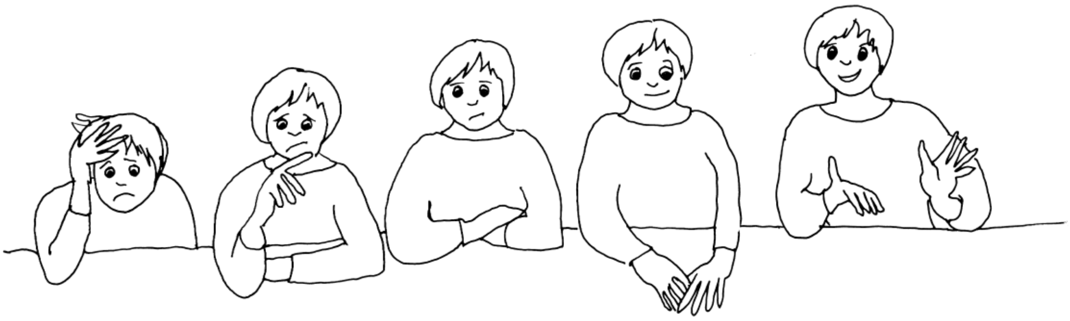


Figure 2. An example page of the AIQ

This week…

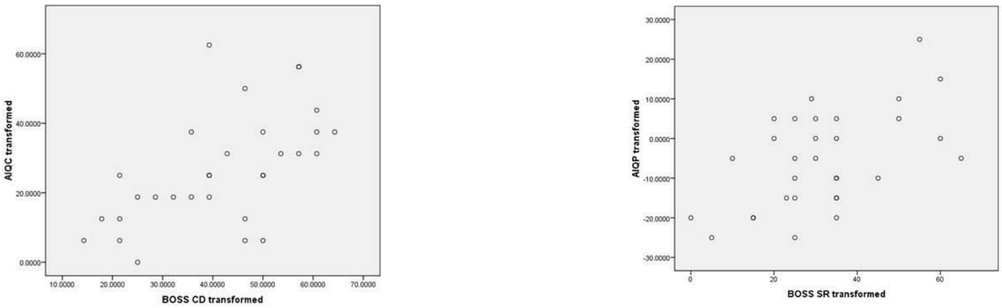
2. How **easy** was it for you to **talk** to a **stranger?**



Impossible
4

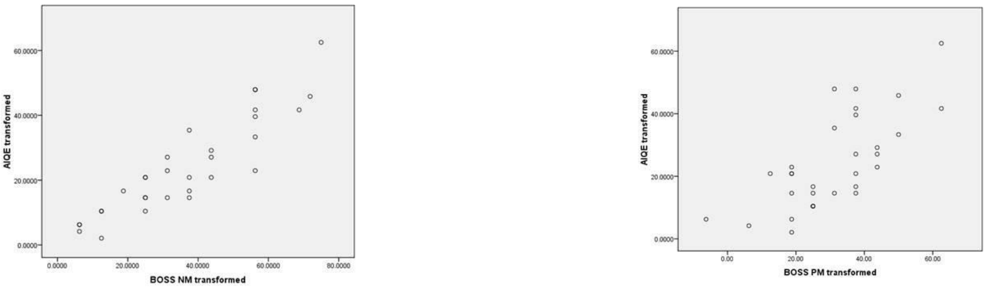
No problem
0

Figure 3. Scatter graphs showing relationships between selected domains of BOSS and AIQ-prototype



3a – Communication scores on BOSS & AIQ-prototype

3b - Social relationship scores on BOSS & participation on AIQ-prototype



3c - Negative emotion scores on the BOSS and AIQ-prototype

3d - Positive emotion scores on the BOSS and AIQ-prototype

Figure 4. Scatter graphs showing relationships between selected domains of BOSS and AIQ-21

4a – Communication scores on BOSS & AIQ-21

4b - Social relationship scores on BOSS & participation on AIQ-21

4c - Negative emotion scores on the BOSS and the AIQ-21

4d - Positive emotion scores on the BOSS and the AIQ-21

Table 1. Type of involvement of PWA in development of the AIQ-21(PWA as *partners* is indicated through *italics* and *shading*)

Stage of development	*Nature of involvement*	*Number PWA involved*	*Amount of contact*	*Contribution of PWA*
CDP	i) In-depth interviews	26	One off interview	Source of constructs important to PWA, research participant
	ii) Group interviews	*12*	*One off groups*	*Advised on format, and scrutiny of constructs*
	iii) Advisory group	*4*	*Monthly meetings over a year*	*Scrutinised and advised on all aspects of the development process (advised on content, format, tone, scoring)*
AIQ-prototype	*i) AIQ Development Group*	*6*	*Four group meetings*	*Chose 21 items from existing 56 CDP to produce the AIQ-prototype. Revisited and refined content, format, tone, scoring.*
	ii) AIQ field testing	18 (11 in Cornwall, 7 in London)	One off meetings	Mostly, data provision, research participant though comments reported on face validity.
	iii) Statistical testing	137 (internal consistency) 90 (responsiveness) 31 (concurrent validity)	One off assessment of AIQ +/- EQ-5D	Data provision, research participant
(*AIQ-prototype* delete 2 items due to redundancy and lack of sensitivity)	No PWA were involved in the decision.			None – decision was based on outcome of statistical results and psychometric analysis
AIQ-21	*Item Selection Panel*	*5*	*One off meeting*	*Selected items reading writing domains to add to the then AIQ-19*
	iii) Statistical testing	20 (internal consistency) + 20 (responsiveness)	One off assessment of AIQ +/- EQ-5D	Data provision, research participant

Table 2. AIQ-prototype Participant Characteristics

Variable	*Concurrent validity (n=31)*	*Internal consistency (n=137)*	*Sensitivity (n=90)*
Ethnicity			
Asian	3	5	3
White	19	116	71
Black	3	10	7
Other	3	4	8
Unknown	3	2	1
Gender			
Male	13	63	47
Female	18	74	43
*Age (years)**			
Mean (sd)	57.9 (14.1)	65.86 (14.60)	64.4 (14.1)
Range	24-77	33-40	35-90
*Time post event (months)***			
Mean (sd)	All > 24 months post event	43.75 (61.28)	44.1 (51.43)
Range		2-500	2-250
*BDAE severity rating ****			
Mean (sd)	3.42 (1.09)	3.06 (1.28)	3.35 (1.27)
Range	1-5	1-5	1-5

**   3 participants (Concurrent validity), and 4 (Internal consistency), whose age is unknown*

*** 10 participants (Internal consistency), and 5 (Sensitivity) whose time post event is unknown*

**** 2 participants (Internal consistency), and 9 (Sensitivity) whose BDAE rating is unknown*

Table 3. Participant numbers across locations during responsiveness testing of AIQ-prototype

Location	*Number of AIQ-prototype participants per location*	*Number of AIQ-prototype administrators per location*
London	48 (53%)	9
Cornwall	27 (30%)	1
Gloucestershire	6 (7%)	1
Sussex	9 (10%)	1
Totals:	**90**	**12**

Table 4. Reasons for dropping out of the AIQ-prototype study

Reasons for drop out	Number dropping out
Unknown	18
Illness	12
Service unavailable - staff illness	12
Did not want to use service	10
Unable to make contact	10
Reported full recovery	4
Too severe to include	4
Unable to access service – e.g. transport issues	3
Moved away	2
No aphasia	1

Table 5. AIQ-prototype domain 3 with each question removed in turn.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
BQ10	16.41	100.398	.718	.577	.908
BQ11	16.93	100.747	.758	.694	.906
BQ12	17.12	99.839	.805	.724	.904
BQ13	17.22	104.108	.701	.548	.909
BQ14	17.35	105.641	.559	.381	.915
BQ15	17.42	102.183	.696	.563	.909
BQ16	17.13	100.487	.711	.554	.908
BQ17	17.13	101.680	.681	.513	.909
BQ18	17.15	104.999	.641	.461	.911
BQ19	18.05	115.416	.341	.252	.921
BQ20	17.26	104.691	.670	.565	.910
BQ21	17.29	106.736	.618	.496	.912

TABLES IN RED TO BE IN SUPPLEMENTARY MATERIAL IF POSSIBLE

Table 6. Inter-Item Correlation Matrix* of the communication domain of AIQ-prototype

	BQ1	*BQ2*	*BQ3*	*BQ4*
BQ1	1.000			
BQ2	.641	1.000		
BQ3	.471	.242	1.000	
BQ4	.490	.486	.562	1.000

*Item - item correlation in boxes indicates the extent to which each item in each domain is correlated.

Table 7. Inter-Item Correlation Matrix* of the participation domain of AIQ-prototype

	BQ5	BQ6	BQ7	BQ8	BQ9
BQ5	1.000				
BQ6	.509	1.000			
BQ7	.373	.401	1.000		
BQ8	.198	.380	.334	1.000	
BQ9	.439	.499	.405	.384	1.000

*Item - item correlation in boxes indicates the extent to which each item in each domain is correlated.

TABLES IN RED TO BE IN SUPPLEMENTARY MATERIAL IF POSSIBLE

Table 8. Inter-Item Correlation Matrix* of the well-being/emotional state domain of AIQ-prototype

	BQ10	BQ11	BQ12	BQ13	BQ14	BQ15	BQ16	BQ17	BQ18	BQ19	BQ20	BQ21
BQ10	1.000											
BQ11	.630	1.000										
BQ12	.663	.784	1.000									
BQ13	.520	.649	.627	1.000								
BQ14	.396	.445	.484	.523	1.000							
BQ15	.527	.592	.618	.467	.465	1.000						
BQ16	.645	.557	.625	.523	.432	.530	1.000					
BQ17	.461	.589	.621	.570	.482	.482	.553	1.000				
BQ18	.533	.491	.475	.500	.361	.539	.548	.460	1.000			
BQ19	.230	.144	.246	.264	.263	.164	.218	.323	.294	1.000		
BQ20	.527	.515	.545	.451	.325	.609	.487	.442	.480	.322	1.000	
BQ21	.494	.487	.540	.414	.300	.465	.444	.409	.409	.368	.645	1.000

*Item – item correlation in boxes indicates the extent to which each item in each domain is correlated.

Table 9. AIQ-prototype sensitivity to change

AIQ-prototype Domain	T1 mean score (s.d)	T2 mean score (s.d)	Mean change	t (d.f.)	95% C.I.	p (2-tailed)
Communication	33.19 (19.33)	25.49 (19.92)	7.7	3.38 (89)	3.18 – 12.23	.001
Participation	29.78 (21.03)	23.22 (17.50)	6.6	2.74 (89)	1.81 – 11.30	.007
Well-being/ emotional state*	38.33 (23.99)	25.19 (18.17)	13.2	5.66 (89)	8.53 - 17.77	.000

* raw scores for positive emotions have already been reversed to ensure consistency of direction of change

Table 10. EQ-5D sensitivity to change

EQ5D domain	**T1 mean score (s.d)**	**T2 mean score (s.d)**	**Mean change**	**t (d.f.)**	**95% C.I.**	**p (2-tailed)**
Mobility	1.61 (.560)	1.58 (.542)	.024	.406 (83)	(-.093, .140)	.686
Self Care	1.40 (.643)	1.43 (.609)	-.036	-.686 (82)	(-.141, .069)	.495
Activities	1.74 (.562)	1.75 (.557)	-.012	-.185 (83)	(-.140, .116)	.854
Pain	1.58 (.605)	1.57 (.645)	.012	.173 (83)	(-.125, .149)	.863
Depression	1.51 (.630)	1.48 (.630)	.036	.505 (83)	(-.105, .176)	.615
Index Score	64.69 (20.518)	64.70 (19.479)	-.012	-.006 (82)	(-4.291, 4.267)	.996

* n=84 (data missing n=6)

Table 11. AIQ-21 Participant characteristics

Variable	*AIQ-21 Item-selection panel (n=5)*	*AIQ-21 Psychometric testing group (n=20)*
Ethnicity		
Asian	0	3
White	4	13
Black	1	4
Other	0	0
Unknown	0	0
Gender		
Male	4	11
Female	1	9
*Age (years)**		
Mean (sd)	63.6 (5.32)	59.26 (13.26)
Range	57-70	34-82
*Time post event (months)***		
Mean (sd)	180.6 (85.86)	90.68 (55.38)
Range	90-320	24-180
BDAE severity rating		
Mean (sd)	3.2 (1.64)	3.05 (1.19)
Range	1-5	1-5

* 1 participant in psychometric testing group whose age is unknown
** 1 participant in psychometric testing group whose time post event is unknown

TABLES IN RED TO BE IN SUPPLEMENTARY MATERIAL IF POSSIBLE

Table 12. Internal consistency* of the AIQ-21 communication items of AIQ-21

Item	*Corrected item-total correlation*	*Cronbach's alpha if item deleted*
Q1: Talking to a friend	0.536	0.703
Q2: Talking to a stranger	0.489	0.704
Q3: Understanding a friend	0.684	0.679
Q4: Understanding a stranger	0.655	0.653
Q5: Writing a letter	0.282	0.789
Q6: Reading a newspaper story	0.483	0.705
		Cronbach's alpha = 0.787

*Item – total correlation in boxes indicates the extent to which each item in each domain is correlated with the domain total

TABLES IN RED TO BE IN SUPPLEMENTARY MATERIAL IF POSSIBLE

Table 13. Internal consistency* of the AIQ-21 participation items of AIQ-21

Item	*Corrected item-total correlation*	*Cronbach's alpha if item deleted*
Q7: Positive things to do	0.346	0.513
Q8: Things you enjoy doing	0.341	0.531
Q9: Friends	0.707	0.285
Q10: Families	0.223	0.686
		Cronbach's alpha = 0.647

*Item – total correlation in boxes indicates the extent to which each item in each domain is correlated with the domain total

TABLES IN RED TO BE IN SUPPLEMENTARY MATERIAL IF POSSIBLE

Table 14. Internal consistency* of the AIQ-21 well-being items of AIQ-21

Item	*Corrected item-total correlation*	*Cronbach's alpha if item deleted*
Q11: Frustrated	0.587	0.875
Q12: Worried	0.615	0.874
Q13: Unhappy	0.720	0.872
Q14: Helpless	0.569	0.877
Q15: Bored	0.740	0.865
Q16: Embarrassed	0.555	0.877
Q17: Angry	0.546	0.884
Q18: Isolated	0.874	0.853
Q19: Stupid	0.678	0.870
Q20: Confident	0.370	0.887
Q21: Hope for the future	0.564	0.878
		Cronbach's alpha = 0.894

*Item – total correlation in boxes indicates the extent to which each item in each domain is correlated with the domain total

List of FIGURES

List of TABLES

*tables to go in online supplementary information if possible