Critical Review:
Outcome Measurement among Child Users of Augmentative and Alternative Communication

Kendra Smith
M.Cl.Sc SLP Candidate
University of Western Ontario: School of Communication Sciences and Disorders

This critical review evaluates the evidence regarding two outcome measures for pediatric Augmentative and Alternative Communication (AAC) users: The Therapy Outcome Measure for AAC (TOM-AAC) and the Family Impact of Assistive Technology Scale for AAC (FIATS-AAC). Studies reviewed employed qualitative, survey and test-retest designs. The TOM-AAC exhibits low external and internal validity, suggesting that practitioners should use caution when using it to quantify extent of quality-of-life effects and change. In contrast, the FIATS-AAC exhibits moderate external validity and moderately high internal validity; however, its greater focus on the family rather the child's quality of life reduces its utility for benchmarking child progress. Both require further psychometric development to be confidently used for pediatric AAC users. A key limitation is that neither directly measures the child's perspective.

Introduction

Within special education services, approximately 2-6% of children and upwards of 12% of preschoolers cannot communicate verbally (Binger & Light, 2006). These children may require Augmentative and Alternative Communication to communicate, which can range from low-tech options, including picture exchange and gestures, to high-tech or battery-powered AAC, such as speech-generating devices or computer-based apps. It enables users to more clearly express needs, interests, develop relationships, and participate in day-to-day activities. In fact, better health outcomes have been linked with patient autonomy in making healthcare decisions by using AAC (Department of Health, 2012).

When AAC is not adequately individualized to the needs of the individual, it can cause frustration, challenging behaviors and abandonment of AAC (e.g., Andzik et al., 2018; Hamm & Mirenda, 2006); whereas, when well-liked, it is linked with higher quality of life in adulthood (Hamm & Mirenda, 2006).

In a 2019 review of the literature, Broomfield and colleagues highlighted that although quality of life measures are commonly used to assess outcomes with AAC, none were specifically designed for AAC. Fundamentally, they noted that these tools have not been adapted to allow for data collection from the AAC user's perspective. Furthermore, individuals with autism spectrum disorder or severe intellectual disability were excluded from the literature search, even though AAC can be crucial to their successful communication.

Satisfaction has been measured in several different ways, including determining how well a new user can navigate a device, level of user anxiety, child-parent interaction style, amount of communicative participation, and more comprehensive ways. I have chosen to compare measures that use the International Classification of Functioning, Disability and Health (ICF) framework to capture the effect of AAC on the individual as a whole.

The ICF framework situates functioning and disability within the context of health condition and environmental factors (WHO, 2013). It captures degree of disability of various conditions across the age span, being designed to understand health states, outcomes, and aetiologies. Pertaining to AAC, the first component, body functions and structures, involves inefficient communication when unaided and related physical limitations (e.g., tone in cerebral palsy). Secondly, activity and participation limitations can include reduced vocational and social participation due to lack of individualization of services, according to Hamm and Mirenda (2006). Finally, the WHO (2013) explains that environmental factors can be physical, social, or attitudinal and can either minimize or exacerbate barriers to functioning.

Johnston and colleagues (2002) explore these environmental factors. Firstly, social interaction can be reduced due to segregation in special education classes. Additionally, policy or funding barriers prevent optimal language development, including long wait times, restrictions on taking AAC devices home, and lack of AAC individualization. Finally, negative attitudes towards AAC held by the family, educators, and peers can affect the child's willingness to use it. By eliminating barriers, the AAC user can be set up for success.

Objective

The objective of this paper is to determine which ICF-rooted measure best gauges outcomes among pediatric AAC users.

Methods

Search Strategy
Databases searched include PubMed, CINAHL, ProQuest Arts & Humanities, Medline OVID, Medline ProQuest, EMBASE, Health and Psychosocial Instruments, and PsycInfo. Search terms included: (AAC or "augmentative and alternative communication") and ("outcome measure" or (("scale" or "questionnaire") and ("effect" or "result" or "impact"))).

Selection Criteria
Limiters included peer-reviewed studies, publications from the past 10 years, and inclusion of pediatric participants (0-18 years). Title and abstract were then reviewed for relevance. Articles using measures not specifically designed for AAC, not using the ICF framework, articles referring to measures that are not publicly available, as well as tools lacking psychometric data were omitted. Reference sections of these papers were scoured as well as Google search engine searches of author names for other versions of the measures. Duplicates of five eligible articles ensued, so the search was halted.

Data Collection
Five articles were retained, involving two measures: The Therapy Outcome Measure for AAC (TOM-AAC) and the Family Impact of Assistive Technology Scale for AAC (FIATS-AAC). The TOM-AAC is supported by a Level 4 study and a Level 5 article; while the FIATS-AAC is supported by a Level 3 study and two Level 4 studies (Howick et al., 2011).

Results

Research Designs
Qualitative research is grounded in subjective experiences, making it difficult to generalize to the population at large, unless the sample is sufficiently representative. For instance, in focus groups, group dynamics could suppress nonconforming opinions.

In contrast, survey research involves having respondents complete questionnaires. It is at risk of nonresponse bias and bias due to item misinterpretation. Random sampling and adequate piloting to establish good validity and reliability characterize quality survey studies.
Finally, a pre-post design involves taking measures prior to and following treatment. It suffers from the fact that maturation could account for effects; it can be improved by having a control condition.

TOM-AAC
Enderby (2014). The expressed rationale was compelling for the need for the TOM-AAC in this article. He was justified in using this measure as it contained all ICF components and has been validated for rehabilitation benchmarking, albeit not specific to AAC. Domains include impairment, activity, participation, and wellbeing of the child and caregiver.

The measure (available in Enderby, 2014; Enderby & John, 2015; and Murphy & Enderby, 2014) starts by appropriately defining AAC, which is important because a few domains require scoring of function with and without AAC. The title clearly indicates the purpose of the measure. Similarly, the instructions are clear. It is comprised of 6-point ratings, each of which are accompanied by a description, which reduces subjectivity in scoring. The ratings are described in 2-5 lines; the measure could be improved by being more concise. The activity subsection is akin to the Communication Function Classification System, available at http://cfcs.us/wp-content/uploads/2014/02/CFCS_universal_2012_06_06.pdf, although the meaning of the scale items from the headings is less clear.

Murphy, Boa, and Enderby (2014). Preliminary psychometrics were established with qualitative and survey research. They sampled a comprehensive group of 19 AAC professionals throughout Scotland (16 speech-language pathologists (SLPs), eight teachers, two rehab engineers, one social worker, occupational therapist (OT), physiotherapist, and nurse), although the sampling method was not reported. Respondents rated vignettes using the scale, producing an interclass correlation coefficient (ICC) above the 0.70 criterion (Cortina, 1993) overall and for all but three of the eight subsections, which indicated some agreement.

Based on the questionnaire administered to the professionals, the authors claim that the items are appropriate, domains are well-understood, and that the questionnaire is quick and easy to use. However, these claims must be tempered because the questionnaire is biased towards positive responses (response options consist of "totally agree", "agree in part", "unsure", and "do not agree") as well as not having any negatively worded questions.

Focus groups conducted with 16 of these professionals resulted in questionnaire rewording and established feasibility. They considered it to be efficient with practice and that it would be useful in monitoring progress.

Further testing is needed to determine if the wording is understandable to other professionals, especially those who have not received training by the authors. It could be made easier to use by providing page numbers for the speech and language descriptors. Additionally, Specific Language Impairment (SLI) should be changed to Developmental Language Disorder (DLD) to reflect current terminology, support advocacy, and interprofessional communication. Other limitations included the small sample size which reduced generalizability, and a lack of factor analysis to mathematically confirm the relevance of items. Because reliability between raters was poor, especially for the expressive subsection, which AAC is designed to improve, I agree with the authors' position that more psychometric development is needed. This study provides somewhat suggestive evidence for the use of the TOM-AAC to measure outcomes among AAC users.

FIATS-AAC

Unlike the TOM-AAC, the FIATS-AAC also considers the family's perspective. Available from https://hollandbloorview.flintbox.com/technologies/ae1ef6f8-878b-4dd8-a45b-9f7f73894fee, it uses a 7-point Likert scale, ranging from "strongly agree" (7) to "strongly disagree" (1). It assesses the following constructs: behavior, caregiver relief, contentment, doing activities, education, energy, communication, family roles, finances, security, self reliance, social versatility, and supervision. Instructions and title are clear. It does not include any negatively worded items, which may allow yea-sayers to inflate the overall score. The full version is lengthy at 89 items, although a shorter version (38 items) reduces respondent burden. Finally, the face-to-face scale is available standalone for a quick 8-item overview of communication. In all cases, scoring is simple, being automatically scored with the provided excel sheet.

For the purposes of evaluation, I will discuss the long version. It strategically places more emotionally upsetting questions pertaining to their child's dependence at the end. Some possibly problematic items include item 11 "my child tells me about her/his day" and item 58 "I have little time to get chores done around the house." Item 11 would be better suited to a frequency scale, and item 58 may not be related to having a child who requires AAC; perhaps the parent simply works several jobs. Furthermore, items are geared for older children; items relating to impact of child dependence could be explained by young age.

Delarosa and colleagues (2012). Because positive family attitudes translate to children being more likely to benefit from the AAC, this pioneering study demonstrates sound rationale in developing a measure

from the family perspective. A combination of appropriate qualitative and survey methods was used and evaluated by a random sample of parents of children aged 6-12 years from the Holland Bloorview database. Items were developed based on a literature search, confirmed in a focus group, then revised based on an adequately comprehensive sample of focus group participants (two AAC researchers, an OT, a SLP, an assistive technology specialist, and two parents of AAC users). Parents reported understanding 87% of items.

Subsequent sampling was adequate in size (n=135), considering that there were 13 dimensions and 5-10 individuals per dimension is a rule of thumb for factor analysis (Frost et al., 2007). Being that it was a mail-out survey, it is not surprising that there was a high non-response bias (64%). Because the researchers had anonymized the survey, they could not follow-up to determine why the survey had not been completed. Internal reliabilities of most domains surpassed the Cronbach alpha criterion of 0.7 (Cortina, 1993), although contentment and family roles marginally did not. Since content experts endorsed them, they should be retained and revised.

Finally, test-retest reliability and convergent validity was determined via a random sample of 30 English-speaking parents from the Holland Bloorview database. All mail-out surveys were returned. Interclass correlation coefficient was well above the 0.7 criterion in all cases, establishing excellent test-retest reliability. The FIATS-AAC was significantly negatively and highly correlated (Cohen, 1992) with the Impact on Family Scale, indicating excellent convergent validity especially concerning parental well-being.

This article provides fairly compelling evidence for the use of the FIATS-AAC to measure AAC-use on family functioning, which can be considered mildly suggestive of child functioning.

Kron and colleagues (2018). The FIATS-AAC, the Participation and Environment Measure for Children (PEM-CY) and the Child Health Questionnaire-Parent (CHQPF28) were distributed by mail to a random hospital sample of 352 families. Of those returned, forty-seven matched inclusion criteria, being 5-12 years old with at least 12 weeks' experience with AAC.

Non-response bias could not be determined due because participants self-assessed eligibility and forms were anonymized. Parents who completed the survey tended to be female and AAC users tended to be male; further replication with improved gender representation would ensure results are more generalizable.

The first six participants to submit their survey and consent were interviewed; this may have introduced bias if these caregivers as a group were less stressed and had more time on their hands, meaning that only one perspective would be elicited, possibly that of caregivers of children with less severe complex communication needs and fewer comorbid conditions. They minimized errors by having authors rate the six interviews on the FIATS-AAC domains while being blinded to the parent-completed FIATS-AAC scores. These scores were then converted into a 3-point scale (positive, negative, neutral) to permit use of the quadratic weighted kappa statistic; however, this data reduction limits the preciseness of the measure of inter-rater reliability.

Appropriate statistical analysis suggests that the PEM-CY and FIATS-AAC are not closely related although the CHQ-PF28 and the FIATS-AAC are. The small correlation between the PEM-CY and FIATS-AAC is not unexpected because the measures assess different constructs; the FIATS-AAC measures a more comprehensive range of constructs and the measures use different scales (frequency or degree). The authors indicate that the larger agreement between the FIATS-AAC and the CHQ-PF28 establishes that although the FIATS-AAC can measure child quality of life, the FIATS-AAC must be tapping into more constructs than child factors alone, since psychosocial factors only accounted for 29% of the variance. The quadratic weighted kappa statistic was 0.39 (CI: 0.22-0.56), indicating fair agreement between researcher and parent (Landis & Koch, 1977). This study provides suggestive evidence for the FIATS-AAC as a measure of child psychosocial functioning.

Ryan and colleagues (2018). Finally, Ryan and colleagues (2018) implemented a case-series pre-post-test design upon receipt of a speech-generating device, within 3 days of receipt, at 6 then 12 weeks after receiving the device. Fifty parent-child pairs aged 3-16 requiring speech generating devices were recruited from Holland Bloorview and Thames Valley Children's Centre; thirty of which were from Toronto, while 20 of which were from London, Ontario. Most were young boys. They used a younger age cut-off to establish psychometrics for a younger demographic. At each time point, the FIATS-AAC was administered over the phone. Beginning at the second phone call, clinicians also rated the children's operational, social, strategic, and linguistic competencies over time.

Test-retest reliability between the first two phone calls was excellent, as the overall ICC exceeded the 0.9 cut-off (Koo & Li, 2016). An analysis of variance indicated a significant main effect of time, with small to medium-magnitude effect sizes at first and second follow-up measured via post hoc t-tests. Effect sizes increased by 45% when post-hoc t-tests included only the participants whose communication had improved according to both parent and clinician.

Authors controlled for researcher error by training research assistants to conduct the interview in a standardized way. A confound is that device delivery times by post and need to reschedule calls slightly affected the amount of exposure a client had to a device between assessments. Additionally, a maturation effect may have accounted for results, as no control condition was used. This study provides suggestive evidence for the ability of the FIATS-AAC to detect change related to AAC use.

Discussion

Both measures enabled assessment of children whose communication ability may be limited. They were available only in English (other than Carloni and colleague's 2020 translation of FIATS-AAC into Italian). Aside from Delarosa's 2012 study, small sample sizes (as defined by Cohen, 1992; Frost et al., 2007) may have reduced the power of the studies, so results should be considered with caution. Additionally, articles published on each were produced by the developers of these tests; replication by other researchers is needed. Furthermore, socioeconomic status was not described in detail in any of the papers, even though having higher income may mean that families would be able to afford devices that may better suit their child's needs, which may be linked with higher client satisfaction with AAC.

They differed in that the TOM-AAC was a clinician-specific measure, whereas the FIATS-AAC is a parent-report measure. Only the FIATS-AAC mentions the populations studied. Inclusion of several disorder areas (e.g., Kron et al., 2018; Ryan et al., 2018) is suggestive of its utility in a wide range of conditions, although limiting sampling to AAC centres limits representativeness of clients who are unable or ineligible to access services. Consequently, clients with financial barriers, caregivers who are unable to take time off work to attend AAC assessment, and the less severe clientele may be underrepresented. As a result, the measure may be less sensitive to aspects of AAC-related quality of life pertaining to those populations.

The TOM-AAC exhibits limited external validity due to no reports of sampling method and no testing on children with AAC; it has limited internal validity due to a relative lack of psychometric validation and poor inter-rater reliability. Improved psychometric development

and research on its sensitivity to detect change to AAC users longitudinally is needed to determine its utility in practice.

In contrast, the FIATS-AAC has been through more testing, of which, most studies are randomized with experimenter controls in place to minimize any researcher error, although reason for nonresponse was not reported. It does not appear to be very sensitive to change following AAC device use possibly due to the greater focus on family rather than child functioning.

Neither the TOM-AAC nor the FIATS-AAC is adapted for any auditory or visual impairments. Additionally, neither permit the child to share his or her experiences directly.

It would be of greater utility to also include ratings from the child's perspective. Swett and colleagues (2020) adapted their Youth Evaluation of Products AAC (YEP-AAC) scale for AAC users by using simple language, pictures to complement the text, and a closed-ended rating scale. They also permitted a family member to read the scale to the child, indicate the child's responses on the record form or abbreviate the 7-point scale into two options. Adaptations for younger children, and children with more severe cognitive or communication deficits should also be further explored.

Clinical Implications

Practitioners can consider using the TOM-AAC to benchmark progress, although they should practice until their intra-rater reliability is high. Interpretation of client functioning should also include other sources of information, as internal validity is low. Caution should be exerted when interpreting ratings from other professionals, as inter-rater reliability is poor. As for the FIATS-AAC, it should be used primarily to attain a holistic picture of family functioning and less so to measure the child's communicative functioning on its own. Additionally, interpretations of a score at one point in time should be taken with caution for younger children, because heightened dependence on adults could be accounted for by age rather than comorbid conditions. For children who have insight and can express it, use of additional scales to capture their perceptions are recommended. Practitioners are urged to stay up-to-date on emerging client-specific measures of functioning.

Conclusion

Once suitable AAC is developed for an individual, it is important to support families in promoting long-term use and improved quality of life. One such way is by monitoring outcomes over time and adjusting device settings when change is needed. This paper evaluated the strengths and caveats of the TOM-AAC and the FIATS-AAC to enable rehabilitation professionals to judiciously select outcome measures. Currently, the TOM-AAC better assesses child-specific quality of life compared to the FIATS-AAC, although its reliability and validity need to be further examined. Adaptations for specific populations, including the unique needs of children with developmental and intellectual disabilities, need to be researched.

References

Andzik, N. R., Schaefer, J. M., Nichols, R. T., & Chung, Y. (2018). National survey describing and quantifying students with communication needs. *Developmental Neurorehabilitation, 21*(1), 40-47. doi: 10.1080/17518423.2017.1339133

Binger, C., & Light, J. (2006). Demographics of preschoolers who require AAC. *Language, Speech, and Hearing Services in Schools, 37*, 200-208. doi: 10.1044/0161-1461(2006/022)

Broomfield, K., Harrop, D., Judge, S., Jones, G., & Sage, K. (2019). Appraising the quality of tools used to record patient-reported outcomes in users of augmentative and alternative communication (AAC): A systematic review. *Quality of Life Research, 28*, 2669–2683. **https://doi.org/10.1007/s11136-019-02228-3**

Carloni, J., Magni, R., Veglio, E., Ryan, S. E., Gherardini, A, & Desideri, L. (2020). Translation and preliminary validation of the Italian version of the Family Impact of Assistive Technology Scale for Augmentative and Alternative Communication (FIATS-AAC.it). *Technology & Disability*, *32*, 129-135. doi: 10.3233/TAD-200261

Cohen J. (1992). A power primer. *Quantitative Methods in Psychology, 112*(1), 155-159.

Cortina, J. M. (1993). What is Coefficient Alpha: An examination of theory and applications? *Journal of Applied Psychology, 78*, 98-104. **http://dx.doi.org/10.1037/0021-9010.78.1.98**

Delarosa, E., Horner, S., Eisenberg, C., Ball, L., Renzoni, A. M., & Ryan, S. E. (2012). Family impact of assistive technology scale:

Development of a measurement scale for parents of children with complex communication needs. *Augmentative and alternative communication, 28*(3), 171-180. doi: 10.3109/07434618.2012.704525

Department of Health. (2012). *The power of information: Putting us in control of the health and care information we need.* London: Department of Health.

Enderby, P, & John, A. (2015). Therapy Outcome Measures for Rehabilitation Professionals. J & R Press.

Enderby, P. (2014). Introducing the therapy outcome measure for AAC services in the context of a review of other measures. *Disability and Rehabilitation: Assistive Technology, 9*(1), 33-40. doi: 10.3109/17483107.2013.823576

Frost, M. H., Reeve, B. B., Liepa, A. M., Stauffer, J. W., & Hays, R. D. (2007). What is sufficient evidence for the reliability and validity of patient-reported outcome measures? *Value in Health, 10*, S94-S105. **https://doi.org/10.1111/j.1524-4733.2007.00272.x**

Hamm, B., & Mirenda, P. (2006). Post-school quality of life for individuals with developmental disabilities who use AAC. *Augmentative and alternative communication, 22*(2), 134-147. doi: 10.1080/07434610500395493

Howick, J., Chalmers, I., Glasziou, P., Greenhalgh, T., Heneghan, C., Liberati, A., … Hodgkinson, M. (2011). The Oxford Levels of Evidence 2. **https://www.cebm.ox.ac.uk/resources/levels-of-evidence/ocebm-levels-of-evidence**

Johnston, S. S., Blue, C., Gevarter, C., Ivy, S., & Stegenga, S. (2020). Opportunity barriers and promising practices for supporting individuals with complex communication needs. *Current Developmental Disorders Reports, 7*, 100-108. **https://doi.org/10.1007/s40474-020-00195-w**

Koo, T. K., & Li, M.Y. (2016). A guideline of selecting and reporting Intraclass Correlation Coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2),155-63. **https://doi.org/10.1016/j.jcm.2016.02.012**

Kron, A. T., Kingsnorth, S., Wright, F. V., & Ryan, S. E. (2018). Construct validity of the family impact of assistive technology scale for augmentative and alternative communication. *Augmentative and alternative communication, 34*(4), 335-347. **https://doi.org/10.1080/07434618.2018**

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Murphy, J., Boa, S, & Enderby, P. (2014). Testing the Reliability and Validity of the Therapy Outcome Measure for AAC. Retrieved from **https://www.talkingmats.com/wp-content/uploads/2014/11/TOM-AAC-Final-report-Oct-2014.pdf**

Ryan, S. E., Shepherd, T. A., Renzoni, A. M., Servais, M., Kingsnorth, S., Laskey, … & Bradley, K. (2018). Responsiveness of a parent-reported outcome measure to evaluate AAC interventions for children and youth with complex communication needs. *Augmentative and Alternative Communication, 34*(4), 348-358. **https://doi.org/10.1080/07434618.2018.1520296**

Swett, C., Surins, H., Welton, C., Shepherd, T. A., Renzoni, A. M., & Ryan, S. E. (2020). Toward greater involvement of youth with complex communication needs in the selection of augmentative and alternative communication devices. *Disability and rehabilitation: Assistive Technology, 15*(1), 92-100. doi: 10.1080/17483107.2018.1534997

World Health Organization [WHO]. (2013). How to use the ICF: A practical manual for using the International Classification of functioning, Disability and Health (ICF): exposure draft for comment. WHO. **https://www.who.int/docs/default-source/classification/icf/drafticfpracticalmanual2.pdf?sfvrsn=8a214b01_4**