JSLHR

Review Article

Measures of Psychological Impacts of Stuttering in Young School-Age Children: A Systematic Review

Monique L. Jones,[a] Ross G. Menzies,[a] Mark Onslow,[a] Robyn Lowe,[a] Sue O'Brian,[a] and Ann Packman[a]

Purpose: Recent research has shown that some school-age children who stutter may have speech-related anxiety. Given this, speech-language pathologists require robust measures to assess the psychological effects of stuttering during the school-age years. Accordingly, this systematic review aimed to explore available measures for assessing the psychological impacts of stuttering in young school-age children and to examine their measurement properties.
Method: The systematic search protocol was registered with PROSPERO (ID: 163181). Seven online databases, in addition to manual searching and screening of reference lists, were used to identify appropriate measures for the population of children who stutter aged 7–12 years. The first two authors independently assessed the measures using the quality appraisal tool described by Terwee et al. (2007).
Results: Despite the comprehensive search strategy, only six measures were identified for quality appraisal. No assessment tool was found to possess adequate measurement properties for the eight assessed domains: content validity, internal consistency, construct validity, reproducibility, reliability, responsiveness, floor and ceiling effects, and interpretability. No measure had clear evidence of responsiveness to clinical change. Based on the criterion defined by the Terwee et al. (2007) appraisal tool, the Communication Attitude Test and the Overall Assessment of the Speaker's Experience of Stuttering for School-Age Children received the highest number of ratings in support of their measurement properties.
Conclusions: The results highlight a lack of available measures in this domain and poor practices in developing and testing measurement instruments. To ensure that clinicians and researchers are equipped with sound measures to meet the mental health needs of this vulnerable population, further research to establish resources is needed.

Communication is a fundamental part of almost all aspects of our everyday life. Stuttering is a communication disorder that significantly impacts people across all ages, races, cultures, and languages (Van Riper, 1982). For those affected, the behavioral features of stuttering can present as repeated movements and fixed postures of the speech organs, as well as superfluous behaviors (Teeson et al., 2003). Repeated movements comprise the repetition of sounds, syllables, and phrases. Fixed postures involve the prolongation of speech sounds or complete cessation of speech in long "blocks" of silence. Superfluous behaviors comprise idiosyncratic verbal and nonverbal behaviors and can include tics, blinking, and grimacing. These disruptions to verbal communication can present alone or in any combination (Teeson et al., 2003).

Stuttering typically develops during the preschool years when young children begin to produce more complex utterances (Packman et al., 1996). The onset of stuttering can be gradual or sudden, the latter accounting for a third of cases (Yairi & Ambrose, 1992). Many children who stutter recover naturally, without the need for speech-language pathology intervention (Yairi, 2004). For preschool children who require therapy, efficacious treatments are available that are supported by randomized clinical trial evidence (De Sonneville-Koedoot et al., 2015; Onslow et al., 2012). However, if stuttering is left untreated before the school-age years or does not respond to treatment in this period, it can become a chronic speech disorder. Persistent stuttering can cause a lifetime of struggle with the negative impacts spanning a vast array of domains, including but not limited to education, occupational attainment, personal relationships, stigma, and mental health. Educational problems

[a]Australian Stuttering Research Centre, Graduate School of Health, University of Technology Sydney, Ultimo, New South Wales, Australia

Correspondence to Mark Onslow: Mark.Onslow@uts.edu.au

Editor-in-Chief: Bharath Chandrasekaran
Editor: Julie D. Anderson

Received August 3, 2020
Revision received December 3, 2020
Accepted February 7, 2021
https://doi.org/10.1044/2021_JSLHR-20-00455

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

can present in the school-age years. C. A. Boyle et al. (1994) showed that children who stutter have poorer school attendance rates and are more at risk of having to complete a further 12 months at the same year level. The negative impact of stuttering on education is also highlighted in a study by O'Brian et al. (2011); the results showed that adults with more severe stutters were less likely to have achieved higher levels of education. Given the effects of stuttering on education, it is not surprising that stuttering can also negatively impact occupational attainment. In a study by Klein and Hood (2004), a large percentage of participants considered that stuttering adversely affects opportunities for gaining employment and hinders chances for promotion. Negative perceptions around stuttering and occupation were greater for males than females. More recently, Gerlach et al. (2018) found that females who stutter were more adversely affected than nonstuttering females, regarding underemployment and yearly remuneration compared to males.

Beyond the school and work place, stuttering also has the potential to impact personal relationships. Results from Van Borsel et al. (2011) suggested that people labeled as stutterers were perceived as less physically attractive by peers without a stutter. The adolescents and young adults in this study also held negative perceptions toward being involved in an intimate relationship with a person who stutters.

Furthermore, in an American study, M. P. Boyle (2018) investigated experiences of stigmatization in 342 adults who stutter. Results from this study found that the majority of participants reported experiencing stigma about their stutter during some stage of life. Participants also commonly reported anticipation of stigma in the future. The broader impacts of the speech disorder outlined above are not exhaustive, as speech-related anxiety can also be associated with stuttering.

While not all people who stutter experience speech-related anxiety, the potential psychological impacts of stuttering are documented in research on adults. Iverach et al. (2009) showed that adults seeking treatment for stuttering, compared with age- and sex-matched community controls, had sixfold odds of having an anxiety disorder. Most notably, compared with controls, they had a 16- to 34-fold odds increase of developing social anxiety disorder (SAD). This condition involves a fear of scrutiny from others in performance-based and social situations. Situation avoidance is common with SAD, as individuals seek to reduce or prevent the likelihood of scrutiny, judgment, embarrassment, and humiliation (American Psychiatric Association, 2013). Not surprisingly, this can lead to underachievement and social isolation across the life of the disorder for those who stutter (Blumgart et al., 2010; Onslow, 2021).

There is now recurring evidence that these psychological problems are not only associated with adults who stutter but with some school-age children as well. A recent Australian study involving 75 stuttering school-age children and 150 nonstuttering school-age children found that 24% of the children with stuttering warranted a diagnosis of SAD, compared with only 5% of the control group (Iverach et al., 2016). Given that there is no consistent evidence of temperamental differences between young children who stutter and those who do not (Kefalianos et al., 2014), it is plausible that social anxiety emerges from early socialization in which stuttering is punished. Langevin et al. (2009) showed that, even during the preschool years, peers may imitate, mock, or walk away from children who stutter, demonstrating that aversive experiences for stuttering may occur from an early age.

The identification of psychological disorders in people who stutter is controversial, and it should be noted that criticisms in this area of research (e.g., Manning & Beck, 2011, 2013) have highlighted that anxiety in social contexts is not an unnatural response for people who stutter. These authors have questioned whether additional psychological labels are helpful or valid. Regardless of this debate, there is no doubt that people who stutter often experience speech-related anxiety. Consequently, clinical tools are needed to help identify anxiety so that the appropriate support and management options can be put in place.

Given the data on speech-related anxiety in those who stutter, there has been a growing recognition that speech-language pathologists (SLPs) need to engage in a more global assessment of clients (Speech Pathology Australia, 2017). The clinical guidelines for the treatment of stuttering by the national association for speech pathologists in Australia state that "clinical assessment of the disorder requires that behavioral and mental health domains be considered" (Speech Pathology Australia, 2017, p. 7). Consequently, SLPs need to be equipped with robust clinical tools to ensure the mental health needs of this population are assessed and dealt with clinically. The well-known psychological sequela of stuttering in adults has led to the development of a range of assessment measures and tools for this population. There are several psychologically based measures reported in peer-reviewed journals for adults who stutter, including but not limited to (a) the Brief Fear of Negative Evaluation Scale, which measures fear of negative scrutiny in social situations (Carleton et al., 2006); (b) the Overall Assessment of the Speaker's Experience of Stuttering (OASES) for adults (Yaruss & Quesal, 2010), which includes scales of quality of life and emotional responding; and (c) the Unhelpful Thoughts and Beliefs About Stuttering Scale, which lists negative automatic thoughts that may drive speech-related anxiety in adults who stutter (Iverach et al., 2011). The Unhelpful Thoughts and Beliefs About Stuttering Scale became the foundation for the development of effective cognitive behavior therapy for adults who stutter (Helgadóttir et al., 2009). Phase I, II, and III trials have shown that cognitive behavior therapy for adults who stutter is capable of eliminating SAD in this population and may also improve fluency (Helgadóttir et al., 2009; Menzies et al., 2016; Menzies, O'Brian, et al., 2019; Menzies, Packman, et al., 2019).

In summary, there are a large and growing number of scales for measuring the psychological impact of stuttering in adults. Given that the potential psychological impacts of the speech disorder are not exclusive to adults who

stutter, a corresponding research effort for children who stutter is also clinically and theoretically important. Hence, the purpose of this review article is a systematic review of available measures for assessing the psychological impact of stuttering in the young school-age population.

Method

The systematic review protocol was submitted in January 2020 and registered with PROSPERO (ID: 163181).

Search Strategy

The systematic search was conducted using seven online databases: EMBASE, PsychINFO, Scopus, CINAHL, The Cochrane Library, Web of Science, and ERIC. Keyword searches included terms pertaining to four concepts: (a) stuttering, (b) child, (c) measures, and (d) measurement properties. To minimize the risk of missing pertinent articles, the authors chose not to narrow the search strategy by attempting to capture all possible terms relating to "psychological impacts" with a fifth concept. A copy of the precise search terms is available upon request.

Alerts were set for each of the databases to identify any new studies for potential inclusion from January to March 2020. The reference lists of included studies were also scanned for eligible articles. No limitations for the date of publication were set.

Inclusion Criteria

Eligible articles included studies that (a) described the development and validation of a measure designed to assess the nonspeech impacts of stuttering on young school-age children (e.g., measures of quality of life, communication attitudes, automatic thoughts, bullying, and fear of negative evaluation), (b) detailed the measurement properties, (c) utilized an English language–speaking stuttering sample that included at least some participants aged from 7 to 12 years, (d) were written in English, and (e) were published in peer-reviewed journals.

Selection Process

Results from the search strategy were exported to a reference management program, Endnote. Duplicates of articles were removed before importing the results to a systematic review data management software program, Covidence. Further duplicates were identified in Covidence. All remaining titles and corresponding abstracts were independently screened by the first two authors. If the titles or abstracts lacked sufficient detail for screening purposes, they were included for full-text review. Conflicting screening results were resolved through a consensus meeting of the first two authors before proceeding to the review of full texts. The interrater agreement for the screening phase was 98%.

Appraisal of Measurement Properties

The quality appraisal tool outlined by Terwee et al. (2007) was used to evaluate the measurement properties of the included studies. This tool has been successfully utilized in the field of psychology to evaluate the adequacy of self-report anxiety measures (e.g., Modini et al., 2015; Stein et al., 2017; Zuccala et al., 2019). The appraisal tool consists of nine quality criteria: content validity, internal consistency, criterion validity, construct validity, reproducibility, reliability, responsiveness, floor and ceiling effects, and interpretability. The scoring system utilizes four symbols for rating each of the measurement properties: "+" indicating a positive rating, "?" indicating an intermediate rating, "−" representing a negative rating, and "0" for when no information is available (Terwee et al., 2007). The criteria for a positive, intermediate, and inadequate rating are defined for each of the measurement properties in Table 1. Three modifications were made to the quality criteria defined by Terwee et al. for this evaluation. First, in the absence of a gold-standard instrument for measuring psychological impacts in young school-age children who stutter, one of the nine quality criteria—criterion validity—was excluded from the present review. Second, in keeping with the previous use of this appraisal tool (Burton et al., 2016; Zuccala et al., 2019), smallest detectable change scores were replaced with evidence of test–retest reliability for the purposes of evaluating agreement. Finally, evidence of mean and standard deviation scores for a minimum of four subgroups was considered adequate for a positive rating on interpretability. This final modification was also made in congruence with prior utilization of the Terwee et al. appraisal tool in the field of psychology (Zuccala et al., 2019; see Table 1 for further details).

Results

Results of the Search Strategy

The initial search yielded a total of 5,036 results with 3,265 duplicates (3,146 in Endnote and a further 119 in Covidence). Thus, a total of 1,771 titles and abstracts were finally screened by the first two authors. Seventeen studies met criteria for full-text screening. Two met the inclusion criteria for a final review of measurement properties, including one additional measure that was identified through the screening of reference lists. Due to the scarcity of articles meeting the full inclusion criteria, the fifth criterion was dropped. That is, measures were included in the review if their psychometric properties were established in test manuals or books rather than peer-reviewed journals describing the development of the measures. This adaptation yielded a further four measures, resulting in a total of six instruments for quality appraisal. The results of the search strategy are presented in Figure 1.

The final measures included in the quality appraisal were the Parent Questionnaire, the Teasing/Bullying Questionnaire for Children Who Stutter (TBQ-CS), OASES for School-Age Children (OASES-S), and three self-report tests

Table 1. A modified version of the quality criteria for measurement properties of health status questionnaires (Terwee et al., 2007).

Property	Definition	Quality criteria
1. Content validity	The extent to which the domain of interest is comprehensively sampled by the items in the questionnaire	(+) A clear description is provided of the measurement aim,the target population, the concepts that are being measured, and the item selection AND target population and (investigators or experts) were involved in item selection (?) A clear description of above-mentioned aspects is lacking OR only target population involved OR doubtful design or method (–) No target population involvement (0) No information found on target population involvement
2. Internal consistency	The extent to which items in a subscale are intercorrelated	(+) Factor analysis performed on adequate sample size (7 times the number of items) AND Cronbach's alpha(s) calculated per dimension and Cronbach's alpha(s) between .70 and .95 (?) No factor analysis OR doubtful design or method (–) Cronbach's alpha(s) < .70 or > .95, despite adequate design and method (0) No information found on internal consistency
3. Criterion validity[a]	Removed	
4. Construct validity	The extent to which the scores relate to other measures in a manner that is consistent with theoretically derived hypotheses concerning the concepts that are being measured	(+) Specific hypotheses were formulated AND at least 75% of the results are in accordance with these hypotheses (?) Doubtful design or method (e.g., no hypotheses) (–) Less than 75% of hypotheses were confirmed, despite adequate design and methods (0) No information found on construct validity
5. Reproducibility		
5.1 Agreement (test–retest)[b]		(+) Test–retest reliability of $r \geq .70$ AND means and standard deviations are presented for X2 time points (?) Test–retest agreement > .70 but means and standard deviations missing OR doubtful design or method (–) Test–retest agreement calculated to be $r < .70$ (0) No information found on test–retest reliability
5.2 Reliability	The extent to which patients can be distinguished from each other, despite measurement error	(+) ICC or weighted kappa ≥ .70 (?) Doubtful design or method (e.g., time interval not mentioned) (–) ICC or weighted kappa < .70, despite adequate design and method (0) No information found on reliability
6. Responsiveness	The ability of a questionnaire to detect clinical changes over time	(+) SDC OR SDC < MIC OR MIC outside the LOA OR RR > 1.96 OR AUC ≥ .70 (?) Doubtful design or method (–) SDC or SDC ≥ MIC OR MIC equals or inside LOA OR RR ≤ 1.96 OR AUC < .70, despite adequate design and method (0) No information found on responsiveness
7. Floor and ceiling effects	The number of respondents who achieved the lowest or highest possible score	(+) ≤ 15% of the respondents achieved the highest or lowestpossible scores (?) Doubtful design or method (–) > 15% of the respondents achieved the highest or lowest possible scores, despite adequate design and methods (0) No information found on floor and ceiling effects
8. Interpretability[c]	The degree to which one can assign qualitative meaning to quantitative scores	(+) Mean and *SD* scores presented for at least four relevant subgroups of patients (?) Doubtful design or method OR less than four subgroups defined (0) No information found on interpretation.

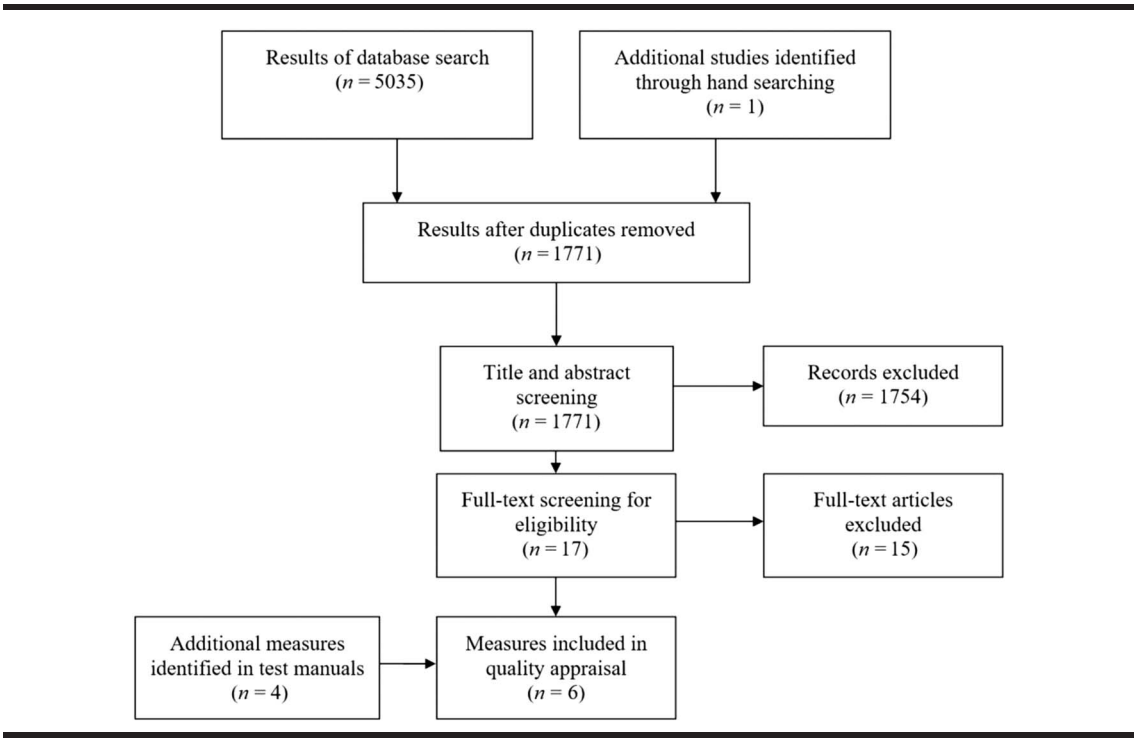
Note. Adapted with permission of Elsevier Science & Technology Journals, from "Quality Criteria Were Proposed for Measurement Properties of Health Status Questionnaires," Caroline B. Terwee et al., *Journal of Clinical Epidemiology,* Vol. 60, No. 1, Copyright © 2007. Permission conveyed through Copyright Clearance Center, Inc. ICC = intraclass correlation; SDC = smallest detectable change; MIC = minimal important change; LOA = limits of agreement; RR = responsiveness ratio; AUC = area under the receiver operating characteristics curve; *SD* = standard deviation.

[a]Criterion validity was removed from the appraisal in the absence of a gold-standard instrument for assessing psychological concomitants in school-age children who stutter. [b]Reproducibility (agreement) was modified so that evidence of test–retest reliability could constitute a positive rating. [c]Interpretability was adapted so that means and *SD* scores for at least four subgroups were sufficient for a positive score.

of the Behavior Assessment Battery for School-Age Children Who Stutter (BAB; the Communication Attitude Test [CAT], the Behavior Checklist [BCL], and the Speech Situation Checklist–Emotional Reaction [SSC-ER]). Details of the included measures, including sample items from the questionnaires, are presented in Table 2 and described below.

The Parent Questionnaire was developed by Erickson and Block (2013). Participants in that study included the parents of children (aged 11–18 years) who stuttered. The parent questionnaire was created to investigate parents' perceptions of the impact of stuttering on their family and their child or adolescent. Specifically, it explores emotional,

Figure 1. Results of the search strategy.



financial, communication, and family impacts. The parent measure consists of seven items, each scored on a 7-point Likert scale.

The TBQ-CS was developed by Langevin et al. (1998) and trialed with children aged 7–15 years. The questionnaire contains 11 items and utilizes a combination of frequency and impact subscales to capture how often teasing or bullying is occurring and the impact of these events on the child.

OASES-S is a school-age version of the popular adult OASES. The OASES-S is designed for children aged 7–12 years. The questionnaire has four sections. The first section asks about the individual's general knowledge about stuttering. The second section concerns reactions to stuttering, the third refers to communication difficulties in daily situations, and the last explores the child's quality of life. A 5-point scale is used to score each of the items (Yaruss & Quesal, 2010).

The BAB is a multifaceted test designed to assess how children feel, think, and cope with their stuttering (Longland, 2009). Normative information on the BAB is provided for children aged 6–15 years (Brutten & Vanryckeghem, 2007). The assessment battery is composed of three subtests, the CAT, the BCL, and the SSC. The CAT is designed to measures thoughts, beliefs, and attitudes toward communication. The scoring system is dichotomous, with children required to answer *true* or *false* to 33 test items (Brutten & Vanryckeghem, 2007). The BCL is a 50-item measure that assesses how children cope in response to their stutter. It explores various behavioral reactions to stuttering, including word and situation avoidance. The SSC has two components covering emotional reactions (SSC-ER) and speech disruption (SSC-SD). The SSC-SD was not included in this systematic review as it does not cover the psychological impacts of stuttering.

Assessment of Measurement Properties

The six included instruments were independently appraised by the first two authors using the Terwee et al. (2007) quality criteria tool. Eight measurement properties were evaluated using the scoring criterion outlined in Table 1. There were no discrepancies between the ratings of the two authors. The ratings are provided in Table 3.

Content Validity

Content validity pertains to whether the test items comprehensively cover the domains of interest (Terwee et al., 2007). In accordance with the criteria outlined by Terwee et al., the articles should clearly stipulate the measurement aims, for whom the measure is specifically designed, exactly what the instrument intends to measure, and how test items were generated. Involvement of the target audience and experts in the field is a requirement for item generation. Evidence of content validity was found for three of the six measures, namely, the OASES-S, the TBQ-CS, and one of the three components of the BAB.

Internal Consistency

Internal consistency refers to whether or not the items on a scale are measuring the same or different constructs

Table 2. Details of included measures.

Measure/article/source	Population	Sample size	Age range	Gender ratio (male to female)	No. of items	Domain of interest	Example items	Range of scores (totals)	Response categories
Parent Questionnaire (Erickson & Block, 2013)	Parents of children who stutter	36	Unknown	Unknown	7	Parent perceived impact on the family and impact on the child	Exact wording of questions not available; examples of areas probed include the child's happiness or well-being and the impact of the stutter on the child's life	0–42	Likert 7-point scale and open-ended responses
Teasing/Bullying Questionnaire for Children Who Stutter (Langevin et al., 1998)	School-age children who stutter	28	7–15 years	24:4	11	Teasing/bullying	Have you been teased/bullied about your stuttering at school? How often in the least year at school were you teased/bullied about your stuttering?	0–10	Likert 4-point scale and frequency/ impact subscales
Overall Assessment of the Speaker's Experience of Stuttering for School-Age Children (Yaruss & Quesal, 2010)	School-age children who stutter	75[a]	7–12 years[a]	63:12[a]	60	Reactions to stuttering, communication in daily situations, quality of life, and general information	When you think about stuttering, how often do you feel sad or upset because of your stutter? Overall, how negatively is your life affected by the fact that you stutter?	52–300	Likert 5-point scale
Communication Attitude Test[b]	School-age children who stutter	M	6–15[c]	111:28[c]	33	Speech-associated attitudes	Kids make fun of the way I talk. Many people do not like the way I talk.	0–33	Dichotomous—true or false
Speech Situation Checklist–Emotional Reaction[b]	School-age children who stutter	139[c]	6–15[c]	111:28[c]	55	Emotional reaction to stuttering across a range of situations	Because of your speech, are you afraid to talk with a new kid in school?	55–275	Likert 5-point scale
Behavior Checklist[b]	School-age children who stutter	139[c]	6–15[c]	111:28[c]	50	Coping responses to deal with the stutter	To help your sounds or words come out without trouble: Do you change sounds or words? Do you pretend you do not know the answer to a question?	0–50	Dichotomous—yes or no

Note. Authors acknowledge that data on these measures are not limited to test manuals. M = Multiple sample sizes reported in the test manual from various studies.

[a]Data pertaining to a final standardization study reported in the 2010 test manual. [b]From the Behavior Assessment Battery for School-Age Children Who Stutter (Brutten & Vanryckeghem, 2007). [c]Participant information for the normative data reported in the 2007 test manual.

Table 3. Overview of ratings for each of the measurement properties.

Questionnaire	Content validity	Internal consistency	Construct validity	Agreement (test–retest)	Reliability	Responsiveness	Floor and ceiling effects	Interpretability
Parent Questionnaire (Erickson & Block, 2013)	–	–	?	0	0	0	0	0
TBQ-CS (Langevin et al., 1998)	+	?	?	0	0	0	0	0
OASES-S (Yaruss & Quesal, 2010)								
Section I: General information	+	–	?	+	0	0	?	?
Section II: Your reactions to stuttering	+	?	?	+	0	0	?	?
Section III: Communication in daily situations	+	?	?	+	0	0	?	?
Section IV: Quality of life	+	?	?	+	0	0	?	?
BAB (Brutten & Vanryckeghem, 2007)								
CAT	+	?	+	+	0	?	+	+
SSC-ER	?	?	?	0	0	0	+	?
BCL	?	?	?	0	0	0	?	+

Note. TBQ-CS = Teasing/Bullying Questionnaire for Children Who Stutter; OASES-S = Overall Assessment of the Speaker's Experience of Stuttering for School-Age Children; BAB = Behavior Assessment Battery for School-Age Children Who Stutter; CAT = Communication Attitude Test; SSC-ER = Speech Situation Checklist–Emotional Reaction; BCL = Behavior Checklist.

(Terwee et al., 2007). Studies were required to report Cronbach's alpha scores of .70–.95 and the results of a factor analysis to receive a positive rating. Information was found on Cronbach's alpha scores for all measures. Scores of adequate magnitude were found for the TBQ-CS, the CAT, the SSC-ER, the BCL, and the three most relevant sections of the OASES-S to this topic (your reactions to stuttering, communication in daily situations, and quality of life). The general information section of the OASES-S, which is not exclusive to psychological impacts of stuttering, and the parent questionnaire reported internal consistency reliabilities of .67. This marginally lower value was acknowledged by the authors of the measures. An intermediate rating was assigned for the TBQ-CS, the CAT, the SSC-ER, the BCL, and three out of four sections of the OASES-S, due to the lack of evidence of an appropriate factor analysis.

Construct Validity

Construct validity concerns the correlation of test scores on measures of related and unrelated constructs. To demonstrate adequate construct validity, Terwee et al. (2007) stipulates that studies should report specific hypotheses and provide evidence that a minimum of 75% of the results are congruent with their assumptions. Information on construct validity was found for all of the appraised measures. The CAT was the only measure to receive a positive rating. The TBQ-CS, the OASES-S, the Parent Questionnaire, the SSC-ER, and the BCL all received intermediate ratings. Intermediate ratings were due to the lack of clear evidence of adequate design or unclear hypotheses. It is worth noting here, as other researchers have done (Zuccala et al., 2019), that the use of the intermediate rating criteria for construct validity defined by Terwee et al. (2007) does involve subjective judgments.

Reproducibility—Agreement

Agreement concerns the stability of a measure on repeated administrations (Terwee et al., 2007). Consistent with more commonly used methods for evaluating agreement (Burton et al., 2016; Zuccala et al., 2019), test–retest reliability was used in this study instead of the smallest detectable change score approach of Terwee et al. (2007). To receive a positive rating, authors were required to report a test–retest reliability of $r > .70$ and to provide means and standard deviations at two points in time (Burton et al., 2016; Zuccala et al., 2019). The OASES-S and the CAT scored a positive rating. The OASES-S also included data on standard error of measurement, providing further information on this measurement property. The other measures received a score of 0 because no information was available for this criterion. It is also worth noting that the authors of the TBQ-CS discussed the need for establishing the test–retest reliability of this measure in further development studies.

Reproducibility—Reliability

Reliability is the capacity of a test to differentiate between individual respondents (Terwee et al., 2007). According to the criterion outlined by Terwee et al. (2007), reliability is evidenced by an intraclass correlation coefficient or weighted kappa of at least .70. No information on intraclass correlation coefficient or weighted kappa was found for any of the measures assessed.

Responsiveness

Responsiveness refers to the sensitivity of a measure to identify changes over the course of time (Terwee et al., 2007). Information on responsiveness could only be found for the CAT. An intermediate rating was assigned for the latter measure, owing to the lack of a clear description of this criterion. Despite the number of measures that scored a 0, it may be worth considering the stage of questionnaire development when judging this criterion. For questionnaires in the early stages of development, such as the TBQ-CS and Parent Questionnaire, appraisal of responsiveness could be considered premature.

Floor and Ceiling Effect

A floor or ceiling effect refers to participant scoring that is skewed toward the maximum or minimum scores on an instrument (Franic & Bothe 2008; Terwee et al., 2007). This effect is not desirable because it may indicate inaccuracies in participants' responses (Franic & Bothe 2008). To receive a positive rating for this measurement property, Terwee et al. (2007) stipulated that highest or lowest possible scores are acceptable for 15% of respondents or less. Information regarding this criterion was found for the OASES-S, the CAT, and the SSC-ER. The latter two measures received a positive rating. While the OASES-S manual stipulated that "8 of the 60 items on the OASES-S (13.3%) exhibited a maximum response of 4 on the 5-point scale in the final validation study" (Yaruss & Quesal, 2010, p. 34), the authors could not ascertain which sections of the OASES-S encompassed these eight items. Consequently, an intermediate rating was assigned.

Interpretability

Interpretability concerns the extent to which test scores can be used to derive clinically meaningful information (Terwee et al., 2007). To receive a positive rating, studies were required to report means and standard deviations for a minimum of four subgroups. Evidence of minimal important change scores was not required for a positive rating, keeping in line with adaptations made by previous research teams (e.g., Zuccala et al., 2019).

Interpretability of test scores was available for the OASES-S and all components of the BAB. Results according to gender and age for children who stutter and those who do not stutter were the most commonly reported subgroups. Subgroups help account for heterogenous data and assist clinicians with making more accurate normative assessments. Overall ratings for these criteria ranged from intermediate to positive, based on the level of evidence available.

Discussion

Measurement instruments to assess the psychological impacts of stuttering in young school-age children are needed to ensure their mental health needs, if present, can be adequately met. Without adequate assessment, clinicians cannot make informed judgments about how to proceed with client care. The results of this systematic review revealed a dearth of available measures whose development was reported in peer-reviewed journals. Of the measures identified in the title and abstract screening, many did not meet the inclusion criterion for full-text screening because the participants were non–English-speaking children. Due to the limited initial findings, the inclusion of measures for appraisal was extended to those flagged in the screening phase for which test manuals were readily available. Even with this leniency, only six instruments were identified, namely, the Parent Questionnaire, the TBQ-CS, the OASES-S, the CAT, the SSC-ER, and the BCL.

Assessment tools are only clinically useful to the extent that they are reliable, valid, and sensitive to change. Consequently, a primary purpose of this study was to critically appraise the properties of the measures identified in the review. The intention of doing so was to identify robust findings that clinicians and researchers could use with children who stutter. Eight of the nine criteria outlined by Terwee et al. (2007) were used to evaluate the properties of the six included measures. As previously detailed, criterion validity was excluded from the evaluation in the absence of a gold-standard instrument for measuring psychological impacts of stuttering in young school-age children. No measure received a positive score for all eight criteria. Four of the six measures received two or fewer positive evaluations. All sections of the OASES-S received two positive ratings, and the CAT received five positive ratings. However, the developers of these two measures did not provide strong evidence of the responsiveness of their measures to clinical change, which is a critical factor when they are used in a clinical setting. Accordingly, at present, it is difficult to recommend any single measure for general use in the population of children who stutter. In broad terms, the psychometric properties of available measures and the strategies used in their development were lacking.

Several potential limitations of this study need to be acknowledged and explored. First, limitations in database coverage could have impacted the results of the study. For example, the CINAHL database only indexes articles in the *Contemporary Issues in Communication Science and Disorders* from 2004 onward. The TBQ-CS discussed in this review was from the 1998 issue of the journal. Consequently, it was not captured in the initial pool of 1,771 results but was identified through manual searching. Having said this, the authors attempted to limit issues of database coverage by searching across multiple databases, seven in total.

A second limitation may be the restriction of scales to those whose development was reported in peer-reviewed journal articles. For example, there may be other scales whose development was reported in conference proceedings and textbooks. Weighing the potential loss of scales in the review against the absence of peer review in their development is not easy. However, the present authors argue that the integrity of the review is strengthened by excluding conference proceedings and textbooks. The method chosen is consistent with several other reviews of scales in other domains (Burton et al., 2016; Stein et al., 2017; Zuccala et al., 2019). Furthermore, as a compromise to include more scales,

we did include measures whose psychometric properties were described in a publicly available manual.

A third limitation may be the extreme rigor of the quality appraisal tool used for this review. The quality criteria outlined by Terwee et al. (2007) have been criticized as being overly strict and at risk of inflating false-negative evaluations (Reneman et al., 2010). This may have contributed to the high number of intermediate ratings in the present review. Having said this, the authors maintain that a rigorous appraisal tool is necessary if we are to increase the quality of scale development in the field of stuttering over time. The Terwee et al. criteria are becoming more widely used, and they highlight the critical considerations that researchers need to be aware of in developing their scales.

Regardless of the potential limitations detailed above, a lack of existing measures to identify psychological impacts of stuttering in young school-age children was evident. The scarcity of available measures may be due to historical beliefs surrounding the causes and consequences of stuttering. It was once argued that drawing attention to a child's speech through self-report measures could worsen stuttering (Brutten & Vanryckeghem, 2007) and that negative psychosocial consequences of stuttering only appeared in adulthood (Guttormsen et al., 2015). Accordingly, the development of nonspeech measures may not have been at the forefront of the research objectives of most scholars in this area. Many treatment approaches used with school-age children who stutter focus almost exclusively on reductions in stuttering severity (Andrews et al., 2016; Koushik et al., 2009; Lincoln et al., 1996; Onslow et al., 1997). Given this, it is not surprising that the development of measures of psychological difficulties in this population have lagged behind.

Clinical and Research Implications

There are several implications that arise from this review. First, it is clear that clinicians and researchers need to be discerning when selecting assessment tools to measure the psychosocial impacts of stuttering on young school-age children. SLPs are encouraged to consider how rigorously measures have been constructed and tested before selecting assessment tools to use in clinical practice or for research purposes. This information should be readily available in test manuals or publications about the measurement tool of interest. Second, given the complexity of factors to consider when evaluating the quality of assessment tools, it is recommended that SLPs use an established framework to guide their evaluations. The present authors chose to use the Terwee et al. (2007) appraisal tool, for reasons previously outlined, but acknowledge that other quality appraisal resources may serve the equivalent purpose. Third, researchers in the field of stuttering are encouraged to direct their efforts toward the development of new and robust measures to assess the impact of stuttering in the vulnerable school-age years. There is clearly a dearth of available instruments. Fourth, it should be acknowledged that formal questionnaires do not replace informal methods of assessment. Informal methods include, but are not limited to, information gained through clinical observation and open channels of communication with the child, their caregiver(s), and teachers throughout the treatment journey.

Conclusions

To our knowledge, this is the first review to have systematically explored and evaluated measures designed to assess the psychological impacts of stuttering in the school-age population. The results highlight lack of available measures in this domain, as well as poor practices in developing and testing measurement instruments. No measure included in this review received a positive rating for all criteria outlined by the Terwee et al. (2007) appraisal tool. Despite this, the CAT and the OASES-S were found to have the most evidence in support of their measurement properties. In our view, the field would benefit if the results of this review spur future research to comprehensively plan, test, and document the development and testing of new measures to assess the psychosocial problems of young school-age children who stutter.

Acknowledgments

This research was supported in part by National Health and Medical Research Council Program Grant 1132370 (awarded to Onslow, Packman, and Menzies). The authors would like to thank Damien Liu-Brennan for his scientific copyediting contribution to this publication. The authors would also like to thank Sarah Su (Graduate School of Health Librarian, University of Technology Sydney).

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Andrews, C., O'Brian, S., Onslow, M., Packman, A., Menzies, R., & Lowe, R. (2016). Phase II trial of a syllable-timed speech treatment for school-age children who stutter. *Journal of Fluency Disorders, 48,* 44–55. https://doi.org/10.1016/j.jfludis.2016.06.001

Blumgart, E., Tran, Y., & Craig, A. (2010). An investigation into the personal financial costs associated with stuttering. *Journal of Fluency Disorders, 35*(3), 203–215. https://doi.org/10.1016/j.jfludis.2010.03.002

Boyle, C. A., Decoufle, P., & Yeargin-Allsopp, M. (1994). Prevalence and health impact of developmental disabilities in U.S. children. *Pediatrics, 93,* 399–403.

Boyle, M. P. (2018). Enacted stigma and felt stigma experienced by adults who stutter. *Journal of Communication Disorders, 73,* 50–61. https://doi.org/10.1016/j.jcomdis.2018.03.004

Brutten, G. J., & Vanryckeghem, M. (2007). *Behavior Assessment Battery for School-Age Children Who Stutter*. Plural.

Burton, A. L., Abbott, M. J., Modini, M., & Touyz, S. (2016). Psychometric evaluation of self-report measures of binge-eating symptoms and related psychopathology: A systematic review of the literature. *The International Journal of Eating Disorders, 49*(2), 123–140.

Carleton, R. N., McCreary, D. R., Norton, P. J., & Asmundson, G. J. G. (2006). Brief Fear of Negative Evaluation Scale–Revised. *Depression and Anxiety, 23*(5), 297–303.

De Sonneville-Koedoot, C., Stolk, E., Rietveld, T., & Franken, M. C. (2015). Direct versus indirect treatment for preschool children who stutter: The RESTART randomized trial. *PLOS ONE, 10*(7), e0133758. https://doi.org/10.1371/journal.pone.0133758

Erickson, S., & Block, S. (2013). The social and communication impact of stuttering on adolescents and their families. *Journal of Fluency Disorders, 38*(4), 311–324. https://doi.org/10.1016/j.jfludis.2013.09.003

Franic, D. M., & Bothe, A. K. (2008). Psychometric evaluation of condition-specific instruments used to assess health-related quality of life, attitudes, and related constructs in stuttering. *American Journal of Speech-Language Pathology, 17*(1), 60–80. https://doi.org/10.1044/1058-0360(2008/006)

Gerlach, H., Totty, E., Subramanian, A., & Zebrowski, P. (2018). Stuttering and labor market outcomes in the United States. *Journal of Speech, Language, and Hearing Research, 61*(7), 1649–1663. https://doi.org/10.1044/2018_JSLHR-S-17-0353

Guttormsen, L. S., Kefalianos, E., & Næss, K. A. B. (2015). Communication attitudes in children who stutter: A meta-analytic review. *Journal of Fluency Disorders, 46,* 1–14. https://doi.org/10.1016/j.jfludis.2015.08.001

Helgadóttir, F. D., Menzies, R. G., Onslow, M., Packman, A., & Brian, S. O. (2009). Online CBT II: A Phase I trial of a standalone, online CBT treatment program for social anxiety in stuttering. *Behaviour Change, 26*(4), 254–270. https://doi.org/10.1375/bech.26.4.254

Iverach, L., Jones, M., McLellan, L. F., Lyneham, H. J., Menzies, R. G., Onslow, M., & Rapee, R. M. (2016). Prevalence of anxiety disorders among children who stutter. *Journal of Fluency Disorders, 49,* 13–28. https://doi.org/10.1016/j.jfludis.2016.07.002

Iverach, L., Menzies, R., Jones, M., O'Brian, S., Packman, A., & Onslow, M. (2011). Further development and validation of the Unhelpful Thoughts and Beliefs About Stuttering (UTBAS) Scales: Relationship to anxiety and social phobia among adults who stutter. *International Journal of Language & Communication Disorders, 46*(3), 286–299.

Iverach, L., O'Brian, S., Jones, M., Block, S., Lincoln, M., Harrison, E., Hewat, S., Menzies, R. G., Packman, A., & Onslow, M. (2009). Prevalence of anxiety disorders among adults seeking speech therapy for stuttering. *Journal of Anxiety Disorders, 23*(7), 928–934. https://doi.org/10.1016/j.janxdis.2009.06.003

Kefalianos, E., Onslow, M., Ukoumunne, O., Block, S., & Reilly, S. (2014). Stuttering, temperament, and anxiety: Data from a community cohort ages 2–4 years. *Journal of Speech, Language, and Hearing Research, 57*(4), 1314–1322. https://doi.org/10.1044/2014_JSLHR-S-13-0069

Klein, J. F., & Hood, S. B. (2004). The impact of stuttering on employment opportunities and job performance. *Journal of Fluency Disorders, 29*(4), 255–273. https://doi.org/10.1016/j.jfludis.2004.08.001

Koushik, S., Shenker, R., & Onslow, M. (2009). Follow-up of 6- to 10-year-old stuttering children after Lidcombe Program treatment: A Phase I trial. *Journal of Fluency Disorders, 34*(4), 279–290. https://doi.org/10.1016/j.jfludis.2009.11.001

Langevin, M., Bortnick, K., Hammer, T., & Wiebe, E. (1998). Teasing/bullying experienced by children who stutter: Toward development of a questionnaire. *Contemporary Issues in Communication Science and Disorders, 25,* 12–24. https://doi.org/10.1044/cicsd_25_S_8

Langevin, M., Packman, A., & Onslow, M. (2009). Peer responses to stuttering in the preschool setting. *American Journal of Speech-Language Pathology, 18*(3), 264–276. https://doi.org/10.1044/1058-0360(2009/07-0087)

Lincoln, M., Onslow, M., Lewis, C., & Wilson, L. (1996). A clinical trial of an operant treatment for school-age children who stutter. *American Journal of Speech-Language Pathology, 5*(2), 73–85. https://doi.org/10.1044/1058-0360.0502.73

Longland, A. (2009). Behavior Assessment Battery for School-Age Children Who Stutter. *International Journal of Language & Communication Disorders, 44,* 1075–1076. https://doi.org/10.3109/13682820802682996

Manning, W., & Beck, J. G. (2011). Comments concerning Iverach et al., Screening for personality disorders among adults seeking speech treatment for stuttering [J. Fluency Disorders 34 (2009) 173–186]. *Journal of Fluency Disorders, 36*(1), 61–65. https://doi.org/10.1016/j.jfludis.2010.09.001

Manning, W., & Beck, J. G. (2013). Personality dysfunction in adults who stutter: Another look. *Journal of Fluency Disorders, 38*(2), 184–192. https://doi.org/10.1016/j.jfludis.2013.02.001

Menzies, R. G., O'Brian, S., Lowe, R., Packman, A., & Onslow, M. (2016). International Phase II clinical trial of CBTPsych: A standalone Internet social anxiety treatment for adults who stutter. *Journal of Fluency Disorders, 48,* 35–43. https://doi.org/10.1016/j.jfludis.2016.06.002

Menzies, R. G., O'Brian, S., Packman, A., Jones, M., Helgadóttir, F. D., & Onslow, M. (2019). Supplementing stuttering treatment with online cognitive behavior therapy: An experimental trial. *Journal of Communication Disorders, 80,* 81–91. https://doi.org/10.1016/j.jcomdis.2019.04.003

Menzies, R. G., Packman, A., Onslow, M., O'Brian, S., Jones, M., & Helgadóttir, F. D. (2019). In-clinic and standalone Internet cognitive behavior therapy treatment for social anxiety in stuttering: A randomized trial of iGlebe. *Journal of Speech, Language, and Hearing Research, 62*(6), 1614–1624. https://doi.org/10.1044/2019_JSLHR-S-18-0340

Modini, M., Abbott, M. J., & Hunt, C. (2015). A systematic review of the psychometric properties of trait social anxiety self-report measures. *Journal of Psychopathology and Behavioral Assessment, 37,* 645–662. https://doi.org/10.1007/s10862-015-9483-0

O'Brian, S., Jones, M., Packman, A., Menzies, R., & Onslow, M. (2011). Stuttering severity and educational attainment. *Journal of Fluency Disorders, 36*(2), 86–92. https://doi.org/10.1016/j.jfludis.2011.02.006

Onslow, M. (2021). *Stuttering and its treatment: Eleven lectures.* https://www.uts.edu.au/asrc/resources

Onslow, M., Jones, M., Menzies, R., O'Brian, S., & Packman, A. (2012). Stuttering. In P. Sturmey & M. Hersen (Eds.), *Handbook of evidence-based practice in clinical psychology: Vol. 1. Child and adolescent disorders* (pp. 185–207). Wiley. https://doi.org/10.1002/9781118156391.ebcp001008

Onslow, M., Packman, A., Stocker, S., van Doorn, J., & Siegel, G. M. (1997). Control of children's stuttering with response-contingent time-out. *Journal of Speech, Language, and Hearing Research, 40*(1), 121–133. https://doi.org/10.1044/jslhr.4001.121

Packman, A., Onslow, M., Richard, F., & van Doorn, J. (1996). Syllabic stress & variability: A model of stuttering. *Clinical Linguistics & Phonetics, 10*(3), 235–263. https://doi.org/10.3109/02699209608985174

Reneman, M. F., Dijkstra, A., Geertzen, J. H. B., & Dijkstra, P. U. (2010). Psychometric properties of chronic pain acceptance questionnaires: A systematic review. *European Journal of Pain, 14*(5), 457–465. https://doi.org/10.1016/j.ejpain.2009.08.003

Speech Pathology Australia. (2017). *Clinical Guideline: Stuttering Management for Speech Pathologists.* The Speech Pathology Association of Australia Ltd. https://www.speechpathologyaustralia.org.au/SPAweb/Members/Position_Statements.aspx?hkey=b1a46941-246c-4609-bacc-1c1b5c52d19d#tp

Stein, J. E., Modini, M., Hunt, C., & Abbott, M. J. (2017). A systematic review of the psychometric properties of trait cognitive self-report measures in social anxiety. *Journal of Psychopathology and Behavioral Assessment, 39,* 147–163. https://doi.org/10.1007/s10862-016-9573-7

Teeson, K., Packman, A., & Onslow, M. (2003). The Lidcombe behavioral data language of stuttering. *Journal of Speech, Language, and Hearing Research, 46*(4), 1009–1015. https://doi.org/10.1044/1092-4388(2003/078)

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60*(1), 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

Van Borsel, J., Brepoels, M., & De Coene, J. (2011). Stuttering, attractiveness and romantic relationships: The perception of adolescents and young adults. *Journal of Fluency Disorders, 36*(1), 41–50. https://doi.org/10.1016/j.jfludis.2011.01.002

Van Riper, C. (1982). *The nature of stuttering* (2nd ed.). Prentice Hall.

Yairi, E. (2004). The formative years of stuttering: A changing portrait. *Contemporary Issues in Communication Science and Disorders, 31,* 92–104. https://doi.org/10.1044/cicsd_31_S_92

Yairi, E., & Ambrose, N. (1992). Onset of stuttering in preschool children. *Journal of Speech and Hearing Research, 35*(4), 782–788. https://doi.org/10.1044/jshr.3504.782

Yaruss, J. S., & Quesal, R. W. (2010). *Overall Assessment of the Speaker's Experience of Stuttering (OASES)*. Pearson Assessments.

Zuccala, M., Menzies, R. E., Hunt, C. J., & Abbott, M. J. (2019). A systematic review of the psychometric properties of death anxiety self-report measures. *Death Studies*. Advance online publication. https://doi.org/10.1080/07481187.2019.1699203