

Applicability of Neuropsychological and Psychometric Tests in Autosomal Recessive Spastic Ataxia of Charlevoix-Saguenay (ARSACS)

Kevin Brassard¹, Julie Bouchard^{1,2}, Geneviève Forgues¹, Alexe Boivin-Mercier¹, & Cynthia Gagnon^{2,3}

¹Université du Québec à Chicoutimi

²Groupe de recherche interdisciplinaire sur les maladies neuromusculaires, CIUSSS du Saguenay-Lac-St-Jean

³Faculté de médecine et des sciences de la santé, Université de Sherbrooke, Site Saguenay - Centre de recherche Charles-Le-Moyne

Autosomal Recessive Spastic Ataxia of Charlevoix-Saguenay (ARSACS) is a degenerative spinocerebellar disease with pyramidal, cerebellar, and neuropathic impairments. Recent studies highlight possible deficits in cognitive functions like language. Psychometric tests selection implies careful consideration due to upper limbs incoordination and dysarthria. The objective of this study is to document the applicability of 37 neuropsychological and 2 psychological tests in 8 individuals with ARSACS aged between 20 and 60 years. All tests were rated on 4 applicability criteria using a 3-level rating scale: *A* for *excellent*; *B* for *acceptable*; *C* for *reconsider*. Most tests posed few or no applicability limits with ARSACS patients. However, certain tests (e.g., Leiter-3 and Raven's Standard Progressive Matrices) are not recommended due to significant issues related to applicability. These results may help clinicians and researchers working with this population to select evaluations and tests applicable in this population.

Keywords: Autosomal Recessive Spastic Ataxia of Charlevoix-Saguenay, metrological properties, test administration, cognition, personality

L'ataxie récessive spastique autosomique de Charlevoix-Saguenay (ARSACS) est une maladie neurodégénérative spinocérébelleuse impliquant des atteintes pyramidales, cérébelleuses et neuropathiques et de possibles déficits dans certaines fonctions cognitives comme le langage. La sélection des tests psychométriques appropriés implique certaines précautions liées à la présence d'incoordination des membres supérieurs et de dysarthrie. L'objectif de cette étude consiste à documenter l'applicabilité de 37 tests neuropsychologiques et de 2 tests psychologiques chez 8 individus atteints d'ARSACS (20-60 ans). Les tests ont été évalués en fonction de 4 critères d'applicabilité en utilisant une échelle à 3 niveaux : *A* pour *excellent*; *B* pour *acceptable*; *C* pour *reconsidérer*. La plupart des tests ont posé peu ou pas de limites liées à l'applicabilité chez les patients atteints d'ARSACS. Cela dit, certains tests (c.-à-d., le Leiter-3 et les Matrices Progressives de Raven) ne sont pas recommandés en raison de problèmes liés à leur applicabilité. Ces résultats pourront venir en aide aux cliniciens et chercheurs qui travaillent auprès de cette population en vue de la sélection d'évaluations et de tests.

Mots clés : Ataxie spastique autosomique de Charlevoix-Saguenay, propriétés métrologiques, administration de test, cognition, personnalité

Autosomal Recessive Spastic Ataxia of Charlevoix-Saguenay (ARSACS) is a rare neuromuscular genetic disease with the highest prevalence in the province of Quebec, Canada.

This research was supported by the Groupe de recherche interdisciplinaire sur les maladies neuromusculaires (GRIMN) affiliated to the CIUSSS du Saguenay-Lac-Saint-Jean. Financial support was provided by the Ataxia Charlevoix-Saguenay Foundation. Special thanks are extended to Janie Gauthier-Boudreau and Stéphane Bergeron for their assistance.

Correspondence regarding this article should be addressed to / La correspondance concernant cet article devrait être adressée à : Kevin Brassard, Université du Québec à Chicoutimi.

555, Boulevard de l'Université,
Chicoutimi, QC, Canada, G7H 2B1
Tel.: 418-720-2092
E-mail: kevin.brassard1@uqac.ca

Following the recent description of more than 140 known mutations worldwide, it is now described as one of the most common forms of recessive spastic ataxia (Synofzik et al., 2013; Thiffault et al., 2013). Approximately 93% of the affected individuals from Quebec are homozygous for one of the two founder variants of the mutation (i.e., c.8844delT; Vermeer, Van de Warrenburg, & Kamsteeg, 2003). It is a degenerative spinocerebellar disease with a slow progression. The first symptoms are noticed at walking onset: family members report problems with balance and frequent falls (Bouchard, Barbeau, Bouchard, & Bouchard, 1978; Bouchard et al., 1998). The evolution of the disease is obvious in the late teens and early twenties. It is characterized by

pyramidal (spasticity and weakness), cerebellar (ataxia), and distal neuropathic (amyotrophy and distal weakness) signs of variable intensity and progression (Bouchard et al., 1978; Bouchard et al., 1998; Gazulla, Mayayo-Sinués, Benavente, Modrego, & Berciano, 2014; Vermeer, Van de Warrenburg, & Kamsteeg, 2003). Besides that, cerebral imagery shows atrophy of the superior cerebellar vermis (Bouchard et al., 1978; Bouchard et al., 1998; Mignarri et al., 2014; Pilliod et al., 2015) and cerebellar hemispheres (Martin, Bouchard, Sylvain, St-Onge, & Truchon, 2007; Pilliod et al., 2015; Vermeer et al., 2003). A thickening of inferior cerebellar vermis is also noticed as well as a progressive cortico-cerebellar atrophy (Bouchard et al., 1998). Generally, cerebellar infringements are known to cause perturbations of sensory-motor functions (Rae-Grant & Parsons, 2014), but increasing evidences suggest they may contribute to cognitive function deficits in tasks where motor skills are not involved (Fancellu et al., 2013; Pinel, 2007; Schmahmann, 2004). However, in order to ensure validity of measurement, selection of psychometric tests for neuropsychological and psychological evaluation with this population should be done with particular attention to upper limbs incoordination, spasticity, ataxia, and neuropathic signs, as they may bias tests interpretation.

Cognitive Functions and Personality in ARSACS

Since its first description in 1978, few studies have investigated cognitive functioning in ARSACS. Bouchard et al. (1978) were the firsts to gather information on intellectual functioning in adults with ARSACS. More than 20 years later, studies were conducted with children and adolescents to explore specific cognitive domains such as visuoperceptive functions, planification abilities, and visual attention (Drolet, 2002; Lévesque, 2004). In this regard, these authors suggested preservation of the first two functions but significant difficulties with visual attention. Drolet (2002) also discusses indicators of decreased processing speed and deficits with visual sequential memory and auditivo-verbal learning. More recently, three collaborative multiple case studies (Boivin-Mercier, n.d.; Boucher, 2017; Desmeules, 2017) were conducted with eight ARSACS patients aged between 20 and 60 years. These studies aimed to explore ARSACS cognitive and personality profiles. Concerning the cognitive profile, Boivin-Mercier (n.d.) evaluated four participants aged between 20 and 39 years. Results from this part of the collaborative studies are not available at this time. For her part, Boucher (2017) investigated cognitive profiles in four participants aged between 40 and 60 years old. Results indicated four affected cognitive functions: processing speed, sustained attention, language, and visual logical

reasoning. For other cognitive functions investigated, the author reported inconsistent results between participants. For instance, performances on working memory tasks diverged significantly depending on sensory modality. Indeed, normative results were noted in visual modality, but deficits were observed in auditory modality. Regarding personality, there seems to have no homogeneous profile among the eight participants (Desmeules, 2017). However, results revealed a tendency regarding the openness trait according to the Five-Factor Model of Personality (Costa & Widiger, 1994). The authors observed a low openness usually reflecting a rigidity of thinking, a low openness to new or different values and ideas, a difficulty in recognizing the emotions of others, and a tendency to pragmatism.

These exploratory studies on cognitive and personality profile in ARSACS thus gave preliminary insights on this matter. Even if validity and fidelity of psychometric tests used in these studies have been demonstrated, practical considerations must be addressed to ensure feasibility of their use in future clinical and research contexts (e.g., administration time, utility of clinical information gathered, adaptation to ARSACS patient's characteristic). Before going into larger studies to describe central nervous involvement, a careful selection of neuropsychological tests is necessary as they require motor or verbal responses (Seron & Van der Linden, 2014) that may be influenced by the presence of incoordination and dysarthria. It is critical in order to minimize interpretative bias and optimize reliability of the measure. When selecting psychological and neuropsychological psychometric tests most adapted to evaluation needs and patients' features, researchers' and clinicians' expertise with ARSACS patients is an important source of information. However, an objective evaluation guided by specific empirical applicability criteria would allow the establishment of selection indicators for psychometric tests used with ARSACS patients.

Applicability of Measurement Tools

The pragmatic and psychometric properties must be considered in the selection of the most appropriate measurement tools in a given context (Andresen, 2000; Auger et al., 2007). Beyond the importance of verifying their validity, Wade (2004) insists on their practical aspects (e.g., time taken, need for training, acceptability to patients). In a literature search on operational criteria used for this purpose, Auger, Demers, and Swaine (2006) proposed the term "applicability" to englobe a panel of pragmatic qualities such as clinical utility, sensibility, feasibility, acceptability, and practicability. In 2007, Auger et al. tested this conceptual framework to assess the

applicability of a multidimensional toolkit for geriatric rehabilitation outcomes. Seedat, James, and Rose (2014) also exploited this model to evaluate 19 tests intended for populations with muscle diseases. The model proposes four major criteria comprising sub-criteria to evaluate the applicability of a measurement tool (cf. Appendix A). *Respondent burden* corresponds to invasiveness (physical or emotional), global effort required, security, refusal, and acceptability from the respondent's point of view. *Examiner burden* comprises different components (administration time, need for training, simplicity of administration, scoring and interpretation, availability, etc.) to be considered as qualities or limitations in the application of the measurement tool. *Score distribution* refers to how appropriate is a scale to represent the diversity of behaviours, opinions, or performance levels in a sample (i.e., normal curve and floor/ceiling effect). Finally, *format compatibility* is defined by the consistency between population specific characteristics and assessment format in order to avoid age, gender, cultural, or disability biases (Andresen, 2000). The present study proposes this conceptual framework to evaluate applicability of psychometric tests used with ARSACS patients. Added to validity and fidelity, data on psychometric tests applicability will increase clinicians' and researchers' confidence in choosing the most suitable neuropsychological and personality tests with this population.

Objectives

The present study is closely attached to the exploratory multiple case studies on cognitive and personality profiles in individuals with ARSACS (Boivin-Mercier, n.d.; Boucher, 2017; Desmeules, 2017). In these studies, a wide variety of psychometric instruments have been used to provide a first extensive picture of cognitive and personality features in individuals with the disease. The objective of this paper is to document the applicability of the instruments used in the context of these exploratory multiple case studies with this specific population for clinical and research purposes. To our knowledge, the present study is the first to assess applicability of neuropsychological and personality tests in a specific population. This will facilitate the selection by the clinicians of the most adapted tests for patients with ARSACS.

Method

Sample

Eight patients with ARSACS were selected from three ongoing collaborative studies (Boivin-Mercier, n.d.; Boucher, 2017; Desmeules, 2017) that recruited twenty patients among a larger subset of people listed

at the Neuromuscular Clinic of the *Centre Intégré Universitaire de Santé et de Services Sociaux (CIUSSS) du Saguenay-Lac-Saint-Jean* (Quebec, Canada). The recruitment was done according to a stratified sample by age and gender to get a more complete picture of the cognitive functions considering the progression of the disease. Thus, one man and one woman were selected in each of the four age categories (21–30; 31–40; 41–50; 51–60). Table 1 presents participants' sociodemographic characteristics (e.g., age, education, occupation, mobility).

Procedure

The present research assessed the exact same tests used in the above-mentioned exploratory multiple case studies. Consequently, initial tests selection depended on the research designs and objectives of these studies. Tests assessed in the present study come from two sources. First, as presented in Table 2, specific subtests included in four psychometric batteries were selected and administered to participants. For instance, 10 of the 15 Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV, one of the four batteries) subtests were initially selected in the exploratory multiple case studies. Thereby, a total of 28 subtests from either one or the other of the four batteries were chosen. Second, 11 individual tests (i.e., not included in a psychometric battery, such as Rey Tangled Lines Test or Hooper Visual Organization Test) were used. Consequently, a total of 39 psychometric tests were administered to the participants during the summer of 2014 by a doctoral student in clinical neuropsychology who was previously trained to their use. Globally, original administration instructions were followed. However, some specific tests required adaptations from the examiner, especially concerning the nonverbal measures (e.g., Leiter-3) or the personality tests, where instructions were read to some participants. Deviations to original instructions were considered in applicability analysis. All participants were assessed at the Neuromuscular Clinic. Considering the number of tests to be administered, the evaluation was spread over three or four half-days, generally over three weeks. The administration order was sometimes adapted depending on completion time, exhaustion, and difficulty level. When required, participants could take breaks to minimize the impacts of fatigue and decrease in motivation.

Description of the Tools

Wechsler Adult Intelligence Scale–Fourth edition (WAIS-IV). The WAIS-IV (Psychological Corporation, 2008) measures general intelligence through the administration of numerous subtests. Ten out of fifteen subtests were used in the present research (Block Design, Digit Span, Matrix

Reasoning, Arithmetic, Symbol Search, Visual Puzzles, Coding, Figure Weights, Cancellation, and Picture Completion) to gather an overview of intellectual functioning according to three of the four specific cognitive domains: Perceptual Reasoning Index (PRI), Working Memory Index, and Processing Speed Index (PSI; Weiss, Saklofske, Coalson, & Raiford, 2010).

Neuropsychological Assessment Battery (NAB).

The NAB is a battery of neuropsychological tests developed for the assessment of a wide array of cognitive skills and cognitive functions in adults (White & Stern, 2003). Six subtests of the NAB (French home translation of the instructions) were used in this research: Dots, Auditory Comprehension, Naming, Shape Learning, Visual Discrimination, and Judgement.

Leiter-3. The Leiter-3 (Roid, Miller, Pomplun, & Koch, 2013) provides a nonverbal measure of intellectual functioning. It measures the cognitive, attentional, and neuropsychological abilities through the exclusive use of nonverbal means of communication. The instructions are delivered using gestures and pantomime. The battery was specifically developed for populations with communication

disorders, hearing and motor impairments, traumatic brain injuries, attentional problems, and learning disabilities (Tsatsanis et al., 2003). Nine subtests were used in the present study: Figure Ground, Form Completion, Classification/Analogies, Sequential Order, Visual Patterns, Forward Memory, Attention Sustained, Reverse Memory, and Nonverbal Stroop.

Delis-Kaplan Executif Function System (D-KEFS). The D-KEFS (Delis, Kaplan, & Kramer, 2001) is a battery of tests used to evaluate executive functions such as reaction time, planning of oculomotor actions, selective attention, cognitive flexibility, and inhibition abilities (Mattioli et al., 2014). In this study, three subtests were administered: the Trail Making Test (TMT), the Tower of London Test (TLT), and the Color-Word Interference Test (CWIT).

Praxis Test of Jason Brown (PTJB). This test evaluates praxis through four parts: intransitive gestures, transitive gestures, whole body gestures, and bimanual coordination. It includes a list of actions the participants must execute (e.g., brushing their hair, cutting paper). Observations as well as clinical judgment of the evaluator are used to determine the level of capacity of the participant (Brown, 1972).

Table 1

Sociodemographic Characteristics of Participants

Participants	1	2	3	4	5	6	7	8
Gender	Male	Female	Male	Female	Female	Male	Female	Male
Age (years)	27	23	33	32	45	49	52	56
Education (years)	8	12-13	14	12-13	9	10	14	12
Marital status	Single	Single	Single	Common-law relationship	Single	Single	Single	Single
Occupation	Part-time worker and student	Part-time worker	At home	Full-time student	At home	At home	At home	At home
Type of house	House	House	House	Semi-detached house	House	Adapted apartment (> six unit)	House	Adapted apartment (> six unit)
Inside mobility	Unaided	Unaided	Unaided	With a walker	With a walker	Wheel-chair	Wheel-chair	Wheel-chair
Outside mobility	Cane	Unaided	With a walker	With a walker	With a walker	Mobility scooter	Wheel-chair	Mobility scooter
Autonomy level	Independent	Independent	Independent	Independent	Partial assistance required	Independent (with 24h assistance available)	Partial assistance required	Complete assistance required
Perception of physical health	Sufficient	Sufficient	Good	Good	Sufficient	Good	Good	Sufficient

Motor-Free Visual Perception Test—Third edition (MVPT-3). The MVPT-3 (Colarusso & Hammill, 2002) assesses an individual's visual perceptual ability without involving motor response. Visual perceptual tasks included in this test are spatial relations, visual discrimination, figure-ground, visual closure, and visual memory. The test provides a single score representing general visual perceptual abilities (Colarusso & Hammill, 1972).

Table 2

Subtests Included in the Four Batteries

Batteries	Subtests
WAIS-IV	Block Design Digit Span Matrix Reasoning Arithmetic Symbol Search Visual Puzzles Coding Figure Weights Cancellation Picture Completion
NAB	Dots Auditory Comprehension Naming Shape Learning Visual Discrimination Judgement
Leiter-3	Figure Ground Form Completion Classification and Analogies Sequential Order Visual Patterns Forward Memory Sustained Attention Reverse Memory Nonverbal Stroop
D-KEFS	Trail Making Test Tower of London Test Color-Word Interference Test

Rey Tangled Lines Test (RTL). Designed by André Rey in 1956, this test measures oculomotor ability. Respondents are required to visually track a line from one side of the page to the other. The line courses following a random trajectory against a field of other lines, some of which constitute further trials of the test (Senior, Kelly, & Salzman, 1999).

Raven's Standard Progressive Matrices (RSPM). The aim of RSPM test (Raven, Raven, & Court, 2000) is to assess visual reasoning. Tasks, ordered by increasing difficulty, are completed using logical reasoning, which evaluates fluid intelligence

(Verguts, De Boeck, & Maris, 1999). The items require participants to analyze graphic elements in a matrix in order to select the correct solution out of eight alternative responses (Loesche, Wiley, & Hasselhorn, 2015).

Conners' Continuous Performance Test II (CPT-II). The CPT-II (Conners, 1994) is a computerized test measuring sustained attention, reaction time, and motor impulsivity (Lange et al., 2013). Targets (letters) appear at inconsistent occurrences. The participant has to press the space bar when a target appears, except when it is an "X".

Hooper Visual Organization Test (HVOT). The HVOT (Hooper, 1958) is a brief screening tool measuring the ability to organize visual stimuli. This test is also used to assess brain dysfunction. It consists in a number of cut up pictures of an object that the participant must mentally rearrange in order to name it precisely (Lopez, Lazar, & Oh, 2003).

Benton Facial Recognition Test (BFRT). The BFRT (Benton, Sivan, Hamsher, Varney, & Spreen, 1983) assesses facial recognition skills using black and white photographs. In the first part of the test, participants must identify a target face out of six options. In the second part of the test, participants have to identify three out of six photographs corresponding to the target face (Schretlen, Pearlson, Anthony, & Yates, 2001).

Wisconsin Card Sorting Test (WCST). The WCST (Heaton, Chelune, Talley, Kay, & Curtiss, 1993) was developed as a measure of executive function, specifically conceptual deduction abilities. It also provides an overview of the capacity to apply a strategy and to adapt behaviour in context of changing conditions (Eling, Derckx, & Maes, 2008). In this test, the respondent has to pair stimuli cards with one of the four basic cards according to a rule known only by the examiner, who can only tell the participant if the match is good or not. The respondent's ability to use environmental feedback and cognitive flexibility are essential to achieve the task (Heaton, 1993).

Continuous Visual Memory Test (CVMT). The CVMT (Larrabee, Trahan, & Curtiss, 1992) measures the ability to learn and remember visually presented information without the use of motor skills. The test is a recognition task that was developed in response to the difficulties that have been identified with many of the traditional visual memory measures (Strong & Donders, 2008).

The Symptom Checklist-90-R (SCL-90-R). The SCL-90-R (Derogatis, 1994) is a screening tool used to detect self-reported psychological distress and psychopathological symptoms (Derogatis & Savitz,

2000). The test consists of three global indices and nine subscales covering clinically relevant psychiatric and psychosomatic symptoms (Derogatis & Melisaratos, 1983). Among other things, the SCL-90 assesses obsessive-compulsive disorder, anxiety, depression, paranoid ideation, and phobic anxiety.

NEO Five-Factor Inventory (NEO-FFI). The NEO-FFI is the short version of the NEO Personality Inventory (Costa & McCrae, 1989; McCrae & Costa, 2010). It is a self-reported questionnaire measuring the Five-Factor Model (FFM) through 60 items consisting of short sentences based upon specific traits known to be related to the five following factors: Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness (McCrae & Costa, 2007).

Applicability Criteria Scoring

The examiner scored applicability with the measurement tool-scoring sheet (cf. Appendix A) while reviewing video of the testing sessions. The scoring sheet consists of a three-level rating scale: *A* for *excellent*, *B* for *acceptable*, or *C* for *reconsider*, which allows the test administrator to assign a score for each of the 15 sub-criteria of the four applicability criteria described earlier: respondent burden, examiner burden, score distribution and format compatibility. An *A* means that the tool fully meets the applicability criterion, while a *C* means failure to meet the requirements. When some difficulties are observed and some adaptations may be recommended to improve measurement tool applicability, an *acceptable* score (i.e., *B*) is attributed (Andresen, 2000; Auger et al., 2007). The scoring sheet was a French translation (cf. Appendix A) of the original one proposed by Auger et al. (2006; cf. Appendix B).

Specific indications for scoring are described for each sub-criterion of the main applicability criterion to support the examiner. For the 11 individual tests, the examiner attributed one score for each of the 15 applicability sub-criteria. The 28 sub-tests included in the 4 batteries were assessed following the same method. However, to simplify presentation of the results and make them more operational, only one score per sub-criterion per battery (the most frequent score assigned to the subtests of a given battery) has been retained for final analysis, for a total of 15 tests (11 individual tests and 4 batteries).

Criteria Analysis for Final Recommendations

Final recommendations from tests applicability assessment were formulated according to the following procedures. First, each test was individually analysed to verify scores (*A*, *B* or *C*) assigned to the 15 sub-criteria. A systematic qualitative analysis was manually performed to assess the nature and extent of eventual applicability limits according to these scores. The 15 tests were then classified into three categories based on the following general guidelines: 1) when no limits were identified, or had few or no real impact on general applicability, tests were *recommended*; 2) when identified limits had a potential negative impact on general applicability, but could easily be minimized or compensated (e.g., equivalent test, adaptations), tests were *recommended with caution*; 3) tests with serious applicability limits, impacts on interpretation, scoring, acceptability, or other parameters that could not be conveniently overcome were classified as *not recommended*.

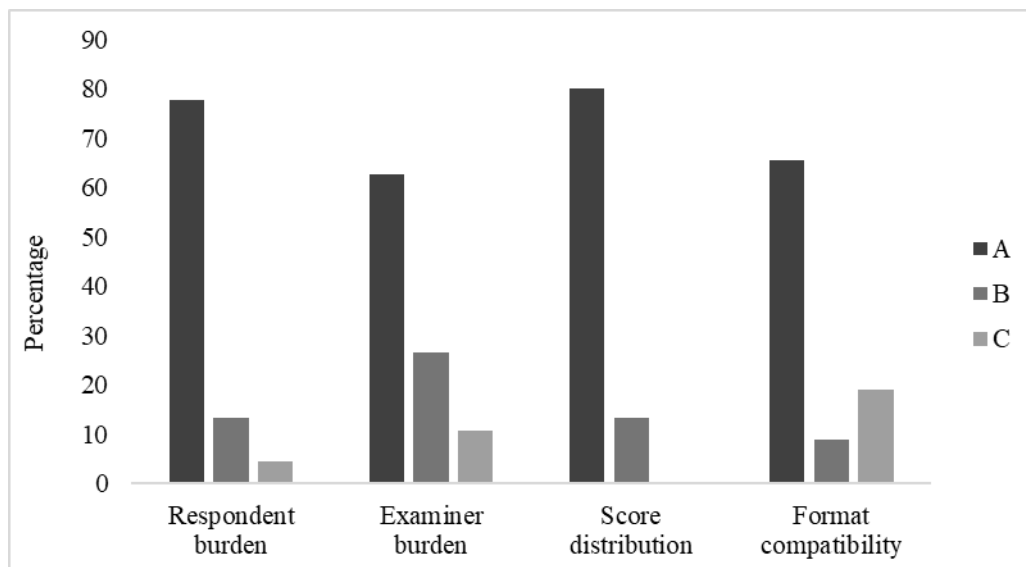


Figure 1. Distribution of the 15 tests' mean percentages of scores per applicability criterion.

Results

Results are presented in two parts. First, a general description of the psychometric battery is provided according to the four main applicability criteria. Secondly, a more specific analysis is presented to describe applicability of tests taken individually.

Scores on Applicability Criteria

Respondent burden. As seen in Figure 1, 77.78% of the tests are rated *A* for sub-criteria of the respondent burden criterion. For the majority of tests, the sub-criteria invasiveness and security of the tests were rated *A*. The PTJB was respectively rated *C* and *B* for these aspects considering motor disabilities and physical requirements of the test. Movements involving lower limbs were hence not assessed to prevent tumble and injury. Five tests were rated *B* on the acceptability sub-criterion (RSPM, CPT-II, Leiter-3, CVMT, and WCST) due to relatively long (but not inappropriate) administration time.

Examiner burden. Overall, 62.67% of the tests receive *A* among the 5 sub-criteria for this applicability criterion (cf. Figure 1). Quality of instructions is excellent for all tests. Complexity of training (93.33%) as well as administration, scoring, and interpretation burden (86.67%) is also excellent for most of the tests. However, Leiter-3 receives *C* on these aspects due to significant difficulties identified by the examiner in the assimilation of the instrument's functioning. The WCST is rated *B* on complexity of administration, scoring, and interpretation considering the scoring system's sophistication. Regarding availability, RTLT and PTJB are rated *excellent*. Otherwise, availability is considered *acceptable* (i.e., *B*) for almost all tests (80.00%) due to material cost, existence of copyright or environmental requirements for administration. Leiter-3 and WCST are rated *C* due to substantial acquisition costs. Administration time is under 10 minutes (which corresponds to *A* in the rating scale) with only 20.00% of the tests (PTJB, MVPT-3 as well as WAIS-IV, NAB, and D-KEFS subtests taken individually). To finish, 46.67% of the tests are rated *B* on this matter whilst the remaining 33.33% are rated *C* (above 20 minutes): WCST, CVMT, Leiter-3, RSPM, and NEO-FFI.

Score distribution. For this criterion, participants' results in standard deviations (*z* scores) were visually examined for each test. When the score's distribution clearly represented an array of behaviours (that is, significant clinical differences could easily be seen between participants), an *A* was attributed. When some differences could be observed with a moderate ceiling or floor effect, the examiner attributed *B*, whereas the absence of differences was scored *C*. In the present case, 80.00% of the tests are rated *A* (cf.

Figure 1). Only two tests get *B* (Leiter-3 and WCST) while data is incompatible with this type of analysis for the PTJB.

Format compatibility. Tests are rated *A* on 65.56% of the 6 sub-criteria on the format compatibility applicability criterion (cf. Figure 1). Adaptation to age, gender, and culture is *excellent* for all tests. However, none of these provide alternative formats (e.g., language, assistance, answer cards) which is why they all obtain *C* on this matter. Format adaptability is rated *A* for most tests, respondents being able to follow original instructions without adaptations except for Leiter-3 where some of them needed further explanations. Physical requirements are generally slight. Obviously, PTJB necessitates excessive physical involvement due to motor disabilities and is thus rated *C*. Only four (26.67%) tests (HVOT, BFRT, PTJB, and NAB) are considered adapted to level of education in terms of cognitive requirements. Seven (46.67%) tests are rated *B*, administration manuals providing no specification in this sense (NEO-FFI, SCL-90-R, CPT-II, Leiter-3, RTLT, D-KEFS, and WAIS-IV). It was not possible to compile data for the other (26.67%) tests for this sub-criterion as administration manuals and corresponding norms were not available anymore at the time of data collection. Finally, difficulty threshold is not an issue for the psychometric tests used in this study (rated *A*) except for PTJB which was too hard for the participants without significant adaptations.

Tests' applicability

Taken as a whole, the 15 tests used in this study get 68.00% of *A*'s, 16.00% of *B*'s, and 12.00% of *C*'s on the 15 applicability sub-criteria. Grades could not be assigned in 4.00% of cases, due to various reasons (e.g., video recording's unavailability, information not available in test's administration manuals). Tests' applicability is presented in three parts according to the distribution for each test (cf. Table 3 for distribution of scores for each test taken individually).

Taken individually, HVOT, BFRT, D-KEFS, and NAB each get 80.00% of *A*'s and 13.33% of *B*'s on applicability sub-criteria.

Eight tests obtain between 66.67% and 73.33% of *A*'s on applicability sub-criteria: RSPM, CPT-II, CVMT, MVPT-3, RTLT, and WAIS-IV. These tests also get 13.33% of *B*'s except for the following: SCL-90-R (20.00%), CPT-II (26.00%), MVPT-3 (6.67% of *B*'s and 6.67% of *C*'s), and WAIS-IV (20.00%).

Finally, three tests individually get between 40.00% and 60.00% of *A*'s on applicability sub-criteria: Leiter-3 (26.67% of *B*'s and 33.33% of *C*'s),

Table 3

Distribution of Applicability Scores by Psychometric Test

Tests	A (%)	B (%)	C (%)	Missing data (%)
HVOT	12 (80.00)	2 (13.33)	1 (6.67)	0 (0.00)
NEO-FFI	10 (66.67)	2 (13.33)	2 (13.33)	1 (6.67)
SCL-90-R	10 (66.67)	3 (20.00)	1 (6.67)	1 (6.67)
RSPM	10 (66.67)	2 (13.33)	2 (13.33)	1 (6.67)
CPT-II	10 (66.67)	4 (26.67)	1 (6.67)	0 (0.00)
Leiter-3	6 (40.00)	4 (26.67)	5 (33.33)	0 (0.00)
BFRT	12 (80.00)	2 (13.33)	1 (6.67)	0 (0.00)
CVMT	10 (66.67)	2 (13.33)	2 (13.33)	1 (6.67)
MVPT-3	11 (73.33)	1 (6.67)	1 (6.67)	2 (13.33)
RTLTL	11 (73.33)	2 (13.33)	1 (6.67)	1 (6.67)
PTJB	9 (60.00)	1 (6.67)	4 (26.67)	1 (6.67)
WCST	7 (46.67)	4 (26.67)	3 (20.00)	1 (6.67)
D-KEFS	12 (80.00)	2 (13.33)	1 (6.67)	0 (0.00)
WAIS-IV	11 (73.33)	3 (20.00)	1 (6.67)	0 (0.00)
NAB	12 (80.00)	2 (13.33)	1 (6.67)	0 (0.00)

PTJB (6.67% of *B*'s and 26.67% of *C*'s), and WCST (26.67% of *B*'s and 20.00% of *C*'s).

Discussion

The objective of this study was to assess the applicability of a psychometric battery for individuals with ARSACS. In the present study, all fifteen tests are rated with a majority (over 60%) of *excellent* scores for applicability sub-criteria. These results are encouraging, some tests having few or no significant applicability limits in individuals with ARSACS. However, other tests present challenges in this regard. In addition to giving a description of these tests regarding their applicability, the present analysis provides specific recommendations concerning whether they should be used or not in the future with this population. Influence of spasticity, upper limbs incoordination, and dysarthria were noted as limiting factors in the selection of neuropsychological and psychological tests in ARSACS (Boivin-Mercier, n.d.; Bouchard et al., 1978; Boucher, 2017; Drolet, 2002). The need to select appropriate outcome measures is directly related to ongoing international efforts to document disease evolution and eventual clinical trial effectiveness according to the U.S. Food and Drug Administration (FDA) drugs development and approval process (U.S. Food and Drug Administration, 2016).

General Observations

As expected, administration time (a sub-criterion implying both respondent's and examiner's burden) is an issue for a number of tests. Indeed, RSPM, Leiter-3's subtests, CVMT, WCST, NEO-FFI, and CPT-II are considered relatively long (10 to 20 minutes). One should note that psychometric tests (particularly neuropsychological and personality tests) are known to have a somehow long administration time (Strauss, Sherman, & Spreen, 2006). For example, CPT-II is a computerized test evaluating sustained attention, among other things. Obviously, evaluation of this function requires prolonged exposure to stimuli in a particular setting. Similarly, CVMT aims to evaluate long-term visual memory and includes a 30 minutes delay (during which other tests involving minimal visual reasoning can be administered). However, administration time for RSPM was over 45 minutes for some subjects over 40 years of age. This represents an unusually long administration time for this test and thus calls into question its use in individuals with ARSACS. Matrix Reasoning (WAIS-IV), a subtest evaluating similar functions, does not present this limit and should be preferred in situations where time is a constraint.

The following considerations concerning examiner's burden are based on the premise that considerable time and efforts are required to harness

administration, scoring, and interpretation procedures for several psychometric tests. It should be noted that showing great knowledge of these aspects is a fundamental pre-requisite for anyone interested in their administration. The great majority of the tests used in the present study include clear instructions and have relatively simple administration, scoring, and interpretation procedures. Leiter-3 and WCST pose specific challenges in this matter and are discussed below. Otherwise, material availability is generally qualified as acceptable due to frequent existence of copyright and relatively high acquisition costs.

As stated above, tests raw scores results were converted into z scores using most appropriate normative sample. In their studies, Boivin-Mercier (n.d.), Boucher (2017), and Desmeules (2017) employed categories based upon deviation's magnitude and cut-off used in clinical setting. This procedure makes it possible to classify the results into five categories: normal, vulnerability, mild deficit, moderate deficit, and severe deficit. Accordingly, most of the tests used in this study allowed discrimination of patient's performance in terms of function's preservation or infringement.

Regarding format compatibility, tests used in this study are considered adapted to age, gender, and culture. However, there is few present alternative formats. For example, NAB is an interesting battery for the evaluation of cognitive functions but no validated French translation is available neither for instructions nor normalisation samples. Although verbal tests were not used, it is impossible to exclude language bias. Despite growing desire to develop local norms and fill that gap, this issue is frequently encountered in Quebec considering its differences (e.g., language, culture) when compared to the rest of North America. Furthermore, most tests have acceptable adaptation to level of education in terms of cognitive requirements. This outcome originates from the nature of the tests and instructions provided in the applicability scoring table: score attributed for this sub-criterion is based on whether or not a given test contains adaptations allowing administration in subjects with lower education or communication levels. Since neuropsychological tests often aim to evaluate cognitive functions directly or indirectly correlated with these features, adaptations in this regard are uncommon. Test items are often presented in increasing order of complexity, a structure that precisely allows the differentiation of preserved and affected functions (Institute of Medicine, 2015). In the present study, four tests (HVOT, BFRT, PTJB, and NAB's subtests) present normative data taking into account education level and are considered excellent in this regard.

Table 4

Recommended Tests According to Applicability Evaluation

Recommendation	Psychometric tests
Recommended	WAIS-IV ^a
	CPT-II
	MVPT-3
	NEO-FFI
	HVOT
	CVMT
	BFRT
	SCL-90-R
	RTLT
	Recommended with caution
WCST	
CWIT subtest (D-KEFS)	
TLT subtest (D-KEFS)	
Not recommended	RSPM
	PTJB
	TMT subtest (D-KEFS)
	Leiter-3

Note. ^aDespite generally good applicability, Block Design, Symbol Search, Coding, and Cancellation subtests from the WAIS-IV involve motor abilities and should thus be considered with caution.

Recommendations

Table 4 presents recommendations from the present study regarding applicability evaluation for the 15 tests used. More details on this matter are presented in the following section. It should be noted that clinicians and researchers must always consider specific objectives (e.g., differential diagnosis, external reference) and environmental constraints (e.g., time allowed, subject's physical state) when it comes to choosing appropriate instruments.

Globally, WAIS-IV has acceptable applicability with differences depending on the subtests. Digit Span, Matrix Reasoning, Arithmetic, Visual Puzzles, Figure Weights, and Picture Completion present no significant issues. On the other hand, subtests involving motor abilities (Block Design, Symbol Search, Coding, and Cancellation) should be treated with great caution. Block Design's items have time limits: results from individuals with ARSACS must not only be considered in terms of final score, but also in terms of success or failure regardless of completion time. For this subtest, clinical judgement and

observations may help refine comprehension (e.g., extent of upper limbs incoordination prevents completion of the task). Evaluation of perceptual reasoning should rely more heavily on subtests like Matrix Reasoning and Visual Puzzles. Subtests measuring processing speed (Symbol Search, Coding, and Cancellation) also imply motor abilities. Their use in ARSACS presents major limits, time limit also being the principal scoring criterion. In her study, Boucher (2017) reports interpretation limits in tests measuring processing speed as they imply motor abilities and hand-eye coordination in a time-limited context, suggesting addressing this limit when choosing psychometric tests in future studies with ARSACS patients. Digit Symbol's subtest from WAIS-R as a Neuropsychological Instrument (WAIS-R-NI; Kaplan, 1991) constitutes an alternative, allowing the establishment of a baseline for direct graphic copy speed. The examiner can then modulate interpretation of processing speed results according to motor difficulties. Despite relatively long administration time, CPT-II is recommended considering simplicity of administration and scoring, good acceptability for patients, and contribution of the results for cognitive processes understanding. CVMT is also an interesting alternative to other visual memory tests such as Rey Complex Figure Test, which strongly involves motor abilities. Regarding HVOT, BFRT, MVPT-3, RTL, SCL-90-R, and NEO-FFI, no significant applicability limits are noted and their use is thus recommended with ARSACS population.

Other tests present mixed results on applicability criteria. As discussed above, WCST's scoring and interpretation system is quite complex, specifically concerning the rating of perseverative responses. This significantly extends required time to analyse and interpret the results and use of this test should consider these features. Future examiner might consider the computerized version of this test, which provides automatic scoring. The use of a keyboard and computer mouse could potentially increase applicability of the test considering motor difficulties in ARSACS. On the other hand, NAB proved to be an interesting instrument: most tests can be administered in ten to twenty minutes and are easy to harness. As stated above, its principal limit is the absence of validated French translation. Its use with ARSACS individuals from Quebec province thus requires a great deal of caution. CWIT and TLT from D-KEFS battery generally present acceptable applicability. Indeed, the former implies minimal motor response, but the examiner should consider potential influence of dysarthria on execution time and take a close look on contrast scores to properly interpret the results. For its part, TLT requires great caution due to possible slowdown related to motor difficulties. Consequently,

deficit in realization time does not necessarily imply a deficit in the evaluated cognitive function (planification and cognitive flexibility).

Finally, results show significant applicability limits with ARSACS individuals in four tests. As discussed above, other tests (e.g., Matrix Reasoning subtest from WAIS-IV) should be preferred to RSPM considering excessive administration time, problems with acceptability, and relatively old normative data. PTJB is the only other test having acceptability issues, physical requirements being too high for some patients who could not complete it. TMT from D-KEFS battery presents similar problems: patients can complete the task, but coordination and motor difficulties are too great, biasing results interpretation. For its part, Leiter-3 was selected to gather cognitive functioning data, since the examiner does not use language to provide instructions with the initial objective to minimize potential bias related to dysarthria and education level. However, in addition to having considerable administration time and to be difficult to harness, certain patients could hardly understand instructions (presented as gestures). Therefore, the use of these four tests is not recommended in individuals with ARSACS.

These results should be interpreted with caution considering some limitations. First, applicability criteria's scoring was completed by the examiner only. In future studies, the use of an inter-judge rating system for the measurement tool-scoring sheet should be considered in order to increase applicability score's objectivity and, therefore, validity (Hogan, 2012). A second limitation is related to the criteria scoring context. As stated above, the measurement tool-scoring sheet was scored by the examiner while reviewing video of the testing session. A scoring in real time during the testing sessions may have led to different results. For instance, acceptability from the respondent's perspective or cognitive requirements as evaluated by the examiner could have been assessed by direct questions after test administration. This supplementary information might have added to the understanding of psychometric tests' general applicability in individuals with ARSACS. One way to address simultaneously the two precedent suggestions would be to ask the examiner to evaluate in real time the applicability with patients using the measurement tool-scoring sheet after test administration, then to ask an independent qualified judge to score the same applicability criteria while reviewing video of the sessions including patients' commentaries. Besides, the present study evaluated subtests included in psychometric batteries on each applicability sub-criterion. For a given battery, only one score (the most frequent among batteries' subtests) was considered to interpret results and

simplify their presentation. In their study, Auger et al. (2007) used a distinct measurement tool-scoring sheet to assess entire batteries including several subtests. Criteria presented in this tool could have been used in the present study to increase consistency of scoring. It is also important to note that tests selection process in Boivin-Mercier (n.d.), Boucher (2017), and Desmeules (2017) had resulted in the exclusion of psychometric tests evaluating language or verbal concepts formation (e.g., WAIS-IV Verbal Comprehension subtests) due to anticipated impact of dysarthria, mainly to avoid interpretative bias and excessive confrontation for patients. Insights from the present study suggest that these tests could be used in future studies involving individuals with ARSACS. Finally, the small sample size (eight participants) may have limited variety of observations on the applicability tool-scoring sheet. Although this number may seem small, it represents about 10% of the patients listed at the Neuromuscular Clinic of the CIUSSS du Saguenay-Lac-Saint-Jean (Quebec, Canada). This is considered a relatively high sample size in the rare neuromuscular disease research field. It would be interesting to re-evaluate psychometric tests applicability in forthcoming studies on cognitive and personality profiles in ARSACS as they should include more participants.

Conclusion

The selection process of measurement tools must consider psychometric properties such as validity, fidelity, and existence of appropriate normative samples. In the same line as Auger et al. (2007), this study supports the relevance to address practical aspects like applicability when choosing most appropriate psychometric tests, particularly with populations presenting specific characteristics that may limit their interpretation. Insights from this study can sustain ongoing research efforts aiming to document cognitive functions in a large sample of individuals with ARSACS. Recommendations may also help clinicians and researchers working with individuals with ARSACS and other ataxias presenting upper limbs involvement and speech disturbances. In all cases, psychological and neuropsychological evaluations should always consider information gathered from other sources like health care professionals (e.g., physicians, audiologists, ophthalmologist, physiotherapist) to exclude eventual influence of other functions deficits in performance on cognitive tests.

References

- Andresen, E. M. (2000). Criteria for assessing the tools of disability outcomes research. *Archives of Physical Medicine and Rehabilitation*, 81, S15-S20. <https://doi.org/10.1053/apmr.2000.20619>
- Auger, C., Demers, L., Desrosiers, J., Giroux, F., Ska, B., & Wolfson, C. (2007). Applicability of a toolkit for geriatric rehabilitation outcomes. *Disability & Rehabilitation*, 29, 97-109. <https://doi.org/10.1080/09638280600731540>
- Auger, C., Demers, L., & Swaine, B. (2006). Making sense of pragmatic criteria for the selection of geriatric rehabilitation measurement tools. *Archives of Gerontology and Geriatrics*, 43, 65-83. <https://doi.org/10.1016/j.archger.2005.09.004>
- Benton, A. L., Sivan, A. B., Hamsher, K. D., Varney, Nils R., & Spreen, O. (1983). *Contributions to neuropsychological assessment: A clinical manual*. Oxford, UK: Oxford University Press.
- Boivin-Mercier, A. (n.d.). *Étude exploratoire des fonctions cognitives chez les individus âgés de 20 à 40 ans atteints d'ataxie récessive spastique de Charlevoix-Saguenay (ARSCS)* (Unpublished doctoral dissertation). Université du Québec à Chicoutimi, Chicoutimi, Canada.
- Bouchard, J. P., Barbeau, A., Bouchard, R., & Bouchard, R. W. (1978). Autosomal recessive spastic ataxia of Charlevoix-Saguenay. *The Canadian Journal of Neurological Sciences*, 5, 61-69. <https://doi.org/10.1017/S0317167100024793>
- Bouchard, J. P., Richter, A., Mathieu, J., Brunet, D., Hudson, T. J., Morgan, K., & Melançon, S. B. (1998). Autosomal recessive spastic ataxia of Charlevoix-Saguenay. *Neuromuscular Disorders*, 8, 474-479. [https://doi.org/10.1016/S0960-8966\(98\)00055-8](https://doi.org/10.1016/S0960-8966(98)00055-8)
- Boucher, A. (2017). *Étude exploratoire des fonctions cognitives chez les individus âgés de 41 à 60 ans atteints d'ataxie récessive spastique de Charlevoix-Saguenay (ARSCS)* (Doctoral dissertation). Retrieved from <https://constellation.uqac.ca/4150/>
- Brown, J. W. (1972). *Aphasia, apraxia, and agnosia: Clinical and theoretical aspects*. Springfield, IL: Charles C. Thomas.
- Colarusso, R. P. & Hammill, D. D. (1972). *Motor-free Visual Perception Test*. Novato, CA: Academic Therapy Pub.
- Colarusso, R. & Hammill, D. (2002). *Motor-Free Visual Perception Test (MVPT-3)* (3rd ed.). Austin, TX: Pro-Ed.
- Conners, C. K. (1994). *Conners' continuous performance test computer program 3.0 user's manual*. Toronto, Canada: Multi-Health Systems, Inc.
- Costa, P. T. Jr. & McCrae, R. R. (1989). *The NEO-PI/NEO-FFI manual supplement*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T. Jr. & Widiger, T. A. (1994). *Personality disorders and the five-factor model of personality*. Washington, DC: American Psychological Association.

- Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System*. San Antonio, TX: Psychological Corporation.
- Derogatis, L. R. (1994). *SCL-90-R: Symptom Checklist-90-R: Administration, scoring, and procedures manual* (3rd ed.). Minneapolis, MN: NCS Pearson.
- Derogatis, L. R. & Melisaratos, N. (1983). The brief symptom inventory: An introductory report. *Psychological medicine*, *13*, 595-605. <https://doi.org/10.1017/S0033291700048017>
- Derogatis, L. R. & Savitz, K. L. (2000). The SCL-90-R and Brief Symptom Inventory (BSI) in primary care. *Handbook of psychological assessment in primary care settings*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Desmeules, A. (2017). *Étude exploratoire de la personnalité et du profil psychologique chez les individus ayant l'ataxie récessive spastique de Charlevoix-Saguenay (ARSCS)* (Doctoral dissertation). Retrieved from <https://constellation.uqac.ca/4241/>
- Drolet, M. (2002). *Étude des fonctions neuropsychologiques des enfants et adolescents atteints de l'ataxie récessive spastique de Charlevoix-Saguenay (ARSCS)* (Master's thesis). Retrieved from <https://constellation.uqac.ca/874/>.
- Eling, P., Derckx, K., & Maes, R. (2008). On the historical and conceptual background of the Wisconsin Card Sorting Test. *Brain and Cognition*, *67*, 247-253. <https://doi.org/10.1016/j.bandc.2008.01.006>
- Fancellu, R., Paridi, D., Tomasello, C., Panzeri, M., Castaldo, A., Genitrini, S., ... Girotti, F. (2013). Longitudinal study of cognitive and psychiatric functions in spinocerebellar ataxia types 1 and 2. *Journal Of Neurology*, *260*, 3134-3143. <https://doi.org/10.1007/s00415-013-7138-1>
- Gazulla, J., Mayayo-Sinués, E., Benavente, I., Modrego, P. J., & Berciano, J. (2014). Ataxia of Charlevoix-Saguenay: MR and clinical results in lower-limb musculature. *The Canadian Journal of Neurological Sciences/Le Journal Canadien des Sciences Neurologiques*, *41*, 37-41. <https://doi.org/10.1017/S0317167100016231>
- Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test Manual: Revised and expanded*. Odessa, FL: Psychological Assessment Resources.
- Heaton, R. K. (1993). *Wisconsin Card Sorting Test: Computer version 2.0*. Odessa, FL: Psychological Assessment Resources.
- Hogan, T. P. (2012). *Introduction à la psychométrie*. Montréal, Canada: Chenelière Éducation.
- Hooper, H. E. (1958). *The Hooper Visual Organization Test: Manual*. Beverly Hills, CA: Western Psychological Services.
- Institute of Medicine. (2015). *Psychological Testing in the Service of Disability Determination*. Washington, DC: The National Academies Press.
- Kaplan, E. (1991). *WAIS-R as a Neuropsychological Instrument (WAIS-R-NI)*. San Antonio, TX: Psychological Corporation.
- Lange, R. T., Iverson, G. L., Brickell, T. A., Staver, T., Pancholi, S., Bhagwat, A., & French, L. M. (2013). Clinical utility of the Conners' Continuous Performance Test-II to detect poor effort in U.S. Military personnel following traumatic brain injury. *Psychological Assessment*, *25*, 339-352. <https://dx.doi.org/10.1037/a0030915>
- Larrabee, G. J., Trahan, D. E., & Curtiss, G. (1992). Construct validity of the Continuous Visual Memory Test. *Archives of Clinical Neuropsychology*, *7*, 395-405. <https://doi.org/10.1093/arclin/7.5.395>
- Lévesque, K. (2004). *Relation entre les troubles affectifs et les difficultés scolaires chez les jeunes atteints d'ataxie récessive spastique de Charlevoix-Saguenay et ayant un quotient intellectuel normal* (Master's thesis). Retrieved from <https://constellation.uqac.ca/681/>
- Loesche, P., Wiley, J., & Hasselhorn, M. (2015). How knowing the rules affects solving the Raven Advanced Progressive Matrices Test. *Intelligence*, *48*, 58-75. <https://doi.org/10.1016/j.intell.2014.10.004>
- Lopez, N. M., Lazar, M. D., & Oh, S. (2003). Psychometric properties of the Hooper Visual Organization Test. *Assessment*, *10*, 66-70. <https://doi.org/10.1177/1073191102250183>
- Martin, M. H., Bouchard, J. P., Sylvain, M., St-Onge, O., & Truchon, S. (2007). Autosomal recessive spastic ataxia of Charlevoix-Saguenay: a report of MR imaging in 5 patients. *American Journal of Neuroradiology*, *28*, 1606-1608. <https://doi.org/10.3174/ajnr.A0603>
- Mattioli, F., Stampatori, C., Bellomi, F., Scarpazza, C., Galli, P., Guarneri, C., ... Capra, R. (2014). Assessing executive function with the D-KEFS sorting test: normative data for a sample of the Italian adult population. *Neurological Sciences*, *35*, 1895-1902. <https://doi.org/10.1007/s10072-014-1857-7>
- McCrae, R. R. & Costa P. T. Jr. (2007). Brief versions of the NEO-PI-3. *Journal of Individual Differences*, *28*, 116-128. <https://doi.org/10.1027/1614-0001.28.3.116>
- Mignarri, A., Tessa, A., Carluccio, M. A., Rufa, A., Storti, E., Bonelli, G., . . . Dotti, M. T. (2014). Cerebellum and neuropsychiatric disorders: Insights from ARSACS. *Neurological Sciences*, *35*, 95-97. <https://doi.org/10.1007/s10072-013-1592-5>
- Pilliod, J., Moutton, S., Lavie, J., Maurat, E., Hubert, C., Bellance, N., . . . Coupry, I. (2015). New

- practical definitions for the diagnosis of autosomal recessive spastic ataxia of Charlevoix-Saguenay. *Annals of Neurology*, 78, 871-886. <https://doi.org/10.1002/ana.24509>
- Pinel, J. P. J. (2007). *Biopsychologie* (6th ed.). Paris, France: Pearson Education.
- Psychological Corporation (2008). *WAIS-IV technical and interpretation manual*. San Antonio, TX: Psychological Corporation.
- Rae-Grant, A. & Parsons, M. W. (2014). Neuropsychology in the clinical setting: conceptual and practical issues. In M. W. Parsons & T. A. Hammeke (Eds.), *Clinical neuropsychology: A pocket handbook for assessment* (3rd ed., pp. 9-30). Washington, DC: American Psychological Association.
- Raven, J., Raven, J. C., & Court, H. H. (2000). *Raven manual: Section 3, Standard Progressive Matrices*. Oxford, England: Oxford Psychologists Press Limited.
- Roid, G. H., Miller, L. J., Pomplun, M., & Koch, C. (2013). *Leiter International Performance Scale* (3rd ed.). Los Angeles, CA: Western Psychological Services.
- Schmahmann, J. D. (2004). Disorders of the cerebellum: Ataxia, dysmetria of thought, and the cerebellar cognitive affective syndrome. *Journal of Neuropsychiatry and Clinical Neurosciences*, 16, 367-378. <https://doi.org/10.1176/jnp.16.3.367>
- Schretlen, D. J., Pearlson, G. D., Anthony, J. C., & Yates, K. O. (2001). Determinants of Benton Facial Recognition Test performance in normal adults. *Neuropsychology*, 15, 405-410. <https://dx.doi.org/10.1037/0894-4105.15.3.405>
- Seedat, F., James, M. K., & Rose, M. R. (2014). Activity rating scales in adult muscle disease: How well do they actually measure? *Muscle Nerve*, 50, 24-33. <https://doi.org/10.1002/mus.24090>
- Senior, G., Kelly, M., & Salzman, L. (1999). Clinical utility of the Rey tangled lines test. *Archives of Clinical Neuropsychology*, 14, 719.
- Seron, X. & Van der Linden, M. (2014). L'anamnèse et l'examen neuropsychologique de base. In X. Seron & M. Van der Linden (Eds.), *Traité de neuropsychologie clinique de l'adulte : Tome 1, Évaluation* (2nd ed., pp. 113-130). Paris, France: De Boeck/Solal.
- Strauss, E., Sherman, E. M., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, normas, and commentary* (3rd ed.). New-York, NY: Oxford University Press.
- Strong, C.-A. H. & Donders, J. (2008). Validity of the Continuous Visual Memory Test (CVMT) after traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, 30, 885-891. <https://doi.org/10.1080/13803390701858224>
- Synofzik, M., Soehn, A. S., Gburek-Augustat, J., Schicks, J., Karle, K. N., Schule, R., . . . Schols, L. (2013). Autosomal recessive spastic ataxia of Charlevoix-Saguenay (ARSACS): Expanding the genetic, clinical and imaging spectrum. *Orphanet Journal of Rare Diseases*, 8, 41. <https://doi.org/10.1186/1750-1172-8-41>
- Thiffault, I., Dicaire, M. J., Tetreault, M., Huang, K. N., Demers-Lamarche, J., Bernard, G., . . . Brais, B. (2013). Diversity of ARSACS mutations in French-Canadians. *The Canadian Journal of Neurological Sciences/Le Journal Canadien des Sciences Neurologiques*, 40, 61-66. <https://doi.org/10.1017/S0317167100012968>
- Tsatsanis, K. D., Dartnall, N., Cicchetti, D., Sparrow, S. S., Klin, A., & Volkmar, F. R. (2003). Concurrent validity and classification accuracy of the Leiter and Leiter-R in low-functioning children with autism. *Journal of Autism and Developmental Disorders*, 33, 23-30. <https://doi.org/10.1023/A:1022274219808>
- U.S. Food and Drug Administration. (2016). *Clinical outcome assessment qualification program*. Retrieved from <https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm284077.htm>
- Verguts, T., De Boeck, P., & Maris, E. (1999). Generation speed in Raven's progressive matrices test. *Intelligence*, 27, 329-345. [https://doi.org/10.1016/S0160-2896\(99\)00023-9](https://doi.org/10.1016/S0160-2896(99)00023-9)
- Vermeer, S., Van de Warrenburg, B. P., & Kamsteeg, E. J. (2003). ARSACS. *GeneReviews*. Retrieved from <https://www.ncbi.nlm.nih.gov/books/NBK1255/>
- Wade, D. T. (2004). Assessment, measurement and data collection tools. *Clinical Rehabilitation*, 18, 233-237. <https://doi.org/10.1191/0269215504cr183ed>
- Weiss, L. G., Saklofske, D. H., Coalson, D., & Raiford, S. E. (2010). *WAIS-IV clinical use and interpretation: Scientist-practitioner perspectives*. Amsterdam, Netherlands: Elsevier Academic Press.
- White, T. & Stern, R. A. (2003). *Neuropsychological Assessment Battery: Psychometric and technical manual*. Lutz, FL: Psychological Assessment Resources.

Received January 19, 2018

Revision received May 26, 2018

Accepted June 22, 2018 ■

Appendix A

Applicability Criteria Measurement Scoring Table Sheet Translation

Critère	Sous-critère	A	B	C
Fardeau du répondant	Caractère invasif	L'instrument n'est pas invasif		L'instrument est invasif
	Sécurité	Aucun risque de blessure physique	Risque faible de blessure physique	Risque important de blessure physique
	Acceptabilité	Court (≤ 15 min.) ET acceptable ($< 5\%$ de refus)	Long (mais de manière appropriée) OU problèmes d'acceptabilité ($< 15\%$ de refus ou quelques plaintes)	Long ET inacceptable ($> 15\%$ de refus; $> 50\%$ de plaintes; difficultés notées par l'examineur)
Fardeau de l'examineur	Temps d'administration	1-10 min.	10-20 min.	> 20 min.
	Qualité des instructions	Instructions complètes, claires, pertinentes, faciles d'accès et faciles à consulter	Instructions disponibles mais incomplètes, questions sans réponse	Instructions minimales
	Entraînement	Aucun ou entraînement simple/court	Spécifique à un professionnel, entraînement court et simple	Spécifique à un professionnel et entraînement long et complexe
	Simplicité de mesure (administration, cotation et interprétation)	Simple d'administration et cotation facile Résultats pertinents et interprétables	Simple d'administration et cotation nécessitant une consultation fréquente du manuel (pour calcul) Interprétation obscure des résultats	Complexe à administrer, ou difficultés de cotation ou interprétation des résultats complexes
	Disponibilité	Coût du matériel minimale Pas de droit d'auteur Paramètres environnementaux simples	Si deux critères ou plus sont répondus	Si un critère est répondu
	Résultats	Normalité de distribution et effets plafond et plancher	Pas de biais de distribution ou pas d'effet plafond/plancher du score total; 0 – 20%	Biais de distribution modéré ou effet plafond/plancher du score total $> 20\%$ -50%
Compatibilité du format	Âge, genre, culture	Preuve d'applicabilité générale pour les sous-groupes d'âge /genre/culture	Preuve d'administration avec des sous-groupes, certains problèmes soulevés (p.ex., aucune adaptation proposée)	Aucune administration avec un sous-groupe
	Format alternatif	Oui (p.ex., langage, Braille, âge)		Non
	Adaptabilité du format	Instructions originales suivies par tous les sujets	$\geq 90\%$ des sujets réalisent le test avec les instructions originales et le reste des participants réalisent le test avec des instructions adaptées	Les instructions originales ne peuvent être suivies ou le matériel ne peut être adapté
	Exigences physiques	Exigence physique minimale		Exigence physique nécessaire pour exécuter le test ou résultats influencés par des difficultés physiques
	Exigences cognitives (humeur, intelligence, communication)	Adapté pour les niveaux d'éducation faibles	Pas de spécification sur le niveau d'éducation	Haut niveau d'éducation requis ou résultats influencés par le niveau cognitif
	Données manquantes dues au seuil de difficulté	$< 10\%$ de données manquantes parce que l'instrument est trop difficile	10 – 15% de données manquantes parce que l'instrument est trop difficile	$> 15\%$ de données manquantes parce que l'instrument est trop difficile

Appendix B

Applicability Criteria Measurement Scoring Table Sheet Original Version

Criteria	Sub-criteria	A	B	C
Respondent burden	Invasiveness	Instrument is not invasive		Invasiveness is an issue
	Safety	No risk of physical injury	Low risk of physical injury	Important risk of physical injury
	Acceptability	Brief (≤ 15 min.) AND acceptable ($< 5\%$ of refusal)	Longer (but appropriately so) OR problems of acceptability ($< 15\%$ of refusals or some complaints)	Long AND unacceptable ($> 15\%$ of refusals; $> 50\%$ of complaints; difficulties noted by the examiner)
Examiner burden	Administration time	1-10 min.	10-20 min.	> 20 min.
	Quality of instruction	Instruction complete, clear, relevant, easy to access and consult	Instruction available but incomplete, questions unanswered	Minimal instruction
	Training	No or simple/short training	Specific professional and simple/short training	Specific professional and complex/long training
	Simplicity of measurement (administration, scoring, interpretability)	Simplicity of administration and easy scoring Results relevant and interpretable	Simplicity of administration and scoring necessitate frequent consultation of manual/calculating Obscure interpretation of results	Complex to administrate, or scoring difficulties, or complex interpretation of results
	Availability	Minimal cost / material No copyrights Light environmental settings	If one criterion meets	If two or more criteria meet
Score	No distribution bias or no ceiling-floor effect of total score; 0 – 20%	Moderate distribution bias or ceiling-floor effect of total score $> 20\%$ - 50%	Severe distribution bias or ceiling-floor effect of total score $> 50\%$	No distribution bias or no ceiling-floor effect of total score; 0 – 20%
Format compatibility	Age, gender, cultural	Evidence of broad applicability for cultural/age/gender subgroups	Evidence of testing with subgroups, some problems underlined (e.g., no adaptation proposed)	No test on any subgroups
	Alternate format	Yes (e.g., language, Braille, age)		No
	Adaptability of format	Original instructions followed for all subjects	$\geq 90\%$ of subjects performed the test with original instructions and the rest performed with adapted instructions	Original instructions could not be followed or material could not be adapted
	Physical Requirement	Minimal physical requirement		Physical requirement necessary to perform test or results influenced by physical impairment
	Cognitive Requirement (mood, intelligence, communication)	Adapted for lower educational level	No specification about educational level	High educational level required or results influenced by cognitive impairment
	Missing data due to difficulty of threshold	$< 10\%$ missing because instrument is too difficult	$10 - 15\%$ missing because instrument is too difficult	$> 15\%$ missing because instrument is too difficult