Test Critique: Childhood Autism Rating Scale-Second Edition (CARS2)

Attallah College of Educational Studies, Chapman University

EDUC601: Assessment and IEP Development

Dr. Randy Busse

Test Critique: Childhood Autism Rating Scale-Second Edition (CARS2)

The original Childhood Autism Rating Scale (CARS) was developed on children and adults referred to the Division Treatment and Education of Autistic and related Communications Handicapped Children (TEACCH) centers in North Carolina. The Childhood Autism Rating Scale, Second Edition (CARS2), by Eric Schopler, Ph.D., Mary E. Van Bourgondien, Ph.D., G. Janette Wellman, Ph.D., and Steven R. Love, Ph.D., was published by Western Psychological Services in 2010. The CARS2 is designed to identify children with Autism Spectrum Disorders (ASD) and distinguish them from those with developmental disabilities. The CARS2 consists of three rating booklets, including the Childhood Autism Rating Scale, Second Edition-Standard Version (CARS2-ST), the Childhood Autism Rating Scale, Second Edition-High-Functioning Version (CARS2-HF), and the Questionnaire for Parents or Caregivers (CARS2-QPC). According to the authors, the CARS2 is not meant as a universal screening measure. It should be used to provide additional information when a diagnosis of ASD has already been made or is highly suspected.

Content and Use

Practical Information and Features

The CARS2 includes a manual with descriptions of the contents, background information, instructions on administering, scoring, interpreting the test, and intervention planning and resources. The CARS2 includes three available forms, the CARS2-ST, CARS2-HF, and CARS2-QPC. The forms must be completed manually with paper and pencil, as there is no electronic version of the CARS2 booklets available. The CARS2 may be used in various settings, including schools, clinics, and intervention programs. As a rating scale system, the CARS2 is

easy to learn and can be administered by a wide range of professionals, provided they have had exposure and training about autism.

In addition to observations, it is recommended that developmental history, comprehensive clinical records, review of prior evaluations, the parent or caretaker interview and reports, behavioral rating areas, and direct interaction with the individual being evaluated are potential types of information the professional should use when formulating ratings (Schopler et al., 2010). The manual clarifies that ratings from the CARS2-ST or CARS2-HF should be used as only one piece of a multifaceted evaluation. The CARS2 by itself does not produce a diagnosis, and factors must be evaluated by experienced professionals who are experts in autism and authorized to make clinical diagnoses. As the CARS2 is a rating system of individual behavior, it is not intended for group administration. The CARS2-QPC is completed by a parent or caregiver and incorporated into the CARS2-ST or CARS2-HF by the professional. After gathering the necessary information and observations, the professional will complete the CARS2-ST or CARS2-HF rating scale independently.

The approximate time to complete each booklet is five to ten minutes. However, there is no time limit to complete the assessment, and the professional, parent, or caregiver may utilize as much time necessary to complete the booklet with accuracy. The complete CARS2 kit includes 25 booklets of each the CARS2-ST, CARS2-HF, CARS2-QPC, and a manual, and can be purchased for $218.30 on the Pearson website. Additional CARS2-ST and CARS2-HF booklets can be purchased in a bulk of 25 for $53.10, and additional CARS2-QPC booklets can be purchased in a bulk of 25 for $37.80.

Characteristics of the Test

The CARS2 is recommended for use with individuals ages two years or older and can assess virtually all functional levels. The CARS2 is not intended as a screener for use in the general population, but rather for providing a brief yet comprehensive summary of information that can be used to help develop diagnostic hypotheses among referred individuals of all ages and functional levels. The individual's behavior should be compared to that of a typically developing individual of the same age. Results from the CARS2 provides information not only on the presence of ASD symptoms but also the frequency, peculiarity, intensity, and duration of the behaviors, thus providing a more comprehensive picture of an individual's functioning. The utility of the CARS2 extends beyond informing diagnosis; results may be used for intervention planning, providing diagnostic feedback to parents, and creating "functional profiles".

As previously mentioned, the CARS2 contains three rating booklets, the CARS2-ST, CARS2-HF, and CARS2-QPC. The original Childhood Autism Rating Scale (CARS) has been retained in its original form and renamed the Childhood Autism Rating Scale, Second Edition-Standard Version (CARS2-ST). The CARS2-ST booklet is recommended for use with individuals with estimated overall IQs of 79 or lower, who have notably impaired communication skills, or who are younger than six years old of age regardless of their estimated IQ. New to the CARS2, the CARS2-HF is a different rating scale intended to identify high-functioning individuals with autism or Asperger's Disorder. It is recommended for use with individuals with estimated IQs of 80 or higher, who have relatively good verbal skills, and are aged six or older. The CARS2-ST and CARS2-HF each include 15 items addressing the following functional areas:

- Relating to People
- Imitation (ST); Social-Emotional Understanding (HF)

- Emotional Response (ST); Emotional Expression and Regulation of Emotions (HF)

- Body Use

- Object Use (ST); Object Use in Play (HF)

- Adaptation to Change (ST); Adaptation to Change/Restricted Interests (HF)

- Visual Response

- Listening Response

- Taste, Smell, and Touch Response and Use

- Fear or Nervousness (ST); Fear or Anxiety (HF)

- Verbal Communication

- Nonverbal Communication

- Activity Level (ST); Thinking/Cognitive Integration Skills (HF)

- Level and Consistency of Intellectual Response

- General Impressions

Both the CARS2-ST and CARS2-HF include a 7-point Likert-type scale. The seven allowable ratings for each item are as follows:

- 1 = Within normal limits for that age.

- 1.5 = Very mildly abnormal for that age.

- 2 = Mildly abnormal for that age.

- 2.5 = Mildly-to-moderately abnormal for that age.

- 3 = Moderately abnormal for that age.

- 3.5 = Moderately-to-severely abnormal for that age.

- 4 = Severely abnormal for that age.

New to the CARS2, the CARS2-QPC rating booklet is intended to be completed by the individual's parent or caregiver and may be used in tandem with the CARS2-ST or the CARS2-HF. The CARS2-QPC is organized into seven sections; the first six sections include four to nine statements for which parents or caregivers are required to rate the individual's functioning on a 5-point Likert-type scale. The 5-point Likert-type scale answer options include "not a problem", "mild-to-moderate problem", "severe problem", "not a problem now, but was in the past", and "don't know". The seventh section allows for open-ended responses to questions about other behaviors not described in the previous sections.

The layout of the CARS2 is intuitive and user friendly. All three rating booklets are written in anchor and response format. The items, questions, statements, and rating values appear neutral and direct, free of biases and anchoring a particular response. All items in the CARS2 are gender-neutral, and there is no evidence of sex bias. The text for rating value descriptions is small but written in a font that is easy to read. Below each rating value and subsequent description, lengthy space is provided for recording related observations of the individual. The booklets are designed to accommodate detailed note-taking regarding behavioral observations, information from diverse sources, or unique information that may be used in interpreting score results, planning interventions, or delivering feedback to parents. Forms are color-coded, which is helpful, given that the front pages of the CARS2-ST and CARS2-HF appear nearly identical at first glance.

Items are labeled according to which domain of functioning they are designed to assess, and descriptions under each rating value provide examples of behavior corresponding to each rating. The Likert-type scales are useful in suggesting specific areas where problems exist and determine exactly where the individual functions relative to the ASD specific population. In the

CARS2-HF, there is a description specific and unique to each item category. As high functioning autism has a particular criterion, it is important to capture more subtle aspects of behavior; this is helpful for professionals when using the CARS2-HF to rate individuals. Also, the individual's behavior should be compared with that of a typically developing individual of the same age. It is necessary for the professional to have substantial background knowledge of human development to complete the CARS2. Also, professionals who have had exposure to and training about autism can acquire the ability to make valid ratings.

Characteristics of the Test Manual

The CARS2 manual includes information on administering, scoring, and interpreting the assessment. Chapter 1 provides an introduction, general description, and intended uses of the CARS2. Chapter 2 and 3 detail how to administer and score the CARS2-ST and CARS2-HF and how to use the CARS2-QPC to inform ratings. Chapter 4 includes guidelines for interpretation, case examples, and how to provide diagnostic feedback to parents. Chapter 5 provides intervention planning and resources based on results as well as case examples. Chapter 6 discusses the development of the CARS method. Lastly, chapter 7 provides information about the psychometric properties of the assessment.

The CARS2 test manual is clear and well organized. The organization of the material is clear to readers to follow material flows nicely. For both the CARS2-ST and CARS2-HF, each item is broken down to help users understand how to make observations and ratings. The manual also provides information on the items most and least likely to receive high ratings for specific age groups and diagnoses. In addition, the chapter about how to provide diagnostic feedback when the diagnosis is or is not ASD is beneficial for professionals to strategically plan before communicating with parents. The manual also provides guidelines for intervention planning

based on item responses. There are detailed descriptions of effective interventions and appropriate goals for each item. Practical case examples are also included in chapters 1 through 5 to illustrate guidelines for "real-life" situations. Lastly, chapters 6 and 7 provide a thorough overview of the test's item development and psychometric properties.

<p style="text-align:center">Standardization Sample and Norms</p>

There were three large normative data samples used for purposes of standardization: the original CARS development sample, the CARS2-ST verification sample, and the CARS2-HF developmental sample. In the 1988 edition, the original CARS development sample consisted of individuals referred for evaluation at one of the five Division TEACCH centers in North Carolina, which was very limited for generalization purposes. The CARS2 developers improved their normative samples for the CARS2-ST and CARS2-HF by basing their participant demographic information on the 2000 U.S. Census data. In addition, the original CARS development participants could only choose between White, Black, and Other as options for their ethnic background. This disparity was appropriately revised in the CARS2-ST and CARS2-HF normative samples in an attempt to accurately represent the U.S. population. The race or ethnicity category of the CARS2-ST and CARS2-HF included Asian/Pacific Island, Black/African American, Hispanic/Latino, Native American, White, and Other.

The CARS2-ST verification sample consisted of 1,034 individuals with a diagnosis of autism in various clinical settings. Ages of individuals in the sample ranged from 2 to 36, with 30% in the 2 to 5 age group, 43% in the 6 to 10 age group, 20% in the 11 to 15 age group, and 7% in the 16 to 36 age group. For research purposes, all age groups in the CARS2-ST normative sample consisted of at least 100 participants, except for the 16 to 36 age group. This is a

limitation as the 16 to 36 age group included 73 participants and should include a minimum of 100 participants for increased statistical significance.

The CARS2-HF normative sample was similar in composition to the CARS2-ST sample, particularly concerning age and gender. Of the 994 total individuals in the CARS2-HF development sample, participants were diagnosed with a variety of disorders including high-functioning autism (n = 248), Asperger's Disorder (n = 231), Pervasive Developmental Disorder – Not Otherwise Specified (PDD-NOS) (n = 95), Attention Deficit Hyperactivity Disorder (ADHD) (n = 179), Learning Disorder (n = 111), other internalizing and externalizing clinical disorders (n = 69). The sample also included 40 general education students and 21 nonautistic individuals placed in a special education classroom. The ages of the sample range from 6 to 57 years old, with 35% in the 6 to 10 age range, 41% in the 11 to 15 age range, and 24% in the 16 to 57 age range. All age groups in the CARS2-HF consisted of at least 100 participants, suitable for research purposes. However, the CARS2 developers did not specify how many participants represent a single age within the age group. This may be a limitation because students of specific ages may not be accurately represented in both normative samples.

Compared to 2000 Census data, the racial and ethnic background of the CARS2-ST verification sample did not accurately represent the U.S. population. In the CARS2-ST sample, 60% of the participants identified themselves as White, 16% as Black/African American, 13% as Hispanic/Latino, 7% as Asian/Pacific Islander, and 4% as Other. Whereas the 2000 Census data identified 77% as White, 13% as Black/African American, 13% as Hispanic/Latino, 4% as Asian/Pacific Islander, and 7% as Other. The Asian/Pacific Islander and Black/African American populations were overrepresented, and the White and Other race or ethnicities were underrepresented. The racial and ethnic background of the CARS2-HF developmental sample

accurately represented the U.S. population compared to 2000 Census data, except the Hispanic/Latino population. The normative sample included 6% of the Hispanic/Latino population, while the 2000 Census sample identified 13% as Hispanic/Latino. Approximately 73% of the sample described their ethnic background as White, 14% as Black/African American, 3% as Asian/Pacific Islander, 1% as Native American, and 3% identified as being from other ethnic backgrounds.

Of the CARS2-ST verification sample and CARS2-HF development sample's demographic characteristics, the geographical regions were categorized as Northeast, Midwest, South, and West. In each region, the state groupings were not specified. The CARS2-ST verification sample's geographical regions do not accurately represent the U.S. population based on 2000 Consensus data. In the normative sample, 29% of the sample was from the Northeast, while the 2000 Census reports 19%, and 11% of the sample was from the Midwest, while the 2000 Census reports 23%. Also, 43% of participants were from the South, while the 2000 Census reports 35%, and 17% of participants were from the West, while the 2000 Census reports 23%. The Northeast and South regions were heavily overrepresented, and the Midwest and West regions were underrepresented in the CARS2-ST normative sample.

Compared to 2000 Census data, the CARS2-HF verification sample's geographical regions did not accurately represent the U.S. population. In the sample, 20% of the participants were from the Northeast, 21% from the Midwest, 48% from the South, and 10% from the West. However, 2000 Census data reported 19% from the Northeast, 23% from the Midwest, 35% from the South, and 23% from the West. The South was extremely overrepresented, and the Northeast was slightly overrepresented. While the West was significantly underrepresented, and the Midwest was somewhat underrepresented.

The CARS2-ST and CARS2-HF normative data samples consisted of skewed distributions for gender to represent the populations of individuals with autism and individuals with high-functioning autism. The CARS2-ST verification sample included 22% of females and 78% of males, and the CARS2-HF development sample included 22% of females and 78% of males. For both normative data samples, the age distributions were similar among females and males. In addition, in the CARS2-ST verification sample, approximately 81% of participants had IQs at or below 79, and 19% had IQs between 80 and 85. This skewed distribution was purposeful as the CARS2-ST is intended for referred individuals with relatively low cognitive functioning. Lastly, all sample participants in the CARS2-HF development sample had an IQ of 80 and above, consistent with the premise that the CARS2-HF is intended for use with individuals of average or better cognitive abilities.

Finally, the CARS2-ST and CARS2-HF standardization samples included the head of household education level as an indicator of socioeconomic status. However, 7% of the CARS2-ST sample and 16% of the CARS2-HF sample did not report on education level, which is not representative of the entire sample. Test users should interpret results related to this sample demographic norm with caution because education and socioeconomic status are not synonymous entities. In the CARS2-ST verification sample, the some college level is significantly underrepresented (13%) when compared to the 2000 Census data (26%), and the less than high school graduate level is underrepresented (13%) when compared to the 2000 Census data (18%). The remaining standardization sample education levels include high school graduate (34%), college graduate (20%), and postgraduate (13%). In the CARS2-HF development sample, less than high school graduate level (8%) was heavily unrepresented compared to the 2000 Census data (18%). Also, the some college level (15%) was significantly

underrepresented compared to the 2000 Census data (26%). The rest of the standardization sample education levels include high school graduate (30%), college graduate (17%), and postgraduate (14%).

Scores and Interpretation

The CARS2 manual provides detailed instructions for scoring the assessment. Scoring the CARS2 is a very straightforward process, and both the CARS2-ST and CARS2-HF rating booklets have the same scoring processes. In the CARS2-ST and CARS2-HF, the individual will be given a rating for each of the 15 items. For each item, the individual has seven possible ratings: "1" indicating within normal limits for that age, "1.5" indicating very mildly abnormal for that age, "2" displaying mildly abnormal for that age, "2.5" showing mildly-to-moderately abnormal for that age, "3" indicating moderately abnormal for that age, "3.5" indicating moderately-to-severely abnormal for that age, and "4" indicating severely abnormal for that age. The total score can range from 15 to 60. On both the CARS2-ST and CARS2-HF, the scoring sheet is located on the front page of the booklet.

After the professional gathers the necessary information and completed observations, the CARS2-ST or CARS2-HF rating scale is conducted independently by the professional. The developers note that scoring should not be calculated or interpreted unless all 15 item areas have been rated, and all ratings have been made based on reasonably reliable sources. Once all ratings are completed, the professional transfers the ratings for all 15 items from the inside of the booklet to the corresponding spaces provided in the Summary section on the booklet's front page. Then, the professional will calculate the sum of the ratings to obtain the total raw score. The total raw score is written in the designated "Total raw score" box. Next, the professional uses the total

raw score to indicate the Severity Group that corresponds with the total raw score and makes a checkmark in the appropriate box.

In the CARS2-ST, the three possible Severity Groups include: Minimal-to-No Symptoms of ASD (scores 15-29; 15-27.5 for ages 13+), Mild-to-Moderate Symptoms of ASD (scores 30-36.5; 28-34.5 for ages 13+), and Severe Symptoms of ASD (scores 37 and higher; 35 and higher for ages 13+). In the CARS2-HF, the three possible Severity Groups include: Minimal Symptoms of ASD (scores 15-27.5), Mild-to-Moderate Symptoms of ASD (scores 28-33.5), and Severe Symptoms of ASD (scores 34 and higher). Both the CARS2-ST and CARS2-HF provide cutoff score values intended to inform professionals of the further need to evaluate the presence of autism. The CARS2 developers suggest that individuals with a total raw score at or above the minimum score of the Mild-to-Moderate Symptoms of ASD Severity Group are categorized as likely to have an Autism Spectrum Disorder, which can include autism, Asperger's Disorder, or PDD-NOS. In that case, an autism diagnosis should be a strong hypothesis in a comprehensive differential diagnosis.

To obtain a standard score in the form of a T-score, the professional circles the value that corresponds to the raw score table's total raw score located on the right side of the Summary section. The T-score and percentile rank are the numbers printed to the left of the value the T-scores indicate symptom levels compared to individuals with autism spectrum diagnoses. It is worth noting that the purpose of the CARS2 is to provide more detailed information about individuals with a diagnosis of ASD or who are highly suspected of having a diagnosis. This must be taken into consideration when interpreting scores on the CARS2, as scores reflect functioning relative to the specific ASD population, rather than typically developing children.

Before interpreting CARS2-ST or CARS2-HF scores, the professional needs to summarize the level of confidence in the accuracy of the information on which ratings are based in each area of behavior (Vaughan, 2011). Ideally, data would be obtained from multiple formats and settings, along with a thorough developmental history. The developers suggest that the professional considers the degree to which information is internally consistent. The professional should always consider the potential impact of any critical gaps or potential biases in the sources that have been drawn upon for determining the ratings.

The manual provides information on items most and least likely to receive high ratings for specific age groups and diagnoses to aid with interpretation. The tables are included consisting of interpretive categories associated with the CARS2-ST and CARS2-HF T-score ranges and raw score ranges. These tables also indicate a standard error of measurement (SEM) value for both the CARS2-ST and CARS2-HF. The SEM provides a way of considering external factors not directly related to characteristics being measured that may influence a test score. The SEM is used to calculate the confidence interval and the "true" score, or the score uninfluenced by any external factors. For the CARS2-ST or CARS2-HF T-scores, the "true" score will fall in a range of 3 T-score points, or 1 SEM, above and below the obtained score 68% of the time, or in a range of 6 T-score points, or 2 SEMs, above and below the obtained score 95% of the time. Incorporating these estimates into test score interpretation is vital for professionals, so that too much emphasis is not placed on a single number as an exact representation of a respondent's characteristics (Vaughan, 2011).

Overall, when scoring the CARS2, the professional is simply responsible for transferring the 15 ratings to the scoring sheet, calculating the sum of the ratings (simple math), and corresponding the total raw score to the T-score and percentile rank. The T-scores and percentile

ranks are already calculated in the table on the scoring page. This suggests that while scoring, the

chances of an error occurring are minor. I found scoring and interpreting the scores of the

CARS2 to be a quick, simple, and unambiguous process.

Psychometric Properties

Reliability

The CARS2 manual defines reliability as the consistency and relative accuracy with

which test results estimate the characteristic a test is intended to measure (Schopler et al., 2010).

All measurements contain error, whether the measurements are of physical attributes, objects, or

behaviors. Reliability helps to estimate the degree to which such error must systematically be

taken into account when interpreting a test score. The test developers determined reliability by

using measures of internal consistency and interrater reliability from both versions of the CARS

standardization sample data.

According to the CARS2 manual, estimates of internal consistency reflect one source of

systematic error that affects the test's reliability (Schopler et al., 2010). It is desirable that a test,

such as the CARS2-ST or CARS2-HF, yield consistent and accurate results across all items

within a test. To examine internal consistency reliability, Cronbach's alpha coefficient was used,

an appropriate measure for the CARS because of its estimate for a multiple response format. The

original CARS measured internal consistency with a sample of 537 individuals referred for their

first diagnostic session. Raters observed the sessions through one-way windows and completed

the ratings immediately following the session. Coefficient alpha was computed for these ratings

to assess the internal consistency of the total scores. The original CARS total raw score alpha

estimate in the sample was .94, which is adequate for clinical and educational purposes.

The CARS measured internal consistency estimates on the CARS2-ST verification sample and CARS2-HF development sample, using alpha scores for the total raw score estimates and 15 domains. In the CARS2-ST verification sample, an internal consistency estimate of .93 was obtained, adequate for clinical and educational purposes. The estimated alpha scores for the 15 domains on the CARS2-ST ranged from .43 to .81, with a median value of .69. For subscales with a limited number of items, a .70 alpha estimate is adequate. In the CARS2-HF development sample, an internal consistency estimate of .96 was achieved, sufficient for clinical and educational purposes. The estimated alpha scores for the 15 domains on the CARS2-HF ranged from .53 to .88, with a median value of .79. On both the CARS2-ST and CARS2-HF, the Levels and Consistency of Intellectual Response domain contained the most measurement error. Overall, professional using the CARS2 should be cautious when using domain scale scores as resources for educational decision making because the scores include a considerable amount of measurement error.

As stated in the CARS2 manual, interrater reliability estimates examine whether similar CARS2-ST or CARS2-HF ratings are given by different professionals based on the same clinical information about a given individual (Schopler et al., 2010). To assess the original CARS interrater reliability, individual item rating values were examined for ratings made by several sets of two independent, trained rater for 280 of the cases in the developmental sample. On the original CARS, a correlation of .84 was obtained for total scores. Interrater reliability estimates centered around a median correlation of .71, with a range from .55 to .93. This indicates an adequate consistency across raters for each item.

For the CARS2-HF, interrater reliability was examined for ratings made by several sets of two independent, trained rater for 239 participants of the developmental sample. A correlation of

.95 was obtained for total scores. The interrater reliability estimates for item ratings centered around a median correlation of .73, with a range from .53 to .93. These results indicate an adequate consistency across raters for each item. In addition, weighted *kappa* estimates were calculated for the pairs of raters in the study. Weighted *kappa* evaluates categorical, an absolute agreement between pairs of raters across the seven rating categories for each item and considers a consistent selection of nearby categories when the absolute agreement is not achieved. For the CARS2-HF, the results indicated a median level of agreement of .73 and a range from .51 to .90. These results suggest an adequate agreement between raters in the study. Interrater reliability estimates were not provided for the CARS2-ST.

As the CARS2 is intended for use by a wide variety of professionals who are not necessarily experts in autism, interrater reliability was also assessed by comparing ratings made by visiting professionals with limited autism experience with ratings made by expert clinical directors (Vaughan, 2011). The correlation of visitors' total scores with directors' total scores on the original CARS was .83 (p < .01). Similarly, interrater reliability estimates for the CARS were calculated for ratings given by different professionals based on various sources of clinical information about referred individuals; these correlations ranged from .73 to .82, indicating good agreement across raters. Overall, the ratings of the CARS2 have been demonstrated to be reliable across different settings, for different sources of information for the same individual, and different raters.

Validity

As a psychometric concept, validity refers to the accuracy and appropriateness of the interpretation of a test score in a given setting. The validity of the internal structure of the CARS2 was assessed by calculating correlations between item ratings and total raw scores. For

both the CARS2-ST and the CARS2-HF, correlations among item ratings were moderate to high. In the CARS2-ST, correlations ranged from .42 to .77, and in the CARS2-HF, correlations ranged from .40 to .79. Factor analyses were conducted for both the CARS2-ST and CARS2-HF. However, neither of the factor analyses were consistent with the autistic triad of communication, social relations, and repetitive/restrictive behaviors (Dawkins et al., 2016).

Of the original CARS instrument, sensitivity and specificity were observed using a total raw score cutoff value of 30. This cutoff value correctly identified 87% of individuals in the development sample as autistic or not autistic. Using the total raw score alone, identifying those with or without an autism diagnosis resulted in a sensitivity value of .88 and a specificity value of .86. Ratings for the CARS2-ST sample were consistent with these findings. Similarly, sensitivity and specificity values for the CARS2-HF for distinguishing high-functioning individuals with ASD from all nonautism groups in the sample were .81 and .87. Rating patterns were compared across groups with different diagnoses, levels of cognitive function, and age. For the CARS2-ST sample, a shift in ratings occurred for those with a diagnosis of autism and a higher level of cognitive functioning. However, these differences were not statistically significant. Differences in rating patterns also existed between older individuals with autism and younger individuals with autism. Analysis of rating patterns on the CARS2-HF revealed substantial differences between individuals with Asperger's disorder and individuals diagnosed with autism.

Concurrent validity was demonstrated by comparing total scores on the CARS2-ST and CARS2-HF with results obtained from other measures of autism. Correlations with the Autism Diagnostic Observation Schedule (ADOS) were .79 for the CARS2-ST total score and .77 for the CARS2-HF total score. These results indicated a somewhat strong relationship between clinician

ratings on two very different measures. The correlation between clinician-generated scores on the CARS2-HF total score and mothers' scores on the Social Responsiveness Scale (SRS) was moderate (.47), as was the correlation between the CARS2-ST total score and the SRS (.38). As expected, the relationship is stronger for higher functioning individuals rated on the CARS2-HF than for individuals rated on the CARS2-ST. The correlation between the CARS2-ST and the Autism Behavior Checklist (ABC) was .67. This correlation was robust, however there was no correlation available between the CARS2-HF and the ABC.

Overall, the validity of the CARS2 was adequately established. However, the test-retest reliability and construct validity could not be concluded for its acceptability, given the lack of available analysis studies (Vaughan, 2011). To draw a more definitive and clinically useful conclusion, more carefully designed studies with minimal bias and comprehensive reporting would be necessary.

Conclusions

One overarching roadblock I encountered while critiquing the CARS2 is the fact that there is very limited outside literature on the CARS2. When searching on the Chapman University Library website, the only article to appear was Vaughan's Review of CARS2. It was difficult to accurately critique the reliability and validity of the CARS2 when limited outside sources are available.

Overall, I believe the CARS2 is a useful tool for supporting the diagnostic process and forming intervention recommendations once a diagnosis has been made. Standardization samples were overall adequate and were broadly representative of the populations the CARS2 is made to assess. However, the CARS2 developers can improve the test's standardization sample to include a more precise, accurate representation of the U.S. population. Reliability was appropriate for

informing diagnosis and research. Validity was adequately established and indicated that the interpretation of scores from the CARS2 is accurate across settings, informants, and age groups. All in all, the CARS2 is a comprehensive measure for the assessment of behavioral problems related to autism.

References

Dawkins, T., Meyer, A. T., & Van Bourgondien, M. E. (2016). The Relationship Between the

Childhood Autism Rating Scale: Second Edition and Clinical Diagnosis Utilizing the

DSM-IV-TR and the DSM-5. *Journal of Autism and Developmental Disorders, 46*(10), 3361.

https://doi-org.libproxy.chapman.edu/10.1007/s10803-016-2860-z

Schopler, E., Van Bourgondien, M. E., Wellman, J., & Love, S. (2010). *Childhood autism rating

scale-second edition (CARS2): manual*. Los Angeles: Western Psychological Services.

Vaughan, C. A. (2011). Review of Childhood Autism Rating Scale (2nd ed). *Journal of

Psychoeducational Assessment, 29*(5), 489–493. https://doi-

org.libproxy.chapman.edu/10.1177/0734282911400873