Description and Critique of the Child Autism Rating Scale- Second Edition (CARS 2)

XXXXXXX

Chapman University

Description and Critique of the Child Autism Rating Scale- Second Edition (CARS 2)

Description of the Test

The *Child Autism Rating Scale-* Second Edition (CARS-2) provides three rating forms designed to identify the behavioral symptoms of Autism Spectrum Disorder (ASD) and aid in the diagnostic process. Authored by Eric Schopler, Ph.D, Mary Bourgondien, Ph.D, G. Janette Wellman, Ph.D and Steven Love, Ph.D, the CARS-2 now offers the improved *Children Autism Rating Scale, Second Edition-* Standard Test (CARS2-ST), formally known as the CARS, and the *Child Autism Rating Scale Second Edition-* High Functioning (CARS2-HF). The CARS2-HF version is more sensitive to individuals with higher IQ scores, better verbal skills, higher functioning social skills, less apparent behavioral symptoms or who fall under the category high functioning Autism or who were formerly diagnosed under Asperger's Syndrome.

The CARS2-ST and CARS2-HF are not intended as screeners for the general population, but as an aid in the diagnostic process of individuals of all ages and functional levels. Their brief, quantitatively specific and reliable summaries provide value to practitioners who are testing suspected individuals with ASD. The test also includes the Questionnaire for Parents or Caregivers (CARS2-QPC) which is an unscored questionnaire intended to acquire more information from parents or caregivers to assist in developing the diagnostic hypothesis of the referred individuals.

Content

The CARS2 consists of three forms (CARS2-ST, CARS2-HF and the CARS2-QPC) and the updated manual. The CARS2 is an untimed assessment appropriate for use in schools, clinics, and intervention programs. It can be administered by professionals such as physicians,

special educators, school psychologists, speech pathologists and audiologists, who all have exposure to and training about Autism and are able to make valid ratings. Aspects of this test can also be administered by parents, caregivers or teachers who are observing the child. Because it is an observational assessment, other people observing the student are able to rate the student. Although this can increase the accessibility to more professionals, it also increases the subjectivity of the test because people can have different ideas of what constitutes abnormal behavior. The observational aspect of the test also puts the test at risk for the Halo or Pitchfork effect if parents are completing the CARS2-QPC and are biased towards representing their child a certain way to gain or prevent a diagnosis. In order to counter this potential affect, the CARS2-ST and CARS2-HF ratings during parent interviews, classroom observations or a case history review, do not produce the diagnosis. In order to obtain a diagnosis, the child's developmental history, medical symptoms and other unique characteristics must be evaluated by experienced professionals who are trained in Autism and able to make clinical diagnoses.

The CARS2-ST should be administered for assessing individuals who have an IQ of 79 or lower, noticeable communication impairments or under the age of six years old. The CARS2-HF should be administered to individuals who have an IQ of 80 or higher, relatively sufficient verbal skills or older than six years of age. The separation of these two tests is a benefit to the second edition of the test because the test can be appropriately used for higher functioning individuals.

A key difference between the CARS2-ST and the CARS2-HF is that information from multiple sources is needed to complete the CARS2-HF. For example, information from a direct observation of the child being rated and an interview with someone who knows the child is needed to complete the CARS2-HF. The growing number of individuals being referred for a

diagnosis of high functioning Autism makes having multiple sources of information crucial to diagnosing properly. This is because the CARS2-HF is intended to help determine if a high functioning individual is demonstrating sufficient symptoms of Autism, which can be difficult and not as straightforward as students who are tested on the CARS2-ST. The CARS2-ST ratings, on the other hand, can be completed based on the information from a single source such as direct observation, parent report, or clinical records. Because the ratings can be gathered under a variety of conditions, multiple sources of error can occur (i.e. Temporal, source, method and setting). Although the ratings are sufficient with one source of information, an administrator should gather multiple points of data, regardless of which test they are administering before giving the final rating score. It should be pointed out that to officially diagnosis individual with Autism more than one source of information is needed.

Characteristics of Test and Manual

Both the CARS2-ST and the CARS2-HF examine 15 items that correspond to 15 domains of behavior. The CARS2-SF items are relating to people, imitation, emotional response, body use, object use, adaptation to change, visual response, listening response, taste smell and touch response, fear or nervousness, verbal communication, nonverbal communication, activity level, level and consistency of intellectual response, and general impressions. The CARS2-HF items are very close to the standard version; however, they differ slightly because it is assumed that the children being tested on the CARS2-HF have a higher level of IQ and more advanced skills. For example, instead of an item on imitation, the CARS2-HF substitutes an item that examines social-emotional understanding. Another example is instead of emotional response on the standard version, the high functioning version examines emotional expression. The last difference in items is the CARS2-HF tests thinking and cognitive integration skills while the

CARS2-ST only examines the activity level of the child. Each of the 15 items are scored as values from 1 through 4. A score of 1 indicates that the child's behavior is within normal limits for his/her age, while a 4 indicates the child's behavior is severely abnormal for his/her age. In addition, there are midpoint values available if the child's behavior is in-between two of the values. Therefore, the breakdown of ratings 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

The child's scores on each of the 15 items are combined to produce a total score. The total score cutoff values are given to determine if a more extensive evaluation for Autism is necessary. The total scores can also be converted into a standard score that is based on the developmental sample.

Even though the CARS2 is vulnerable to subjectivity of the scorer, because the test is designed to be a supplemental aspect of the evaluation process, the child's score on the CARS2 should not determine the diagnosis. Therefore, the test allows for a broad understanding of where the child's behaviors fall within a variety of signs and symptoms of Autism. Lastly, an advantage of the CARS2 is that the forms describe a range in extensiveness and intensity of symptoms and behaviors to help the parent feel confident that their child's behaviors are represented in the evaluation.

The manual for the CARS2 is very user friendly and is well organized. To further explain the scoring aspect of the test, the manual breaks down each of the items and clearly describes the definition of the item, the aspects to take into consideration when scoring, and examples of behaviors that coincide with the 1 through 4 ratings. The manual also lays out a full case example that administrators can follow and use as a guide when administering the test themselves. The case example shows what the administrator's rating booklet should look like with raw scores, observation notes and correlation to severity group. Finally, the CARS2 item ratings in areas of difficulty can be linked to specific intervention planning and strategies. This means the CARS2 goes further than suggesting if the child has Autism or not, showing the comprehensiveness of the test. Overall, the test and manual are comprehensive and user friendly, furthering the effectiveness of the CARS2.

Standardization and Sample Norms

The CARS-2 provides sample sizes for all three sets of data. In the original CARS data from 1988, developers used a sample of n=1,606. Although this could seem sufficient on the surface, when looking closer, it is found that more than half of that sample were under five years of age, even though the test was normed for ages two and older. In addition, the ethnic background of these 1,606 adolescents was lacking vital information. The ethnic information of the participants was divided into three categories: Black, White and Other with percentages of each being 28%, 62% and 3% respectively. The original CARS also did not list any geographical locations of where participants came from, making it impossible to know if this test is representative of the population.

The CARS-2 Standard Test (CARS-2ST) created in 2010, improved on the sample issues from the original CARS. The total population tested for the CARS-2ST was n=1,034 with a more

even distribution of ages, strengthening the CARS2 generalizability. The CARS-2ST population included 308 participants with ages between two and five years old, 446 with ages between six and ten years old, 207 between the ages of 11 and 15 and finally 73 participants were between 16 and 36 years old. This is a much more even sample size than the original CARS. It should also be pointed out that over two thirds of the sample size was male participants. This is, however, reflective of population of people who have Autism, as males are more likely to have Autism. The CARS-2ST also provided a more descriptive ethnic background table and a geographic table which were lacking in the original CARS. The ethnic background table shows 68 participants (7%) were Asian, 167 (16%) were Black, 131 (13%) were Hispanic and 626 (60%) were White. When compared to the 2000 Census, this sample is a good representation of the country's population. The geographical locations of the sample were also listed. This table showed that 305 participants (29%) were from the Northeast, 115 (11%) were from the Midwest, 440 (43%) were from the South and 174 (17%) were from the West. All four regions of the US were represented which further strengthens the generalizability of the CARS-2.

It is also worth noting that the entire sample had an IQ lower than 85. This was intentionally skewed because the CARS-2 is intended to test children who are referred for having low cognitive functioning. However, as mentioned earlier, the CARS-2 also has a version for High Functioning individuals. The sample size for this test was n=994 who all have an IQ higher than 80. The breakdown of diagnosis for the CARS-2HF includes four categories. These categories are High Functioning Autism, Asperger's, PDD NOS and Other.

Scores and Interpretation

The scoring of the CARS2 is a rating scale of abnormal behavior. Due to this, one must take into strong consideration the age of the child being assessed. In addition, the frequency,

duration, intensity and the peculiarity of the behavior should be considered. In order to account for potential error, the manual clearly defines each item and gives detailed description of what the behavior can look like for each item and each score (1-4) within each item. The observations to get a score for each item can occur in one of the various settings including the psychological testing environment, classroom observation, parent reports or clinical records. The CARS2 allows for all observations to come from a single setting, making the convenience of the CARS2 a strength. This can, however, also be a limitation because allowing scores to come from one source can lead to error or a misrepresentation of particular behaviors. Another potential limitation is that in order to use the detailed examples explaining the behaviors for each item, an administrator would have to memorize the examples or bring the manual with them when observing the child. However, once the observations are done and the scores have been selected, the rating booklet is very straightforward and user-friendly.

As explained in the manual and reporting booklet, the scores gathered from any of these settings on the 15 items are totaled to make the child's Total raw score. Then the Total raw score is linked with a severity grouping based on the cutoff score which is predetermined to be a score of 30. This score was determined because the average score in the developmental sample of n= 1,034, was 38.5 with a standard deviation of 8.4. The raw score can be converted into a T-score and percentile rank by the straightforward table in the rating booklet.

Before interpreting the score, it is important to consider the full level of confidence and accuracy of the information that the ratings scores were based on. Ideally, the information on the behavior of the child would come from multiple sources who know the child and from multiple settings over a long developmental period. In addition, when documenting observations from parents or teachers, it is important to consider their biases because one person's view of peculiar

behavior can be different from someone else's view. Once the information has been determined to be appropriate and accurate, then the scores from the CARS2-ST and CARS2-HF can be interpreted as a supplement to an evaluation.

Psychometric Properties

Reliability

Similar to the original CARS, the internal consistency for the CARS2-ST and the CARS2-HF remained robust in the second edition. By using Cronbach alpha, the internal consistency of the standard form and high-functioning version came out to be .93 and .96, respectively. These numbers reveal a strength of the test because the internal consistency is above the minimum threshold (.90) for clinical use. This means the CARS2 can justifiably determine Autism in a clinical or educational application. Item to item consistency ranged from .43 to .81 for the CARS2-ST and from .53 to .88 for the CARS-HF. These findings show that the test, as a whole, measures Autism and not unrelated aspects of behavior. This justifies the test combining all 15 ratings into a single Total score which determines if a more extensive evaluation should be considered. To assess interrater reliability for the original CARS, individual ratings were examined by several sets of two independent trained raters. The interrater reliability ranged from .55 to .93, with an estimate of .84. For the CARS2-HF, the estimate interrater reliability was .95, with ranges from .53 to .93. The estimated score of .95 reflects there was a very high agreement between the raters of the test. The overall average of the interrater reliability estimates center on a median of .71 for the CARS and .73 for the CARS2-HF. Because an estimate average of .70 is adequate, both the original CARS and the CARS2-HF are above the minimum threshold. Unfortunately, these estimates were not supplied for the CARS2-ST. This representss as a limitation of the test. Another limitation of the CARS2 is that it was stated that

the interrater reliability was examined by "trained raters", but did not provide any other information about who these raters were or their credentials.

The CARS was originally created for experts in Autism, while the CARS2 was designed to be used by people who have limited experience in Autism. As a result, test makers wanted to make sure the CARS2 was reliable for non-experts of the field, as these are the types of people who will be administering the test. To do this, they examined the ratings made by professionals who are not experts in Autism, such as medical students, pediatric residents, interns, special educators, speech pathologists or audiologists. These professionals from different disciplines were given a brief introduction to the test, asked to read the manual and view a 30-minute training tape, and then provide their own ratings. These ratings were compared to ratings by the clinical directors observing the same taped session. They found a high significant correlation (r= .83, p <.01) and a 92% agreement between the two groups. These numbers show a strength to the CARS2. The small sample size of only 18 cases, however, is a limitation to the study. In addition, there was lack of information on what constitute someone as an expert or non-expert in Autism. Providing information on the criteria for determining and distinguishing between an expert and a non-expert, would have been helpful.

To show the CARS2 has stable ratings over time, studies assessing the test-retest reliability for the CARS were discussed in the manual. The test-retest examination accounts for any temporal source of error, because an effective test for Autism should show a child having Autism in the first test and in the second test because Autism doesn't go away. For the CARS, in 91 cases, a test-retest was done one year apart and found a resulting correlation of .88 which means that they were not statistically different. Because of the high correlation of .88, the CARS shows it has good stability over time. Other independent studies have also shown the CARS with

stability over time. Perry, Condillac, Freeman, Dunn-Geier and Belair (2005) studied 47 cases of retest stability of the CARS and found a 3-month retest correlation of .77, which indicates good stability. Although the stability over time of the CARS could be applied to the CARS2, information specifically for the CARS2 would be helpful in determining the effectiveness of the CARS2.

To account for variation of scores due to measurement error, the Standard Error of Measurement (SEM) was discussed for the CARS2-ST and CARS2-HF. The SEM for the CARS2-ST Total raw score is 0.68, meaning that true Total raw score is likely to fall above or below one point of the obtained score. For the CARS2-ST T-score, the SEM is 2.7, meaning the true T-score likely falls, above or below, within 3 T-scores of the original T-score. Similar to the standard version, for the CARS2-HF, the SEM for the Total raw score is 0.73, meaning the true Total raw score falls above or below one point of the obtained score. For the CARS2-HF T-score, the true score will fall above or below 3T scores because the SEM is 2.8. Although these errors are relatively low, it should be pointed out that these scores were tested with a 68% confidence interval. Although a 68% confidence band is adequate for clinical purposes, to better understand the error of measurement, a higher confidence band should be used.

Validity

The validity of a test shows that a test is actually effective in identifying what it is meant to identify. Therefore, validity evidence will show that the CARS2-ST and CARS2-HF can both accurately identify children with Autism. The internal structure of the CARS2-ST and the CARS2-HF found moderate to high correlations among item ratings. Correlations of item ratings with the total raw score were found to be uniformly higher for each item than for inter-item ratings. Correlation among items for the CARS2-ST ranged from .42 to .77 and ranged from .40

to .79 for the CARS2-HF. This is compared to the correlation of item ratings with the total raw score which averaged to be r = .81 and r = .86 for the CARS2-ST and the CARS2-HF, respectfully.

Validity was also assessed by comparing the ratings from the original CARS of individuals in the development sample to expert classifications. This was done to determine the relationship of the scores to three categories: not autistic, autistic with mild-to-moderate level of symptoms, and autistic with severe level of symptoms. This analysis showed the sensitivity of the test to be .88, meaning the test was able to correctly identify individuals with Autism. The CARS also showed to have a specificity of .87, meaning the test was able to correctly identify those without Autism at an adequate level. These analyses were based off a cut off score of 30. The score of 30 was determined to be the cutoff because the average CARS2ST raw score for the sample n=1,034 was 38.5 with a standard deviation (SD) of 8.4. This supports the cutoff of 30 because it is one SD away from the mean. With this cutoff, the test correctly identified 87% children with Autism. When deciding to change the cutoff raw score of 30, it must be taken into consideration that lowering a test's cutoff score to identify cases of Autism increases the test's sensitivity, but decreases its specificity. On the other side, raising the test's cutoff threshold in order identify cases decreases the test's sensitivity and increases its specificity.

Finally, when assessing for criterion-related validity, ratings from the original development sample of the CARS correlated with independent clinical ratings of Autism symptoms at a correlation of r= .84 (p< .001). Criterion-related validity means the test is valid against other measures testing something similar. Therefore, this correlation indicates that the scores of the CARS have a high validity when compared to clinical ratings. According to Mayes, et al. (2009), when comparing the Checklist for Autism Spectrum Disorder, the CARS, and other

tests of Autism for congruent validity, the diagnostic agreement between CARS and Checklist scores was 98%. Other studies mentioned in the manual show findings from Eaves and Milner (1993) reporting 76% of the items from the CARS2-ST and the Autism Behavior Checklist score to be correlated at .67. This correlation is very close to the desired .7 correlation for congruent tests. The manual also mentions that the 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 also indicate a strong relationship between the two tests' ratings. All of these studies show the validity of the CARS to be very high, suggesting it to be an acceptable assessment measurement of Autism.

Conclusion

Overall, the CARS2 is adequate tool for assessing if a child has Autism. Because of the potential subjectivity of the test, it should not be used as the only data to support a formal diagnosis. Nevertheless, due to the user-friendly rating booklet and manual, it is an excellent tool for many types of administrators who see children in many different settings. The fact that the CARS2 is well-researched and found to be both reliable and valid, supports that it is an appropriate tool to use when observing children with Autism. A highlight of the CARS2 is that it provides not only an interpretation that a child has Autism, but also includes intervention techniques that can be a useful resource for administrators to give teachers or parents, or use themselves. Though there are some limitations to the CARS2, such as the 68% confidence interval and the subjectivity, the test is a very appropriate tool to use as an aspect of the diagnosing process.

References

Mayes, S. D., Calhoun, S. L., Murray, M, J., Morrow, J. D., Yurich, K., Mahr, F., Cothren, S., Purichia, H., Bouder, J. N., Petersen, C. (2009). Comparison of Scores on the Checklist for Autism Spectrum Disorder, Childhood Autism Rating Scale, and Gilliam Asperger's Disorder Scale for Children with Low Functioning Autism, High Functioning Autism, Asperger's Disorder, ADHD, and Typical Development. *Journal of Autism Development Disorders, 39,* 1682–1693. doi: 10.1007/s10803-009-0812-6.

Perry, A., Condillac, R. A., Freeman, N. L., Dunn-Geier, J., & Belair, J. (2005). Multi-site study of the Childhood Autism Rating Scale (CARS) in five clinical groups of young children. *Journal of Autism and Developmental Disorders, 35*(5), 625-634.

Schopler, E., Van Bourgondien, E., Wellman, G. & Love, S. (2011). The Child Autism Rating Scale: Test Review. *Journal of Psychoeducational Assessment, 29*(5), 489–493.