


Assessment of Cognitive Outcome Measures in Teenagers with 15q13.3 Microdeletion Syndrome

Emeline Crutcher^{1,2,3} · May Ali^{2,3} · John Harrison⁴ · Judit Sovago⁵ · Baltazar Gomez-Mancilla^{5,6} · Christian P. Schaaf^{1,2,3} 

© Springer Science+Business Media New York 2016

Abstract 15q13.3 microdeletion syndrome causes a spectrum of cognitive disorders, including intellectual disability and autism. We aimed to determine if any or all of three cognitive testing systems (the KiTAP, CogState, and Stanford–Binet) are suitable for assessment of cognitive function in affected individuals. These three tests were administered to ten individuals with 15q13.3 microdeletion syndrome (14–18 years of age), and the results were analyzed to determine feasibility of use, potential for improvement, and internal consistency. It was determined that the KiTAP, CogState, and Stanford–Binet are valid tests of cognitive function in 15q13.3 microdeletion patients. Therefore, these tests may be considered for use as objective outcome measures in future clinical trials, assessing change in cognitive function over a period of pharmacological treatment.

Keywords Neuropsychiatric disease · Autism · Intellectual disability · Cognitive tests · Cognitive function

✉ Christian P. Schaaf
schaaf@bcm.edu

¹ Translational Biology and Molecular Medicine, Baylor College of Medicine, Houston, TX, USA

² Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA

³ Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, 1250 Moursund St, Suite 1325, Houston, TX 77030, USA

⁴ Alzheimer Center, VU Medical Center, Amsterdam, The Netherlands

⁵ Novartis Pharmaceutical Corporation, Basel, Switzerland

⁶ McGill University, Montreal, Canada

Introduction

15q13.3 microdeletion syndrome (OMIM #612001) is caused by small deletions in the extremely unstable q13.2–q13.3 region of chromosome 15, and is implicated in multiple neurodevelopmental disorders (Tropeano et al. 2014; Lowther et al. 2014). This chromosomal region contains seven genes and is flanked by a breakpoint (BP) on each side. The BPs are marked by clusters of low copy repeat (LCR) elements, which are vulnerable to inversion and subsequent non-allelic homologous recombination (NAHR), resulting in deletion of the involved region (Gillentine and Schaaf 2015). Most clinical cases of this syndrome present with microdeletions occurring between BP4 and BP5 (Tropeano et al. 2014). About 80 % of patients with this syndrome have one or more neuropsychiatric diagnoses, with 57.5 % diagnosed with developmental disability/intellectual disability (ID), 10.9 % diagnosed with autism spectrum disorder, 15.9 % diagnosed with speech problems, and 6.5 % diagnosed with attention deficit hyperactivity disorder (ADHD) (Lowther et al. 2014). The neuropsychiatric phenotypes of 15q13.3 microdeletion syndrome have been proposed to be caused by haploinsufficiency of *CHRNA7*, which is one of the seven genes in the region affected by the microdeletions (Gillentine and Schaaf 2015).

CHRNA7 codes for the $\alpha 7$ subunits composing the $\alpha 7$ homopentameric nicotinic acetylcholine receptor ($\alpha 7$ nAChR), which is expressed throughout the brain (Schaaf 2014). Here, the receptors are involved in neuronal calcium signaling, which helps to mediate synaptic plasticity, learning, and memory (Gillentine and Schaaf 2015). The heterozygous deletion of this gene, as present in 15q13.3 microdeletion cases, has been proposed to lead to a decreased number of functional receptors, causing altered

calcium signaling and, consequently, deficits in cognitive function. Agonists of $\alpha 7$ nAChR have been shown to improve cognitive function in patients with schizophrenia during phase 1 clinical trials (Freedman 2014). However, the therapeutic potential of such treatments in individuals with 15q13.3 microdeletion syndrome has not been assessed to date.

As is true for other groups of individuals with ID, there is a lack of well-established, objective tests validated for assessment of cognitive function in 15q13.3 microdeletion patients. This makes it difficult to accurately track changes in cognitive function of these patients throughout life. In addition, the efficacy of potential treatments in respective clinical trials is assessed primarily through subjective questionnaires. This results in a large placebo effect, and likely contributes to the low number of approved treatments for individuals with ID (Sandler 2005).

Here, we assess performance of three test systems for cognitive function in a cohort of 15q13.3 del patients, aiming to identify potential outcome measures for clinical trials to be considered in this patient population.

Methods

Participants

Ten study participants were recruited for this study. Individuals were recruited from our own patient registry, based on a known diagnosis of 15q13.3 microdeletion syndrome. That diagnosis had been made by clinical chromosome microarray analysis prior to this study. A total of 18 families were contacted, with affected individuals between 12 and 21 years of age. The first 10 individuals to respond

were consented for enrollment under a research protocol approved by the Institutional Review Board. Clinical details for these individuals are listed in Table 1.

Procedures

Components of three test systems were administered to each subject: KiTAP (test of attentional performance in younger children), CogState, and Stanford–Binet (5th edition). Both the KiTAP and CogState were administered by the principal investigator (C.S.). These tests contain multiple components the user can choose to administer based upon the cognitive function they wish to assess. Components administered to these subjects, and cognitive functions assessed by each, are detailed in Table 2 for the KiTAP and Table 3 for the CogState. Outcome measures assessed for each component are those listed as primary outcome measures by each test’s instruction booklet and are noted in Table 2 for the KiTAP and Table 3 for the CogState. Data generated from each component of the KiTAP and the CogState tests were analyzed using Microsoft Excel. The Stanford–Binet test (5th edition) was administered by certified, research-reliable psychologists. The test administrator, following the instructions given in the test’s instruction booklet, calculated scores for the Stanford–Binet.

For the KiTAP and CogState components (except as noted below), regarding the “Feasibility” assessment (Table 4), individuals “Able to Perform Test” were determined to be those who were able to respond to the test stimuli adequately enough to generate a score for the test component indicated. The subjects determined to be unable to perform the test component got zero answers correct for that component. “Basal Effect” represents subjects that

Table 1 Clinical information of 15q13.3 deletion subjects

Subject ID	Sex	Age at time of visit	Coordinates of deleted region ^b	Inheritance
1	M	14 years 11 months	chr15:30971530–32445252	Maternal
2	M	14 years 2 months	chr15:30304529–33316817	Unknown (adopted)
3 ^a	M	18 years 9 months	chr15:29213945–32439113	Maternal
4 ^a	M	15 years 5 months	chr15:29213945–32439113	Maternal
5	M	14 years 0 months	chr15:32081775–32449185	Maternal
6	M	18 years 11 months	chr15:31292475–32445252	Unknown (adopted)
7	M	14 years 8 months	chr15:29224601–32445252	Unknown (not maternal)
8	M	18 years 8 months	chr15:30971530–32925141	Unknown (adopted)
9	M	16 years 9 months	chr15:31292475–32925142	De novo
10	F	15 years 0 months	chr15:31073735–32446830	Maternal

^a Subjects 3 and 4 are siblings

^b Mutation coordinates based on UCSC Genome Build hg19. Deleted region represents the minimal interval deleted, as determined by the first and last deleted oligonucleotide on chromosome microarray analysis

Table 2 Primary outcome measures and cognitive domains assessed by test components of the KiTAP

Test component administered	Primary outcome measure	Cognitive domain assessed
The witch	Mean reaction time	Psychomotor speed
The owls	Missed signals	Attention
Ghost's ball	Missed signals	Working memory
Dragon's house	False reactions	Executive functions

Table 3 Primary outcome measures and cognitive domains assessed by test components of the CogState

Test component administered	Primary outcome measure	Cognitive domain assessed
Detection	Mean reaction time to correct responses	Psychomotor speed
Identification	Mean reaction time to correct responses	Attention
One-back	Correct responses	Working memory
International shopping list task	Correct responses	Verbal episodic memory
International shopping list task (delayed same day)	Correct responses	Verbal episodic memory delayed—same day
International shopping list task (delayed 24 h)	Correct responses	Verbal episodic memory delayed—24 h later
One card learning	Correct responses	Visual episodic memory
Groton maze learning test	Number of errors made	Visual episodic memory

responded incorrectly to greater than 50 % of the test component stimuli. This was set under the assumption that individuals with <50 % accuracy are essentially guessing. Percent incorrect was calculated using the equation:

$$\% \text{ incorrect} = \left(\frac{\text{incorrect responses}}{\text{incorrect responses} + \text{correct responses}} \right)$$

“Ceiling Effect” indicates the subjects that made no errors in response to stimuli on that test component. For the international shopping list task (ISLT) and ISLT delayed (same day and 24 h) components of the CogState, and all Stanford–Binet components, “Able to Perform Test” were determined to be those who attempted to respond to the questions presented, and “Basal Effect” represents subjects who generated the lowest possible score for that component.

For the “Proof of Potential for Improvement” assessment (Table 5), mean percentile rankings for the KiTAP outcome measures were obtained through converting the T-scores calculated by the testing software to percentile rankings using Appendix B in the KiTAP users manual. The normed values for this calculation were from test scores generated by healthy 8–10 years olds. Mean percentile rankings for the CogState components were obtained by calculating Z-scores for each individual and outcome, using normed data obtained from a true age-matched cohort. These Z-scores were then converted to percentile rankings using the table found on MedFriendly (“MedFriendly: Standard Score to Percentile Conversion,” n.d.) Z scores were calculated using the formula below.

$$Z = \left(\frac{\text{subject score} - \text{mean norm score}}{SD \text{ norm}} \right) \times (\pm 1)$$

For both the KiTAP and the CogState tests, mean percentile rankings were calculated for all outcome measures with normed data available for each component. For the CogState, mean percentile rankings were calculated only for components and outcome measures with true age-matched norms available. Mean percentile rankings for the Stanford–Binet were calculated by the test administrator, and are based on standard scores obtained by comparison to normed data from a true age matched cohort.

To assess the internal consistency of each test’s components, Cronbach’s alpha or Kuder Richardson 20 was calculated (Table 6). The test used to calculate this value was dependent upon the outcome variable assessed for each test component. Cronbach’s alpha calculation was used to determine internal consistency in outcome variables with response measured on a scale i.e. reaction time (RT). Kuder Richardson 20 calculation was used to determine internal consistency in outcome variables with dichotomous responses i.e. correct or incorrect response. Test components with a value at or above 0.70 are considered to have an acceptable level of internal consistency. All internal consistency calculations were performed by assessing response via the primary outcome variable for each component, as determined by the testing instructions for each test. If more than one primary outcome variable was listed for a component, the first variable listed was used for the calculation.

Table 4 Feasibility of test components in 15q13.3 deletion subjects

Test	Component administered	Able to perform test component (%)	Basal effect (%)	Ceiling effect (%)
KiTAP ^a	The witch	100	0	100
	The owls	100	20	10
	Ghost's ball	100	20	0
	Dragon's house	100	0	10
CogState ^a	Detection	80	0	20
	Identification	100	20	30
	One-back	100	40	10
	International shopping list task	100	0	0
	International shopping list task delayed (same day)	100	0	0
	International shopping list task delayed (24 h)	100	10	0
	One card learning	100	30 ^c	0
	Groton maze learning test	100	0	0
Stanford–Binet ^b	Non-verbal IQ	100	0	0
	Verbal IQ	100	0	0
	Full-scale IQ	100	0	0
	Fluid reasoning	100	0	0
	Knowledge	100	0	0
	Quantitative reasoning	100	0	0
	Visual-spatial	100	0	0
	Working memory	100	0	0

^a n = 10 for all components

^b n = 9 for all components (one subject was unable to take the Stanford–Binet due to scheduling issues)

^c All subjects close to showing a basal effect for this component

Results

Based on the results in Table 4, individuals with 15q13.3 deletions are capable of performing all administered components of both the KiTAP and CogState test, with minimal basal and ceiling effects. All individuals were able to perform all test components of both the KiTAP and the CogState, except for the detection component of the CogState. For this component, only 8 of the 10 individuals were able to generate a score (Table 4). Components for both tests exhibited minimal basal effects in our cohort, with only two components displaying basal effects above 20 %. Ceiling effects were also minimal for components of both tests, with all components displaying ceiling effects over 10 % having a primary outcome measure of mean RT rather than number of correct or incorrect answers. Table 5 shows definite potential for score improvement on components of all three tests in our 15q13.3 deletion cohort. As expected, the mean percentile rankings of our cohort compared to age-matched norms for both the CogState and the Stanford–Binet test components were very low, with all but one falling below the 20th percentile (Table 5). For the KiTAP test

components, the percentile rankings were mostly below the 50th percentile. Based on the standard cut-off value of 0.70 for acceptable internal consistency, all evaluated test components in both the KiTAP and the CogState have an acceptable level of internal consistency in our cohort (Table 6).

Discussion

This is the first description of validation of these three test systems, assessing various elements of cognition, including memory, and attention, in a cohort of individuals with 15q13.3 microdeletion syndrome. Both the KiTAP and CogState have been used to assess cognitive function in ADHD populations, but use of the CogState has never been verified for use in subjects with ASD or ID (Jucaite et al. 2014; Mollica et al. 2004; Yamashita et al. 2011; Rothmann et al. 2014; Hellwig-Brida et al. 2011; Kaufmann et al. 2010). Assessment of feasibility for use of the KiTAP in a Fragile X syndrome (FXS) cohort has been previously investigated, and it was found that this test is a valid assessment of cognitive function in FXS patients (Knox

Table 5 Potential for improvement in 15q13.3 deletion subjects

Test	Component administered	Outcome measures analyzed	Average percentile ranking
CogState ^a	Detection	Mean reaction time to correct responses	2
	Identification	Mean reaction time to correct responses	3
	One-back	Mean reaction time to correct responses	3
KiTAP ^b	One card learning	Accuracy of performance	3
	The witch	Median reaction time	37
		Standard deviation of reaction time	32
	The owls	Omissions	46
		Errors	57
	Ghost's ball	Omissions	44
		Errors	48
	Dragon's house	Errors	39
		Median reaction time	64
		Total index	48
Speed accuracy index		27	
Stanford–Binet ^a	Non-verbal IQ	Score	17
	Verbal IQ	Score	9
	Full-scale IQ	Score	12
	Fluid reasoning	Score	38
	Knowledge	Score	6
	Quantitative reasoning	Score	13
	Visual-spatial	Score	11
Working memory	Score	18	

^a Age matched (true age)

^b Compared to 8–10 years old normative population

Table 6 Internal consistency in the KiTAP and CogState for 15q13.3 deletion subjects

Test	Component administered	Outcome measure analyzed	Internal consistency score
CogState	Detection	Mean reaction time	0.997 ^b
	Identification	Mean reaction time	0.820 ^b
	One-back	Correct	0.975 ^c
	One card learning ^a	Correct	0.734 ^c
	Groton maze learning	Errors	0.847 ^c
KiTAP	The witch	Mean reaction time	0.894 ^b
	The owls	Missed signals	0.959 ^c
	Ghost's ball	Missed signals	0.689 ^c
	Dragon's house	False reactions	0.901 ^c

^a Removed P9 due to subject receiving too few stimuli

^b Cronbach's Alpha statistical test used

^c Kuder–Richardson 20 statistical test used

et al. 2012). Although the 5th edition of the Stanford–Binet included some subjects diagnosed with ASD in its normative population, its ability to accurately assess cognitive function in these individuals has not been validated (Grondhuis and Mulick 2013; Ozonoff et al. 2005; Silverman et al. 2010).

Impact of Similarities and Differences Between the KiTAP, CogState, and Stanford–Binet Cognitive Batteries

Although these three test systems all assess cognitive function, the KiTAP and CogState assess more ‘plastic’

cognitive functions, while the Stanford–Binet is the gold standard for IQ testing and assesses cognitive functions that are more resistant to change. While the functions assessed by the KiTAP and the CogState are nearly identical, there is little overlap between these functions and those assessed by the Stanford–Binet. All three test systems are composed of multiple components, each meant to assess a different cognitive domain. However, unlike the KiTAP and CogState, the Stanford–Binet has a single outcome reading for each component. The Stanford–Binet is hand-scored by the test administrator, and is given a numerical score between 0 and 2, based upon the subject’s response. Importantly, the difficulty level at which you start each component of the Stanford–Binet depends on performance in a placement section at the beginning of the test, which dictates the starting position of each component for that individual. The testing for each individual will continue until that subject scores all zeros on a sub-section of a component. Because of this, different subjects may be responding to different questions on each component, based upon where their starting level was determined to be, and how far past the starting point they make it. The Stanford–Binet, like other IQ tests, is considered a poor choice for measuring clinical outcomes over the period of a standard trial treatment (Aman et al. 2004). This is partially due to the 0–2 type scoring method of these tests, which creates low sensitivity, and makes it almost impossible to detect small changes in cognitive function (Aman et al. 2004). The KiTAP and CogState are more sensitive to change due to the broad range of outcome measures collected for each component. Additionally, each component of these tests administers the same type of stimulus repeatedly over the period of the test component, and collects response data for each outcome measure upon each stimulus. This allows a more accurate assessment of how a subject is responding to a particular stimulus, as opposed to each stimulus being different from the next, and scored separately, which is the case for the Stanford–Binet. Since both the KiTAP and CogState are computerized, the possibility of human bias or error affecting the results is greatly reduced compared to a test that is hand-scored, such as the Stanford–Binet. Although standardized IQ tests such as the Stanford–Binet may not be the best measure of change in cognitive function over the period of a clinical trial, they are still useful for long term assessments, as well as for baseline comparisons to performance on less characterized cognitive tests, such as the KiTAP and CogState.

Although the KiTAP and CogState test components measure very similar cognitive functions, the method by which each test’s stimuli evoke a response is quite different. The KiTAP was specifically designed to engage children, and each test component has a theme based around the main premise of being in a haunted castle. Each component

contains colorful animations serving as the response stimulus. Additionally, each component’s theme is different from the others, with themes including colorful caricatures of ghosts, witches, and dragons. Conversely, the CogState was not designed to engage children specifically, and each test component’s response stimulus is similar. While each component of the CogState asks a different question (i.e. is the card black vs. have you seen this card before), the stimulus presented remains the face of a random playing card. This type of repetitive stimulus across multiple test components presents the challenge of not being able to hold the attention of test takers as well as the more engaging KiTAP. This is especially true for adolescent populations, or any populations having attention deficits. Considering that our population was composed mostly of teenagers (average age = 15.6 years), and ADHD is one of the known clinical manifestations of 15q13.3 deletion, we conclude that the CogState may not be as good an assessment of cognitive function as the KiTAP in that particular population. This is supported by the fact that the two components displaying a basal effect over 20 % both come from the CogState test. Additionally, the percentile rankings of the CogState components are much lower than either the KiTAP or the Stanford–Binet. Although some of the difference between the CogState and KiTAP mean percentile rankings can be explained by age differences in the sources of normed values (age matched vs. 8–10 years olds), it is unlikely that the 4–8 years difference between normed values and actual age of test subjects in the KiTAP has such a profound effect on difference in percentile ranking. If this were the case, we would expect to see much more similar percentile rankings between components of the CogState and Stanford–Binet, since percentile rankings for both were calculated from age-matched norms. The lack of variance in mean percentile ranking between components of the CogState when compared to variances between components of the other two tests reinforces the concept that monotony in the CogState stimuli may result in less efficient distinction of different cognitive domains when compared to the KiTAP and Stanford–Binet. A larger cohort of 15q13.3 deletion individuals, as well as a healthy, age-matched control cohort for all components of the CogState, and KiTAP tests would help determine if indeed one of these tests is superior to the other for measuring cognitive function in 15q13.3 deletion patients.

Individuals with 15q13.3 Deletion are Capable of Obtaining Valid Scores on Components of the KiTAP, CogState, and Stanford–Binet Cognitive Batteries

When validating cognitive tests in a population with cognitive abnormalities, it is important to ensure that the

subjects have the ability to respond to the test stimulus adequately enough to generate a valid score. Inability to generate a score on a test component could be a result of unwillingness to participate, or physical or mental disabilities resulting in a lack of ability to respond to the test stimulus appropriately. Based on the observation that all subjects were able to perform all but one test component, we conclude that our cohort is capable of responding to the stimuli presented by each test. For the one CogState test component that only 8 of 10 subjects were able to perform, it is likely that the two non-responding subjects were simply not engaged with this particular test component as they were both able to respond to all other test components with similar stimuli. It is also essential to verify that there is not a large basal or ‘floor’ effect, as is often the case when such populations take tests created to assess cognitive function in the normal population. The presence of a basal effect indicates that some individuals in the cohort being studied are generating a score at or below the guessing rate. This could indicate that the individual did not understand the test component instructions, or that they were not fully engaged in taking the test during that component. Ultimately, a large basal effect for a test component indicates that component is not a good measure of cognitive function in that group of individuals, likely because the task requires a higher level of function than the average individual in the group is capable of. The overall percentage of basal effect across test components of all three tests is low, and indicates that most of these test components are at an appropriate level of difficulty for our cohort. The presence of some basal effect in our subject population is expected due to the wide range and severity of phenotypes in 15q13.3 deletion patients. There were two components of the CogState presenting with a basal effect higher than 20 %. It is possible that these components are testing cognitive function at a level unsuitable for many of these subjects. However, it would be ideal to test these components on a larger cohort to determine if these results hold true over a more inclusive population of 15q13.3 deletion patients. Verifying a low ceiling effect is important in order to ensure that the cohort being studied does not perform so well on the test components that there is no room to improve scores. It is important to note that ceiling effect is based off of correct answers, as mentioned in the “[Methods](#)” and “[Results](#)” section. Therefore, for those test components whose primary outcome measure does not depend on correct response (i.e. mean RT), a large ceiling effect holds no meaning. This is the case in our cohort for all test components with a ceiling effect over 10 % (Table 2). Based on this, we can conclude that the test components assessed in this study are testing cognitive function at a level that challenges our cohort, and that improvement upon initial test scores is possible.

Ample Potential for Improvement in 15q13.3 Deletion Cohort

To assess potential for improvement of test scores upon increased cognitive function in our 15q13.3 deletion cohort, we calculated the average baseline percentile rankings for our cohort across test components for all outcome measures having available normed values. The Stanford–Binet has been thoroughly standardized, with norms available for age matched standardization of each subject’s scores. This has been done to a lesser extent for the KiTAP and CogState. For the KiTAP, a complete set of norms is available for age groups 6–7 and 8–10. However, no norms for age groups older than that are available. In the case of the CogState, a complete set of norms for primary outcome measures is available for age groups 18 and older, but only incomplete norms exist for ages younger than 18. While other assessments of test validity can be performed, this makes it difficult to accurately determine the average percentile ranking of our 15q13.3 deletion cohort. Some argue that these percentile rankings are not valid in ID populations (Knox et al. 2012). While this may be accurate in some cases, the main point of assessing these rankings in our case is to show that there is definite room for improvement in the scores generated on the respective test components in individuals with 15q13.3 microdeletion syndrome. Comparing the scores generated by our cohort to those generated by healthy controls, we show that this is indeed the case. Both for the CogState and the Stanford–Binet tests, all of the mean percentile rankings obtained by the 15q13.3 deletion cohort are below the 50th percentile. The CogState mean percentile rankings are especially low (~3rd percentile), while the Stanford–Binet rankings are more variable depending on the skill being measured (6th–38th percentile). This is to be expected since the Stanford–Binet was developed primarily to assess IQ and reasoning ability, while the CogState was developed to assess more interactive cognitive functions such as psychomotor speed and attention. For the KiTAP, most of the mean percentile rankings are still below average (<50th percentile). However, it is important to remember that these percentile rankings compare our cohort to a normal cohort aged 8–10. Since our subjects were 14–18 years old (true ages) at the time of testing, it is reasonable to expect that they will score higher when compared to 8–10 years olds than when compared to an age-matched cohort. The fact that the large majority of primary outcomes show percentile scores below average leaves room for improvement, in particular when considering clinical trials of pharmacologic intervention. Although a change in IQ scores (as measured by the Stanford–Binet) may not be seen over the course of a treatment trial period, it is likely that an effective treatment could be identified by

monitoring changes in cognitive functions with the more sensitive CogState and KiTAP tests.

Acceptable Levels of Internal Consistency Across Components of the KiTAP and CogState Cognitive Batteries

Internal consistency is a measure of intra-individual agreement between different items/questions testing the same general function. A value of 1 represents perfect internal consistency, and any value above 0.70 indicates that the test component has an acceptable level of internal consistency.

Cronbach's alpha is the statistical test most often used to determine internal consistency across a group of questions meant to measure the same function. However, this test was created to measure how consistently individuals answered questions with the response being measured on a scale. For several test components we wished to assess, the primary outcome measure is not measured on a scale, but rather dichotomous (number of correct responses or number of errors). For these cases, Kuder–Richardson Formula 20 was used.

Since each component of these tests is essentially asking the same question over and over again, we would expect to see an alpha value close to 1 for components which the subjects understood instructions well, and were able to respond similarly to each question/stimulus.

Internal consistency statistics were not performed on Stanford–Binet components due to the nature of scoring (not every child starts at the same place, therefore questions will be different between children).

Based on the internal consistency statistics presented in Table 6, it appears that each evaluated test component in both the KiTAP and the CogState has acceptable internal consistency in our cohort, with all but two components achieving a score above 0.80. There is only one test component, “sustained attention” from the KiTAP, which is on the border of the acceptable range. This value rounds up to 0.7 from 0.69, and would potentially be higher upon evaluation of a larger group of subjects.

Conclusions

Based on the data presented, we conclude that the KiTAP, CogState, and the Stanford–Binet are valid measures of cognitive function in individuals with 15q13.3 microdeletion syndrome. Along with previous studies in individuals with FXS, the data presented herein provide encouragement that some cognitive tests established in healthy patients may be used to assess cognitive function in individuals with psycho-cognitive abnormalities stemming

from a range of genetic or environmental causes. The 15q13.3 microdeletion syndrome has great potential for therapeutic intervention, with haploinsufficiency of *CHRNA7* being considered a likely contributor to the overall phenotype, and promising alpha7 agonist drugs and positive allosteric modulators being developed (Schaaf 2014). This pilot study provides important groundwork that will be useful for sample size calculations and consideration of quantifiable outcome measures to be considered in clinical trials for individuals with 15q13.3 microdeletion syndrome.

Acknowledgments We are indebted to the patients and their families for their willingness to participate in our study. This work was supported in part by the Doris Duke Charitable Foundation Grant # 2011034 and in part as an Investigator Initiated Research Protocol by Novartis Pharmaceuticals Corporation. The project was supported in part by IDRC Grant Number 1U54 HD083092 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development. Cores: Tissue culture core, translational core. CS is generously supported by the Joan and Stanford Alexander Family.

Disclosure John Harrison has previously received consultancy payments from the Novartis Pharmaceutical Corporation.

Author Contributions All listed authors have made contributions to the work discussed in this article. All authors have reviewed the completed manuscript, and given feedback, resulting in the final submitted product. Specific contributions of each author are listed below. Emeline Crutcher: Analyzed data collected during the study and wrote the manuscript discussing the results. May Ali: Clinical research coordinator—established and maintained contact with the patients' families throughout the study to coordinate travel arrangements and clinical visits for administration of each test. John Harrison: Assisted with study design, and provided valuable input throughout the data analysis process. Judit Sovago: Assisted with study design. Baltazar Gomez-Mancilla: Assisted with study design. Christian P. Schaaf: Primary investigator for this study. Planned, designed, and conducted the study.

References

- Aman, M. G., Novotny, S., Samango-Sprouse, C., Lecavalier, L., Leonard, E., Gadow, K. D., et al. (2004). Outcome measures for clinical drug trials in autism. *CNS Spectrums*, 9(1), 36–47.
- Freedman, R. (2014). α 7-nicotinic acetylcholine receptor agonists for cognitive enhancement in schizophrenia. *Annual Review of Medicine*, 65(September 2013), 245–261.
- Gillentine, M. A., & Schaaf, C. P. (2015). The human clinical phenotypes of altered *CHRNA7* copy number. *Biochemical Pharmacology*. doi:10.1016/j.bcp.2015.06.012.
- Grondhuis, S., & Mulick, J. (2013). Comparison of the Leiter International Performance Scale-Revised and the Stanford–Binet Intelligence Scales, 5th Edition, in children with autism spectrum disorders. *American Journal on Intellectual and Developmental Disabilities*, 118(1), 44–54.
- Hellwig-Brida, S., Daseking, M., Keller, F., Petermann, F., & Goldbeck, L. (2011). Effects of methylphenidate on intelligence and attention components in boys with attention-deficit/hyperactivity disorder. *Journal of Child and Adolescent Psychopharmacology*, 21(3), 245–253.

- Jucaite, A., Öhd, J., Potter, A. S., Jaeger, J., Karlsson, P., Hannesdottir, K., et al. (2014). A randomized, double-blind, placebo-controlled crossover study of $\alpha 4\beta 2^*$ nicotinic acetylcholine receptor agonist AZD1446 (TC-6683) in adults with attention-deficit/hyperactivity disorder. *Psychopharmacology*, *231*(6), 1251–1265.
- Kaufmann, L., Zieren, N., Zotter, S., Karall, D., Scholl-Bürgi, S., Haberlandt, E., et al. (2010). Predictive validity of attentional functions in differentiating children with and without ADHD: A componential analysis. *Developmental Medicine and Child Neurology*, *52*(4), 371–378.
- Knox, A., Schneider, A., Abucayan, F., Hervey, C., Tran, C., Hessel, D., et al. (2012). Feasibility, reliability, and clinical validity of the Test of Attentional Performance for Children (KiTAP) in Fragile X syndrome (FXS). *Journal of Neurodevelopmental Disorders*, *4*(1), 2. doi:10.1186/1866-1955-4-2.
- Lowther, C., Costain, G., Stavropoulos, D. J., Melvin, R., Silversides, C. K., Andrade, D. M., et al. (2014). Delineating the 15q13.3 microdeletion phenotype: A case series and comprehensive review of the literature. *Genetics in Medicine*, *17*(2), 149–157.
- MedFriendly: Standard Score to Percentile Conversion. (n.d.). Retrieved from <http://www.medfriendly.com/standardscoretopercenileconversion.html>. Accessed 2 June 2015.
- Mollica, C. M., Maruff, P., & Vance, A. (2004). Development of a statistical approach to classifying treatment response in individual children with ADHD. *Human Psychopharmacology*, *19*(7), 445–456.
- Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*(3), 523–540.
- Rothmann, K., Hillmer, J., & Hosser, D. (2014). Evaluation of the Musical Concentration Training with Pepe (MusiKo mit Pepe) for children with attention deficits. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, *42*(5), 325–335.
- Sandler, A. (2005). Placebo effects in developmental disabilities: Implications for research and practice. *Mental Retardation and Developmental Disabilities Research Reviews*, *11*(2), 164–170.
- Schaaf, C. P. (2014). Nicotinic acetylcholine receptors in human genetic disease. *Genetics in Medicine*, *16*(9), 649–656.
- Silverman, W., Mizejeski, C., Ryan, R., Zigman, W., Krinsky-McHale, S., & Urv, T. (2010). Stanford-Binet and WAIS IQ differences and their implications for adults with intellectual disability (aka mental retardation). *Intelligence*, *38*(2), 242–248.
- Tropeano, M., Andrieux, J., Vassos, E., & Collier, D. A. (2014). Clinical utility gene card for: 15q13.3 microdeletion syndrome. *European Journal of Human Genetics*, *22*(11), 1338.
- Yamashita, Y., Mukasa, A., Anai, C., Honda, Y., Kunisaki, C., Koutaki, J., et al. (2011). Summer treatment program for children with attention deficit hyperactivity disorder: Japanese experience in 5 years. *Brain and Development*, *33*(3), 260–267.