**This is a post-peer-review, pre-copyedit version of an article published in the *Journal of Developmental and Physical Disabilities*. The final authenticated version is available online at: https://doi.org/10.1007/s10882-022-09871-8**

# Salivary Cortisol and Stereotypy in Minimally Verbal Children with Autism: A Pilot Study

Marie-Michèle Dufour[1,2], Marc J. Lanovaz[1,2,3] et Pierrich Plusquellec[1,3]

[1]École de psychoéducation, Université de Montréal

[2]Institut universitaire en déficience intellectuelle et en trouble du spectre de l'autisme

[3]Centre de recherche de l'Institut universitaire en santé mentale de Montréal

**Correspondence**

Correspondence concerning this article should be addressed to Marie-Michèle Dufour, École de Psychoéducation, Université de Montréal, C.P. 6128, succursale Centre-Ville, Montreal, QC, Canada, H3C 3J7. Email: marie-michele.dufour@umontreal.ca

**Statements and declarations**

This work was supported by a doctoral scholarship from the Fonds de recherche du Québec – Société et culture (#191314). The authors have no relevant financial or non-financial interests to disclose. They have no competing interests to declare that are relevant to the content of this article.

**Author contributions**

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Marie-Michèle Dufour. The first draft of the manuscript was written by Marie-Michèle Dufour and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

## Abstract

*Background:* Several studies have reported conflicting results when assessing associations between stress and repetitive behaviors in individuals with autism spectrum disorder (ASD). Some factors that may have caused these discrepant results include the monitoring of a single broad category for repetitive behaviors, the heterogeneity of the participants, and the use of indirect measures.

*Method:* To address the prior issues, our study explored the relationship between salivary cortisol and direct observation measures of stereotypy in four minimally verbal children with ASD. To this end, we combined an alternating-treatment design with multiple regression analyses to examine the interaction between the two variables.

*Results:* The analyses indicated that the mean value of cortisol was negatively associated with global and motor stereotypy. No significant relation was found between mean value of cortisol and vocal stereotypy.

*Conclusion:* These results highlight the complex relationship between stress and stereotypy and emphasize the relevance of conducting research on a larger scale, which would have a direct impact on our understanding of a core feature of ASD.


*Keywords:* autism, salivary cortisol, stereotypy, stress

**Salivary Cortisol and Stereotypy in Minimally Verbal Children with Autism: A pilot study**

Repetitive behaviors represent a core feature of autism spectrum disorder (ASD; American Psychiatric Association, 2013). Stereotypy is one form of repetitive behaviors that is often classified in two broad categories: vocal and motor stereotypy. Vocal stereotypy involves repetitive vocalizations that persist in the absence of social consequences (Ahearn et al., 2007; Lanovaz & Sladeczek, 2011) whereas motor stereotypy is defined by patterned, repetitive, coordinated, rhythmic and non-reflexive physical behaviors (Mahone et al., 2004). According to a recent review, at least 50% of individuals with ASD display one or more forms of motor stereotypy (Melo et al., 2020). This proportion approaches 90% when vocal forms of stereotypy are included (Chebli et al., 2016).

Researchers have proposed three main hypotheses to explain the maintenance of stereotypy in ASD: (1) the self-stimulatory hypothesis, (2) the manifestation of stress hypothesis, and (3) the self-regulation hypothesis. The self-stimulatory hypothesis considers that stereotypy is maintained by automatic positive reinforcement (Lovaas et al., 1987; Rapp & Vollmer, 2005). That is, stereotypy generates forms of stimulation that function as positive reinforcement for the behavior. Hence, individuals with ASD who engage in stereotypy access this reinforcement. The large body of research on functional analysis showing that stereotypy tends to occur most when individuals with ASD have the least environmental stimulation supports this hypothesis (Querim et al., 2013).

The proponents of the stress hypothesis argue that stereotypy is a manifestation of a physical state including stress and excitement (Muthugovindan & Singer, 2009). Several studies have found a positive correlation between cortisol basal levels and stereotypy, as measured by questionnaires, in individuals with ASD (Bitsika et al., 2015; de Vaan et al., 2020; García-

Villamisar & Rojahn, 2015; Lydon, Healy, Roche et al., 2015; Yang et al., 2015). Specifically, individuals with higher levels of cortisol tend to show higher levels of stereotypy. Some of these authors hypothesized that repetitive behaviors could be a manifestation of the internal state of the individual such as a high level of physical activation. Recently, de Vaan et al. (2020) also found that higher cortisol response was related to more stereotypic behaviors. For its part, Bitsika et al., (2015) found a positive association between cortisol basal levels and stereotypy in children, but not in adolescents, with ASD. Although these studies associate cortisol with stereotypy, the underlying mechanisms responsible for this association remain unknown.

The self-regulation hypothesis has also garnered some support. Groden et al. (1994) created an adaptive and maladaptive theoretical model that contrasts individuals who adopt appropriate coping strategies with individuals who do not. In their model, the authors interpreted stereotypy as a coping strategy within a population that does not have appropriate coping mechanisms. One study supporting this hypothesis found that ASD children who presented a higher level of stereotypy had lower levels of cortisol (Gabriels et al., 2013). After controlling for age, intelligence quotient (IQ), and the type of repetitive behaviors, the authors found that the children who were in the high repetitive behaviors group showed 36% lower diurnal salivary cortisol in comparison with those who were in the low repetitive behavior group. This result suggested that repetitive behavior could play a self-regulation function in this population.

Three hypotheses may explain the discrepant results across these studies. First, repetitive behaviors may take on many forms such stereotypy, self-injurious behavior, tics, compulsions, and restricted interests (Bodfish et al., 2000). With the exception of Lydon et al. (2015), all prior studies have combined these different forms together. The different forms may serve different function, which may explain the inconsistent results. Second, the level of functioning of the children varied across studies. It is possible that repetitive behaviors may serve different functions in verbal and minimally verbal children. Finally, prior studies have all used indirect measures of repetitive behaviors (i.e., questionnaires). Being reported by caregivers, these measures may be somewhat biased or at least not as precise as direct observation measures (van de Mortel, 2008). To our knowledge, no study has evaluated the association between cortisol and repetitive behavior using observational measures.

To address the previous limitations, our study examined the association between cortisol levels and one form of repetitive behavior (i.e., stereotypy) measured using direct observation in minimally verbal children with ASD. Taking into consideration the innovative character of this study in terms of types of measurements, research design, and target population, this study was designed as a preliminary investigation of the association between salivary cortisol and direct observation measures of stereotypy in minimally verbal children with ASD.

## Method

This study follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines (von Elm et al, 2007).

### Participant and Settings

Four children between the ages of 5 and 9 years old participated in this study. To be included in the study, the children had to: (a) already have a diagnosis of ASD provided by an independent multidisciplinary team, (b) present at least one form of stereotypy, and (c) have a score of 3.5 or more on the verbal communication subscale of the Childhood Autism Rating Scale (CARS-2; Schopler et al., 2010). This score indicates that speech was absent or, if present, took the form of peculiar language (jargon or echolalia; Schopler et al., 2010). All participants presented severe symptoms of ASD according to the CARS-2 (T-score > 49). Ana and Oli (fictitious names) did not have means of communication other than informal sounds. Matt and Jim (fictitious names) used nonfunctional words (i.e., echolalia). No participant was taking any psychotropic medication that could have impacted neuroendocrine functioning. Table 1 presents the characteristics of the participants. Our university research ethics board approved the present study and we obtained consent from the legal guardian of each participant. All sessions took place in each participant's home.

## Measures

### *The Childhood Autism Rating Scale (CARS-II)*

The CARS-II (Schopler et al., 2010) is a 15-item observation-based rating system that provides the severity of the autistic symptomology. The CARS-II is also used in the diagnostic process. For children, a score of 30 represents the cut-off for a diagnosis of ASD. The psychometric properties of this tool have been evaluated and found to be adequate (Parkhurst & Kawa, 2018). This measure was not used for diagnostic purposes (which was conducted by an independent multidisciplinary team), but only to characterize our participants and determine eligibility (i.e., minimally verbal status).

### *Saliva Collection*

The first author and research assistants collected sixty samples of saliva from each participant for a total of 240 saliva samples. Four saliva samples were taken on each day that an observation took place (before, during and after an observation session). To prevent teaching undesirable behaviors to the participants (e.g., spitting) and to facilitate data collection, we used the same procedure as described by Dozier et al. (2006). The saliva samples were taken with Salivabio children's swab®. The data collector briefly introduced the swab into the child's mouth in order to slightly humidify. She then dipped the swab in a small bag containing 0.8 g of flavored beverage crystal (Tang orange flavor drink mix®). The data collector put the swab back in the child's mouth until it was wet with saliva. She then put the wet swab in the designated storage tube and placed it in a freezer for preservation prior to all samples being sent out to the lab. A study reported no significant differences between samples collected with and without low levels (0.8g) of drink mix (Gordon et al., 2005). All saliva collections took place in the afternoon, which represents a reliable procedure that reduced inter-participant variability considering the nycthemeral rhythm of cortisol secretion (Sharpley et al., 2016).

The lab used a high sensitivity enzyme immune assay kit from Salimetrics State College, PA to analyze the salivary cortisol concentrations. The samples were brought to room temperature to be centrifuged at 15,000 xg (3,000 rpm) for 15 min. The sensitivity of this assay ranged from 0.012 - 3 µg/dl. The lowest limit of detection was 0.007 µg/dl. All samples with a CV% >15 were rerun, unless the difference between the two values was below .03 µg/dl. To examine the cortisol level, we used the mean value level of cortisol for each session. This decision was based on two arguments: (1) by using the mean value, we eliminated the impact of a missing value on our four cortisol samples, and (2) according to the 5% trimmed mean test, outliers had no effect on this indicator whereas they do on areas under the curve with respect to ground ($AUC_g$).

### *Direct Observation Measures*

The first author videotaped each session, and manually coded each video recording stereotypy second-by-second with the support of two research assistants. Table 2 presents the behavioral definitions for each form of stereotypy emitted by the participants. If stereotypy was present at any time during a second, the observer coded the behavior as occurring during the second. When coding, the observer also made the distinction between vocal and motor stereotypy. We also created a global stereotypy variable which was the total percentage of time that the participant emitted stereotypy, vocal and motor combined.

For the analyses, a percentage was calculated by dividing duration in which the participant was engaging in stereotypy by the total duration of the session and multiplying the quotient by 100. A second observer measured stereotypy on 33% of the recordings. The mean second-by-second interobserver agreement was .96 (range: .93-.98) and the mean kappa interobserver agreement was .86 (range: .68-.96).

**Procedures**

Each child participated in 15, 30-min observation periods. The weekly frequency varied slightly depending on their availability but was on average one to two sessions per week. For each session, four cortisol measures were taken: 10 min prior to the beginning of the session, 15 min after it started, at the end of the session, and 20 min after it ended. To make the levels of stress vary during our observations, we designed three conditions inspired from the stress ingredients of Dickerson and Kemeny (2004): control, unpredictability, and novelty. Each child participated in five sessions of each condition for a total of 15 sessions per participant. The conditions varied randomly to avoid order effects. We used an alternating-treatment design to compare levels of stereotypy across the three conditions. In the week preceding the first session, the trainer visited each participant's home twice for a 30-min play period. The purpose of these visits was for the child to habituate to the data collector's presence and feel less stressed by the presence of a new person. During these visits, the data collector also observed the child and completed the CARS-II.

### *Control*

In the control condition, the participants had access to preferred items, received continuous attention from the data collector, and no demands were made. The trainer stayed with the child, participated in play, and ignored all forms of stereotypy.

### *Unpredictability*

In the first 5 min of this condition, the participants had access to their preferred item. Then, the trainer removed that item and offered a less favored toy for the next 10 min (preferred and less favored toys were identified by the parents). Then, the participants were asked to perform two different tasks for the next 10 min (2 x 5 min). Tasks included activities like pairing identical images, puzzles, and classifying objects according to their color. For the last 5 min, the data collector offered another non-preferred item to the child. The order of the task and the less favored item presentation varied during each session in order to maintain the unpredictability component.

### *Novelty*

In this condition, participants were allowed to interact with their preferred item for the first 5 min. Then, the trainer removed that item and offered a new toy for the next 10 min. After that, the participants were asked to perform a new task (10 min). Finally, another new toy was presented for the last 5 min of the session. All toys or tasks presented in this condition were items with which the child had never interacted before. Figure 1 shows the structure of the sessions.

## Analysis

To examine patterns across conditions, the first author graphed stereotypy and cortisol levels on alternating-treatment graphs. We applied a rule developed by Lanovaz et al. (2019) to identify any significant differences between conditions. Prior to the main analysis, we performed an overall inspection of our data as well as an assessment of the normality of our variables. Inspection of the data revealed some missing data (no missing data for the stereotypy; 1.2% of missing data for cortisol). Because we decided to use the mean indicator of cortisol, these missing data had no or few effects on our analyses. The three cortisol missing samples were due to the collection of an insufficient volume of saliva for the analyses. Our inspection revealed that cortisol data were positively skewed, as it is frequently the case with cortisol distributions (Adam & Kumari, 2009). We tried to apply a natural logarithmic transformation in order to normalize the distribution. Since the transformation did not produce the expected effect (i.e., normalized the distribution, produced different statistical outcome), we kept the original data.

We ran a non-parametric Spearman rho correlation test to provide an overview of the association between our two variables. Furthermore, we used a multiple regression test to assess the association with mean cortisol value as the independent variable and percentage of stereotypy as the dependent variable. Given that our alternating-treatment graphs showed no distinction between the types of conditions for neither stereotypy nor cortisol, we used the number of sessions ($n = 60$) as the number of observations, and we controlled for the non-independence of our data (considering that each participant had 15 sessions).

The analysis was performed using Mplus. The type=COMPLEX function with clustering was applied to consider the interdependence of the data. This function produces standard errors that reflect the nature of the clustering and that is robust to non-independence and non-normality (McNeish et al., 2017). As our study was preliminary, we considered $p < .05$ as significant and $p < .10$ as marginally significant. Because the chi-square ($\chi 2$) is sensitive to sample size, model fit was assessed using the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), and the Tucker-Lewis Index (TLI; Hu & Bentler, 1999; Marsh et al., 2005; Yu, 2002). RMSEA values smaller than 0.06 suggest acceptable and excellent model fit. Values above .95 for the CFI and TLI also indicate an adequate and excellent model fit. To examine whether our results were explained by between- or within-participant patterns, we generated correlation graphs that illustrated cortisol and stereotypy data.

## Results

Figure 2 shows the variation of mean values of cortisol (left panels) and the percentage of global stereotypy (right panels) across the three conditions. No condition generated a distinct level of cortisol for the four participants. Nevertheless, the patterns showed enough variability to conduct further analyses. Figure 3 presents variations in percentage of motor and vocal stereotypy across conditions. The three conditions were undifferentiated in terms of cortisol levels and all stereotypy levels (global, motor and vocal). We applied the rule developed by Lanovaz et al. (2019) and the results showed that there was no significant difference between the three conditions. Hence, we treated all the 60 sessions as being undifferentiated in terms of condition for the upcoming analyses.

Table 3 presents the Spearman rho correlations between all the variables and Figure 4 offers a more detailed perspective of these correlations. The upper panel of the Figure 4 shows the relationship between the percentage of time engaged in global stereotypy and the mean value of cortisol for each participant. The middle panel presents the relationship with motor stereotypy, and the lower panel shows the relationship with vocal stereotypy. Each form of data point refers to a different participant, and each point refers to a different session. The mean value of cortisol was negatively associated with the percentage of global stereotypy ($r_s$ = -.392, $p$ < .01). The mean value of cortisol was also negatively associated with the percentage of motor stereotypy ($r_s$ = -

.508, $p < .01$). No significant correlation was found between mean value of cortisol, and vocal stereotypy.

The multiple regression using Mplus revealed that the model fitted the data well (RMSEA = .00, CFI = 1.00, TLI = 1.00). We tested three models with different dependent variables: percentage of global stereotypy, percentage of motor stereotypy only, and percentage of vocal stereotypy. Independent variables remained the same across the models (mean value of cortisol). Table 4 displays the standardized regression coefficient (β) and *p* values. In the first model, the mean cortisol value was negatively related to the percentage of global stereotypy ($\beta = -.404$, $p < 0.1$). In the second model, mean cortisol values was negatively related to the percentage of motor stereotypy ($\beta = -.416$, $p < 0.1$). Unlike the first two models, the mean value of cortisol did not contribute significantly to the third model (percentage of vocal stereotypy).

## Discussion

Altogether, our exploratory results indicate that the mean level of cortisol was negatively associated with global and motor stereotypy. Although our design does not allow us to establish a directional link, our results are consistent with those of Gabriels et al. (2013) who showed that children with ASD who presented higher levels of stereotypy had lower cortisol levels. The negative correlation between cortisol, motor and global stereotypy tends to support the hypothesis according to which stereotypy could serve a self-regulation function. These results conflict with previous studies which identified positive correlation between cortisol levels and stereotypy (Bitsika et al., 2015; de Vaan et al., 2020; García-Villamisar & Rojahn, 2015; Lydon, Healy, Roche et al., 2015; Yang et al., 2015). Further studies, with larger samples are needed to clarify the causal relationship between theses variables.

The current study did not identify an association between cortisol and vocal stereotypy. This result highlights the possibility that the relationship between stress and stereotypy can vary depending on the form of stereotypy. Previous studies had highlighted the possibility that stereotypy has different functions depending on the internal state of the individual, environmental characteristics, or both. The results of our experiment add another layer of complexity to this hypothesis in showing that different forms of stereotypy could serve different functions.

To our knowledge, the pilot study is the first to use direct observation measures to monitor one category of repetitive behaviors (stereotypy). It is also the first to combine within- and between-participant analyses to explore the relation between cortisol and stereotypy. Our results highlight the complexity of the relationship between biological measures and stereotypy. We believe that direct observation measures are central to elucidating this complex relationship and our study is a first step in this direction. The inconsistency in the literature highlights the need to develop effective measures that will identify the multiple functions of stereotypy. To address this issue, a large amount of data is needed to take care of interindividual variability in ASD children. An inherent limit of direct observation measures are the resources needed to collect the data (human and financial). Artificial intelligence algorithms would be a well-suited tool to automate the measurement of those behaviors (Cantin-Garside et al., 2020; Dufour, Lanovaz & Cardinal, 2020).

Our study has some limitations that should be noted. First, our conditions did not yield differential stress levels, which limited the range of cortisol values studied. Second, we did not conduct an experimental functional assessment to assess the environmental variables associated with stereotypy due to the limited amount of data available. Examining environmental variables in the future may uncover novel association between physiological measures and specific behavioral functions. Finally, the small sample size of our pilot study and the nature of our design could have an impact on the external validity of our results. Our statistical analyses demonstrated an association between cortisol and stereotypy, but the visual analysis of Figure 4 tends to show that these relationships could be explained by between-participant differences rather than within-participant patterns, which should be investigated further in future studies. It would also be relevant to investigate ways to identify patterns between physiological measures and participant characteristics. Machine learning algorithms would also be well suited to investigate these new avenues for research. For example, researchers may use machine learning algorithms to determine which combination of physiological measures best predict engagement in stereotypy. The current study did not have sufficient data to apply this methodology.

The unique characteristics of our sample, the direct observation measures of stereotypy and the exploration of different forms of stereotypy may explain part of our results and underline the relevance of pursuing this line of research with a larger sample of minimally verbal ASD children. Future studies should use direct observation measures with a larger sample, which will allow the assessment of directionality. Considering that the relationship between stress and stereotypy seems to vary across developmental periods (Bitsika et al., 2015), emotional states (Lydon et al., 2013; Willemsen-Swinkels et al., 1998) and forms of the stereotypy, we highly recommend that future studies be conducted with samples that are as homogenous as possible. Due to the complex nature of the association between stress and stereotypy, we also suggest that future research combine multiple physiological measures to obtain a broader representation of this dynamic.

## Compliance with Ethical Standards

The authors have no competing interests to declare that are relevant to the content of this article. University of Montreal research ethics board approved the present study and we obtained consent from the legal guardian of each participant. The appropriate ethical guidelines were followed in the conduct of the research study.

## References

Adam, E. K., & Kumari, M. (2009). Assessing salivary cortisol in large-scale, epidemiological research. *Psychoneuroendocrinology, 34*(10), 1423-1436. https://doi.org/10.1016/j.psyneuen.2009.06.011

Ahearn, W. H., Clark, K. M., MacDonald, R. P., & Chung, B. I. (2007). Assessing and treating vocal stereotypy in children with autism. *Journal of Applied Behavior Analysis, 40*(2), 263–275. https://doi.org/10.1901/jaba.2007.30-06

American Psychiatric Association. (2013*). Diagnostic and statistical manual of mental disorders (5th ed.)*. American Psychiatric Publishing.

Bitsika, V., Sharpley, C. F., Agnew, L. L., & Andronicos, N. M. (2015). Age-related differences in the association between stereotypic behaviour and salivary cortisol in young males with an Autism Spectrum Disorder. *Physiology & Behavior, 152,* 238-243. https://doi.org/10.1016/j.physbeh.2015.10.010

Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties of repetitive behavior in autism: Comparisons to mental retardation. *Journal of Autism and Developmental Disorders, 30*, 237–243. https://doi.org/ 10.1023/a:1005596502855

Cantin-Garside, K. D., Kong, Z., White, S. W., Antezana, L., Kim, S., & Nussbaum, M. A. (2020). Detecting and classyfing self-injurious behavior in autism spectrum disorder using machine learning techniques. *Journal of Autism and Developmental Disorders, 50,* 4039-4052. https://doi.org/ 10.1007/s10803-020-04463-x

Chebli, S. S., Martin, V., & Lanovaz, M. J. (2016). Prevalence of stereotypy in individuals with developmental disabilities: A systematic review. *Review Journal of Autism and Developmental Disorders, 3*(2), 107-118. https://doi.org/10.1007/s40489-016-0069-x de Vaan, G., Beijers, R., Vervloed, M. P. J., Knoors, H., Bloeming-Wolbrink, K. A., de Weerth, C., & Verhoeven, L. (2020). Association between cortisol stress levels and autism symptoms in people with sensory and intellectual disabilities. *Frontiers in Education, 5*, 540387. https://doi.org/ 10.3389/feduc.2020.540387

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*(3), 355-391. https://doi.org/10.1037/0033-2909.130.3.355

Dozier, M., Manni, M., Gordon, M. K., Peloso, E., Gunnar, M. R., Stovall-McClough, K. C., Eldreth, D., & Levine, S. (2006). Foster children's diurnal production of cortisol: An exploratory study. *Child Maltreatment, 11*(2), 189-197. https://doi.org/10.1177/1077559505285779

Dufour, M.-M., Lanovaz, M. J., & Cardinal. P. (2020). Artificial intelligence for the measurement of stereotypy. *Journal of the Experimental Analysis of Behavior, 114*(3), 368-380. https://doi.org/10.1002/jeab.636

Gabriels, R. L., Agnew, J. A., Pan, Z., Holt, K. D., Reynolds, A., & Laudenslager, M. L. (2013). Elevated repetitive behaviors are associated with lower diurnal salivary cortisol levels in autism spectrum disorder. *Biological Psychology, 93*(2), 262-268. https://doi.org/10.1016/j.biopsycho.2013.02.017

García-Villamisar, D., & Rojahn, J. (2015). Comorbid psychopathology and stress mediate the relationship between autistic traits and repetitive behaviours in adults with autism: Mediators of repetitive behaviour in autism. *Journal of Intellectual Disability Research, 59*(2), 116-124. https://doi.org/10.1111/jir.12083

Gordon, M. K., Peloso, E., Auker, A., & Dozier, M. (2005). Effect of flavored beverage crystals on salivary cortisol enzyme-immunoreactive assay measurements. *Developmental Psychobiology, 47*(2), 189-195. https://doi.org/10.1002/dev.20081

Groden, J., Cautela, J., Prince, S., & Berryman, J. (1994). The Impact of stress and anxiety on individuals with autism and developmental disabilities. In E. Schopler & G. B. Mesibov (Éds.), *Behavioral issues in autism* (p. 177-194). Springer US. https://doi.org/10.1007/978-1-4757-9400-7_9

Hu, L., & Bentler, P. M. (1999). Cut-off criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55. https://doi.org/10.1080/10705519909540118

Lanovaz, M. J., Cardinal, P., & Francis, M. (2019). Using a visual structured criterion for the analysis of alternating-treatment designs. *Behavior Modification, 43*(1), *115-131*. https://doi.org/10.1177/0145445517739278

Lanovaz, M. J., & Sladeczek, I. E. (2011). Vocal stereotypy in children with autism: Structural characteristics, variability, and effects of auditory stimulation. *Research in Autism Spectrum Disorders, 5*(3), 1159-1168. https://doi.org/10.1016/j.rasd.2011.01.001

Lovaas, I., Newsom, C., & Hickman, C. (1987). Self-stimulatory behavior and perceptual reinforcement. *Journal of Applied Behavior Analysis, 20*(1), 45-68. https://doi.org/10.1901/jaba.1987.20-45

Lydon, S., Healy, O., & Dwyer, M. (2013). An examination of heart rate during challenging behavior in autism spectrum disorder. *Journal of Developmental and Physical Disabilities, 25*(1), 149-170. https://doi.org/10.1007/s10882-012-9324-y

Lydon, S., Healy, O., Roche, M., Henry, R., Mulhern, T., & Hughes, B. M. (2015). Salivary cortisol levels and challenging behavior in children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 10*, 78-92. https://doi.org/10.1016/j.rasd.2014.10.020

Mahone, E. M., Bridges, D., Prahme, C., & Singer, H. S. (2004). Repetitive arm and hand movements (complex motor stereotypies) in children. *The Journal of Pediatrics, 145*(3), 391-395. https://doi.org/10.1016/j.jpeds.2004.06.014

Marsh, H. W., Hau, K.-T., & Kong, C.-K. (2002). Multilevel causal ordering of academic self-concept and achievement: Influences of language of instruction (English compared with Chinese) for Hong Kong Students. *American Educational Research Journal, 39*(3), 727-763. https://doi.org/10.3102/00028312039003727

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods, 22*(1), 114-140. https://doi.org/10.1037/met0000078

Melo, C., Ruano, L., Jorge, J., Pinto Ribeiro, T., Oliveira, G., Azevedo, L., & Temudo, T. (2020). Prevalence and determinants of motor stereotypies in autism spectrum disorder: A systematic review and meta-analysis. *Autism, 24*(3), 1-22. https://doi.org/10.1177/1362361319869118

Muthugovindan, D., & Singer, H. (2009). Motor stereotypy disorders. *Current Opinion in Neurology*, *22*(2), 131-136. https://doi.org/10.1097/WCO.0b013e328326f6c8

Parkhurst, J., & Kawa, J. M. (2018). Childhood Autism Rating Scales. In J. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of Clinical Neuropsychology* (pp. 1-3). Springer International Publishing. https://doi.org/10.1007/978-3-319-56782-2_1530-2

Querim, A. C., Iwata, B. A., Roscoe, E. M., Schlichenmeyer, K. J., Viruès Ortega, J., & Hurl, K. E. (2013). Functional analysis screening for problem behavior maintained by automatic reinforcement. *Journal of Applied Behavior Analysis*, *46*(1), 47-60. https://doi.org/10.1002/jaba.26

Rapp, J. T., & Vollmer, T. R. (2005). Stereotypy I: A review of behavioral assessment and treatment. *Research in Developmental Disabilities, 26*(6), 527-547. https://doi.org/10.1016/j.ridd.2004.11.005

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *The Childhood Autism Rating Scale* ($2^{nd}$ ed.). Western Psychological Services.

Sharpley, C. F., Bitsika, V., Andronicos, N. M., & Agnew, L. L. (2016). Is afternoon cortisol more reliable than waking cortisol in association studies of children with an ASD? *Physiology & Behavior, 155*, 218-223. https://doi.org/10.1016/j.physbeh.2015.12.020

van de Mortel, T. F. (2008). Faking it: Social desirability response bias in self-report research. *Australian Journal of Advanced Nursing, 25*(4), 40-48.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gotzsche, P. C., & Vandenbroucke, J. P. (2008). The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Journal of Clinical Epidemiology*, 61(4), 344-349. doi:10.1016/j.jclinepi.2007.11.008

Willemsen-Swinkels, S. H. N., Buitelaar, J. K., Dekker, M., & van Engeland, H. (1998). Subtyping stereotypic behavior in children: The association between stereotypic behavior, mood, and heart rate. *Journal of Autism and Developmental Disorder, 28*, 547-557. https://doi.org/10.1023/A:1026008313284

Yang, C.-J., Tan, H.-P., Yang, F.-Y., Wang, H.-P., Liu, C.-L., He, H.-Z., Sang, B., Zhu, X.-M., & Du, Y.-J. (2015). The cortisol, serotonin and oxytocin are associated with repetitive behavior in autism spectrum disorder. *Research in Autism Spectrum Disorders, 18*, 12-20. https://doi.org/10.1016/j.rasd.2015.07.002

Yu, C.-Y. (2002). Evaluation of model fit indices for latent variable models with categorical and continuous outcomes. *Unpublished dissertation*. Retrieved from Mplus website http://www.statmodel.com/download/Yudissertation.pdf

**Table 1**
*Participant Characteristics*

| Participants | Age | Sex | CARS-II T-score | Verbal communication score (CARS-II) |
|---|---|---|---|---|
| Ana | 6 | F | 61 | 4 |
| Matt | 9 | M | 55 | 3.5 |
| Jim | 5 | M | 58 | 3.5 |
| Oli | 5 | M | 49 | 4 |

**Table 2**

*Operational Definitions of Forms of Stereotypy Observed in the Study*

| Participants | Stereotypy | Form (M = motor, V = vocal) | Behavioral Definition |
|---|---|---|---|
| All | Repetitive movements | M | Two or more rapid back and forth movements of the same body part (without a clear functional purpose) |
| | Rocking | M | Two or more rapid back and forth movements of the torso |
| | Vocal stereotypy | V | Acontextual sounds or words produced by the vocal apparatus |
| | Toy stereotypy | M | Repetitively activating a toy that produces light or sounds |
| | Jumping | M | Repetitively having both feet leave the ground while standing |
| Ana | Object stereotypy | M | Two or more back and forth movements of an object |
| Matt | Rubbing | M | Rubbing any body part with hand for at least 2 consecutive seconds |
| | Scratching | M | Contact between the child’s nails and any body part |
| | Visual stimulation | M | Putting his forearm or the crest of his elbow on his forehead or eyes |
| Jim | Ear touching | M | Two or more consecutive contacts between the child’s ears and any part of his hand |
| Oli | Arm flapping | M | Two or more consecutive up and down arm movements |
| | Face rubbing | M | Touching his face with any part of an object |

**Table 3**

*Spearman's Rho Correlations*

| | Cortisol (mean) | Global stereotypy (%) | Motor stereotypy (%) | Vocal stereotypy (%) |
|---|---|---|---|---|
| Cortisol (mean) | 1.00 | -.392** | -.508** | .129 |
| Global stereotypy (%) | -.392** | 1.00 | .818** | .318* |
| Motor stereotypy (%) | -.508** | .818** | 1.00 | -.098 |
| Vocal stereotypy (%) | .129 | .318* | -.098 | 1.00 |

**$p$<.01 *$p$<0.05

**Table 4**

*Summary of Multiple Regression Analyses for Stereotypy Regressed on Mean Cortisol Values*

| | *β* | *S.E.* | *p* |
|---|---|---|---|
| **Model 1** | | | |
| Global stereotypy (%) | | | |
| Cortisol (mean) | -0.404 | 0.148 | 0.006 |
| **Model 2** | | | |
| Motor stereotypy (%) | | | |
| Cortisol (mean) | -0.416 | 0.160 | 0.009 |
| **Model 3** | | | |
| Vocal stereotypy (%) | | | |
| Cortisol (mean) | -0.006 | 0.210 | 0.164 |

**Figure 1**

*Number of Sessions for Each Conditions (Left) and Structure of Sessions, Including the Duration and the Four Measures of Cortisol (Right)*

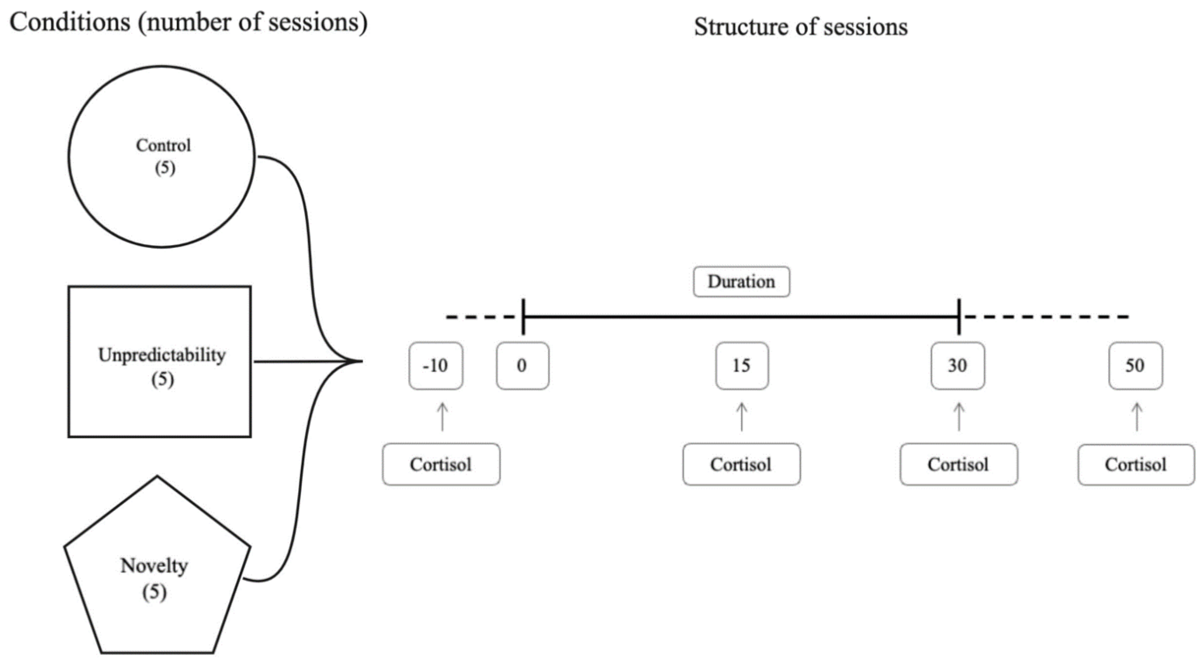

**Figure 2**

*Level of Mean Value of Cortisol (Left Panels) and Percentage of Time Engaged in Global Stereotypy (Right Panels) Across Three Conditions: Control, Novelty and Unpredictability*

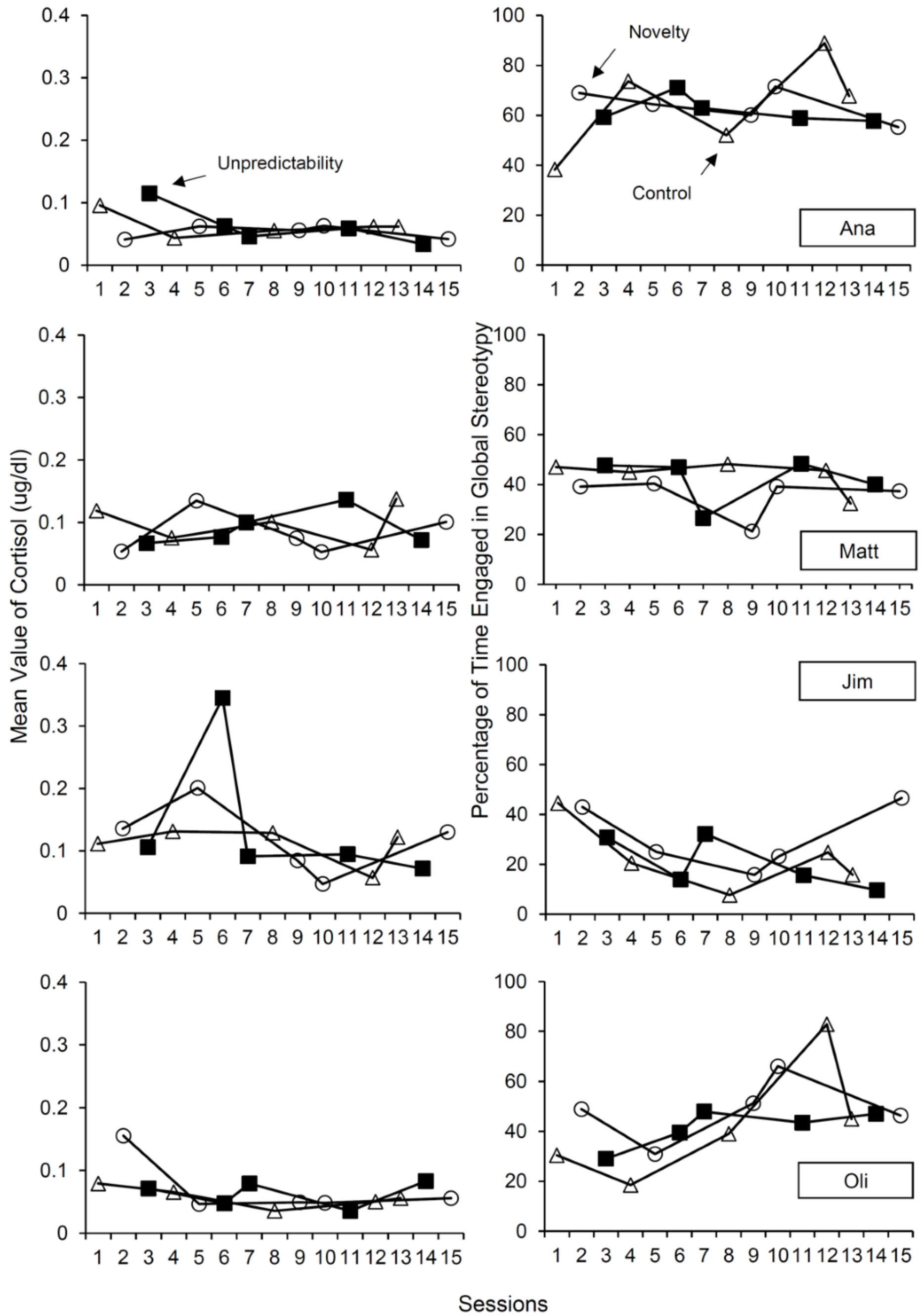

**Figure 3**

*Percentage of Time Engaged in Motor Stereotypy (Left Panels) and Vocal Stereotypy (Right Panels) Across Three Conditions: Control, Novelty and Unpredictability*

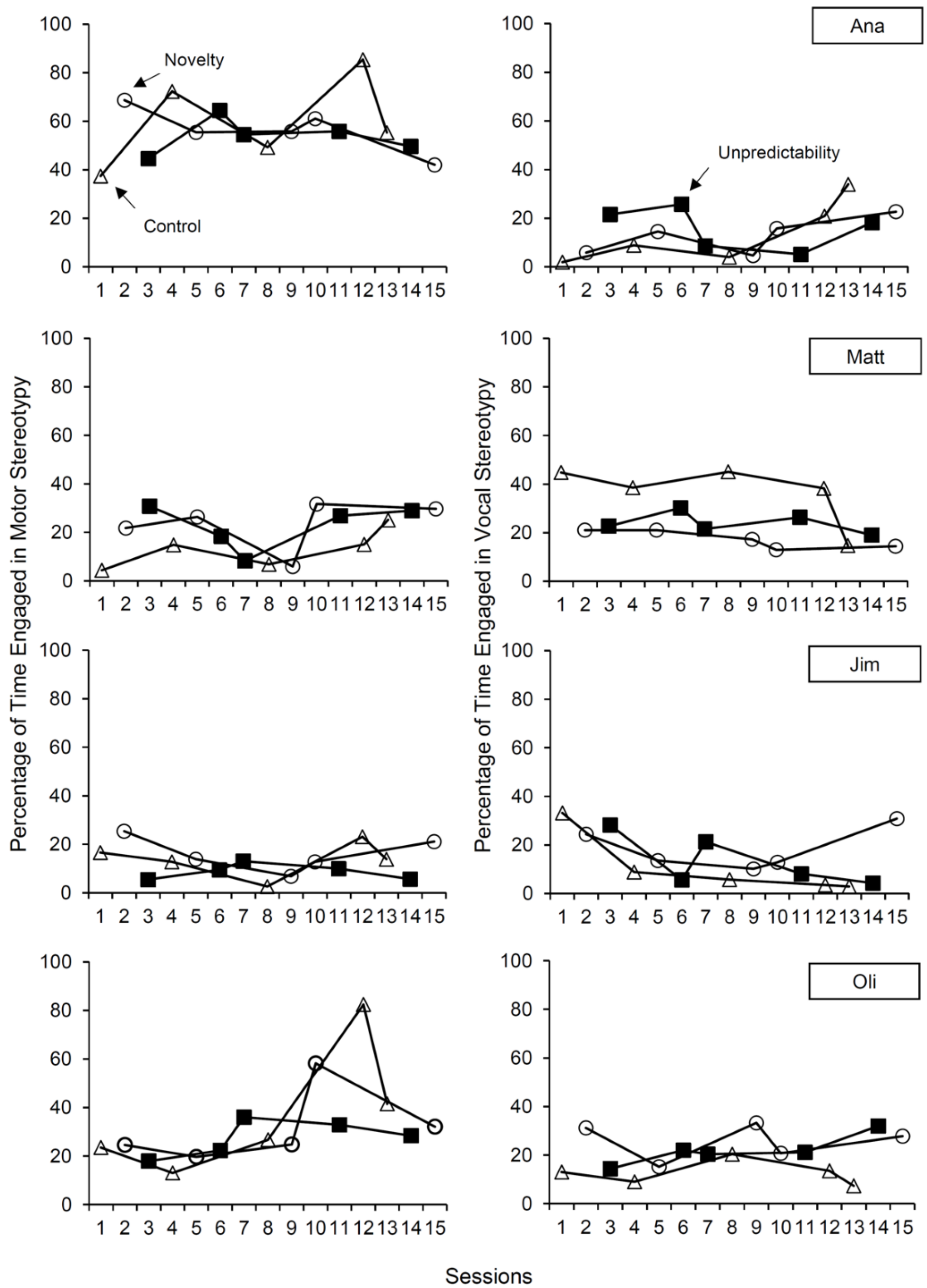

**Figure 4**

*Mean Value of Cortisol (ug/dl) Compared to Percentage of Time Engaged in Global Stereotypy, Motor Stereotypy and Vocal Stereotypy*