



Salivary oxytocin after play with parents predicts behavioural problems in preschool children

Nuno Torres^{a,*}, Daniel Martins^{b,e}, Lúgia Monteiro^c, António J. Santos^a, Brian E. Vaughn^d, Manuela Veríssimo^a

^a ISPA - Instituto Universitário de Ciências Psicológicas Sociais e da Vida, William James Research Center, Rua Jardim do Tabaco, no34, 1149-041 Lisbon, Portugal

^b King's College London, Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience, 16 De Crespigny Park, London SE5 8AF, United Kingdom

^c Instituto Universitário de Lisboa (ISCTE-IUL), CIS-IUL, Av. das Forças Armadas, 1649-026 Lisboa, Portugal

^d Auburn University, Department of Human Development and Family Science Auburn, AL 36849, United States

^e NIHR Maudsley Biomedical Research Centre, South London and Maudsley NHS Trust, London, UK

ARTICLE INFO

Keywords:

Oxytocin
Saliva
Children
Radioimmunoassay
Behavioural problems

ABSTRACT

Background: Oxytocin (OXT) has attracted research interest for its potential involvement in many of the behavioural problems observed in childhood. Due to its logistical advantages, saliva is an attractive fluid to quantify neuropeptides in children. Salivary OXT has been suggested as a potential biomarker for psychopathology during childhood. However, several questions still remain about the extent to which, and under what conditions, concentrations of OXT in saliva can be reliably measured and are related to behavioural problems in preschool age children.

Methods: Seven samples of saliva from 30 preschool children (17 girls) were collected in five different days at their homes. Three of the samples were collected by the children's parents at baseline daily routine conditions, and four of the samples were collected by researchers during two home-visits: before and after two 15-minute dyadic play sessions (one with mothers and one fathers) between each individual parent and the child. Oxytocin concentrations were quantified by Radioimmunoassay with prior extraction. Children's behavioural problems were assessed by the Caregiver-Teacher Report Form (C-TRF) questionnaire, completed by the child's preschool teacher.

Results: Salivary OXT measured in baseline samples could not predict any of the behavioural problems measured by the C-TRF. However, when measured after playing with parents salivary OXT showed a stronger pattern of negative correlations, specially with the depression and opposition scales of the C-TRF. Furthermore, salivary OXT was unlikely to be reliably measured using single sampling, but acceptable reliabilities were achieved when averaging several samples. Finally, the single measures of salivary OXT evoked after an episode of play with parents showed better reliabilities than collected at baseline.

Conclusion: Measurements of OXT evoked after positive affect interactions with parents seem to capture aspects of the OXT system in young children that might be relevant for understanding the role of this system in children's social behaviour.

1. Introduction

Over the past decades the hypothalamic neuropeptide oxytocin (OXT) has attracted considerable attention as a physiological underpinning of social bonds in mammals (Leeds et al., 2018; Nelson and Panksepp, 1998). Accumulating evidence has demonstrated the importance of OXT in the development of both normative and deviant social behaviour and cognition (Bachner-Melman and Ebstein, 2014; Caldwell,

2017; Jurek and Neumann, 2018; Stein, 2009). Several studies have reported that clinical aspects of some mental disorders, particularly the impairments in social functioning, co-vary with OXT levels measured both in the brain and in peripheral fluids such as plasma, urine and saliva (Demirci et al., 2016b; Cochran et al., 2013; Lebowitz et al., 2017; Levy et al., 2015; Sasaki et al., 2016). Additionally the OXT system has been identified as a possible target for the development of pharmacological treatments of core socio-emotional difficulties present in several

* Correspondence to: William James Centre for Research, ISPA-Instituto Universitário, Rua Jardim do Tabaco, 34, 1149-041 Lisbon, Portugal.
E-mail addresses: ntorres@ispa.pt, nmtorr2@gmail.com (N. Torres).

<https://doi.org/10.1016/j.psyneuen.2021.105609>

Received 16 February 2021; Received in revised form 19 November 2021; Accepted 24 November 2021

Available online 25 November 2021

0306-4530/© 2021 Elsevier Ltd. All rights reserved.

disorders (Miller et al., 2017; Taylor et al., 2014).

However, a recent systematic review on the differences between psychiatric patients and healthy controls found that of all the peripheral fluids assessed in a broad range of mental disorders, only serum OXT in anorexia nervosa showed significant differences (Rutigliano et al., 2016). The authors suggested that the lack of significant results, in psychiatric populations, was associated with high methodological heterogeneity, and scarce reliability of the methods used, which need to be validated and standardized.

Most studies have focused on the role of OXT in adults. A smaller number of studies have examined the associations between children's OXT concentrations in CSF, plasma, urine and saliva, and psychopathological/behaviour problems such as ADHD, OCD, Depression, Anxiety, and Conduct problems (for a review see Torres et al. (2018)). Even though the evidence was mixed, several studies have reported significant associations between OXT levels and symptoms in children (Demirci et al., 2016a, 2016b; Carson et al., 2014; Lebowitz et al., 2017; Levy et al., 2015; Oztan et al., 2018; Salzberg and Swedo, 1992; Sasaki et al., 2016). Work still has to be done on the methodological homogeneity of the studies, their reliability assessment and standardization (Rutigliano et al., 2016). Disparate instruments and methods, and low or unknown reliability and validity of the type of measurements used, makes it more challenging to generate broad conclusions. The field would benefit from more standardized assessments both of OXT levels and of problem behaviours (Torres et al., 2018).

1.1. Peripheral oxytocin as a proxy of central oxytocin?

Many questions remain about the extent to which the OXT concentration in peripheral fluids can be a proxy of central OXT levels. A recent meta-analysis of associations between central and peripheral OXT in animals and humans (Valstad et al., 2017) found significant correlation [$r = .29$, 95% CI (.14; .42), $p < 0.0001$]. This correlation was moderated by "experimental context": there were stronger associations: 1) after induced emotional states, such as stress ($r = .49$, $p = .001$), and also 2) after intranasal OXT administration ($r = .66$, $p < .0001$), but there was no significant correlation under baseline. The fact that there is evidence of a synchrony between central and peripheral OXT after some specific emotional events, i.e. "evoked", such as stress but not at baseline (de Jong et al., 2015; Engelmann et al., 1999; Valstad et al., 2017) is crucial to take into account when developing valid methods of OXT assessment based on peripheral fluids (Uvnäs-Moberg et al., 2015). It is plausible that evoked measures capture aspects of the system that might be more relevant than baseline ones to explain interindividual variability in behaviour. Moreover, as for other neuroendocrinal systems, it is also possible that evoked measurements might present higher reliabilities than baseline ones, maximizing the potential utility of such measurements in a number of research and clinical setups. Hence, in order to turn peripheral OXT concentrations into a biomarker we might need to collect measures when the central and peripheral release of OXT is likely to be synchronized (Rutigliano et al., 2016).

Most studies about the coordination of central and peripheral OXT have focused on human and non-human adult individuals, and the knowledge about it in the young is still a largely unexplored field. Kojima et al. (2012) found that there was a strong, positive correlation between peripheral (serum) and central (hypothalamic) OXT in pups of rats 1) at the end of a 1-h separation from the mother (stressful event) and also 2) during the reunion with the mother (pleasant event). They found no correlation between central and peripheral OXT in the pups at baseline. One study in humans found a very strong significant correlation between central (in CSF) and peripheral (in plasma) OXT in children (Carson et al., 2014). Additionally, these authors found that both plasma ($r = -.92$, $p = .026$, $n = 10$) and CSF ($r = -.91$, $p = .033$, $n = 10$) concentrations predicted trait-anxiety scores in children. The blood and CSF were collected under hospitalization for serious health risk, which can be considered a stressful situation. These results, although

preliminary and pre-clinical suggest that peripheral measurements as much as central ones can detect OXT's association with psychopathology symptoms in children.

1.2. Salivary measurements of OXT

Saliva is a particularly attractive fluid to quantify OXT during childhood (Leeds et al., 2018; Uvnäs-Moberg et al., 2015). Several methods for measuring salivary OXT have been already reported and supported saliva as a non-invasive source to monitor central neuroendocrine functions, both in animals and humans (Carter, 1998; Daughters et al., 2015; de Jong et al., 2015; Feldman et al., 2010a; Leeds et al., 2018; MacLean et al., 2018). One very recent study not only found significant correlations between central and salivary OXT, but also that OXT in saliva was more strongly correlated with central OXT than plasma (Martin et al., 2018). Furthermore, a recent study with apes suggested that saliva measurements can be suitable for assessing OXT levels' changes following specific events which included both stressful and hedonic events, such as the death of a conspecific, play, and breeding (Leeds et al., 2018).

1.3. Reliability of OXT measurements

Questions about the reliability of single measurements of peripheral OXT make it yet a controversial topic. Martins et al. (2020) have recently reported that plasmatic and salivary OXT are unlikely to be reliably measured using single samples in healthy adult men and suggested the collection of multiple samples per individual as a way of circumventing this low reliability problem. Although there is a scarcity of reliability assessment of salivary OXT measures, some studies in children have reported strong test-retest correlations. For instance, Feldman et al. (2010b) sampled salivary OXT from 55 infants of 4-6 months old and their parents, before and after a 15-min play interaction. OXT was significantly correlated from pre- to post-interactions in both the parents, $r = .55$, $p < .001$, and the infants, $r = .50$, $p < .001$. In another later study Feldman et al. (2013) assessed salivary OXT of 50 three-year-old children with typical healthy development, which was collected twice in the same day: at baseline and after social interactions. The baseline and post-interaction assessments of child salivary OXT were strongly correlated ($r = .74$, $p < .001$). While these studies show stability of salivary oxytocin in children, they cannot provide information about the absolute agreement of two sets of equivalent measurements, needed to assess reliability. Even less is known about the reliability of salivary OXT measurements at non-baseline conditions, such as after stressful or hedonic events, the later which we addressed in this study.

In the present study we tested the extent to which oxytocin concentrations in saliva were associated with common forms of behavioural problems at these ages. Additionally we tested if those associations vary as function of 1) the evoked emotional stimulus (after playing with parents, versus basal collection), and 2) collected by researchers on a novel situation versus collected by parents in a routine situation. We also aimed to estimate the reliability of several oxytocin measurements in saliva samples collected from pre-school children at different time-points. All measurements and procedures used are standardized and widely used in previous research, and are easily compared and replicated with different samples. In line with current recommendations in the field, we relied on a highly sensitivity method for OXT quantification, radioimmunoassay after extraction, which is currently considered the gold-standard technique for measuring OXT in biological fluids.

1.4. Objectives

- 1- Estimate the reliability of OXT measurements in saliva samples collected from pre-school age children, both at baseline and evoked after playing with parents in a standardized situation.
- 2- Assess if salivary OXT, measured from samples collected at

children's home at different time points, predicts common forms of behavioural and psychopathology problems of young children.

3- Assess whether correlations between salivary OXT and children's behavioural and psychopathology problems differ as a function of whether salivary OXT was measured at baseline and after a playful episode with both parents.

2. Methods

2.1. Participants and recruitment

We recruited the families of 30 preschool children (17 girls) (Mean age = 60.9 months; $SD = 9.5$ months), and their preschool teachers. All Children lived with both their heterosexual parents. At the time of the study all children were attending day-care; with the time spent in these settings ranging from 5 to 9 h a day ($M = 7.8$, $SD = 1.1$). Mothers were 32–51 years old ($M = 39.1$, $SD = 5.3$), and fathers 33–52 years old ($M = 40.3$, $SD = 5.3$). Mothers' education status ranged from 4 to 21 academic years ($M = 12.3$, $SD = 3.8$), and father's education level ranged from 4 to 19 academic years ($M = 11.2$ $SD = 3.2$). Mothers worked outside of the home full time in 80% of the families and fathers in 90% of the families. All preschool teachers had a university degree in early education. To be included in the study, children should be at the preschool age (at least ~36 months of age), show no evidence of major neurodevelopmental disorders (e.g., cerebral palsy, foetal alcohol syndrome, Down's syndrome, Autism Spectrum Disorders), been born at term, and be living in bi-parental dual earner families.

A power analysis calculation based on the number of saliva samples we collected per individual (seven) for each set suggested that with a sample size of 30 children we should had been able to detect an acceptable ICC of .70 if it did exist with 80% of statistical power. Furthermore, for ICCs of at least .70, the same sample size would allow us to detect a moderate correlations of $r = .60$ between salivary OXT and behavioural problems in our sample with 80% of statistical power (Martins et al., 2020).

2.2. Ethical approval

The study was approved by the Research Ethics Committee of ISPA – Instituto Universitário, and was conducted according with the *Ethical principles of psychologists and code of conduct* of the American Psychological Association. All parents and teachers provided a written informed consent, and children assented to participate.

2.3. Protocol

A total of seven samples of saliva, from each child, were collected (i.e., a total of 210 saliva samples). The seven samples were collected over five different days, always at the child's home, between 5 and 6 pm, and were collected on average over two weeks (between 7 and 36 days, depending on the availability of each family). We collected at least 3.5 ml of saliva on each sample. After being collected, the saliva was immediately stored in a freezer (-20 °C), as per standard procedures. Samples were initially frozen at the family's home and then transported in a portable freezer to the lab, where they were stored at -80 °C until being sent for analysis in an outsource service (RIAgnosis, Munich, Germany). Transport was made in solid carbon at an average temperature of -80 °C. Samples were collected by passive drooling to a 5 ml plastic polypropylene tube. In a previous study, measurements from passive drool saliva samples provided more accurate estimations of hormonal levels, even after intermediary processing steps, including freezing, thawing, and centrifugation (Robles et al., 2013).

Three of samples were collected by the parents themselves without a presence of the researchers, in baseline child-routine conditions. The three samples collected by the parents were collected on 2 weekdays and on 1 weekend day. Since the preschool children have a significantly

more structured routine on weekdays (rising hours, meals timetables, school activities, time of returning home, and going to bed, etc.), while during the weekends each child has different and individual routine depending on each family practices, we considered important to include both weekend and weekdays in the assessment of salivary OXT and its respective reliability. To minimize variability, we instructed the parents that before collecting saliva the children should 1) be fasting for at least 2 h, 2) not been exposed to direct strong sunlight without head protection, and 3) had not been taking any medication for the past three days. Additionally, to minimize potential variability related to circadian rhythms, all samples were collected at the same timeframe (i.e., between 5 and 6 p.m.). Parents were explicitly reminded that the time and condition of saliva samples should be the same for both weekday and weekend collections. Before saliva collection, parents were trained by the researchers on a precise collection protocol. Parents viewed a procedure-oriented video to standardize the training. They were also instructed to complete a questionnaire during and after each saliva sampling episode, detailing the following information: date, time, time of last meal/snack, how many hours the child was exposed to direct sun, medication intake and its time (if any), and last time the child brushed his/her teeth.

The rest of the four samples were collected by the researchers which were strangers to the children, during two home-visits. These samples were collected before and after a dyadic play task with the child's parents at their home. Since we wanted to have the assessment of saliva samples both at baseline and post-play as equivalent as possible, we designed the collection of saliva to be made by the researchers before (baseline) and post-play (evoked). With these systematic and controlled variations of saliva sampling we could build several different composites and assess their reliability, as well as their association with the behavioural-problems scales assessed independently at school.

One of the play tasks was performed with the father and the other with the mother, on separate days, which were counterbalanced across participants. The time interval between the sessions was on average 9.5 days (minimum of 3 and maximum of 15 days). The play sessions were completed with each parent on different days yielding a total of four additional saliva samples. One sample was collected 10 min before each play session, and a second sample 15 min after. The researchers completed the same saliva-sampling checklist that the parents used. If the requirements for saliva collection were not met, the task was cancelled and scheduled for another day. This happened in two instances: 1) once, a child ate some cookies when researchers arrived at the family's home, and 2) in the other instance, a child was taking antihistamines to control hay fever symptoms.

2.4. Quantification of salivary OXT

Salivary OXT was quantified by Radioimmunoassay (RIAgnosis, Munich, Germany), after extraction. In summary, for each sample, 300 μ l of saliva was evaporated (SpeedVac, ThermoScientific Inc, Waltham, MA, USA), and 50 μ l of assay buffer was added followed by 50 μ l antibody (raised in rabbits against OXT). After a 60-min pre-incubation interval, 10 μ l 125I-labeled tracer (PerkinElmer, Waltham, MA, USA) was added and samples were allowed to incubate for 3 days at 4 °C. Unbound radioactivity was precipitated by activated charcoal (Sigma-Aldrich, St Louis, MO, USA). Under these conditions, an average of 50% of total counts are supposed to bind with < 5% non-specific binding. The detection limit of this assay was determined to be in the 0.5 pg/sample range, depending on the age of the tracer, with typical displacements of 20–25% at 2 pg, 60–70% at 8 pg and 90% at 32 pg of the standard neuropeptide. In this study, nine (4.4%) of the quantified samples were below the detection range and for this reason were excluded from further analysis. Even if the assay has high sensitivity and can measure reliably concentrations as low as 0.1 pg/sample, here we decided to be conservative, since the sensitivity of the assay also depends on the age of the tracer. Cross-reactivity with arginine vasopressin (AVP), ring

moieties and terminal tripeptides of both OXT and AVP and a wide variety of peptides comprising 3 (-melanocyte-stimulating hormone) up to 41 (corticotrophin releasing factor) amino acids are < 0.7% throughout. The intra- and inter-assay variabilities are estimated in < 10%. Saliva samples were analysed in different batches; however, all samples from an individual were always assayed in the same batch. Serial dilutions of saliva samples containing high levels of endogenous OXT run strictly parallel to the standard curve indicating immuno-identity.

2.5. Aggregation of salivary OXT measurements in composite scores

Most studies investigating associations between OXT concentrations in peripheral fluids and behaviour have relied on measurements from samples collected at one single occasion. In this study, we collected seven samples of saliva from each child and measured their OXT levels. This protocol was implemented to 1) assess the reliability of salivary OXT and 2) investigate whether “evoked” OXT after a play episode might offer more reliable estimates of the physiology of the OXT system and capture aspects of this system that might perform better in predicting behavioural problems in children.

There were two intentional variations in the context of saliva sampling: 1) baseline – vs. post playful task; 2) collected by parents vs collected by researchers, whose effect on the variability OXT concentrations we intended to assess. Hence we tested the reliability of four chosen sets of measures, namely: 1) the total three baseline samples collected by parents – one of which was collected on a weekend day and two that were collected on two week days; 2) the baseline saliva samples during the week days only (i.e., excluded the weekend day here) collected by parents; 3) the samples collected by the researchers before the playful interaction task (2 samples: one with father and one with mother); 3) the evoked samples collected by the researchers after the playful interaction task (2 samples: one with father and one with mother).

In Fig. 2, we present a scheme that describes the process of aggregation of sample sets. The detailed context and settings of each of the seven original sampling sets is presented in Fig. 1.

2.6. The standardized Dyadic Play Session

The standardized Dyadic Play Session consists of a well-established and widely used procedure of dyadic play between the parent and the child, that can be easily replicated. It is intended to provide an index of the quality of parenting behaviours while in close physical proximity with the child, and is based on the NICHD Study of Early Child Care

Mother-Child Interaction “Three Boxes Procedure” (NICHD Early Child Care Research Network, 2002). This protocol has been used in large-scale studies, has excellent training materials, good psychometric properties, and, while brief, produces robust scores predictive of later growth in both cognitive and socioemotional domains (e.g., (Linberg et al., 2016)). The child and parent were asked to play with a set of toys, each placed inside a separate bag labelled “1,” “2,” or “3,” as they would normally do. Parents were told to divide the time between the toys as they saw fit. The only rule given was that they should play with the toys in the established numerical order. The toys were selected to be challenging and interesting for the child, and to offer dyads the opportunity to engage in different types of interactions. In this study the toys used were: 1) Mr Potato-Head; 2) puzzles with animals, and 3) a book. For this study the child was sat on the parent’s lap during the task to promote skin to-skin contact, which has been shown in previous studies to be associated with the functioning of the OXT system both in parents and children (Scatliffe et al., 2019). Sessions were randomly counter-balanced between mothers and fathers. The analysis of the video recordings data will be disseminated separately. In this mentioned study we have analysed both the father and the mother sessions using several scales including the Ainsworth sensitivity rating scale (Ainsworth et al., 1974), which was shown to moderate OXT salivary levels after the same play procedure with mothers and in a similar sample of Portuguese pre-schoolers (Baiao et al., 2019). In the present study there were no significant differences between fathers and mothers in the Ainsworth sensitivity scale. This suggests that the behaviour of fathers and mothers in the play sessions were fairly equivalent in terms of sensitivity, which is a well-studied socio-emotional dimension of attachment-theory. In addition to this, differences between the children’s OXT concentrations after playing with the father and with the mother were also tested and no significant differences were found, showing that the two play sessions were fairly equivalent in terms of eliciting OXT responses.

2.7. Assessment of behaviour problems

Children’s behavioural problems were assessed independently by their preschool teacher using the Caregiver-Teacher Report Form (C-TRF) (Achenbach and Rescorla, 2000). The C-TRF is a standardized measure which provides ratings for 120 problem behaviour items. This is one of the most used measures of children behavioural problems world-wide, and has been validated for 36 world languages, therefore easily replicable in further research using samples of different countries. Test-retest reliability for all the scales used in this study was > .75 (Achenbach and Rescorla, 2000). We used the validated Portuguese

	Saliva sample set n°						
<i>Sampling conditions</i>	1	2	3	4	5	6	7
Day	weekend	week	week	week	week	week	week
Who collected?	parents	parents	parents	researcher	researcher	researcher	researcher
Play session (Ps)	No	No	No	Yes, Before Ps	Yes, After Ps	Yes, Before Ps	Yes, After Ps
Participants of Play session	-	-	-	Father and child	Father and child	Mother and child	Mother and child

Fig. 1. Saliva samples’ collection context.

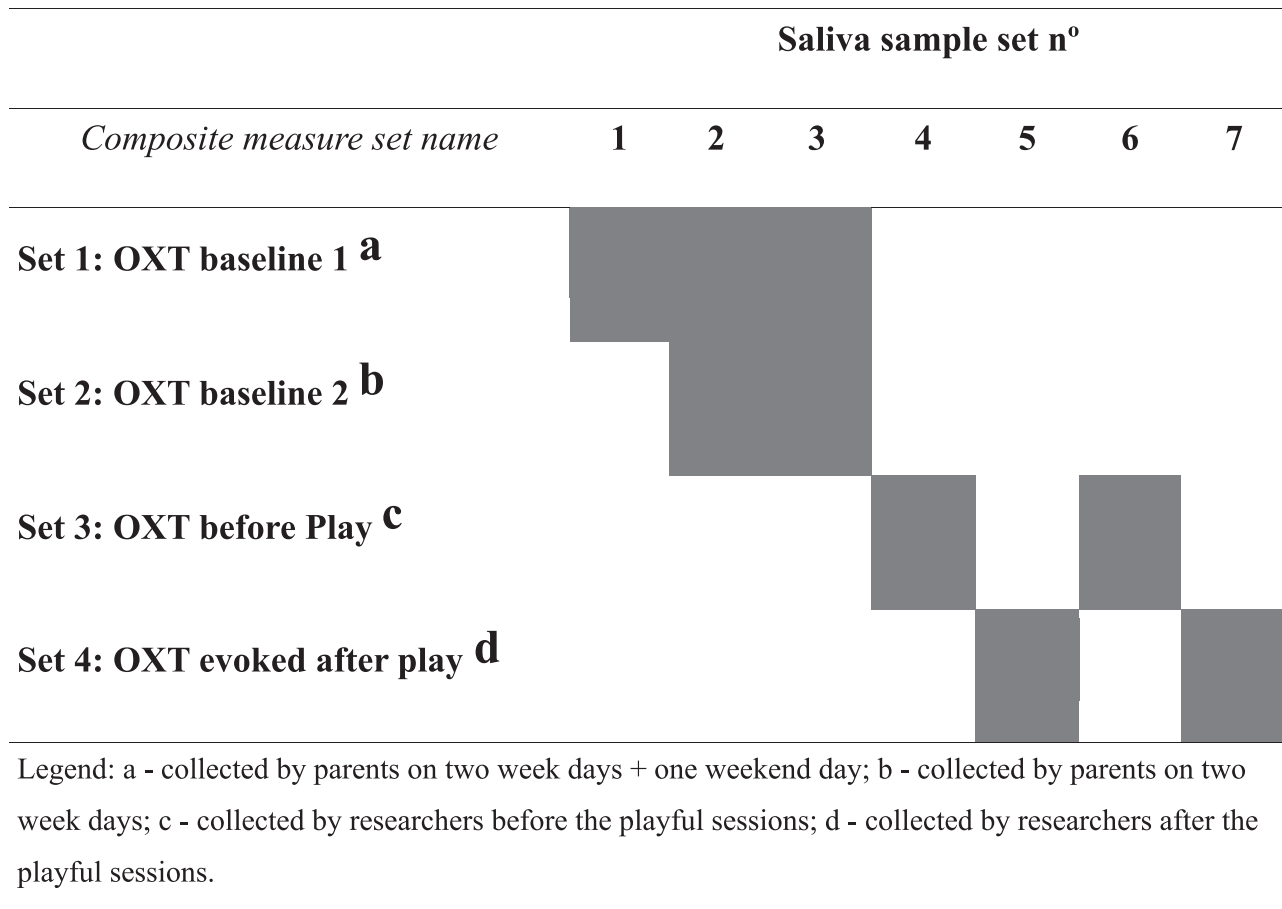


Fig. 2. Aggregation of OXT measures' composite scores.

version and the Portuguese population norms (Rescorla et al., 2012). To be able to complete the questionnaire the teacher should have known the child for at least 6 months, before the beginning of the study. Teachers rated how true each item was at the moment, or was within the past 6 months, using a 3-points scale. The C-TRF yields scores on internalizing, externalizing, and total problems, as well as sub-scores on seven empirically based syndromes (Emotionally Reactive; Anxious/Depressed; Somatic Complaints; Withdrawn; Attention Problems; Aggressive Behaviour), and scores on five DSM-IV oriented scales (Depressive Problems; Anxiety Problems; Autism Spectrum Problems; Attention Deficit/Hyperactivity Problems; Oppositional Defiant Problems).

For this study, only the DSM-IV-oriented scales were used, plus the Internalization and Externalization scales of the C-TRF. These DSM-IV-oriented scales can be more easily compared with other studies using different psychopathology measures that assess DSM taxonomy. The externalization and internalization scales have been used in more than 75.000 articles in developmental psychology research (Achenbach et al., 2016).

2.8. Data analysis plan

1) Firstly, we performed a series of Intraclass Correlation Coefficient (ICC) to assess the reliability of the single measurements and of the four sets of composite OXT measurements. ICCs were estimated using a two-way mixed model, absolute agreement. 2) Secondly, we examined whether our four composite OXT scores were associated with children's behavioural problems in the preschool setting. We tested the Pearson product-moment correlations between the four composite OXT measures and the C-TRF scales, using the percentile bootstrap method (1000

samples). 3) Finally, we used the *r*-to-*z* transform test to compare the magnitude of the correlations between the different composite OXT scores and the behaviour problems scales. A Log transformation was performed to obtain a normal distribution in all OXT measures.

The statistical analyses were performed using the software SPSS 24 (IBM, Armonk, NY, USA).

3. Results

The OXT concentrations of all samples ranged from .29 pg/ml to 6.12 pg/ml (*M* = 1.33; *SD* = .53). As stated above we did not include any sample below 0.5 pg/ml (which were less than 5% of all samples) in the calculations of reliability or in the analyses of correlations with behavioural problems. As shown in Table 1, the product moment inter-correlations among the seven sets of samples were all positive, ranged from .24 to .76. The average correlation was .51. Differences between OXT concentrations after playing with the father and after playing with the mother were tested, using the Mann-Whitney *U* test, no significant differences were found (*z* = -.56, *p* = .61).

The Pearson correlations between the composite scores were as follows: OXT baseline 1 had *r* = .93 (*p* < .001) with OXT baseline 2, *r* = .60 (*p* < .001) with OXT Before Play, and *r* = .73 (*p* < .001) with OXT evoked after Play. OXT baseline 2 had *r* = .68 (*p* < .001) with OXT Before Play and *r* = .72 (*p* < .001) with OXT evoked After Play. Finally, OXT Before Play had *r* = .66 (*p* < .001) with OXT evoked After Play.

3.1. Results for single measures and average measures ICC

We performed a series of Intraclass Correlation Coefficient (ICC) analyses to determine the degree of reliability of the salivary

Table 1
Quantities and Correlations between OXT concentrations in all seven saliva samples.

	1	2	3	4	5	6	7
Saliva measure 1		.69**	.54**	.24	.62**	.36*	.45**
Saliva measure 2			.49**	.25	.71**	.52**	.52**
Saliva measure 3				.39*	.51**	.53**	.44*
Saliva measure 4					.32	.63**	.69**
Saliva measure 5						.39*	.68**
Saliva measure 6							.76**
Saliva measure 7							
Average OXT (SD) in pg/ml,	1.45 (1.02)	1.19 (0,50)	1.40 (0,57)	1.22 (0,52)	1.00 (0,39)	1.36 (0,71)	1.10 (0,38)

* $p < .05$, ** $p < .01$.

measurements of OXT. The ICC is a statistical estimate that measures the extent of agreement between at least two quantitative measurements. Besides from measuring the extent of agreement, ICC is also designed to measure the degree of reliability of quantitative measurements. The ICC calculates coefficients for *single measures* and *average measures*. While the *single measures* coefficient applies to the estimated reliability of a single measurement, the *average measures* coefficient applies to the arithmetic mean of K measurements. According to [Koo and Li \(2016\)](#) the ICC calculation in SPSS, is based on the terminology of [McGraw and Wong \(1996\)](#), which differentiates between ICCs that apply to single measurements (e.g., the ratings of judges), and ICCs that apply to average measurements (e.g., the average rating for k judges). This means that reliability of average measures apply to the composite score of aggregated measurements/observations, and not to a single measurement; this concept is similar to that of Cronbach’s Alpha, which is a particular case of ICC average measures, and can be calculated by applying the Spearman-Brown formula ([Koo and Li, 2016](#); [McGraw and Wong, 1996](#)). The difference between single and average measures reflects precisely differences in the unit of reliability. ICC *single measures* refers to the reliability of one single measure, if this measure was to be used individually. However, ICC *average measures* refers to the reliability of the average of more than one measures collected in the same individual across time-points. Therefore, and very importantly ICC average measures cannot be used to claim reliability of a single measure.

All ICC reliabilities for single measures were not satisfactory ($r < .70$) in all sets of samples, as can be seen in [Table 2](#). Only the single measure for OXT evoked after the dyadic play ($r = .68$) presented a nearly acceptable ICC. Contrastingly, all ICC average measures estimates achieved satisfactory levels ($r > .70$) in all sets. The highest ICC was obtained in the evoked OXT collected after the dyadic play task.

For completeness, we also calculated an ICC for all the five baseline samples, independently of who collected them (parents and researchers). Results for single measures were $ICC = .314$ ([CI95%.090;.617]; $p < .001$) and for average measures were $ICC = .696$ ([CI95% .330;.889]; $p < .001$). These low reliabilities suggest that increasing the number of samples collected on baseline settings

Table 2
ICC values and confidence intervals of salivary OXT.

Composite scores	Average Measures	Single Measures	F	p
OXT baseline 1 a	.738 [0.509 – 0.870]	.484 [0.257–0.691]	4.92	.001
OXT baseline 2 b	.736 [0.535 – 0.836]	.483 [0.133 – 0.718]	4.81	.001
OXT baseline before play c	.747 [0.403 – 0.893]	.596 [0.252 – 0.806]	3.95	.001
OXT evoked after play d	.807 [0.545 – 0.918]	.676 [0.374 – 0.849]	5.18	.001

Legend: a – 3 saliva samples, collected by parents on two week days + one weekend day; b – 2 saliva samples, collected by parents on two week days; c – 2 saliva samples, collected by researchers before the playful sessions; d – 2 saliva samples, collected by researchers after the playful sessions

does not substantially improve reliability of salivary OXT. However, it is expected that increasing the number of samples on the exact same conditions will likely result in higher reliabilities of composite measures.

3.2. Associations between salivary OXT and behavioural problems in preschool

We tested the Pearson product-moment correlations (with bootstrap, percentile method with 1000 samples) between the four composite OXT scores and the five DSM-Oriented C-TRF scales, plus the internalization and externalization scales. Since multiple correlations can increase the probability of errors type I, we applied a correction for multiple comparison using the Benjamini-Hochberg (B-H) procedure. This approach has greater power than the Hochberg or Bonferroni procedures, and its power remains relatively stable as the number of comparisons becomes large ([Williams, Jones & Tukey, 1999](#)). We defined each of the four OXT composite sets’ series of correlations with the C-TRF scales as belonging to one familywise error rate, following the rationale of [Bender and Lang \(2001\)](#). A correlation was significant after correction for multiple comparisons, if it reached a False Discovery Rate of 5% ($FDR = .05$) in the B-H procedure models.

In [Table 3](#) we present the bivariate uncorrected p value for each correlation and we mark with an * the coefficients of which the probability was lower than the critical significance level of 0.05 after correction for multiple comparisons with the B-H procedure. We initially found four significant negative associations between salivary OXT scores and behavioural problems scores, as shown in [Table 2](#). However, after correction of critical p -values for multiple comparisons, only two of them showed to be below critical level of significance. The significant associations were present only with the evoked OXT measurements, i.e., after playing with parents.

In order to test if the correlations between behavioural problems and OXT baseline before play, OXT evoked after play, were significantly different from the same correlations but when saliva was collected by parents in a baseline routine situation (OXT baseline 1 and OXT baseline 2), we used an r to z test.

For the depressive problems, the correlation with OXT before play was significantly stronger than those with both OXT baseline 1 ($z = 2.071$, $p = 0.019$) and OXT baseline 2 ($z = 2.177$; $p = .015$). The evoked OXT after play correlation was also significantly stronger than those with both OXT baseline 1 ($z = 2.443$, $p = .007$) and OXT baseline 2 ($z = 2.548$; $p = .005$).

For the anxiety problems scale, the correlation with OXT before play was also significantly stronger than those with both OXT baseline 1 ($z = -2.196$; $p = .014$) and OXT baseline 2 ($z = -1.93$; $p = .027$).

All the other comparisons between correlations did not reach statistical significance. Additionally, we also compared the correlations between sets 3 and 4 (i.e., OXT baseline before Play versus OXT evoked after Play) with all the problems scales, and none of them reached statistical significance.

Table 3
Correlations between salivary OXT and the C-TRF behavioural problems' scales.

	Set 1 OXT baseline 1	Set 2 OXT baseline 2	Set 3 OXT baseline before Play	Set 4 OXT evoked after Play
Depressive problems	-0.01 [- 0.32, 0.39] p = .958	-0.03 [- 0.42, 0.39] p = .874	-0.37 [- 0.64, - 0.03] p = .042	-0.43 * [- 0.73, - 0.07] p = .009
Anxiety Problems	.10 [- 0.23, 0.51] p = .692	.05 [- 0.33, 0.45] p = .795	-.31 [- 0.59, - 0.04] p = .095	-.19 [- 0.57, 0.33] p = .157
Autism Spectrum Problems	-.02 [- 0.42, 0.52] p = .916	.05 [- 0.44, 0.47] p = .795	-.29 [- 0.63, - 0.03] p = .12	-.19 [- 0.61, 0.42] p = .157
Attention Deficit/ Hyperactivity Problems	-.03 [- 0.34, 0.31] p = .893	-.08 [- 0.46, 0.32] p = .674	-.09 [- 0.42, 0.22] p = .636	-.24 [- 0.58, 0.14] p = .101
Opposition problems	-.23 [- 0.51, 0.16] p = .112	-.17 [- 0.47, 0.19] p = .369	-.18 [- 0.56, - 0.23] p = .341	-.40 * [- 0.72, - 0.09] p = .014
Externalization problems	-.13 [- 0.43, 0.25] p = .510	-.12 [- 0.34, 0.26] p = .527	-.13 [- 0.50, 0.27] p = .493	-.34 [- 0.64, 0.08] p = .033
Internalization problems	.08 [- 0.33, 0.54] p = .674	.12 [- 0.36, 0.56] p = .527	-.27 [- 0.56, 0.11] p = .149	-.26 [- 0.66, 0.29] p = .082

95% Confidence intervals in parenthesis.

Legend

p – uncorrected bivariate probability value.

* probability lower than the critical significance level of 0.05 after corrected for multiple comparisons with the B-H procedure, False Discovery Rate of 5% (FDR = .05).

4. Discussion

In this study, we aimed firstly to characterize the reliability of OXT measurements in our saliva samples. Additionally we tested if young children's' salivary OXT was associated with behavioural problems that are evident in their child-care settings as reported by their preschool teachers. Finally, and based on previous literature suggesting higher synchrony of central and peripheral OXT after emotional events, we aimed to assess if salivary OXT after a playful episode with the parents had a different magnitude of associations with behavioural problems than at baseline levels.

We found significant negative correlations between salivary OXT and depression and oppositional behaviour problems. Furthermore, we found a coherent pattern of stronger associations between OXT levels and behaviour problems for “evoked” measures, i.e. after a playful session with parents, than at baseline, which was statistically significant specifically in the depression problems scale only. These results suggest that evoked measurements after a hedonic situation may identify physiological states that are more relevant from a psychopathological point of view than baseline measurements. Presently we cannot offer a conclusive description of the underlying processes involved. Previous research showed significantly greater central-peripheral correlations after various types of emotional events: a) after separation and after reunion with the mother in rat pups (Kojima et al., 2012), b) after experimentally induced stress (Valstad et al., 2017). In the light of these data, we can speculate and hypothesize that a) the stress induced by presence of strangers in the house (the researchers) and b) the playful and physical proximity with the parents, might have induced a synchronic physiological response of central-peripheral release of OXT that

rendered the measures more relevant to predict social-behaviour in the preschool setting. In other words: It is plausible that the presence of two researchers strange to the child inside their house might induce a stress response in the children that could trigger the OXT negative feedback process. This initial response is likely mitigated over time as non-hostile interactions between researchers and the family take place (both researchers were developmental psychologists, highly experienced in dealing with children). Our assumption here is that the play situation with parents, which is the moment closer in time to the “evoked” measure we collected, is typically a hedonic one. However, we cannot make definitive conclusions about what mechanisms might account for our higher reliability findings.

Although we did not recruit a clinical sample, our findings are globally congruent with previous reports on the associations between OXT in central and peripheral fluids and Depression, Anxiety, and Conduct problems in children (Carson et al., 2014; Demirci et al., 2016a, 2016b; Levy et al., 2015; Salzberg and Swedo, 1992; Sasaki et al., 2016).

These results are interesting for two reasons. First, although causality cannot be determined, if this hormone levels are indeed altered in children with problem behaviours, then it is possible that these measurements may provide some sort of biomarker for psychopathology. Secondly the results suggest that the dysregulation of the OXT system may be associated with manifestation/exacerbation of psychopathological dimensions across several different diagnoses.

Our findings also highlight that in order to capture such associations using salivary OXT concentrations: 1) sample collection should be done after emotional responses are evoked. In our case the evoked situation was a hedonic, playful situation between each parent and the child, likely to evoke a positive socio-emotional state; 2) sample collection should be performed by researchers in a controlled protocol and not by parents under their normal daily routine (even if the later might sound more appealing to minimize perceived stress in the children associated with the interaction with a stranger). The reasons underlying these differences between parents and researchers' collection remain unknown. It is possible that because researchers are strangers to the child, their presence introduces an emotional stress element that is relevant for the neurobiology of the oxytocin response.

The results also suggest that the Salivary OXT measured after an episode of dyadic play with parents yielded higher levels of reliability. The reliability indexes for single measurements calculated in this study are in the range of those already described for other hormones such as vasopressin or prolactin in adults (Koenig et al., 1993; Quintana et al., 2017), suggesting low reliability in single measures may be a feature common to other hypothalamic peptides. The intra-assay variability of the method used for OXT quantification was estimated to be < 10% and is unlikely to account for the low reliabilities here reported. As far as we are aware, no study to date had inspected whether these baseline fluctuations affect saliva OXT concentrations. We minimized sources of variability between sessions related to circadian rhythms and impact of food intake by collecting data at the same time on each day and instructing parents not to allow the child to eat prior to the assessments.

Additionally, we found that composite scores of more than one OXT sample proved to be satisfactorily reliable. In this study salivary OXT was unlikely to be very reliably measured in single assessments. However acceptable high levels of reliability (> .70) were achieved through the averaging of at least two samples. This result supports a previous study that found low reliabilities of single measures of both saliva and plasma oxytocin in healthy adult men (Martins et al., 2020). Although saliva sampling is non-invasive and convenient to perform, collecting several samples per individual has increased costs associated with the quantification of multiple samples per subject and may not be possible for all protocols. For these reasons, we believe that identifying circumstances where saliva OXT may be measured in a more reliable way in single sampling is worth exploring. One such specific circumstance may be the quantification of evoked responses associated with specific events, such as the dyadic interaction play task used in this study, after

which higher reliability was obtained. Indeed, higher reliabilities of hormonal responses in evoked measurements have also been reported for other systems, such as the neuroendocrine response to stress (Coste et al., 1994).

These are, however, preliminary findings and should not be used to reify the clinical value of salivary OXT as biomarker for high-risk to psychopathology in children. Replications in larger scale and in clinical samples will be required. In any case, we note that the low reliability of single measurements underlying our recommendation of multiple samples per subject may make baseline saliva OXT a less attractive target for a screening biomarker in the clinical routine.

Some limitations should be acknowledged. Firstly, our sample size was relatively small. However, an a priori power analysis calculation based on the number of samples we collected per individual suggest we should have been able to detect an acceptable ICC if it did exist. As the power analysis was made a priori to data collection and analysis, as stated, it was based on the total number of saliva samples per subject we planned to have (seven saliva samples). In any case, as modeled and presented by Bujang and Baharum (2017), the number of cases/subjects needed to detect ICCs = .7 is: a) with 3 observations $N = 6$, b) with 2 observations is $N = 10$. Hence, for any of the situations (either 2 or 3 observations/saliva samples), the number of cases we had in the study ($N = 30$) was adequate. In what concerns the correlations between OXT and behavioural problems, our study was sufficiently powered only to detect correlations of medium magnitude for a minimal ICC of 0.7; for smaller correlations or ICCs, our study is not adequately powered, which makes these findings preliminary and in need of replication with larger samples.

Our composite scores for OXT levels after play (and the respective average measures ICC estimates), were computed from interactions with the father and the mother, and hence were not a constant: the two sessions involved a different parent (mother, father). The child may have different profile of interactions with each parent (namely the attachment security of the child may be different across parents). Therefore our reliability estimates most likely represent the lower limit of the true reliability of these measures. However, previous studies with similar samples of Portuguese pre-schoolers living in sociodemographic similar families have shown that children tend to be securely attached to both father and mother (e.g. (Monteiro et al., 2010)). Future studies of reliability should either i) collect saliva before and after an interaction with the same parent or ii) control for pertinent socio-emotional interaction profiles with both parents.

Declaration of interests

None.

Acknowledgements

We would like to acknowledge the precious contributions to this paper by the late Jaak Panksepp, he will always be an inspiration for us and long remembered by many. This research has been supported in part by BIAL Foundation 158/12 and UIDB/04810/2020. The authors thank all the members of the research team for their assistance and support.

References

- Achenbach, T.M., Rescorla, L.A., 2000. *Manual for the ASEBA Preschool Forms & Profiles*. University of Vermont, Research Center for Children, Youth, & Families, Burlington, VT.
- Achenbach, T.M., Ivanova, M.Y., Rescorla, L.A., Turner, L.V., Althoff, R.R., 2016. Internalizing/externalizing problems: review and recommendations for clinical and research applications. *J. Am. Acad. Child Adolesc. Psychiatry* 55 (8), 647–656. <https://doi.org/10.1016/j.jaac.2016.05.012>.
- Ainsworth, D.S., Bell, S.M., Stayton, D., 1974. Infant-mother attachment and social development: socialization as a product of reciprocal responsiveness to signals. In: Richards, M.P.M. (Ed.), *The Integration of a Child into a Social World*. Cambridge University Press, pp. 99–135.
- Bachner-Melman, R., Ebstein, R.P., 2014. The role of oxytocin and vasopressin in emotional and social behaviors. *Handbook of Clinical Neurology*, 124. Elsevier, pp. 53–68.
- Baião, R., Fearon, P., Belsky, J., Baptista, J., Carneiro, A., Pinto, R., Nogueira, M., Oliveira, C., Soares, I., Mesquita, A.R., 2019. Child's oxytocin response to mother-child interaction: the contribution of child genetics and maternal behaviour. *Psychoneuroendocrinology* 102, 79–83. <https://doi.org/10.1016/j.psyneuen.2018.11.022>.
- Bender, R., Lange, S., 2001. Adjusting for multiple testing—when and how? *J. Clin. Epidemiol.* 54 (4), 343–349. [https://doi.org/10.1016/s0895-4356\(00\)00314-0](https://doi.org/10.1016/s0895-4356(00)00314-0).
- Bujang, M.A., Baharum, N., 2017. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. *Arch. Orofac. Sci.* 12, 1–11.
- Caldwell, H.K., 2017. Oxytocin and vasopressin: powerful regulators of social behavior. *Neuroscientist* 23 (5), 517–528. <https://doi.org/10.1177/1073858417708284>.
- Carson, D.S., Berquist, S.W., Trujillo, T.H., Garner, J.P., Hannah, S.L., Hyde, S.A., Sumiyoshi, R.D., Jackson, L.P., Moss, J.K., Strehlow, M.C., Cheshier, S.H., Partap, S., Hardan, A.Y., Parker, K.J., 2014. Cerebrospinal fluid and plasma oxytocin concentrations are positively correlated and negatively predict anxiety in children. *Mol. Psychiatry* 2014, 1–6. <https://doi.org/10.1038/mp.2014.132>.
- Carter, C.S., 1998. Neuroendocrine perspectives on social attachment and love. *Psychoneuroendocrinology* 23, 779–818.
- Cochran, D., Fallon, D., Hill, M., Frazier, J.A., 2013. The role of oxytocin in psychiatric disorders: a review of biological and therapeutic research findings. *Harv. Rev. Psychiatry* 21 (5), 219–247. <https://doi.org/10.1097/HRP.0b013e3182a75b7d>.
- Coste, J., Strauch, G., Letrait, M., Bertagna, X., 1994. Reliability of hormonal levels for assessing the hypothalamic-pituitary-adrenocortical system in clinical pharmacology. *Br. J. Clin. Pharmacol.* 38 (5), 474–479.
- Daughters, K., Manstead, A.S., Hubble, K., Rees, A., Thapar, A., van Goozen, S.H., 2015. Salivary oxytocin concentrations in males following intranasal administration of oxytocin: a double-blind, cross-over study. *PLoS One* 10, e0145104.
- Demirci, E., Ozmen, S., Oztop, D.B., 2016a. Relationship between impulsivity and serum oxytocin in male children and adolescents with attention-deficit and hyperactivity disorder: a preliminary study. *Noro Psikiyatr. Ars.* 53 (4), 291–295.
- Demirci, E., Ozmen, S., Kilic, E., Oztop, D.B., 2016b. The relationship between aggression, empathy skills and serum oxytocin levels in male children and adolescents with attention deficit and hyperactivity disorder. *Behav. Pharmacol.* 27 (8), 681–688. <https://doi.org/10.1097/FBP.0000000000000234>.
- Engelmann, M., Ebner, K., Landgraf, R., Holsboer, F., Wotjak, C.T., 1999. Emotional stress triggers intrahypothalamic but not peripheral release of oxytocin in male rats. *J. Neuroendocrinol.*
- Feldman, R., Gordon, I., Zagoory-Sharon, O., 2010a. The cross-generation transmission of oxytocin in humans. *Horm. Behav.* 58 (4), 669–676.
- Feldman, R., Gordon, I., Schneiderman, I., Weisman, O., Zagoory-Sharon, O., 2010b. Natural variations in maternal and paternal care are associated with systematic changes in oxytocin following parent-infant contact. *Psychoneuroendocrinology* 35 (8), 1133–1141. <https://doi.org/10.1016/j.psyneuen.2010.01.013>.
- Feldman, R., Gordon, I., Influx, M., Gutbir, T., Ebstein, R.P., 2013. Parental oxytocin and early caregiving jointly shape children's oxytocin response and social reciprocity. *Neuropsychopharmacology* 38 (7), 1154–1162. <https://doi.org/10.1038/npp.2013.22>.
- de Jong, T.R., Menon, R., Bludau, A., Grund, T., Biermeier, V., Klampfl, S.M., Jurek, B., Bosch, O.J., Hellhammer, J., Neumann, I.D., 2015. Salivary oxytocin concentrations in response to running, sexual self-stimulation, breastfeeding and the TSST: the Regensburg Oxytocin Challenge (ROC) study. *Psychoneuroendocrinology* 62 (2015), 381–388. <https://doi.org/10.1016/j.psyneuen.2015.08.027>.
- Jurek, B., Neumann, I.D., 2018. The oxytocin receptor: from intracellular signaling to behavior. *Physiol. Rev.* 98 (3), 1805–1908. <https://doi.org/10.1152/physrev.00031.2017>. PMID: 29897293.
- Koenig, K.L., Toniolo, P., Bruning, P.F., Bonfrer, J.M., Shore, R.E., Pasternack, B.S., 1993. Reliability of serum prolactin measurements in women. *Cancer Epidemiol. Biomark. Prev.* 2 (5), 411–414.
- Kojima et al. (2012) Achievement. *Journal of Educational and Behavioral Statistics.* 24 (1):42-69. doi:10.3102/10769986024001042.
- Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 15, 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>.
- Lebowitz, E.R., Silverman, W.K., Martino, A.M., Zagoory-Sharon, O., Feldman, R., Leckman, J.F., 2017. Oxytocin response to youth-mother interactions in clinically anxious youth is associated with separation anxiety and dyadic behavior. *Depress. Anxiety* 34 (2), 127–136. <https://doi.org/10.1002/da.22585>.
- Leeds, A., Dennis, P.M., Lukas, K.E., Stoinski, T.S., Willis, M.A., Schook, M.W., 2018. Validating the use of a commercial enzyme immunoassay to measure oxytocin in unextracted urine and saliva of the western lowland gorilla (*Gorilla Gorilla Gorilla*). *Primates* 59, 499–515. <https://doi.org/10.1007/s10329-018-0678-3>.
- Levy, T., Bloch, Y., Bar-Maisels, M., Gat-Yablonski, G., Djalovski, A., Borodkin, K., et al., 2015. Salivary oxytocin in adolescents with conduct problems and callous-unemotional traits. *Eur. Child Adolesc. Psychiatry* 24 (12), 1543–1551. <https://doi.org/10.1007/s00787-015-0765-6>.
- Linberg, A.S., Hachul, C., Roßbach, H.-G., 2016. Video-based assessment and rating of parent-child interaction within the National Educational Panel Study. In: Blossfeld, H.-P., Maurice, J., Bayer, M., Skopek, M.J. (Eds.), *Methodological Issues of Longitudinal Surveys*. Springer.
- MacLean, E.L., Gesquiere, L.R., Gee, N., Levy, K., Martin, W.L., Carter, C.S., 2018. Validation of salivary oxytocin and vasopressin as biomarkers in domestic dogs. *J. Neurosci. Methods* 293, 67–76.

- Martin, J., Kagerbauer, S.M., Gempt, J., Podtchaske, A., Hapfelmeier, G., Schneider, A., 2018. Oxytocin levels in saliva correlate better than plasma levels with concentrations in the cerebrospinal fluid of patients in neurocritical care. *J. Neuroendocrinol.* 30, 5. <https://doi.org/10.1111/jne.12596>.
- Martins, D., Gabay, A., Mehta, M.A., Paloyelis, Y., 2020. Are single peripheral measurements of baseline oxytocin in saliva and plasma reliable biomarkers of the physiology of the oxytocin system in humans? *eLife* 9, e62456. <https://doi.org/10.7554/eLife.62456>. PMID: 33306025; PMCID: PMC7732341.
- McGraw, K.O., Wong, S.P., 1996. Forming inferences about some intraclass correlation coefficients. *Psychol. Methods* 1(1), 30–46.
- Miller, J.L., Tamura, R., Butler, M.G., Kimonis, V., Sulsona, C., Gold, J.A., et al., 2017. Oxytocin treatment in children with Prader-Willi syndrome: a double-blind, placebo-controlled, crossover study. *Am. J. Med. Genet. A* 173 (5), 1243–1250. <https://doi.org/10.1002/ajmg.a.38160>.
- Monteiro, L., Verissimo, M., Vaughn, B.E., Santos, A.J., Torres, N., Fernandes, M., 2010. The organization of children's secure base behaviour in two-parent Portuguese families and father's participation in child-related activities. *Eur. J. Dev. Psychol.* 7 (5), 545–560. <https://doi.org/10.1080/17405620902823855>.
- Nelson, E., Panksepp, J., 1998. Brain substrates of infant – mother attachment: contributions of opioids, oxytocin, and norepinephrine. *Neurosci. Biobehav. Rev.* 22, 437–452.
- NICHD Early Child Care Research Network, 2002. Early child care and children's development prior to school entry results from the NICHD study of early child care. *Am. Educ. Res. J.* 39, 133–164.
- Oztan, O., Jackson, L.P., Libove, R.A., Sumiyoshi, R.D., Phillips, J.M., Garner, J.P., et al., 2018. Biomarker discovery for disease status and symptom severity in children with autism. *Psychoneuroendocrinology* 89, 39–45. <https://doi.org/10.1016/j.psyneuen.2017.12.022>Get (rights and content).
- Quintana, D.S., Westlye, L.T., Smerud, K.T., Mahmoud, R.A., Djupesland, P.G., Andreassen, O.A., 2017. Reliability of basal plasma vasopressin concentrations in healthy male adults. *Acta Neuropsychiatr.* 29 (5), 315–321. <https://doi.org/10.1017/neu.2016.67>.
- Rescorla, L.A., Achenbach, T.M., Ivanova, M.Y., Bilenberg, N., Bjarnadottir, G., Denner, S., Verhulst, F.C., 2012. Behavioral/emotional problems of preschoolers: caregiver/teacher reports from 15 societies. *J. Emot. Behav. Disord.* 20 (2), 68–81. <https://doi.org/10.1177/1063426611434158>.
- Robles, T.F., Sharma, R., Harrell, L., Elashoff, D.A., Yamaguchi, M., Shetty, V., 2013. Saliva sampling method affects performance of a salivary α -amylase biosensor. *Am. J. Hum. Biol.* 25 (6), 719–724. <https://doi.org/10.1002/ajhb.22438>.
- Rutigliano, G., Rocchetti, R., Paloyelis, Y., Gilleen, J., Sardella, A., Cappucciati, M., Palombini, E., Dell'Osso, L., Caverzasi, E., Politi, P., McGuire, P., Fusar-Poli, P., 2016. Peripheral oxytocin and vasopressin: biomarkers of psychiatric disorders? A comprehensive systematic review and preliminary meta-analysis. *Psychiatry Res.* 241, 207–220.
- Salzberg, A.D., Swedo, S.E., 1992. Oxytocin and vasopressin in obsessive-compulsive disorder. *Am. J. Psychiatry* 149 (5), 713–714, 120.
- Sasaki, T., Hashimoto, K., Oda, Y., Ishima, T., Yakita, M., Kurata, T., et al., 2016. Increased serum levels of oxytocin in treatment resistant depression in adolescents (TRDIA) group. *PLoS One* 11 (8), e0160767. (<https://doi.org/10.1371/journal.pone.0160767>).
- Scatliffe, N., Casavant, S., Vittner, D., Cong, X., 2019. Oxytocin and early parent-infant interactions: a systematic review. *Int. J. Nurs. Sci.* 6 (4), 445–453 <https://doi.org/10.1016/j.ijnss.2019.09.009>.
- Stein, D.J., 2009. Oxytocin and vasopressin: social neuropeptides. *CNS Spectr.* 14 (11), 602–606. <https://doi.org/10.1017/S1092852900023841>.
- Taylor, A.E., Lee, H.E., Buisman-Pijlman, F.T., 2014. Oxytocin treatment in pediatric populations. *Front. Behav. Neurosci.* 8, 360. <https://doi.org/10.3389/fnbeh.2014.00360>.
- Torres, N., Martins, D., Santos, A., Prata, D., Verissimo, M., 2018. How do hypothalamic nonapeptides shape youth's sociality? A systematic review on oxytocin, vasopressin and human socio-emotional development. *Neurosci. Biobehav. Rev.* 90 (1), 309–331. <https://doi.org/10.1016/j.neubiorev.2018.05.004>.
- Uvnäs-Moberg, K., Handlin, L., Peterson, M., 2015. Self-soothing behaviors with particular reference to oxytocin release induced by non-noxious sensory stimulation. *Front. Psychol.* <https://doi.org/10.3389/fpsyg.2014.01529>.
- Valstad, M., Alvares, G.A., Egknud, M., Matziorinis, A.M., Andreassen, O.A., Westlye, L. T., et al., 2017. The correlation between central and peripheral oxytocin concentrations: a systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* 78, 117–124. <https://doi.org/10.1016/j.neubiorev.2017.04.017>.
- Williams, V.S.L., Jones, L.V., Tukey J.W. (1999) Controlling Error in Multiple Comparisons, with Examples from State-to-State Differences in Educational.