# Salivary oxytocin after play with parents predicts behavioral problems in preschool children

**Nuno Torres** ( ntorres@ispa.pt )
Instituto Universitario de Ciencias Psicologicas Sociais e da Vida, William James Research Center

**Daniel Martins**
King's College London, Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience

**Ligia Monteiro**
Instituto Universitário de Lisboa (ISCTE), CIS-IUL

**António J. Santos**
Instituto Universitario de Ciencias Psicologicas Sociais e da Vida, William James Research Center

**Brian E. Vaughn**
Auburn University, Department of Human Development and Family Studies

**Manuela Veríssimo**
Instituto Universitario de Ciencias Psicologicas Sociais e da Vida, William James Research Center

**Research Article**

**Keywords:** oxytocin, saliva, children, radioimmunoassay, behavioural problems

**Posted Date:** February 15th, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-220873/v1

# 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 collected by the parents could not predict any of the behavioural problems measured by the C-TRF. However, when collected by the researchers, salivary OXT before playing with parents correlated negatively with the C-TRF depression and anxiety subscales. Additionally there was a richer and stronger pattern of negative correlations between the salivary OXT measured after playing with parents and the depression, opposition, externalization and total problem 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 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 are relevant for understanding the role of this system in children's social behaviour.


# Introduction

Over the past decades the hypothalamic neuropeptide oxytocin (OXT) has attracted considerable attention as a physiological undepinning of social bonds in mammals (Leeds et al, 2018; Nelson & Panksepp, 1998). Accumulating evidence has demonstrated the importance of OXT in the development of both normative and deviant social behaviour and cognition (Bachner-Melman & Ebstein, 2014; Caldwell, 2017; 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, Ozmen, Kilic, & Oztop, 2016; 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 disorders (Miller et al., 2017; Taylor, Lee, & Buisman-Pijlman, 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 (Grazia 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, Martins, Santos, Prata & Veríssimo, 2018). Even though the evidence was mixed, several studies have reported significant associations between OXT levels and symptons in children (Demirci et al., 2016a; Demirci et al., 2016b; Carson et al., 2014; Lebowitz et al., 2017; Levy et al., 2015, Oztan et al., 2018; Salzberg & Swedo, 1992; Sasaki et al., 2016). Work still has to be done on the methodological homogeneity of the studies, their reliability assessment and standartization (Grazia 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)

## 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 significative 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 2) after intranasal OXT administration ($r$ = .66, $p$ < .0001), butthere was no significant correlation under baseline. The fact that there is evidence of a sincrony beetwen 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 to explain interindividual variance in behaviour than baseline ones. 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 (Grazia 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 (stressfull event) and also 2) during the reunion with the mother. 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. Additionally the authors found that both plasma ($r = -.92$, $p = .0262$, n = 10) and CSF ($r = -.91$, $p = .0335$, 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 stressfull situation. These results, although prelimirary and pre-clinical suggest that peripheral measurements as much as central ones can detect OXT's association with psychopatology symptoms in children

## Salivary measurements of OXT

Saliva is a particularly attractive fluid to quantify OXT during childhood (Leeds, 2018; Uvnäs-Moberg et al., 2015). *S*everal methods for measuring salivary OXT have been already reported and supported saliva as a noninvasive source to monitor central neuroendocrine functions, both in animals and humans (Carter et al., 2007; Daughters et al. 2015; DeJong, et al., 2015; Feldman et al., 2010; Leeds, 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 stongly 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 stressfull and hedonic events, such as the death of a conspecific, play, and breeding (Leeds et al., 2018).

## Reliability of OXT measurements

Questions about the reliablity of single measurements of peripheral OXT make it yet a controversial topic. Martins and colleagues (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 and colleagues (2010) sampled salivary OXT from 55 infants of 4 to 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, and colleagues (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 strongy 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. Even less in known about the reliability of salivary OXT measurements at non-baseline conditions, such as after stressful or hedonic events, the later which we adressed in this study.

In the present study we tested the extent to which oxytocin concentratios 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 situarion *versus* collected by parents in a routine situation. We also aimed to estimate the reliability/stability 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, hence 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, radioimunoassay after extraction, which is currently considered the gold-standard technique for measuring OXT in biological fluids.

### Objetives

1- Test if a series of salivary OXT levels samples, collected at children's home at different time points, would predict common forms of behavioural and psychopathology problems of young children.

2 - Assess the correlations between children's behavioral problems assessed independently by the preeschool teacher and their salivary OXT levels both at baseline and after a playfull episode with both parents

3 - Estimate the stability and 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.

### Data Analysis Plan

1) Firstly, we performed a series of Intraclass Correlation Coefficient (ICC) analyses to determine the degree of reliability of the four sets of salivary OXT measurements. ICCs were estimated using a two-way mixed model, absolute agreement. 2) Secondly, we examined whether our four aggregated OXT measures were associated with children's behavioural problems in the preschool setting. We tested the Pearson product-moment correlations between the four aggregated 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 aggregated OXT measures 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).

## Results

The OXT concentrations ranged from .29 pg/ml to 6.12 pg/ml ($M$ = 1.33; $SD$ = .53). 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.

| 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/l,* | 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)* |
| * $p < .05$ ** $p < .01$ | | | | | | | |

### *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 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, consistency and stability 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 arithmetical mean of *N* measurements. This means that in *average measures* reliability is dependent on the aggregated series of measurements/observations, and not on a single measurement (Martins et al., 2020).

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. Contrastingly, all ICC reliabilities for single measures where not satisfatory ($r < .70$) 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.

| Table 2. ICC values and confidence intervals for aggregated salivary OT measures | | | | |
|---|---|---|---|---|
| *Aggregated measure* | 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 - collected by parents on two week days + one weekend day ; b - collected by parents on two week days; c - collected by researchers before the playful sessions; d - collected by researchers after the playful sessions

### Associations between salivary OXT and Behavioural Problems in preschool

We tested the Pearson product-moment correlations (with percentile bootstrap method with 1000 samples) between the four aggregated OXT measures and the DSM-Oriented C-TRF scales, plus the internalization, externalization and total problems scales. We found significant negative associations between salivary OXT levels and behavioural problems scores in four scales of the C-TRF, as shown in Table 2. The significant associations were mostly present with the evoked OXT measurements, i.e., after playing with parents. No significant associations were found in the baseline 1 and baseline 2 situations: when researchers were not present and saliva was collected by parents.

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

| | OXT baseline 1 | OXT baseline 2 | OXT baseline before Play | OXT evoked after Play |
|---|---|---|---|---|
| **Depressive problems** | - .01 [-0.32, 0.39] | - .03 [-0.42, 0.39] | - .37* [-0.64, 0.03] | - .43** [-0.73, 0.07] |
| **Anxiety Problems** | .10 [-0.23, 0.51] | .05 [-0.33, 0.45] | - .31* [-0.59, -0.04] | - .19 [-0.57, 0.33] |
| **Autism Spectrum Problems** | - .02 [-0.42, 0.52] | .05 [-0.44, 0.47] | - .29 [-0.64, -0.03] | - .19 [-0.61, 0.42] |
| **Attention Deficit/ Hyperactivity Problems** | - .03 [-0.34, 0.31] | - .08 [-0.46, 0.32] | - .09 [-0.42, 0.26] | - .24 [-0.58, 0.14] |
| **Opposition problems** | - .23 [-0.51, 0.16] | - .17 [-0.47, 0.19] | - .18 [-0.56, -0.23] | - .40* [-0.72, 0.00] |
| **Externalization problems** | - .13 [-0.43, 0.25 | - .12 [-0.34, 0.26] | - .13 [-0.50, 0.27] | - .34* [-0.64, 0.08] |
| **Internalization problems** | .08 [-0.33, 0.54] | .12 [-0.36, 0.56] | -.27 [-0.56, 0.11] | - .26 [-0.66, 0.29] |
| **Total problems** | - .03 [-0.42, 0.42] | - .04 [-0.38, 0.38] | - .22 [-0.53, 0.15] | - .32* [-0.64, 0.09] |

*Confidence intervals in parenthesis*

Legend

***p* < .01 ***** *p* < .05

In order to test if the correlations between OXT baseline before play, OXT evoked after play and behavioral problems , 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).

For the total problems scale, the correlation with OXT evoked after play was stronger but only marginally significantly different from those with OXT baseline 1 (z=-1.565; p1313= .059) and OXT baseline 2 (z = -1.512; p = .065). All the other comparisons between correlations did not reach statistical significance.

# Discussion

In this study, we tested if young childrens' salivary OXT was associated with behavioural problems that are evident in their child-care settings as reported by their preschool teachers. Additionally, and based on previous literature suggesting higher sincrony 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 behavioral problems than at baseline levels. Finally we aimed to characterize the reliability of OXT measurements in our saliva samples.

We found significant negative correlations between salivary OXT and internalization problems (depression and anxiety), externalization problems (opositional behavior), and the composite scales of Externalizing and Total problems . Furthermore, we found a coherent pattern of significantly stronger associations between OXT levels and behavior problems for “evoked” measures, i.e. after a playfull session with parents, than at baseline, specifically in the depression and Total problems scales. These results suggest that evoked measurements after an 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 syhncronic physiological response of central-peripheral release of OXT that rendered the measures more relevant to predict social-behaviour in the preschool setting.

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 (Demirci et al., 2016a; Demirci et al., 2016b; Carson et al, 2014; Levy et al., 2015; Salzberg & 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 behaviors, 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/consistency.The reliability indexes 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.

Additionaly we found that aggregated measures of more than one OXT sample proved to be satisfactorily reliable. Salivary OXT was in this study 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 reliably 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, Strauch, Letrait, & Bertagna, 1994).

These are, however, preliminary findings and should not be used to reify the clinical value of saliva 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, a power analysis calculation based on the number of samples we collected per individual suggest we should had been able to detect an acceptable ICC if it did exist. Also, our reliability findings only apply to 1) the specific method of saliva sampling used (passive drooling) with immediate freezing of the saliva at - 20Cº; and 2) to the method of quantification we used, radioimmunoassay, including prior extraction.

# Methods

## 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 hours 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 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)

## 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 assent to participate.

## Protocol

A total of seven samples of saliva, from each child, were collected (i.e., a total of 210 saliva samples). The seven sets of 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 3.5ml of saliva on each sample. After being collected, the saliva was immediately stored in a freezer (- 20Cº), 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 - 80Cº until being sent for analysis in an outsource service (RIAgnosis, Munich, Germany). Transport was made in solid carbon at an average temperature of - 80Cº. Samples were collected by passive drooling to a 5ml 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. To minimize variability, we instructed the parents that before collecting saliva children should 1) be fasting for at least 2 hours, 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 time-frame (i.e., between 5 and 6 pm). 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. 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 of the father and the mother 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 mins before each play session, and a second sample was collected 15 mins after the play session. 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 anti-histamines to control hay fever symptoms.

## Quantification of salivary OXT

Salivary OXT was quantified by Radioimmunoassay (RIAgnosis, Munich, Germany), after extraction. In summary, for each sample, 300 ul of saliva was evaporated (SpeedVac, Thermoscientific Inc, Waltham, MA, USA), and 50 ul of assay buffer was added followed by 50 ul antibody (raised in rabbits against OXT). After a 60-min pre-incubation interval, 10 ul 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. 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.

## Aggregation of salivary OXT measurements

Most studies investigating associations between OXT concentrations in peripheral fluids and behaviour rely on measurements from samples collected at one single occasion. In this study, we collected seven different 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. Since there were two intentional variations in the context of saliva sampling (baseline- *vs.* post playful task, collected by parents *vs* collected by researchers), whose effect on the variability OXT concentrations we intended to assess, we tested the reliability of four chosen sets of measures, namely: a) 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; b) the baseline saliva samples collected by parents during the week days only (i.e., excluded the weekend day here); c) the samples collected by the researchers *before* the playful interaction task (2 samples: one with father and one with mother); d) the evoked samples collected by the researchers *after* the playful interaction task (2 samples: one with father and one with mother ).

In Figure 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 Figure 1.

## 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. 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, Hachul, & Roßbach, 2016). 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). The child and parent are presented with three bags with toys, and asked to spend ~15 minutes playing. More specifically, they were asked to play with three different sets of toys, each placed inside a separate bag labelled "1," "2," or "3." The parent was told that she/he had 15 minutes to play with the three toys and that the only restriction was that they should play with the toys in numerical order, beginning with bag #1 and ending with bag #3. Sessions were randomly counterbalanced between mother and father.

The micro-analysis of the video recordings of these sessions data will be disseminated separately.

## Assessment of behaviour problems

Children's behavioural problems were assessed independently by their preschool teacher using the Caregiver-Teacher Report Form (C-TRF) (Achenbach & Rescorla, 2000). The C-TRF is a standardized measure which provides ratings for 120 problem behaviour items. This is one of the most commonly 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 nationalities. Test-retest reliability for all of the scales used in this study was > .75 (Achenbach & Rescorla, 2000). We used the validated Portuguese version and the Portuguese population norms (Rescorla, Achenbach et al., 2012). To be able 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 also 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, Externalization and Total problem 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,

internalization and total problem scales have been used in more than 75.000 articles in developmental psychology research (Achenbach et al., 2016).

# Declarations


### *Acknowledgements*

This research has been supported in part by BIAL and FCT Grant UID/PSI/04810/2013. The authors thank all the members of the research team for their assistance and support. 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.


### *Competing interests*

The authors declare no competing interests.

# References


Achenbach, T. M., & Rescorla, L. A., 2000. Manual for the ASEBA Preschool Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

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. Journal of the American Academy of Child & Adolescent Psychiatry 55, 8: 647-656. Doi:https://doi.org/10.1016/j.jaac.2016.05.012

Bachner-Melman, R., Ebstein, R.P., 2014. The role of oxytocin and vasopressin in emotional and social behaviors. Handbook of clinical neurology.124:53-68.

Brambilla, D. J. McKinlay, S. M. McKinlay, J. B. Weiss, S. R., Johannes, C. B. Crawford, S. L. Longcope, C., 1996. Does collecting repeated blood samples from each subject improve the precision of estimated steroid hormone levels? J Clin Epidemiol 49,3, 345-350. https://doi.org/10.1016/0895-4356(95)00569-2

Baskaran C., Plessow, F., Silva, L., Asanza, E., Marengi, D., Eddy, K.T, Sluss, P.M, Johnson, M.L, Misra, M., Lawson, E.A., 2017. Oxytocin secretion is pulsatile in men and is related to social-emotional functioning. Psychoneuroendocrinology. 85:28-34 https://doi.org/10.1016/j.psyneuen.2017.07.486

Caldwell, HK., 2017. Oxytocin and Vasopressin: Powerful Regulators of Social Behavior. Neuroscientist. 23(5):517-528. https://doi.org/10.1177/1073858417708284

Carson, DS , Berquist, SW , Trujillo, TH , Garner, JP, Hannah, SL, Hyde, SA, Sumiyoshi, RD, Jackson, LP, Moss, JK, Strehlow, MC, Cheshier, SH, Partap, S, Hardan, AY and Parker, KJ (2014) Cerebrospinal fluid and plasma oxytocin concentrations are positively correlated and negatively predict anxiety in children *Molecular Psychiatry* (2014), 1–6. doi: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." *Harvard Review of Psychiatry*, *21*(5), 219–247. http://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), 474-9.

Daughters K, Manstead AS, Hubble K, Rees A, Thapar A, van Goozen SH (2015) Salivary oxytocin concentrations in males following intranasal administration of oxytocin: a double-blind, cross-over study. PLoS One 10:e0145104

de Jong, TR, Menon, R., Bludau, A., Grund, T., Biermeier, V., Klampfl, SM., Jurek, B,. Bosch, OJ, Hellhammer, J. and Neumann, ID (2015) Salivary oxytocin concentrations in response to running, sexualself-stimulation, breastfeeding and the TSST: The RegensburgOxytocin Challenge (ROC) study. *Psychoneuroendocrinology* 62 (2015) 381–388 . dx.doi.org/10.1016/j.psyneuen.2015.08.027

Demirci, E., Ozmen, S., Oztop, D.B., 2016. 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 DB., 2016 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):681-8. doi.org/10.1097/FBP.0000000000000234

Engelmann, M.,. Ebner, K ,. Landgraf, R ,. Holsboer, F and Wotjak, C. T. (1999) Emotional Stress Triggers Intrahypothalamic But Not Peripheral Release of Oxytocin in Male Rats *Journal of Neuroendocrinology, 1999, Vol.*

Feldman, R., Weller, A., Zagoory-Sharon, O., Levine, A., 2007. Evidence for a neuroendocrinological foundation of human affiliation: plasma oxytocin levels across pregnancy and the postpartum period predict mother-infant bonding. Psychol. Sci. 18, 965–970.

Feldman R, Gordon I, Zagoory-Sharon O. The cross-generation transmission of oxytocin in humans. Horm Behav. 2010;58(4):669-76.

Feldman, R., Gordon, I., Schneiderman, I., Weisman, O., & Zagoory-Sharon, O. (2010). Natural variations in maternal and paternal care are associated with systematic changes in oxytocin following parent-infant contact. Psychoneuroendocrinology, 35(8), 1133–1141. doi:10.1016/j.psyneuen.2010.01.013

Feldman. R., Gordon. I., Influs, M., Gutbir, T., Ebstein, R.P., 2013. Parental oxytocin and early caregiving jointly shape children's oxytocin response and social reciprocity. Neuropsychopharmacology, 38(7):1154-62. doi: 10.1038/npp.2013.22

Gordon, I., Zagoory-Sharon, O., Leckman, J. F., & Feldman, R. (2010). Oxytocin and the Development of Parenting in Humans. Biological Psychiatry, 68(4), 377–382. http://doi.org/10.1016/j.biopsych.2010.02.005

Gröschl, M., 2008. Current Status of Salivary Hormone Analysis. Ann Biol Clin (Paris) . 67. 493-504. https://doi.org/10.1684/abc.2009.0357.

Horvat-Gordon M, Granger DA, Schwartz EB, Nelson VJ, Kivlighan KT (2005) Oxytocin is not a valid biomarker when measured in saliva by immunoassay. Phys Behav 84:445–448

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 Biomarkers Prev. 2(5):411-4.

Lancefield, K. S., Raudino, A., Downs, J. M., & Laurens, K. R., 2016. Trajectories of childhood internalizing and externalizing psychopathology and psychotic-like experiences in adolescence: A prospective population-based cohort study. Dev. Psychopathol. , 28(2), 527–536. http://doi.org/10.1017/S0954579415001108

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 clinicallyanxious 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, PM, Lukas, KE, Stoinski, TS, Willis, MA, Schook. MW (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 (2018) 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. The Example of the National Educational Panel Study.* Springer

MacLean EL, Gesquiere LR, Gee N, Levy K, Martin WL, Carter CS (2018) Validation of salivary oxytocin and vasopressin as biomarkers in domestic dogs. J NeuroSci Meth 293:67–76

J. Martin S. M. Kagerbauer J. Gempt A. Podtschaske A. Hapfelmeier G. Schneider (2018) Oxytocin levels in saliva correlate better than plasma levels with concentrations in the cerebrospinal fluid of patients in neurocritical care *Journal of neuroendocrinology* 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? *BioRXiv*;doi:https://doi.org/10.1101/2020.07.14.202622

Miettunen, J., Murray, G. K., Jones, P. B., Mäki, P., 2014. Longitudinal associations between childhood and adulthood externalizing and internalizing psychopathology and adolescent substance use. Psychological Medicine, 44, 8 , 1727-1738.. doi.org/10.1017/S0033291713002328

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, placebocontrolled, crossover study. Am. J. Med. Genet. A 173 (5), 1243–1250. https://doi.org/10.1002/ajmg.a.38160

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.2017.12.022Get rights and content

Grazia, R. , Matteo R., Yannis, P., James, G., Alberto, S., Marco, C., 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 Research* 241, 207–20.

Prevost, M., Zelkowitz, P., Tulandi, T., Hayton, B., Feeley, N., Carter, C. S., Joseph, L., Pournajafi-Nazarloo, H., Yong Ping, E., Abenhaim, H., Gold, I., 2014. Oxytocin in pregnancy and the postpartum: relations to labor and its management. Front Public Health , 2, 1. https://doi.org/10.3389/fpubh.2014.00001

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):315-21. 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. http://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. 2013 Nov-Dec;25(6):719-24. https://doi.org/10.1002/ajhb.22438 .

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 doi:10.1371/journal.pone.0160767 .

Salzberg, A.D., Swedo, S.E., 1992. Oxytocin and vasopressin in obsessive-compulsive disorder. Am J Psychiatry ;149(5):713-4. 120.

Stein, D.J., 2009. Oxytocin and vasopressin: social neuropeptides. CNS Spectr. 14(11):602-6. 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. and Veríssimo, 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. and Petersson, M. (2015) Self-soothing behaviors with particular reference to oxytocin release induced by non-noxious sensory stimulation Frontiers in Psychology. doi: 10.3389/fpsyg.2014.01529

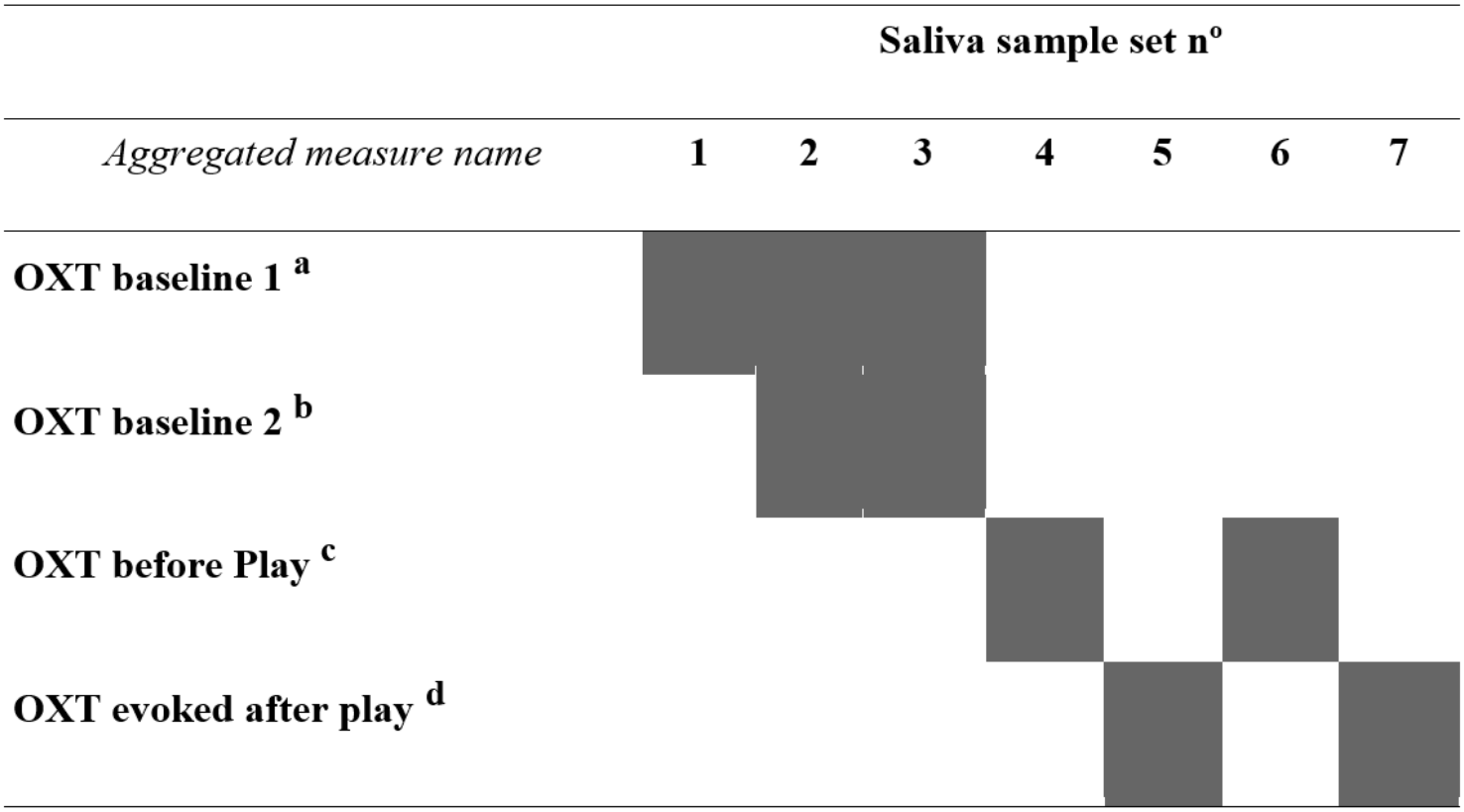
Valstad, M., Alvares, G.A., Egknud, M., Matziorinis, A.M., Andreassen, O.A., Westlye, L.T., et al., 2017. The correlation between central and peripheral oxytocin concentrations: a systematic review and meta-analysis. Neurosci. Biobehav. Rev. 78, 117–124. https://doi.org/10.1016/j.neubiorev.2017.04.017

# Figures

| | Saliva sample set nº | | | | | | |
|---|---|---|---|---|---|---|---|
| *Sampling conditions* | **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 |

**Figure 1**

Saliva samples' collection context




## Figure 2

Composition of the aggregated OXT measures. Legend: a- collected by parents on two week days + one weekend day; b-collected by parents on two week days; c-collected by researchers before the playful sessions; d: collected by researchers after the playful sessions