KU LEUVEN

faculteit psychologie en
pedagogische wetenschappen

Laboratorium voor Biologische Psychologie

Profiling early-life social development in laboratory mice

Carmen Winters

Proefschrift aangeboden tot het verkrijgen van de
graad van Doctor in de Psychologie

Promotor: Prof. Dr. Rudi D'Hooge
Co-promotor: Prof. Dr. Guy Bosmans

2023

Contents

Started making it, had a breakdown, bon appétit.

- James Acaster

List of abbreviations

AMBER	Automated Maternal Behavior during Early Life in Rodents
ASD	Autism spectrum disorder
BAMBI	Bidirectional Automated Mother-pup Behavioral Interaction test
BLA-BMA	Basolateral/basomedial amygdala
COLAB	Google Colaboratory
DAS	Deep Audio Segmenter
DC	Direct current
DLC	DeepLabCut
DMSO	Dimethylsulfoxide
DSM-5	Diagnostic and statistical manual of mental disorders, 5th edition
EQ	Equalization
GD	Gestational day
GLM	General linear model
HPA	Hypothalamic-pituitary-adrenal
LC	Locus coeruleus
LMT	Live Mouse Tracker
MC3/4R	Melanocortin3 and -4 receptor
MC4R	Melanocortin4 receptor
MPOA	Medial preoptic area
Nacc	Nucleus accumbens
NE	Norepinephrine
OT	Oxytocin
OTR	Oxytocin receptor
OB	Olfactory bulb
PBS	Phosphate buffered saline
PFC	Prefrontal cortex
PND	Postnatal day
px	pixels
PRT	Pup retrieval test
PVN	Paraventricular nucleus
SD	Standard deviation
SimBA	Simple Behavioral Analysis
THIQ	N-[(1R)-1-[(4-Chlorophenyl)methyl]-2-[4-cyclohexyl-4-(1H-1,2,4-trazol-1-ylmethyl)-1-piperidinyl]-2-oxoethyl]-1,2,3,4-tetrahydro-3-isoquinolinecarboxamide
USV	Ultrasonic vocalization
vBST	Ventral bed nucleus of the stria terminalis
VP	Ventral pallidum
VPA	Valproic acid
VTA	Ventrl tegmental area

1 General introduction

Early-life social interactions are vital for the survival and development of mammals. The early developmental period is characterized by rapid brain growth and plasticity which promotes the formation of neural circuits and mechanisms underlying social, cognitive and emotional processes. During this period of high susceptibility, experiences and interactions with the environment can have a profound impact and as such set the stage for future behavioral patterns. Research has consistently shown that positive early-life experiences promote healthy social, emotional and cognitive development, while adverse experiences can have long-lasting negative effects.

1.1 Early-life environment

As mammals, social life starts at birth and being social is vital for infant survival. Parental care is the main strategy to promote offspring survival by providing food, warmth and protection (1-3). Although the shape, form and provider (mother, father or alloparent) of parental caretaking behaviors varies across mammals, the mother's capacity to lactate and high maternal motivation results in her playing a central role (1-3). More specifically, nourishment is embedded in the maternal body and thus for offspring survival a close proximity should be maintained for a substantial period of early development (4). Neonates possess a species-specific set of socially adaptive tools, such as active orientation towards the nipples or crying, which are intended to elicit maternal care and proximity (1-2). These innate and reflexive behaviors emerge almost simultaneously with neuroendocrine changes in the maternal brain that make her particularly receptive to infant signals (4-5). As such, mother and offspring continuously interact by performing stereotyped behavioral sequences towards each other (6-10). The establishment of such a dyadic neurobehavioral system is regulated by neuroendocrine mechanisms that reorganize the brain of both the mother and her offspring (9-11).

The origins of this line of inquiry can be traced back to the work of Nobel Prize-winning ethologist Konrad Lorenz who demonstrated that early-life social experiences with the caretaker shape infant biology and behavior (9, 12). Lorenz proposed that young precocial birds require exposure to conspecifics within hours of hatching to develop adaptive filial attachments and appropriate later mate prefences (13). Building on this work, multiple studies described almost simultaneously the disruptive effects of absent maternal sensory stimulation on infants' physical, social, emotional and cognitive development in rodents, nonhuman monkeys and human children (14-17). Conversely, studies showed the beneficial effects of stimulating or nurturing environments by promoting HPA axis development and long-term resilience (18-19). Seminal work of Bowlby and Ainsworth has provided the field with a framework for understanding the significance of secure attachment relationships in infancy and their impact on later emotional wellbeing and interpersonal relationships (20-21).

In subsequent decades, research in the field of early-life social experiences has expanded and diversified, incorporating various methodologies and disciplines. Longitudinal studies in humans have tracked individuals from infancy into adulthood to investigate the links between early social experiences and outcomes such as physical and mental health, academic achievement and social functioning (22-28). Animal studies, on the other hand, have provided valuable insights into the cellular, neural and endocrine mechanisms underlying the effects of early social experiences on brain development and behavior (13, 29-36). Lesion and knockout studies have provided compelling evidence supporting the existence of a well-conserved mother-infant interaction circuits across mammalian species (3, 37). Advances in neuroscience and genetics further opened up new avenues of research for researchers to examine the complex interplay between genetic factors, early-life experiences and adult outcomes (38-39). Neuroimaging studies have, for example, contributed to a better understanding on how early-life experiences influence neural connectivity and organization of brain regions involved in social, emotional and cognitive processing (40-41).

The neurodevelopmental hypothesis of psychopathology posits that disturbances in early brain development can contribute to the emergence of various psychiatric disorders later in life (13, 29). Both genetic and environmental factors can affect brain maturation, hereby leading to the activation of pathologic neural circuits later in life which may increase the likelihood of clinical symptoms (13, 29, 42). Genetic and environmental factors, as well as gene-environment interactions can influence processes involved in brain development including neurogenesis, neuronal migration, synaptogenesis and myelination. Perturbations in these processes can result in atypical brain connectivity, neurotransmitter imbalances, and altered neural circuitries.

Social challenges early in life are a common characteristic or predictor for psychopathologies including autism spectrum disorders (ASDs) and schizophrenia (43). Epidemiological studies conducted in individuals with autism spectrum disorders or schizophrenia suggest that early-life experiences such as abuse, neglect, maternal stress or obstetric complications are associated disruptions in the normal trajectory of neurodevelopment (13). Further evidence for this theory is found in the fact that several genes linked to susceptibility for psychopathology such as neuroligins or SHANK3 play crucial roles in neurodevelopment (13). Additionally, evidence of linkage and association studies suggest that exposure to suboptimal environmental factors such as stress, in combination with individual genetic susceptibility can increase the likelihood of developing psychopathology throughout life.

Evidence has shown that certain genetic variations can increase an infant's susceptibility to social difficulties and impairments in early-life social interactions. For example, variations in genes associated with social cognition, empathy and attachment, such as those involved in the oxytocin and dopamine

signaling, have been associated with altered responses to social stimuli and difficulties in communication (44-46). Especially in mother-infant interactions such genetic predispositions can play a significant role on further social development as various factors come into play. Dynamic interactions between mother and offspring involves that both parties shape each other (47). A suboptimal genetic infant predisposition may prevent a mother-directed approach in an infant and elicit less optimal interactions. As such, the infant might exacerbate the phenotypic expression of atypical sociocommunicative behaviors in risk carriers, ultimately increasing the likelihood to develop a psychopathology. Conversely, a suboptimal genetic predisposition of the mother can affect the sensitivity, motivation, quality and quantity and of maternal caregiving behaviors. It also should be noted that a mother passes half of her genes to her offspring, impacting their genetic predisposition for social behavior. Therefore, it is difficult to disentangle genetic and environmental factors within the complex interplay between mother and infant during early life social development.

1.2 Dam-pup interactions in laboratory rodents

The connecting infant

Already in utero infant behavioral development is directed towards its first extrauterine encounter with its mother. Fetuses engage in a set of spontaneous motor movements preparing them to approach and orient towards the mother and more specifically mammary region (48). Further, development of infant's sensory systems during the third trimester of pregnancy underlies prenatal learning of maternal sensory stimuli such as maternal odor and sounds (7, 17, 48). Once outside the uterus, a rapid learning process is necessary to fine-tune extrauterine recognition and preference (48, 49). The infant brain is equipped with neurobiological mechanisms promoting preference learning and blocking aversion learning (12, 48-49). Preference learning in rodents involves hyper-functioning of the infants' locus coeruleus (LC) releasing high levels of norepinephrine (NE) in the olfactory bulb (OB, 49). This mechanism induces neural plasticity in the olfactory bulb by preventing habituation to repeated olfactory stimulation. NE is both necessary and sufficient for both learning the dams' natural odor and for the learning-induced behavioral and neural changes in pups although the piriform cortex has been suggested assigning positive value to a learned odor (49). Conversely, attenuated amygdala plasticity in infant rats reduces the ability to acquire learning of aversion and fear (49).

The signaling infant

Rodent pups are underdeveloped at the time of birth and require extensive maternal care to mature over the course of two weeks. However, they are equipped with a limited albeit effective physical and behavioral repertoire to establish maternal proximity and elicit maternal caretaking (7, 17).

Tactile and olfactory infant cues are primary communicative forms requiring physical proximity (3). Infants are highly attracted towards maternal odors and specifically odors from the mammary area

and these odors actively elicit infant approach responses (7, 49). Ventral stimulation of the pups in form of mouthing and nuzzling incite the dam to hover over her pups and subsequently to nursing behaviors (3). Further, suckling-induced nipple stimulation is an important tactile sensation for the dam to establish mother-infant interaction. Additionally, elimination of tactile inputs to the perioral region of dams disrupts both the onset and maintenance of maternal behavior (3). Perioral tactile stimulation is involved in maternal retrieval, licking and even nursing (3).

In rodents, infant olfactory cues are important to stimulate neural and behavioral responses in the dam. In nulliparous female rats, odors associated with afterbirth and pups is aversive and prevents them from being maternal and inducing anosomia eliminates this aversiveness and results in rapid onset of maternal behavior (3). This shift naturally occurs when females undergo the hormonal changes associated with pregnancy, birth and lactation. Further, dodecyl propionate is an olfactory substance secreted by rat pups' anogenital area and is highly attractive to the dam (50). When anogenital smears are applied to the heads of rat pups and their anogenital areas are cleaned, dams exclusively lick the pups' heads and surgical removal of the preputial pup glands completely disorganizes maternal licking patterns (50).

Further, auditory infant cues are more distal communicative forms and are among the most effective infant signals promoting proximity as they strongly elicit maternal behaviors (3). When mouse pups are struggling in the nest, they emit low-frequency wriggling and these calls elicit maternal licking, adjustments in suckling position and nest building. Further, pups emit ultrasonic vocalizations (USVs) when they are separated from the dam and littermates. Pup USVs are in the 30-90 kHz range and emerge within hours after birth. The number of USVs increases during the first postnatal week and rapidly decreases until weaning (51-52). These isolation-induced calls have a communicative function as they strongly elicit maternal search and retrieval behavior in dams (3, 7, 53-55). The frequency of pup vocalizations is affected by various environmental factors including situation of maternal isolation, temperature, tactile stimulation and exposure to male odor (56). Evidence suggests that infant USV emission is a social behavior conveying information about endogenous and exogenous sensory experiences disrupting the infants' homeostatic balance of arousal (57-60). Indeed, infant distress due to maternal separation is associated with an increased emission of distress calls and reduced levels of oxytocin and opioids. Conversely, subsequent reunion with the mother again decreases the emission of infant distress vocalizations and increases oxytocin and opioid levels (61-62). As such, infant USVs are extensively studied as an indicator of early communicative capability and numerous studies have reported changes in ultrasonic communication in mouse models of neurodevelopmental disorders and in particular autism spectrum disorder. In these rodent models, alterations in USV quantity and quality, such as duration, peak amplitude and frequence, have been reported (63-66).

The receiving mother

To establish effective caretaking behavior, the dam should integrate various sensory pup stimuli and recognize it as a conspecific young in need. Parturient physiological events promote a reorganization of the maternal brain in the processing of infant stimuli. This critical shift allows infant stimuli to gain access to attraction mechanism and depresses their access to avoidance mechanisms. In rats and mice, this recognition process of infant stimuli is nonselective meaning that generic infant stimuli, rather than stimuli of a specific infant, can elicit maternal care. The presence of the newborn including their odor, taste, touch and sound are crucial for the maintenance of maternal behavior in rodent dams. If parturient dams are prohibited to interact with their neonate pups, their maternal responsiveness declines over the first postpartum week (7). Indeed, infant suckling is known to stimulate both peripheral and central oxytocin release, as well as central dopamine release in the dam which is believed to positively influence brain regions involved in processing infant cues (7, 49, 67).

However, maternal behavior in rats and mice is not regulated by one single sensory input. Previous research on rats has demonstrated that blind, deaf, anosmic, or anaptic lactating females still take care for their pups in a manner comparable to control dams (68-69). The combined elimination of vision, olfaction and/or tactile system did not completely abolish maternal behavior, but did result in more severe defects. This indicates that maternal behavior is established under multisensory control although the anatomical and neural mechanisms underlying the multisensory integration and the importance of different sensory modalities is still poorly understood (7, 69). It also should be noted that the role of these pup cues in maternal behaviors can differ between rats and mice (7). For example, olfactory stimuli of pups appear to play a more crucial role in induction and maintenance of maternal behavior in mice than in rats. Also, maternal experience plays an important role in mice and experienced dams are less dependent on olfactory cues compared to primiparous dams (7).

The motivated mother

Pup recognition through sensory cues does not necessarily initiate parental care in adult mice (69). Behaviorally, effective maternal caregiving requires maternal attraction or at least maternal tolerance of her offspring to invest substantially in offspring wellbeing (3). Generally, naïve females do not naturally exhibit an attraction towards infants or nurture them. Naïve females that have no prior experience with pups even tend to ignore, avoid or even attack infants. However, a notable contrast occurs when females reach the late stages of pregnancy or give birth, as they develop an attraction towards pup stimuli and engage in maternal behaviors. This transition into motherhood involves a reorganization of sensory, motor and integrative systems primed by hormones present during late pregnancy, parturition and lactation (29). This reorganization of the maternal brain involves a suppression of the central avoidance circuit and activation of the central approach circuit (29, 37, 70).

Evidence from preference tasks even suggest that early postpartum rat dams show a higher preference for pup-associated stimuli over cocaine-associated stimuli (71).

In both these processes the medial preoptic area (MPOA) has been identified as a critical brain region. Rodent studies indicate that infant cues stimulate the MPOA and ventral bed nucleus of the stria terminalis (vBST) to activate ventral tegmental area (VTA) dopaminergic (DA) signaling in the nucleus accumbens (NAcc). In the NAcc, DA signaling is associated with motivation and approach to reward-related stimuli (7, 29, 70). Moreover, GABAergic inhibition from the NAcc to the ventral pallidum (VP) further establishes a tendency to approach by inhibition of avoidance and defensive behaviors. The basolateral/basomedial amygdala (BLA-BMA) and the prefrontal cortex (PFC) provide the NAcc and the VP simultaneously with the valence of infant stimuli (7, 29, 70). All together, this induces a state of maximal responsiveness to infant sensory stimulation which is maintained for the duration of lactation and in some species even beyond (3).

Behaviorally, various behavioral test assays aim to precisely measure specific behavioral components of maternal motivation. The Pup Retrieval Test (PRT) is the most widely used behavioral assay to quantify maternal responsivity and motivation in rodent research (72). Fundamentally, one or more pups are gently taken from the nest and placed outside the nest. The latency to sniff and retrieve each pup into the nest is recorded. The PRT has demonstrated its relevance in understanding the role of neurotransmitters, neuropeptides, neural circuits and genetic pathways involved in maternal behaviors. For example, the use of the PRT in various experimental studies provided evidence which supports the hypothesis that maternal responsiveness is fostered by dual outputs from the MPOA (i.e. depression of the aversion system and stimulation of the approach system, 69).

The acting mother

The primary components of maternal behaviors in rats and mice are similar and include nestbuilding, pup retrieval, nursing, licking and grooming, and defense of the young (73). Knockout and intervention studies have demonstrated that each of these components has its own regulatory mechanisms (73). Maternal behavior starts during late pregnancy with the preparation of a brood nest. Whereas nonpregnant rodents produce a flat nest, late pregnant females build a higher and complexer brood nest (69). Up to 2-3 weeks of age, mouse and rat pups have poor thermoregulatory abilities and the construction of a brood nest is crucial for successful rearing (56). Mothers provided with limited nesting and bedding materials exhibit reduced and fragmented maternal care and in some circumstances even harmful behaviors (69). Further, Noirot (74) reported that exposing naïve female mice to pup USVs induced nestbuilding behavior.

Nursing is the behavior through which te mother provides milk to her young. During this behavior, the

dam assumes a relatively calm and immobile state while exposing her nipples through the pups. During the first 3 weeks, nursing accounts for nearly 92% of maternal behavior and initially mothers spend almost all her time curled around the pups in the nest (56). Naturally, nursing activity drops around the age of 23-25 days in undisturbed mice under usual laboratory conditions (56). The time spent nursing appears to be affected by factors such as litter size, litter composition and maternal experience (56). Evidence has shown that pup wriggling calls are important pup stimulus to elicit nursing behavior and offspring to bilaterally deafened dams showed impaired weight development (56). In rats, various nursing postures are described and arched-back nursing is most commonly observed during the first week of lactation, whereas nursing in a passive position is seen mainly after the second week of lactation (69). It has been reported that suckling stimulation from pups promotes maternal nursing behavior (75-76).

Licking and grooming are another form of early-life tactile contact and is a primary component of maternal caregiving behaviors. Anogenital licking, however, is a fundamental pattern of maternal behavior as it contributes to sexual differentiation and stimulates reflexive defecation and urination (77). Research examining the longterm impact of maternal care has indicated the quantity of licking and grooming as a critical feature of the mother-infant interaction that shapes emotionality and HPA response to stress.

When a pup is separated from the nest or a mother relocates her nest, pup retrieval takes place: the dam orients and moves towards the pup, frequently sniffing it before gently picking it up with the incisors and carrying it to the nest site where she deposits it (73). Several studies provided evidence that pup USVs are potent to elicit maternal orienting and search behavior by playing pup USVs from ultrasonic speakers (78-81). However, dam perception of infant USVs requires a specific acoustic profile to induce subsequent approach behavior (79, 82). Pup retrieval is considered a robust maternal behavior although it can be significantly impacted by stress and other environmental factors (72, 69).

The receiving infant

Maternal caregiving behaviors provide essential somatosensory stimuli for the infants' brain development. A wellknown example are a series of studies showing that maternal tactile stimulation affects later life stress responses. Maternal licking and grooming are crucial as a series of rat studies have shown licking and grooming is essential for the regulation of the infant hypothalamic-pituitary-adrenal (HPA) axis whereas maternal absence leads to long-term disruptions in HPA responsiveness (19, 30, 33, 84-85). Moreover, a study showed that the quality of early-life social interactions such as low or high levels of maternal licking and grooming in rats impact various neurobehavioral infant systems through physiological, epigenetic and neuroendocrine mechanisms (86).

1.3 Assessing animal behavior

The study of animal behavior has been limited by researchers' ability to identify, record, and interpret ethologically significant behavioral changes in real-time. Most popular laboratory tests used to assess maternal or infant sociocommunicative abilities require ethological assessments and involve human labeling (69, 87). Maternal behavior, for example, is often assessed in the home cage under undisturbed conditions and includes observation of a limited amount of well-characterized behaviors such as retrieval, nursing, pup licking or grooming, dam self-care. Manual scoring relies on subjective interpretation and has several drawbacks (88-89). First, subtle individual differences in scoring can affect reliability and comparability of results. Moreover, manual scoring is labor-intensive whereas generally only a restricted set of predefined behaviors is scored. As maternal behavior involves dynamic phenotypes, manual scoring may ignore the rich behavioral repertoire displayed or overlook subtle nuances or variations in behavior (90).

The recent emergence of machine learning tools has revolutionized the field of animal behavior research by addressing these limitations. Machine learning techniques provide researchers with powerful tools for automated behavior recognition, offering increased accuracy, consistency, and real-time analysis (88-89). They also enable noninvasive monitoring and facilitate the integration of data from multiple sources and modalities, such as video and audio recordings. This integration allows for a more comprehensive understanding of animal behavior, its underlying mechanisms, and the complex relationships between different behavioral parameters. The advancements and accessibility of machine learning techniques have proven transformative, mitigating the disadvantages associated with manual scoring. These tools can achieve at least humanlike accuracy and, in many cases, surpass manual scoring in terms of speed and efficiency (89, 91). Researchers can now analyze and interpret behavioral data more effectively, leading to faster and data-driven discoveries in the field of animal behavior.

Another important benefit of deep learning tools is its use in observations conducted in unstructured environments. Although accurate animal tracking tools have been used for decades, they generally require highly constrained experimental testing setups (89). Deep learning focuses on training artificial neural networks with multiple layers to learn and extract meaningful patterns from complex data. These techniques can employ data augmentation strategies to generate synthetic variations of training data to increase robustness of models to handle diverse conditions, viewpoints and lighting variations encountered in unstructured environments (92). As such, researchers are able to study dam and pups in a more naturalistic and less stressful environment. Additionally, such homecage testing reduces the need for specialized experimental setups and dedicated observation spaces. Especially during the early-life period, reducing handling stress and maintaining a naturalistic environment as much as

possible is importance since dam and pup behavior affect one another (93-94).

Deep learning tools are increasingly becoming available and utilized in the field of animal ethology. However, the development of specialized deep learning models tailored to analyze and interpret maternal or pup behaviors is limited (95). Due to the subtle and complex nature of dam-offspring behaviors, deep learning methodologies provide a powerful tool to study this field and contribute significantly to a deeper understanding of the intricate mechanism underlying maternal care and its implications for offspring development.

1.4 References

1. Numan, M. (2020). The parental brain: mechanisms, development, and evolution. Oxford University Press.
2. Bosch, O. J., & Young, L. J. (2017). Oxytocin and Social Relationships: From Attachment to Bond Disruption. *Behavioral Pharmacology of Neuropeptides: Oxytocin*, 97–117. https://doi.org/10.1007/7854_2017_10
3. Lonstein, J. S., Lévy, F., & Fleming, A. S. (2015). Common and divergent psychobiological mechanisms underlying maternal behaviors in non-human and human mammals. *Hormones and Behavior*, *73*, 156–185. https://doi.org/10.1016/j.yhbeh.2015.06.011
4. Wilson, E. O. (2000). *Sociobiology: the new synthesis* (25th anniversary ed.). Belknap Press.
5. Feldman, R. (2016). The Neurobiology of Human Attachments. *Trends in Cognitive Sciences*, *21*(2), 80–99. https://doi.org/10.1016/j.tics.2016.11.007
6. Johnson, K. (2013). Maternal-infant bonding: a review of literature. *International Journal of Childbirth Education*, *28*(3).
7. Nagasawa, M., Okabe, S., Mogi, K., & Kikusui, T. (2012). Oxytocin and mutual communication in mother-infant bonding. *Frontiers in Human Neuroscience*, *6*, 31–31. https://doi.org/10.3389/fnhum.2012.00031
8. Broad, K., Curley, J., & Keverne, E. (2006). Mother–infant bonding and the evolution of mammalian social relationships. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476), 2199–2214. https://doi.org/10.1098/rstb.2006.1940
9. Feldman, R. (2012). Parenting Behavior as the Environment Where Children Grow. In *The Cambridge Handbook of Environment in Human Development* (pp. 535–567). Cambridge University Press. https://doi.org/10.1017/CBO9781139016827.031
10. Jansen, J., de Weerth, C., & Riksen-Walraven, J. M. (2008). Breastfeeding and the mother–infant relationship—a review. *Developmental review*, *28*(4), 503-521.
11. Nagasawa, M., & Kikusui, T. (2021). Neuroendocrine mechanisms of social bonds and separation stress in rodents, dogs, and other species. In *Neuroscience of Social Stress* (pp. 3-22). Cham: Springer International Publishing.
12. Sullivan, R., Perry, R., Sloan, A., Kleinhaus, K., & Burtchen, N. (2011). Infant Bonding and Attachment to the Caregiver: Insights from Basic and Clinical Science. *Clinics in Perinatology*, *38*(4), 643–655. https://doi.org/10.1016/j.clp.2011.08.011
13. Marco, E. M., Macrì, S., & Laviola, G. (2011). Critical age windows for neurodevelopmental psychiatric disorders: evidence from animal models. *Neurotoxicity research*, *19*, 286-307.
14. Harlow, H. F., & Zimmermann, R. R. (1959). Affectional response in the infant monkey: Orphaned baby monkeys develop a strong and persistent attachment to inanimate surrogate mothers. *Science*, *130*(3373), 421-432.
15. Spitz, R. A. (1951). The psychogenic diseases in infancy: An attempt at their etiologic classification. *The psychoanalytic study of the child*, *6*(1), 255-275.
16. Robertson, J., & Robertson, J. (1971). Young children in brief separation: A fresh look. *The psychoanalytic study of the child*, *26*(1), 264-315.
17. Fleming, A. S., O'Day, D. H., & Kraemer, G. W. (1999). Neurobiology of mother–infant interactions: experience and central nervous system plasticity across development and generations. *Neuroscience and Biobehavioral Reviews*, *23*(5), 673–685. https://doi.org/10.1016/S0149-7634(99)00011-1
18. Weininger, O., McClelland, W. J., & Arima, R. K. (1954). Gentling and weight gain in the albino rat. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *8*(3), 147.
19. Levine, S. (2001). Primary social relationships influence the development of the hypothalamic–pituitary–adrenal axis in the rat. *Physiology & Behavior*, *73*(3), 255–260. https://doi.org/10.1016/S0031-9384(01)00496-6
20. Ainsworth, M. D. S. (1978). The bowlby-ainsworth attachment theory. *Behavioral and brain sciences*, *1*(3), 436-438.
21. Bowlby, J. (1979). The bowlby-ainsworth attachment theory. *Behavioral and Brain Sciences*, *2*(4), 637-638.

22. Perez, C. M., & Widom, C. S. (1994). Childhood victimization and long-term intellectual and academic outcomes. *Child abuse & neglect*, *18*(8), 617-633.
23. Widom, C. S. (1999). Posttraumatic stress disorder in abused and neglected children grown up. *American journal of psychiatry*, *156*(8), 1223-1229.
24. Widom, C. S., DuMont, K., & Czaja, S. J. (2007). A prospective investigation of major depressive disorder and comorbidity in abused and neglected children grown up. *Archives of general psychiatry*, *64*(1), 49-56.
25. Johnson, J. G., Cohen, P., Dohrenwend, B. P., Link, B. G., & Brook, J. S. (1999). A longitudinal investigation of social causation and social selection processes involved in the association between socioeconomic status and psychiatric disorders. *Journal of abnormal psychology*, *108*(3), 490.
26. Johnson, J. G., Smailes, E. M., Cohen, P., Brown, J., & Bernstein, D. P. (2000). Associations between four types of childhood neglect and personality disorder symptoms during adolescence and early adulthood: Findings of a community-based longitudinal study. *Journal of personality disorders*, *14*(2), 171-187.
27. Danese, A., Moffitt, T. E., Harrington, H., Milne, B. J., Polanczyk, G., Pariante, C. M.,... & Caspi, A. (2009). Adverse childhood experiences and adult risk factors for age-related disease: depression, inflammation, and clustering of metabolic risk markers. *Archives of pediatrics & adolescent medicine*, *163*(12), 1135-1143.
28. Javakhishvili, M., & Widom, C. S. (2021). Childhood maltreatment, sleep disturbances, and anxiety and depression: a prospective longitudinal investigation. *Journal of applied developmental psychology*, *77*, 101351.
29. Rogers, F. D., & Bales, K. L. (2019). Mothers, Fathers, and Others: Neural Substrates of Parental Care. *Trends in Neurosciences (Regular Ed.)*, *42*(8), 552–562. https://doi.org/10.1016/j.tins.2019.05.008
30. Champagne, F. A., Francis, D. D., Mar, A., & Meaney, M. J. (2003). Variations in maternal care in the rat as a mediating influence for the effects of environment on development. Physiology & Behavior, 79(3), 359–371. https://doi.org/10.1016/S0031-9384(03)00149-5
31. Meaney, M. J. (2001). Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations. *Annual review of neuroscience*, *24*(1), 1161-1192.
32. Barr, C. S., Newman, T. K., Becker, M. L., Parker, C. C., Champoux, M., Lesch, K. P.,... & Higley, J. D. (2003). The utility of the non-human primate model for studying gene by environment interactions in behavioral research. *Genes, Brain and Behavior*, *2*(6), 336-340.
33. Caldji, C., Tannenbaum, B., Sharma, S., Francis, D., Plotsky, P. M., & Meaney, M. J. (1998). Maternal Care during Infancy Regulates the Development of Neural Systems Mediating the Expression of Fearfulness in the Rat. *Proceedings of the National Academy of Sciences - PNAS*, *95*(9), 5335–5340. https://doi.org/10.1073/pnas.95.9.5335
34. Liu, D., Diorio, J., Tannenbaum, B., Caldji, C., Francis, D., Freedman, A.,... & Meaney, M. J. (1997). Maternal care, hippocampal glucocorticoid receptors, and hypothalamic-pituitary-adrenal responses to stress. *Science*, *277*(5332), 1659-1662.
35. Diamond, M. C. (2001). Response of the brain to enrichment. *Anais da Academia Brasileira de Ciências*, *73*, 211-220.
36. Kuhn, C. M., & Schanberg, S. M. (1998). Responses to maternal separation: mechanisms and mediators. *International Journal of Developmental Neuroscience*, *16*(3-4), 261-270.
37. Stolzenberg, D. S., & Mayer, H. S. (2019). Experience-dependent mechanisms in the regulation of parental care. *Frontiers in neuroendocrinology*, *54*, 100745.
38. Champagne, F. A. (2010). Early adversity and developmental outcomes: Interaction between genetics, epigenetics, and social experiences across the life span. *Perspectives on Psychological Science*, *5*(5), 564-574.
39. Singh-Taylor, A., Molet, J., Jiang, S., Korosi, A., Bolton, J. L., Noam, Y.,... & Baram, T. Z. (2018). NRSF-dependent epigenetic mechanisms contribute to programming of stress-sensitive neurons by neonatal experience, promoting resilience. *Molecular psychiatry*, *23*(3), 648-657.
40. Luby, J. L., Tillman, R., & Barch, D. M. (2019). Association of timing of adverse childhood experiences and caregiver support with regionally specific brain development in adolescents. *JAMA network open*, *2*(9), e1911426-e1911426.
41. Herzberg, M. P., & Gunnar, M. R. (2020). Early life stress and brain function: Activity and connectivity

associated with processing emotion and reward. *NeuroImage*, *209*, 116493.
42. Luby, J. L., Baram, T. Z., Rogers, C. E., & Barch, D. M. (2020). Neurodevelopmental optimization after early-life adversity: cross-species studies to elucidate sensitive periods and brain mechanisms to inform early intervention. *Trends in neurosciences*, *43*(10), 744-751.
43. Mäntymaa, M., Puura, K., Luoma, I., Vihtonen, V., Salmelin, R. K., & Tamminen, T. (2009). Child's behaviour in mother-child interaction predicts later emotional and behavioural problems. *Infant and Child Development*, *18*(5), 455–467. https://doi.org/10.1002/icd.633
44. Skuse, D. H., & Gallagher, L. (2011). Genetic influences on social cognition. *Pediatric research*, *69*(8), 85-91.
45. Rodrigues, S. M., Saslow, L. R., Garcia, N., John, O. P., & Keltner, D. (2009). Oxytocin receptor genetic variation relates to empathy and stress reactivity in humans. *Proceedings of the National Academy of Sciences*, *106*(50), 21437-21441.
46. Harrison, A. J., Gamsiz, E. D., Berkowitz, I. C., Nagpal, S., & Jerskey, B. A. (2015). Genetic variation in the oxytocin receptor gene is associated with a social phenotype in autism spectrum disorders. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *168*(8), 720-729.
47. Mandy, W., & Lai, M.C. (2016). Annual Research Review: The role of the environment in the developmental psychopathology of autism spectrum condition. *Journal of Child Psychology and Psychiatry*, *57*(3), 271–292. https://doi.org/10.1111/jcpp.12501
48. Hofer, M. A. (2005). The psychobiology of early attachment. *Clinical Neuroscience Research*, *4*(5), 291–300. https://doi.org/10.1016/j.cnr.2005.03.007
49. Landers, M. S., & Sullivan, R. M. (2012). The Development and Neurobiology of Infant Attachment and Fear. *Developmental Neuroscience*, *34*(2-3), 101–114. https://doi.org/10.1159/000336732
50. Brouette-Lahlou, I., Amouroux, R., Chastrette, F., Cosnier, J., Stoffelsma, J., & Vernet-Maury, E. (1991). Dodecyl propionate, attractant from rat pup preputial gland: characterization and identification. *Journal of chemical ecology*, *17*, 1343-1354.
51. Hammock, E. A. (2018). Oxytocin and vasopressin systems in the development of social behavior. In *Routledge international handbook of social neuroendocrinology* (pp. 504-529). Routledge.
52. Premoli, M., Pietropaolo, S., Wöhr, M., Simola, N., & Bonini, S. A. (2023). Mouse and rat ultrasonic vocalizations in neuroscience and neuropharmacology: State of the art and future applications. *European Journal of Neuroscience*.
53. Gaub, S., & Ehret, G. (2005). Grouping in auditory temporal perception and vocal production is mutually adapted: the case of wriggling calls of mice. *Journal of Comparative Physiology A*, *191*, 1131-1135.
54. Dulac, C., O'Connell, L. A., & Wu, Z. (2014). Neural control of maternal and paternal behaviors. *Science*, *345*(6198), 765-770.
55. Valtcheva, S., Issa, H. A., Bair-Marshall, C. J., Martin, K. A., Jung, K., Zhang, Y., Buzsáki, G., Kwon, H., & Froemke, R. C. (2021). Neural circuitry for maternal oxytocin release induced by infant cries. *BioRxiv*, 2021-03.
56. Weber, E. M., & Olsson, I. A. S. (2008). Maternal behaviour in Mus musculus sp.: an ethological review. *Applied Animal Behaviour Science*, *114*(1-2), 1-22.
57. Zeskind, P. S., Mcmurray, M. S., Garber, K. A., Neuspiel, J. M., Cox, E., Grewen, K. M., Mayes, L. C., & Johns, J. M. (2011). Development of translational methods in spectral analysis of human infant crying and rat pup ultrasonic vocalizations for early neurobehavioral assessment. *Frontiers in Psychiatry*, *2*, 56–56. https://doi.org/10.3389/fpsyt.2011.00056
58. Blumstein, D. T. (2013). Evolution of Emotional Communication: From Sounds in Nonhuman Mammals to Speech and Music in Man (Review of *Evolution of Emotional Communication: From Sounds in Nonhuman Mammals to Speech and Music in Man*). *Quarterly Review of Biology*, *88*(4), 342–343. The University of Chicago Press. https://doi.org/10.1086/673795
59. Ehret, G. (2005). Infant rodent ultrasounds - A gate to the understanding of sound communication. *Behavior Genetics*, *35*(1), 19–29. https://doi.org/10.1007/s10519-004-0853-8
60. Ashbrook, D. G., Roy, S., Clifford, B. G., Riede, T., Scattoni, M. L., Heck, D. H., Lu, L., & Williams, R. W. (2018). Born to cry: A genetic dissection of infant vocalization. *Frontiers in Behavioral Neuroscience*, *12*, 250–250.

https://doi.org/10.3389/fnbeh.2018.00250
61. Lingle, S., Wyman, M. T., Kotrba, R., Teichroeb, L. J., & Romanow, C. A. (2012). What makes a cry a cry? A review of infant distress vocalizations. *Current Zoology*, *58*(5), 698–726. https://doi.org/10.1093/czoolo/58.5.698
62. Insel, T. R., & Winslow, J. T. (1991). Central administration of oxytocin modulates the infant rats response to social isolation. *European Journal of Pharmacology*, *203*(1), 149–152. https://doi.org/10.1016/0014-2999(91)90806-2
63. Scattoni, M. L., Gandhy, S. U., Ricceri, L., & Crawley, J. N. (2008). Unusual repertoire of vocalizations in the BTBR T+tf/J mouse model of autism. *PloS One*, *3*(8), e3067–e3067. https://doi.org/10.1371/journal.pone.0003067
64. Peñagarikano, O., Lázaro, M. T., Lu, X. H., Gordon, A., Dong, H., Lam, H. A.,... & Geschwind, D. H. (2015). Exogenous and evoked oxytocin restores social behavior in the Cntnap2 mouse model of autism. *Science translational medicine*, *7*(271), 271ra8-271ra8.
65. Wöhr, M., Silverman, J. L., Scattoni, M. L., Turner, S. M., Harris, M. J., Saxena, R., & Crawley, J. N. (2013). Developmental delays and reduced pup ultrasonic vocalizations but normal sociability in mice lacking the postsynaptic cell adhesion protein neuroligin2. *Behavioural brain research*, *251*, 50-64.
66. Sungur, A. Ö., Schwarting, R. K. W., & Wöhr, M. (2016). Early communication deficits in the Shank1 knockout mouse model for autism spectrum disorder: Developmental aspects and effects of social context. *Autism Research*, *9*(6), 696–709. https://doi.org/10.1002/aur.1564
67. Strathearn, L. (2011). Maternal Neglect: Oxytocin, Dopamine and the Neurobiology of Attachment: Maternal Neglect: Oxytocin, Dopamine and Attachment. *Journal of Neuroendocrinology*, *23*(11), 1054–1065. https://doi.org/10.1111/j.1365-2826.2011.02228.x
68. Beach, F. A., & Jaynes, J. (1956). Studies of Maternal Retrieving in Rats. Iii. Sensory cues involved in the lactating female's response to her young 1. *Behaviour*, *10*(1), 104-124.
69. Kuroda, K. O., Tachikawa, K., Yoshida, S., Tsuneoka, Y., & Numan, M. (2011). Neuromolecular basis of parental behavior in laboratory mice and rats: With special emphasis on technical issues of using mouse genetics. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *35*(5), 1205–1231. https://doi.org/10.1016/j.pnpbp.2011.02.008
70. Feldman, R. (2015). The adaptive human parental brain: implications for children's social development. *Trends in Neurosciences (Regular Ed.)*, *38*(6), 387–399. https://doi.org/10.1016/j.tins.2015.04.004
71. Pereira, M., & Morrell, J. I. (2011). Functional mapping of the neural circuitry of rat maternal motivation: Effects of site-specific transient neural inactivation. *Journal of neuroendocrinology*, *23*(11), 1020-1035.
72. Hahn, M. E., & Lavooy, M. J. (2005). A review of the methods of studies on infant ultrasound production and maternal retrieval in small rodents. *Behavior genetics*, *35*, 31-52.
73. Yoshihara, C., Numan, M., & Kuroda, K. O. (2018). Oxytocin and parental behaviors. *Behavioral pharmacology of neuropeptides: oxytocin*, 119-153.
74. Noirot, E. (1974). Nest-building by the virgin female mouse exposed to ultrasound from inaccessible pups. *Animal Behaviour*, *22*(2), 410-420.
75. Stern, J. M., & Johnson, S. K. (1990). Ventral somatosensory determinants of nursing behavior in Norway rats. I. Effects of variations in the quality and quantity of pup stimuli. *Physiology & behavior*, *47*(5), 993-1011.
76. Stern, J. M., & Azzara, A. V. (2002). Thermal control of mother–young contact revisited: hyperthermic rats nurse normally. *Physiology & behavior*, *77*(1), 11-18.
77. Lévy, F., & Keller, M. (2009). Olfactory mediation of maternal behavior in selected mammalian species. *Behavioural brain research*, *200*(2), 336-345.
78. Sewell, G. D. (1970). Ultrasonic communication in rodents. *Nature*, *227*(5256), 410-410.
79. Ehret, G., & Haack, B. (1982). Ultrasound recognition in house mice: key-stimulus configuration and recognition mechanism. *Journal of comparative physiology*, *148*, 245-251.
80. Farrell, W. J., & Alberts, J. R. (2002). Stimulus control of maternal responsiveness to Norway rat (Rattus

norvegicus) pup ultrasonic vocalizations. *Journal of comparative psychology*, *116*(3), 297.
81. Smotherman, W. P., Bell, R. W., Starzec, J., Elias, J., & Zachman, T. A. (1974). Maternal responses to infant vocalizations and olfactory cues in rats and mice. *Behavioral biology*, *12*(1), 55-66.
82. Ehret, G. (1992). Categorical perception of mouse-pup ultrasounds in the temporal domain. *Animal Behaviour*, *43*(3), 409-416.
83. Champagne, F. A., Francis, D. D., Mar, A., & Meaney, M. J. (2003). Variations in maternal care in the rat as a mediating influence for the effects of environment on development. *Physiology & behavior*, *79*(3), 359-371.
84. Weaver, I. C., Diorio, J., Seckl, J. R., Szyf, M., & Meaney, M. J. (2004). Early environmental regulation of hippocampal glucocorticoid receptor gene expression: characterization of intracellular mediators and potential genomic target sites. *Annals of the New York Academy of Sciences*, *1024*(1), 182-212.
85. Champagne, D. L., Bagot, R. C., van Hasselt, F., Ramakers, G., Meaney, M. J., de Kloet, E. R., Joels, M., & Krugers, H. (2008). Maternal Care and Hippocampal Plasticity: Evidence for Experience-Dependent Structural Plasticity, Altered Synaptic Functioning, and Differential Responsiveness to Glucocorticoids and Stress. *The Journal of Neuroscience*, *28*(23), 6037–6045. https://doi.org/10.1523/JNEUROSCI.0526-08.2008
86. Fish, E. W., Shahrokh, D., Bagot, R., Caldji, C., Bredy, T., Szyf, M., & Meaney, M. J. (2004). Epigenetic programming of stress responses through variations in maternal care. *Annals of the New York Academy of Sciences*, *1036*(1), 167-180.
87. Fonseca, A. H., Santana, G. M., Bosque Ortiz, G. M., Bampi, S., & Dietrich, M. O. (2021). Analysis of ultrasonic vocalizations from mice using computer vision and machine learning. *Elife*, *10*, e59161.
88. Goodwin, N. L., Nilsson, S. R., & Golden, S. A. (2020). Rage Against the Machine: Advancing the study of aggression ethology via machine learning. *Psychopharmacology*, *237*, 2569-2588.
89. von Ziegler, L., Sturman, O., & Bohacek, J. (2021). Big behavior: challenges and opportunities in a new era of deep behavior profiling. *Neuropsychopharmacology*, *46*(1), 33-44.
90. Stevenson, P. K., Casenhiser, D. M., Lau, B. Y., & Krishnan, K. (2021). Systematic analysis of goal-related movement sequences during maternal behaviour in a female mouse model for Rett syndrome. *European Journal of Neuroscience*, *54*(2), 4528-4549.
91. Sturman, O., von Ziegler, L., Schläppi, C., Akyol, F., Privitera, M., Slominski, D.,... & Bohacek, J. (2020). Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions. *Neuropsychopharmacology*, *45*(11), 1942-1952.
92. Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature neuroscience*, *21*(9), 1281-1289.
93. Reis, A. R., De Azevedo, M. S., De Souza, M. A., Lutz, M. L., Alves, M. B., Izquierdo, I.,... & Lucion, A. B. (2014). Neonatal handling alters the structure of maternal behavior and affects mother–pup bonding. *Behavioural brain research*, *265*, 216-228.
94. de Sá Couto-Pereira, N., Ferreira, C. F., Lampert, C., Arcego, D. M., Toniazzo, A. P., Bernardi, J. R.,... & Dalmaz, C. (2016). Neonatal interventions differently affect maternal care quality and have sexually dimorphic developmental effects on corticosterone secretion. *International Journal of Developmental Neuroscience*, *55*, 72-81.
95. Mathis, M. W., & Mathis, A. (2020). Deep learning tools for the measurement of animal behavior in neuroscience. *Current opinion in neurobiology*, *60*, 1-11.

2 Objectives and hypotheses

Early-life social behaviors play a pivotal role in shaping an organism's subsequent neurobehavioral development. Dyadic mother-infant interactions are a central component of this developmental process, and disruptions in these mother-infant interactions have been linked to an elevated risk of impaired physical, socioemotional, and cognitive development. Therefore, preclinical research is essential to gain a deeper understanding of how these interactions affect biopsychological mechanisms that contribute to both typical and atypical sociocommunicative development. However, to establish robust and replicable research findings, accurate and precise assessment tools are imperative.

Objective 1: Establish the detailed assessment of maternal care in laboratory mice

Maternal care is generally the primary and most consistent form of social interaction that infants experience early in life. As such, maternal care behaviors provide a fundamental framework for development of social and communicative skills. It is hypothesized that deviations or disruptions in maternal care behaviors, due to maternal or infant psychopathology, can serve as an indicator of potential sociocommunicative difficulties in infants (1). Therefore, assessing maternal behavior is essential to obtain insights into infant sociocommunicative development and to help disentangle the complex dyadic nature of mother-infant interactions. For this reason, the Pup Retrieval Test (PRT) was adapted to accurately investigate maternal behavior by establishing novel PRT outcome parameters. This involved the definition of a PRT protocol, automated pose estimation and automated behavioral categorization of maternal behaviors. Results related to objective 1 are described in *Chapter 3: Automated procedure to assess pup retrieval in laboratory mice*.

Objective 2: Establish an automated behavioral test assay to assess bidirectional early-life interactions between maternal behavior and pup vocalization in mouse dam-pup dyads

Maternal retrieval behavior is elicited by ultrasonic infant vocalizations, as previously reported (2-6). Due to the bidirectional nature of early-life mother-infant interactions, it is important to differentiate between maternal or infant psychopathology to understand (a)typical developmental trajectories. However, existing assays only take into account one dyadic member, hereby potentially overseeing biological and/or environmental elements of this complex bidirectional interaction. To fill this gap, a novel automated homecage behavioral method, BAMBI (Bidirectional Automated Mother-pup Behavioral Interaction test), was created to investigate complex phenotypes related to early-life social development. Pup vocalizations are expected to elicit maternal care, whereas deviations in vocalization ability are expected to lead to deviations in maternal retrieval behavior. Results related to objective 2 are described in *Chapter 4: BAMBI: A new method for automated assessment of bidirectional early-life interaction between maternal behavior and pup vocalization in mouse dam-pup dyads*.

Objective 3: Characterize the impact of sociocommunicative infant abilities on maternal behavior and manipulate these effects through endogenous oxytocin stimulation.

Previous research reported that mice prenatally exposed to valproic acid (VPA) exhibit sociocommunicative deficits in form of quantitative and qualitative differences in USV emission (7-10). However, the impact of these deficits on early-life dam-infant interactions remains largely unexplored. Further, the oxytocin (OT) system has been reported to modulate infant USV emission (11). Therefore, Chapter 5 focuses on the effect of deviant sociocommunicative abilities in mice prenatally exposed to VPA on maternal retrieval behavior. Moreover, we study whether these deviant sociocommunicative abilities can be manipulated in a dose-dependent manner by targeting the endogenous OT system through a melanocortin-4 receptor agonist THIQ. Dams of pups prenatally exposed to VPA are expected to exhibit deviant retrieval behavior, and infant THIQ administration is expected to affect this maternal care behavior in a dose-dependent manner. Results related to objective 3 are described in *Chapter 5: Beyond the Squeaks: understanding the interplay between deviant sociocommunicative behavior, maternal responses and the oxytocin system*.

2.1 References

1. Numan, Esposito, G., Hiroi, N., & Scattoni, M. L. (2017). Cry, baby, cry: expression of distress as a biomarker and modulator in autism spectrum disorder. *International Journal of Neuropsychopharmacology*, *20*(6), 498-503.
2. Gaub, S., & Ehret, G. (2005). Grouping in auditory temporal perception and vocal production is mutually adapted: the case of wriggling calls of mice. Journal of Comparative Physiology A, 191, 1131-1135.
3. Dulac, C., O'Connell, L. A., & Wu, Z. (2014). Neural control of maternal and paternal behaviors. Science, 345(6198), 765-770.
4. Lonstein, J. S., Lévy, F., & Fleming, A. S. (2015). Common and divergent psychobiological mechanisms underlying maternal behaviors in non-human and human mammals. Hormones and Behavior, 73, 156–185. https://doi.org/10.1016/j.yhbeh.2015.06.011
5. Nagasawa, M., Okabe, S., Mogi, K., & Kikusui, T. (2012). Oxytocin and mutual communication in mother-infant bonding. Frontiers in Human Neuroscience, 6, 31–31. https://doi.org/10.3389/fnhum.2012.00031
6. Valtcheva, S., Issa, H. A., Bair-Marshall, C. J., Martin, K. A., Jung, K., Zhang, Y., Buzsáki, G., Kwon, H., & Froemke, R. C. (2021). Neural circuitry for maternal oxytocin release induced by infant cries. *BioRxiv*, 2021-03.
7. Kuo, H. Y., & Liu, F. C. (2017). Valproic acid induces aberrant development of striatal compartments and corticostriatal pathways in a mouse model of autism spectrum disorder. The FASEB Journal, 31(10), 4458-4471.
8. Dai, Y. C., Zhang, H. F., Schön, M., Böckers, T. M., Han, S. P., Han, J. S., & Zhang, R. (2018). Neonatal oxytocin treatment ameliorates autistic-like behaviors and oxytocin deficiency in valproic acid-induced rat model of autism. Frontiers in cellular neuroscience, 12, 355.
9. Gzielo, K., Potasiewicz, A., Hołuj, M., Litwa, E., Popik, P., & Nikiforuk, A. (2020). Valproic acid exposure impairs ultrasonic communication in infant, adolescent and adult rats. European neuropsychopharmacology, 41, 52-62.
10. Barrett, C. E., Hennessey, T. M., Gordon, K. M., Ryan, S. J., McNair, M. L., Ressler, K. J., & Rainnie, D. G. (2017). Developmental disruption of amygdala transcriptome and socioemotional behavior in rats exposed to valproic acid prenatally. *Molecular Autism*, *8*(1), 42–42. https://doi.org/10.1186/s13229-017-0160-x
11. Hammock, E. A. (2018). Oxytocin and vasopressin systems in the development of social behavior. In Routledge international handbook of social neuroendocrinology (pp. 504-529). Routledge.

3 Automated procedure to assess pup retrieval in laboratory mice*

* Adapted from Winters, C., Gorssen, W., Ossorio-Salazar, V. A., Nilsson, S., Golden, S., & D'Hooge, R. (2022). Automated procedure to assess pup retrieval in laboratory mice. Scientific reports, 12(1), 1663. https://doi.org/10.1038/s41598-022-05641-w

3.1 Abstract

All mammalian mothers form some sort of caring bond with their infants that is crucial to the development of their offspring. The Pup Retrieval Test (PRT) is the leading procedure to assess pup-directed maternal care in laboratory rodents, used in a wide range of basic and preclinical research applications. Most PRT protocols require manual scoring, which is prone to bias and spatial and temporal inaccuracies. This study proposes a novel procedure using machine learning algorithms to enable reliable assessment of PRT performance. Automated tracking of a dam and one pup was established in DeepLabCut and was combined with automated behavioral classification of "maternal approach", "carrying" and "digging" in Simple Behavioral Analysis (SimBA). Our automated procedure estimated retrieval success with an accuracy of 86.7%, whereas accuracies of "approach", "carry" and "digging" were estimated at respectively 99.3%, 98.6% and 85.0%. We provide an open-source, step-by-step protocol for automated PRT assessment, which aims to increase reproducibility and reliability, and can be easily shared and distributed.

3.2 Introduction

Maternal care determines an infant's physical and functional development (1-4). Discriminative maternal care during the postpartum period is ensured by the formation of some sort of mother-infant bond (4-5). Human studies that investigate maternal care and its many effects on infant development are complicated by a variety of practical and ethical factors, lack of control over environmental and genetic background, inaccessibility of brain samples (6). Animal models, particularly rodents, provide a valid research tools as many neural and hormonal mechanisms of care and bonding are homologous between mammalian species (6-7).

The Pup Retrieval Test (PRT) is the most widely used assay to assess maternal care in fundamental and preclinical rodent research (8). In its most basic design, it quantifies the mother's retrieval response to the removal of a pup from the nest, a sequence of pup-directed sensorimotor responses elicited by (multi)modal stimuli from the infant, and processes by the mother (4, 9-13). The test has been used to study the impact of pharmacological and environmental interventions on maternal care (14-18). Finally, it has been used in research on disorders that affect mother-infant interaction, such as autism, fetal alcohol syndrome, chronic prenatal stress, postpartum depression, schizophrenia (19-24).

It was established that differences in methodology, environment and experimenter intervention influence behavior and its assessment (25). Typical PRT protocols rely on manual scoring, which

provides only basic information and is prone to spatial and temporal inaccuracies as well as bias (26). Some authors resorted to limited end-point registration to simplify manual PRT scoring (27-28). Components such as “maternal latency to start retrieving the pup”, “time necessary to complete retrieval”, and “efficacy of the retrieval” have been scored during real-time observation or from video recordings, but pup retrieval is a dynamic, interactive behavior that contains more information than these variables (28). Also, scoring these and other components from videos, or with video tracking software, requires laborious indexing of the behavioral components on a frame-by-frame basis (29-30). Scoring programs may not allow such frame-based analyses, and scoring tends to be a slow and laborious process. Furthermore, manual analyses require well-defined rules to define start, duration and end of behavioral components (31). These rules are researcher and context specific, which affects transferability between laboratories (29-30, 32-33). Further, temporal accuracy is especially relevant for PRT behaviors that are triggered by pup ultrasonic vocalizations and have a millisecond profile (9).

Precise and standardized recording of the maternal response would be most reliably achieved by automated procedures (34). Manual PRT scoring is often inaccurate, involving subjective judgement of the completion of a retrieval event, imprecise definition of nest borders, etc. We therefore present a novel procedure using machine learning algorithms and open-source software to enable reliable, automated PRT registration. Recent advances in motion capture and deep-learning allow extraction of behavioral variables from recorded videos without elaborate recording hardware (35). Tracking and behavioral analysis tools reach at least human-like accuracy, and even outperform manual scoring with unprecedented rapidity (27, 36). We established automated tracking of a dam and one pup using DeepLabCut software, and classified variables such as “maternal approach”, “carrying” and “digging” using Simple Behavioral Analysis (SimBA). Specifically, this paper aims to (1) introduce a dataset for a dam-pup PRT tracking network, including operational definitions and underlying rationale, (2) establish SimBA classification of ethologically relevant PRT behaviors, (3) show that an automated procedure is able to quantify PRT accurately, (4) create a SimBA add-on to evaluate PRT results using an easy-to-use graphical user interface, (5) provide a user-friendly, step-by-step protocol to replicate or expand the present study. All annotated images and videos, tracking models and behavioral classifiers are available at https://doi.org/10.17605/OSF.IO/RWHTD.

3.3 Results

3.3.1 Video pre-processing.

C57BL6J mice were subjected to the PRT on postnatal day 5 (P5). Videos were recorded in the home cage using a Foscam C2 IP-camera (EUport, Wageningen) from an overhead camera perspective (top-down). Per dam, one video file was obtained containing six single pup retrieval trails. This one overall video was splint into six single pup trial videos. To decrease file size and increase efficient neural network training, videos were cropped around the top edges of the home cage and grey-scaled.

3.3.2 Tracking dataset and training

DeepLabCut 2.1.10.4 (DLC, 35) was used to create a dam-pup tracking algorithm in the pup retrieval protocol. 592 Frames were randomly extracted from 38 single trial videos, using the k-means algorithm in DLC. Next, 14 body parts were manually annotated with high stringency, with seven body parts on the dam and seven on the isolated pup (see **Figure 1a**). Occluded body parts in the established training dataset were simulated by using the body part configuration from a similar frame in which the occluded animal was visible as a template. The operational definitions of these body parts are shown in **Supplementary File 1 Table S1**.

The dam-pup network was trained in Google Colaboratory (https://colab.research.google.com/) using the ResNet-50 architecture in DeepLabCut with a batch size of 4 images. A 95:5 train/test ratio was used, meaning that 95% of the 592 annotated frames was used to train the neural network and 5% is used to evaluate network performance. Training of neural networks means that parameters of a mathematical model is iteratively optimized to increase prediction accuracy on the training dataset. As the aim of the dam-pup model is to generalize well in the future, the model should not be overfitted. Overfitted models will only "memorize" initial data meaning they will perform poorly on unseen videos recorded with slightly different animals, lighting or camera distance (37). Therefore, training was stopped after 70,000 iterations as optimization minimized and model parameters of every 1000th iteration were stored.

3.3.3 Performance evaluation of the trained network.

Network performance was evaluated by the computation of a train and test error as measured by the average Euclidian difference between the pixel coordinates from manual annotations and DLC estimations. Prior to selection, four criteria were defined to evaluate the best performing model. First, the model should not be overfitted and thus the best performing model with the lowest number of iterations should be chosen. Second, mean test and train error over all body parts should be as low as possible. Third, especially maternal nose and ears should have a low mean error as these will be important for the behaviors in the PRT. Fourth, the nest can occlude body parts and consequently

impact tracking accuracy. Thus, mean Euclidian errors should decrease after increasing the likelihood threshold and thus calculating mean test or train errors for DLC estimations above a defined probability. The best performing dam-pup model was created after 50,000 iterations. The estimated error averaged over all 14 body parts was 2.3 pixels (px) for the training dataset and 10.44 px for the test dataset (**Figure 1b**). The mean Euclidian errors per body part and using different likelihood thresholds are shown in **Supplementary File 2 Table S2**.

3.3.4 Distance standardization and definition of Regions of Interest.

Euclidean pixel distances were standardized to millimeter distances using Simple Behavioral Analysis (SimBA, (31)). Here, the length between the top lids of the home cage (**Figure 1c**) is used to define a standardized distance of 265 mm. The mean errors per body part were standardized to a metric scale and ranged between 2.2 and 4.5 mm as shown in Figure 1b. Hereafter, per video two regions of interest were defined in SimBA: the nest and the core nest. The nest is defined as a polygonal that encloses the entire nesting site (**Fig. 1d, red**), whereas the core nest is defined as a circle that encloses only the site of the nest in which the pups are (**Fig. 1d, pink**).

3.3.5 Quality control of DeepLabCut tracking output.

Data are tracked from DLC in SimBA to train behavioral classifiers. Since tracking inaccuracies (such as impossible locations or movements) can complicate classifier training, they should be corrected. Outliers were defined by using the median distance between nose and spine1: median distance multiplied by 2.5 was the criterion to define movement outliers, median distance multiplied by 4 defined location outliers.

3.3.6 Random forest classifiers.

SimBA was used to develop three random forest models to predict maternal approach, maternal carrying and maternal digging. In total, 27683 frames of seventeen videos were manually annotated using the SimBA event logger for the absence (‘0’) or presence (‘1’) of each of the three behaviors. A subset of nine videos was used to train the approach and carry models, while the digging model was built using ten videos, yielding respectively 10157 and 18207 frames (**Table 1**). All three random forest models were trained by creating 2000 decision trees from the training data. Frames containing the behaviors of interest were imbalanced compared to frames absent of these behaviors. Therefore, the majority class (i.e. absent behavior) was randomly under-sampled at ratios shown in **Table 1**. Based on precision-recall curves, an optimal classifier performance was chosen on the basis of f1. To balance the number of false-positive and false-negative classifications, we chose the discrimination threshold at the highest f1 value. These discrimination thresholds and minimum bout lengths can be found in **Table 1.**

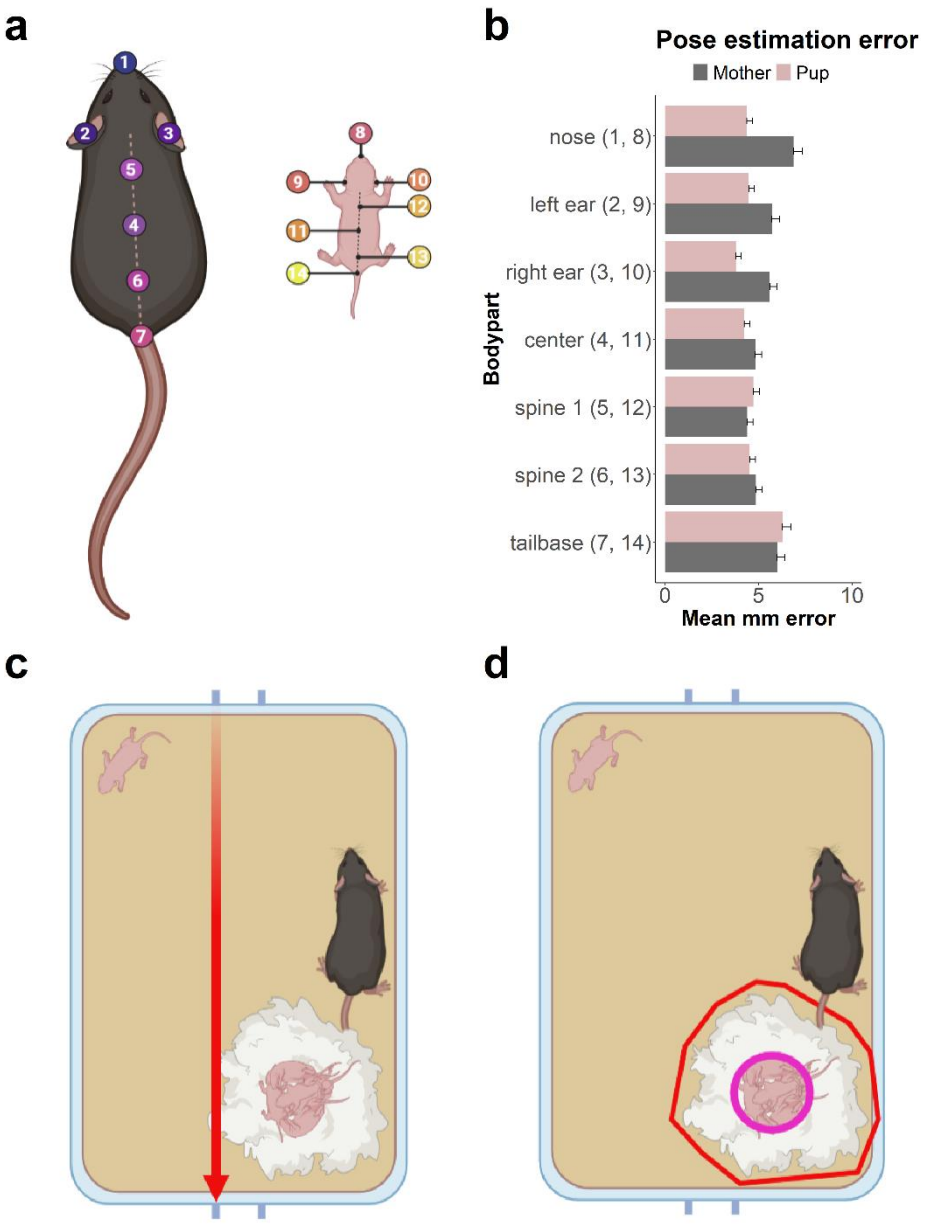


Figure 1. Pose estimation and region of interest (ROI) definition. (**a**) Schematic representation of the 14-bodypart pose configuration used for the PRT protocol in DeepLabCut (image was created with BioRender.com). (**b**) Mean millimeter error per body part and subject for the DeepLabCut PRT pose estimation model. Means were calculated without the use of an estimation p-cutoff. In Supplementary File 2 Table S2, the mean mm errors are shown with different p-cutoff values. (**c**) Illustration of distance on a Type-II mouse cage. We used the length between opposite cage tag hinges as standard length. The pink line indicates a distance of 267 mm. (**d**) Example of region of interest (ROI) definition in SimBA. The yellow area indicates the nest site and the pink circular area the core nest. The red markers do not display any ROI, but indicate that the top corners of the home cage need to be visible in the cropped video.

Classifier	# Annotated frames	Behavior present (%)	Random under-sample ratio	Test set: frames present	Test set: frames absent	Discrimination threshold	Minimum bout length (ms)
Approach	10,157	8.7	8.5	604	5302	0.47	500
Carry	10,157	13.9	16	828	4443	0.47	200
Digging	18,207	32.3	2	930	1802	0.24	1000

Table 1. Summary statistics on data used for behavioral classification.

	Manual behavior absent (0)	Manual behavior present (1)
Approach		
Automated behavior absent (0)	24,676	233
Automated behavior present (1)	157	2618
Carry		
Automated behavior absent (0)	26,155	120
Automated behavior present (1)	76	1333
Dig		
Automated behavior absent (0)	14,412	925
Automated behavior present (1)	3218	9129
Retrieval status	**Manual not retrieved (0)**	**Manual retrieved (1)**
Automated not retrieved (0)	16	2
Automated retrieved (1)	6	36

Table 2. Confusion matrices of manual and automated classification of retrieval status and maternal approach,carry and digging behaviors.

3.3.7 Performance of random forest classifiers.

After training, an estimation was made for the model which estimates the difference between the original input and the predicted output. This, however, does not give an indication of the translatability of the model to an independent dataset. Here, seventeen videos were manually annotated frame-by-frame and were joined into an independent dataset. Using the Caret package in R (38), the performance of the three classifiers was evaluated.

The independent dataset for carry and approach consisted of 17526 frames (approach present = 13.9 %, carry present = 8.7%, **Table 1**). Confusion matrices for carry, approach and digging behavior are shown in **Table 2**. The approach and carry classification accuracy was respectively 98.6% (95% CI = 98.5, 98.7) and 99.3% (95% CI = 99.2, 99.4). The true negative rate or specificity was 91.8% for approach and 91.7% for carry. The true positive rate or sensitivity was 99.4% for approach and 99.7% for carry. For the evaluation of the digging classifier, the independent dataset contained the frames of seven annotated videos with in total 9476 frames (digging present = 44.1%, Table 2). The digging classifier accuracy was 85.0% (95% CI = 84.6, 85.5), with a sensitivity of 81.8% and a specificity of 90.8%.

Besides the frame-by-frame validation of these behaviors, the total count and duration of each behavior per video were validated using Pearson correlations. **Table 3** shows the correlation matrices of maternal approach, carry and digging behaviors scored with three different methods: manual with no defined ROI, manual with defined ROI and our proposed automated scoring with ROI. Correlations were moderate to high (r=0.51-0.90) between all scoring methods. Moreover, scoring with defined ROI showed high internal correlations for approach, carry and dig respectively (r=0.62,0.77,0.90). These

results suggest that defining an ROI has substantial impact on the PRT output. Predefined ROI are currently not used in PRT analysis, leading to a possible bias. Furthermore, the high correlations between both manual and automated scorings with ROI indicate that our automated model is appropriate for the estimation of the included maternal behaviors.

3.3.8 Manual pup retrieval estimations.

Raw retrieval videos were observed and categorized manually into “pup retrieved” or “pup not retrieved” classes. A pup trial was classified as retrieved as the mother carried it back to the nest (38). In these videos the nest ROI was not visible. Successful retrieval trials were labeled ‘1’ and unsuccessful retrieval trials were labeled ‘0’. Additionally, the time was manually estimated for when the pup was carried back into the nest. Unsuccessful retrieval trials are assigned the maximum trial time (here: 90s).

	Manual_noROI	Manual_ROI	Automated_ROI
Approach			
Manual_noROI	1.00	0.88	0.85
Manual_ROI	0.88	1.00	0.90
Automated_ROI	0.85	0.90	1.00
Carry			
Manual_noROI	1.00	0.64	0.51
Manual_ROI	0.64	1.00	0.62
Automated_ROI	0.51	0.62	1.00
Dig			
Manual_noROI	1.00	0.55	0.62
Manual_ROI	0.55	1.00	0.77
Automated_ROI	0.62	0.77	1.00

Table 3. Correlational matrices of maternal approach, carry and digging behaviors scored with three different methods: manual with no defined ROI, manual with defined ROI and automated scoring with ROI.

3.3.9 Automated pup retrieval estimations.

In SimBA, an add-on was created to evaluate results of the automated pup retrieval test, taking into account eventual tracking inaccuracies (**Figure 2**). Here, retrieval is labeled as successful if at least one body part of a pup is present in the nest ROI and carry behavior was observed in the three seconds before entering. Similar as in the manual estimations, successful or unsuccessful retrieval events are scored respectively ‘1’ and ‘0’. Moreover, the output gives the estimated time of retrieval as well as the latency, total duration and counts for each behavior (carry, approach and dig).

3.3.10 Performance evaluation of automated retrieval estimations.

The prediction results created with SimBA were compared to the manually scored videos. Using the Caret package in R (38), confusion matrices were obtained and analyses (**Table 2**). The accuracy of the retrieval success was estimated at 86.7% (95% CI = 75.4, 94.1). The sensitivity and specificity of the predictions were respectively 72.7% and 94.7%. Next, manual time estimations and the predicted time estimations were correlated with r=0.72. The prediction errors were examined in detail and could be classified into two types: (1) “boundary errors”, when the automated prediction was generally good,

but the discrepancy with the manual estimation was related to ROI definition, (2) “Nest shift”, when the pup was predicted to be retrieved due to a moved nest but was without any doubt not retrieved in the nest in reality. Of the eight videos with a prediction error, two errors were explained by a shifted nest and six by boundary errors.

3.3.11 PRT protocol, tracking models and behavioral classifiers.

Our standardized PRT protocol can be found in **Supplementary File 5**, whereas our annotated images and videos, tracking models and behavioral classifiers are available on: https://doi.org/10.17605/OSF.IO/RWHTD. This protocol and these models are made available to the scientific community to ensure a more standardized analysis and more translatable results across researchers. Although our models and behavioral classifiers were constructed for specific settings (e.g. C57BL6J mice, home cage), they can be enlarged with their own data to obtain working models for different settings, as explained in our protocol (**Supplementary File 5**).

3.4 Discussion

Pup retrieval is an essential feature of maternal care in mice, and PRT performance has consequently been used to study parental behavior in various research applications. The present paper proposes a novel, automated machine learning-based procedure to assess PRT performance, and identify performance variables in recorded PRT video files. Although our procedure uses machine learning, only basic equipment is required in combination with easy-to-use, open-access software. We used DLC to construct an accurate model for tracking a dam and one isolated pup synchronously. The model was able to track both the mother and the pup with high accuracy (mean body part tracking error between 2.2 and 4.5 mm). DLC has been used by others to track user-defined features on interacting animals, precisely and simultaneously, in simple and complex setups (37, 39). However, to date, neither DLC or other available software could be used to track a dam and her pup simultaneously, and track multiple body parts on each animal. The lack of such a tool has hindered the analysis of dyadic dam-pup interactions and manual scoring has remained the only possibility to quantify behaviors (28). Manually annotating every frame in a one-hour video has been shown to require 22 working hours (37). Additionally, manual scoring has been confounded by observer bias and observational drift, and had low reproducibility.

Manually scoring of successful pup retrieval is complicated by imprecise operationalization of retrieval success. We defined "successful retrieval" as "the experimental female (…) retrieves the displaced pup back to the nest", based on Marlin and colleagues (12). However, the nest is not a predefined area, and it therefore depends on opinion whether "maternal transport to the periphery of the nest" is enough to label the trial successful. As mentioned, this ambiguity of underlying rules increases the chance of observer bias and complicate reproducibility and reliability (29-30, 32-33). The use of pre-defined ROIs during manual scoring already increased confidence in retrieval estimations, and we showed that our automated procedure yielded comparable results to manual scoring. Retrieval success was estimated with an accuracy of 87% (Table 2). One limitation, however, is that the nesting site is predefined and therefore fixed, but a mother may actually move the nesting site during the trial, and thus inflate retrieval scores. It might therefore be interesting to automate ROI definition in the future also.

Using automated behavioral classification, we were able to predict "maternal approach", "carry" and "digging" with accuracies of 98.6%, 99.3% and 85.0%, respectively. These results are in line with studies using similar tools (31, 44). Further, the time estimated by our three classifiers for every behavior over the trials was highly correlated between automated and manual scoring for "approach" (r=0.85-0.88), "carry" (r=0.51-0.64) and "dig" (r=0.55-0.77). These results indicate automated scoring yields comparable results to the golden standard of manual scoring. Important characteristics for

transferability of behavioral classifiers are whether the features are position or rotation dependent (37). SHAP-analysis (data not shown) showed that our “carry” and “approach” classifiers are grossly explained by a combination of dam-pup proximity and simultaneous movement, making them more

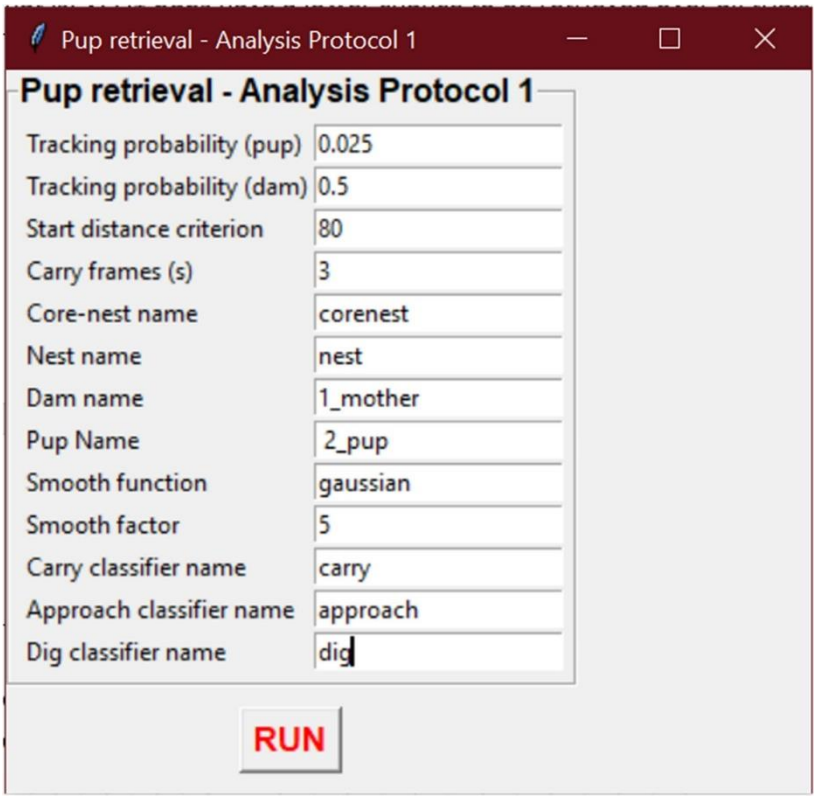


Figure 2. Pup retrieval analysis add-on in SimBA. A new module was built into SimBA for the automated analysis of the pup retrieval test, while performing a post-analysis quality control. Here, some thresholds are specified to resolve potential problems in tracking accuracies affecting the PRT outcome. The output provides the estimated retrieval success (retrieved vs not retrieved), time of retrieval and latency, total duration and counts for each behavior (carry, approach dig).

likely to be transferable. However, “digging” was mostly explained by maternal position relative to the nest and core nest, indicating that “digging” might be less transferable. Unfortunately, we cannot assess how well these classifiers would perform in a different data space. Transferability of a constructed model to a new data space that looks different from previously encountered data has been a serious obstacle in deep learning approaches (40). Here, we maximized transferability of the dam-pup model by avoiding to over-fit the network, or over-standardize cage environment (e.g., nest composition or camera position, (41)). However, only recordings of black C57BL6J dams retrieving their pups on P5 in the home cage were used in the present study, but researchers often use different subjects (e.g., white CD-1 mice), different developmental ages (42) or different test environments (e.g., T-shaped extension to the home cage, (8, 43)). A step-by-step protocol is provided in **Supplementary**

File 5 for transparency and reproducibility, explaining how to expand the basic neural networks. Operational definitions of behavioral classifications are provided in **Supplementary File 4, Table S4**. By sharing all established datasets and models on OSF (https://osf.io/rwhtd/), the scientific community can achieve more generic models.

Typically, PRT read-outs have been limited to parameters such as “retrieval success” and “time to retrieve”. Sequential underlying behavioral components may be ignored by approaching maternal retrieval as a single behavior. However, retrieval comprises “maternal perception of pup distress”, “search” as well as “approach” components, and transporting the pup back to the nest (28). By taking separate behaviors into account, retrieval vs non-retrieval can be more accurately analyzed (for example, by an inability to locate the pup). While the number of animals and workload of sampling stays equal, researchers can answer a much wider range of questions by taking these behaviors into account. In its current form, our automated PRT procedure included only on maternal “approach”, “carry” and “digging”, but other behaviors such as “nursing”, “pup licking”, “self-grooming”, and “disruptive digging on the pup” might ethologically relevant as well (45). The flexibility of the used software allows to include other behavioral parameters, and the advantage of working with open-access software is that new features can be readily implemented (46). As a final note, the use of unsupervised neural networks might be an interesting future direction. In contrast to supervised ones, unsupervised neural networks are not trained on data annotated by human researchers, which could even further avoid human bias, and identify behavioral components that relate to specific brain activity (37).

3.5 Material and Methods

3.5.1 Animals & Ethics Statement

Breeding pairs of primiparous C57BL/6JRj mice (8-10 weeks old) were purchased from Janvier Labs (Le Genest-Saint-Isle, France). Adult mice were group-housed and maintained for time-controlled breeding in standard type II cages. Males were only present in the home cage the night of mating and females were housed individually from gestation through weaning (P28). Mice were kept at 12/12 hour light-dark cycle (lights on at 7 AM), water and food ad libitum, conditioned rooms (22°C, humidity 30%). On day of birth (P0), pups were sexed and nests were reduced to 6 pups with a balanced male:female ratio. All animal procedures were approved the Animal Ethics Committee of KU Leuven (P028/2018), in accordance with European Community Council Directive 86/609/EEC, the ARRIVE guidelines and the ILAR Guide to the Care and Use of Experimental Animals.

3.5.2 Pup retrieval test setup and protocol

Dams were tested on P5 and were transported to the test room at 8 AM. PRT trials were performed between 9 AM and 10 AM. A detailed step-by-step protocol can be found in **Supplementary File 5**. The

PRT test was performed in the home cage inside a Styrofoam box to create a visually isolated environment. A Foscam C2 IP-camera (EUport, Wageningen) was connected to a laptop (Windows 10 as operating system) using an online interface and set at a height of 50 cm above the setup to record from an overhead camera perspective. Recording settings were as follows: a resolution of 1280 x 720 pixel, 10-30 frames per second (fps) and room lighting. Pups were placed in a clean glass cup, which was preheated to 35°C using a heat pad. Every dam performed six semi-randomized retrieval trials (i.e. 3 male and 3 female trials) with a maximum time of 90 seconds per trial. A trial started as the mother was in the nest and a pup was placed in the most distant corner relative to the nest. If the pup is not retrieved within 90 seconds, the pup was placed back into the core of the nest.

3.5.3 Image pre-processing

Video recordings were pre-processed using SimBA. First, spatial dimensions were cropped to fit the upper corners of the home cage (Fig. 1d). Second, videos were shortened to create separate trial videos (e.g. Mother-ID_Trial1). The start of the video was the first frame after placing the pup in the corner and where the researcher's hand is not visible anymore. The end of the video was determined as 90 seconds after the first frame. Third, greyscale was applied to the videos. Finally, videos were formatted to .mp4 video format.

3.5.4 Automated body part tracking using DeepLabCut

Pose-estimation (tracking) data, which is necessary to create features in SimBA, forms the basis of the accurate classification of behavioral patterns in video recordings (31). We defined a minimalist 14-body part pose configuration for the mother and pup together, necessary to classify different maternal behaviors. DeepLabCut 2.1.10.4 (35) was used to develop a PRT mother-pup tracking framework. From 38 single trial video recordings, frames were extracted for labeling using k-means clustering. In total 592 frames were labeled as in Fig. 1a and **Supplementary File 1 Table S1**. Of these labeled frames, 95 percent were used for model training, whereas the remaining 5 percent was used to evaluate network performance. The mother-pup network was trained using ResNet-50 with a batch size of 4 images via Google Colaboratory (https://colab.research.google.com/).

3.5.5 Automated behavioral classification in SimBA

For maternal approaching, carrying and digging, random forest classifiers were constructed in SimBA. In total, 27683 frames of seventeen video recordings of PRT behavior were extracted and manually annotated for each of the three behaviors. To classify carry and approach, a subset of 10157 frames from ten videos were used to construct classifiers. The digging classifier was constructed using 18207 frames from ten different videos. Operational definitions can be found in **Supplementary File 2 Table S4**. All models were built with the following settings: n_estimators = 2000, RF_criterion = entropy, RF_max_features = sqrt, RF_min_sample leaf = 1, and random undersampling (approach = 8.5, carry =

16, digging = 2). A 75:25 ratio was used to split the annotated frames into a train:test dataset. Discrimination thresholds and minimum durations can be found in **Table 1**.

3.5.6 Manual PRT scoring and behavior detection

Raw retrieval videos were classified '1' if retrieval was successful and '0' if retrieval was unsuccessful. A pup trial was classified as retrieved as the mother carried it back to the nest. In these raw videos the nest ROI was not visible. Next, the latency to being retrieved was manually estimated per video and unsuccessful retrieval trials are assigned the maximum trial time (here: 90s).

3.5.7 Automated pup retrieval estimations

In SimBA, an add-on was constructed to assess pup retrieval test performance. Here, retrieval is scored as successful if at least one body part of a pup is present in the nest ROI and carry behavior was observed in the three seconds before entering. Similar as in the manual estimations, successful or unsuccessful retrieval events are scored '1' and '0' respectively. Moreover, the output gives the estimated time of retrieval as well as the latency, total duration and counts for each behavior (carry, approach and dig).

3.5.8 Computer software and hardware

A laptop equipped with an Intel Core i5-8350U CPU and 8 GM RAM was used for image annotations in DeepLabCut (http://www.mackenziemathislab.org/deeplabcut), behavioral classification in SimBA (https://github.com/sgoldenlab/simba) and data processing in R and Python. The mother-pup network was trained using Google Colaboratory (https://colab.research.google.com/).

3.5.9 Data availability

All our annotated images and videos, tracking models and behavioral classifiers are available on: https://doi.org/10.17605/OSF.IO/RWHTD. For further inquiries, please contact the corresponding author.

3.6 Supplementary files

3.6.1 Supplementary File 1 Table S1.

Operational definitions of body parts for labeling. A mother-pup tracking framework was developed by labeling the following seven body parts on both the dam and the pup.

Body part	Definition
Nose	Trivial, see Figure 1a
Ear left	Trivial, see Figure 1a
Ear right	Trivial, see Figure 1a
Center	Spinal location underneath the rib cage. This spinal location functions as a hinge as the mouse changes its direction of movement.
Spine 1	Spinal location where the shoulders are articulated to the spine.
Spine 2	Spinal location where the hips are articulated to the spine.
Tail base	Place on the body of the mouse where the tail starts.

3.6.2 Supplementary File 2 Table S2.

Pose estimation prediction errors (in pixels) with different p-cutoff values compared to the ground truth (manual annotations). For every DLC prediction a likelihood is calculated and a p-cutoff can be defined to filter unreliable predictions. The table shows that a higher confidence DLC prediction results in more precise feature position estimation.

	P-cutoff	**None**	**0.1**	**0.5**	**0.9**	**None**	**0.1**	**0.5**	**0.9**
Individual	Bodypart	**Training dataset**				**Test dataset**			
Dam	nose	4.91	4.91	4.31	3.02	15.59	15.59	12.61	7.22
Dam	left ear	4.19	4.19	3.37	3.05	12.95	12.95	7.71	4.6
Dam	right ear	3.88	3.85	3.78	2.92	12.69	12.69	7.48	3.47
Dam	center	3.89	3.89	3.43	2.95	10.94	10.94	9.34	8.57
Dam	spine 1	4.22	4.22	3.76	3.1	9.96	9.96	9.84	5.56
Dam	spine 2	4.09	4.09	3.63	2.91	11.02	9.06	9.14	7.57
Dam	tail base	4.44	4.41	3.95	2.83	13.6	10.44	9.38	6.48
Pup	nose	3.74	3.74	3.64	3.38	9.9	7.03	5.58	3.45
Pup	left ear	3.38	3.38	3.38	3.24	10.14	6.96	6.65	3.84
Pup	right ear	3.47	3.47	3.46	3.29	8.6	5.72	5.63	4.37
Pup	center	3.22	3.22	3.21	3.13	9.59	5.89	3.76	3.69
Pup	spine 1	3.96	3.96	3.96	3.56	10.7	5.47	5.47	4.27
Pup	spine 2	3.19	3.19	3.13	3.05	10.26	5.44	4.42	4.31
Pup	tail base	3.7	3.7	3.65	3.26	14.26	7.52	6.13	4.81

3.6.3 Supplementary File 3 Table S3.

Pose estimation errors with different p-cutoff values. A likelihood is calculated for every DLC prediction. A P-cutoff can be used to filter confident predictions. We calculated Mean error (mm) based on the test data set. The table shows that a higher confidence DLC prediction results in more precise feature position estimation.

	P-cutoff	None	None	0.1	0.1	0.5	0.5	0.9	0.9
Individual	Bodypart	Mean error (mm)	sd	Mean error (mm)	sd	Mean error (mm)	sd	Mean error (mm)	sd
Dam	center	4.8	0.3	4.9	0.8	4.1	0.3	3.8	0.3
Dam	left ear	5.7	0.4	5.8	1.0	3.4	0.2	2.0	0.1
Dam	nose	6.9	0.5	7.0	1.2	5.6	0.4	3.2	0.2
Dam	right ear	5.6	0.4	5.7	1.0	3.3	0.2	1.5	0.1
Dam	spine 1	4.4	0.3	4.5	0.8	4.3	0.3	2.5	0.2
Dam	spine 2	4.9	0.3	4.1	0.7	4.0	0.3	3.3	0.2
Dam	tail base	6.0	0.4	4.7	0.8	4.1	0.3	2.9	0.2
Pup	center	4.2	0.3	2.7	0.4	1.7	0.1	1.6	0.1
Pup	left ear	4.5	0.3	3.1	0.5	2.9	0.2	1.7	0.1
Pup	nose	4.4	0.3	3.2	0.5	2.5	0.2	1.5	0.1
Pup	right ear	3.8	0.3	2.6	0.4	2.5	0.2	1.9	0.1
Pup	spine 1	4.7	0.3	2.5	0.4	2.4	0.2	1.9	0.1
Pup	spine 2	4.5	0.3	2.4	0.4	1.9	0.1	1.9	0.1
Pup	tail base	6.3	0.4	3.4	0.6	2.7	0.2	2.1	0.1

3.6.4 Supplementary File 4 Table S4.

Behavioral operational classifiers: pup retrieval

Classifier	**Description**	**Start frame**	**Duration of behavior**	**End frame**
Approach	The dam moves towards the pup and sniffs it.	First frame when the dam elongates the body towards the pup to carefully sniff it.	Uninterrupted sniffing of the pup	First frame when the dam moves the head away from the pup or tilts head to take the pup into her mouth
Carry	The dam takes the pup into its mouth and transports it.	First frame when the dam tilts its head to take the pup into her mouth	Uninterrupted transporting of the pup using the mouth	First frame that the pup is no longer in the mouth of the mother.
Dig	The dam uses snout, front- or hindlimbs to displace bedding material.	3.88	Uninterrupted manipulation of the bedding material.	First frame that the mother is stop interacting with the bedding material

3.6.5 Supplementary File 5.

Protocol for automated PRT

Procedure

Preparation ● Timing 5 d

1 On the day of birth (postnatal day 0, P0), reduce the litter size to a maximum of 6 pups if necessary. Each nest should include at least two pups of each gender (47-48). Nests are left undisturbed until P5 as researcher interference can affect cage dynamics in early-life (49-50).

! CAUTION Experiments using rodents must conform to local and national regulations. All animal studies and experimental procedures presented here were approved by the animal ethics committee of the University of Leuven (Belgium).

! CAUTION Some experimental interventions could cause interference before P5. Matched control procedures are always required, but it should be considered that parent-infant bonding might be affected in a more complex fashion. Therefore, conclusions should take possible interactions into account.

2 On the day of testing, clean the glass cup with 70% ethanol outside the test room to reduce odor cues and ensure proper disinfection. Prepare a pen, a recording sheet, balance and timer.

! CAUTION Check whether the video device contains enough memory for a video file of approximately 30 min (~200Mb/nest).

3 One hour prior to testing, transport the home cage with the mother and the pups to the test room for habituation to reduce stress.

4 While the mice are in the test room, silently prepare the set up. Firstly, preheat a heat pad to 35°C and place a clean glass cup onto it. Attach the thermometer probe to the bottom of the glass cup with a clean piece of tape. Secondly, set up the video sampling devices.

CRITICAL STEP The behavioral test room should be quiet, and temperature-controlled to avoid environmental confounds.

5 Set video recording parameters: our videos are recorded with a Foscam, top-down, 50 cm above the cage, greyscale at 10-30 fps and 1280x720 resolution.

CRITICAL STEP Behavioral classification is performed in SimBA with a temporal lower limit for approach, carry and digging (resp. 200, 500, 200 ms). We recommend recording at 25-30 fps.

6 Leave test room quietly.

7 At least five minutes prior to testing, place the first home cage to be tested into the Styrofoam box to create a visually isolated environment. Ideally, the mother should acclimatize for 60 minutes to the test setup. However, multiple nests might have to be subjected to the PRT and test time should be as standardized as possible (51).

! CAUTION Circadian fluctuations in general physiology and hormones should be minded. Therefore, testing should be fixed in time (51), and the experimenter should predefine an upper limit of testable nests per day.

Data sampling ● Timing 20-40 min per nest

8 Start video recording device.

CRITICAL STEP Before touching the pup, make sure to wear new gloves. Take some home cage bedding and rub it on your gloves.

9 Remove one pup from the core nest and place it in the heated glass cup. Record the sex of the pup. The pup stays in the heated cup until the mother is back on the nest. When the mother is back to the nest, place the pup at the most distant corner of the cage. Start timer.

CRITICAL STEP Do not rush while removing a pup. Make no sudden unexpected movements when your hand is in the home cage. Try to standardize the movement to place the pup in the cage (*e.g.* introduce the pup from the left side of the cage).

10 Trials have a fixed duration of 90s in which the mother can retrieve the pup. If the pup is not back in the nest within 90 seconds, place it back into the core of the nest. Between pups of the same nest, the glass cup is not cleaned since pup odors reduces stress in the pup.

! CAUTION Pups are not marked since odor is important in mother-infant recognition. Avoid if possible, although the same pup could be tested multiple times.

11 Repeat Steps 9-10 six times per mother, with three male trials and three female trials in pseudorandom order.

12 End the recording and name the collected video ‘Mother-ID_Date’.

13 Remove all pups from the nest and weigh them individually.

14 Take the home cage back to the animal room. In case of multiple nests, make sure to clean the glass cup outside the test room.

Data processing ● Timing

15 Pre-process video recordings using SimBA or another program of choice.

- Crop the spatial dimensions to fit the upper corners of the home cage (Fig. 1d)
- Shorten the videos to create separate trial videos (*e.g.* Mother-ID_Trial1). The start of your video should be the first frame after the 'beep' of the timer and otherwise after the first frame after placing the pup in the corner and the researcher's hand is not visible anymore. The end of your video is 90 seconds after the first frame.
- Apply greyscale if videos are sampled in color.
- The video format should be either .mp4 or .avi.

CRITICAL STEP Make sure the experimenter's hand is not in the videos since this would disrupt the model for tracking and behavioral classification.

Data processing: animal tracking without a GPU ● Timing ~ 60 min

16 Download our DeepLabCut tracking model from the OSF-page (https://osf.io/rwhtd/) in the folder 'Dropbox: automatedPRT/Tracking – DeepLabCut'.

17 Organize the data as follows. On Google Drive, use your 'My drive'-folder as the root directory to store the 'PupRetrievalTest-CW'-folder (/content/drive/My Drive/PupRetrievalTest-CW).

18 In the PupRetrievalTest-CW'-folder, go to the 'videos'-subfolder. Upload the videos you want to analyze.

19 In the OSF depository, go to the folder 'Notebook' and run the notebook 'DeepLabCut tracking: CW-tracking model' on Google Colab by pressing 'Open in Colab' → 'Runtime' → 'Change runtime type' → select 'GPU'.

Box 1 | Expanding the DLC model: ● Timing 1-2d
Procedure

CRITICAL Pose-estimation can be adapted to experiments involving paternal, nulliparous females, brown-coat mice or mice with head posts (for optogenetics, electrophysiology and/or drug infusion).

1. After the application of the developed neural network, pose-estimation could be unsatisfactory due to:

- Jumps: one or more body parts jump a range of pixels from the last frame
- Fitting: one or more body parts violate a state-space model fit to the time series.

2. Install DLC preferentially in Anaconda environment. (https://github.com/DeepLabCut/DeepLabCut/blob/master/docs/installation.md)

3. Download the 'PupRetrievalTest-CW'-folder from the Google drive.

   ! CAUTION The 'PupRetrievalTest-CW'-folder should contain the .csv tracking files created in Step 24 of this protocol.

4. In this folder, open the config.yaml file and change the project_path to the correct directory on your computer.

5. Load the project in DLC and navigate to the 'Extract outlier frames'-tab. Select 6-10 some videos with poor tracking and specify the video format (.mp4, .avi). Set shuffle to '1' and the training set index to '0'. Define which algorithm you want to use and click 'ok'.

   ! CAUTION By default, 15 frames violating the chosen algorithm will be extracted. We advise to extract at least 90-150 extra frames (thus 6-10 videos) to expand the network.

6. In your Anaconda environment, you will be asked whether you want to proceed. Type 'yes' and press enter. A notification will appear in the Anaconda environment saying that the frames are in a subdirectory under labeled-data.

7. Go to the ‘Refine’ tab in DeepLabCut and click ‘launch’ to access these folders. Open the folders one-by-one and correct labels if necessary. The pose configuration is shown in Fig. 1a.

   ! CAUTION Per animal, 7 points are annotated, which were chosen to have visually defined landmarks based on murine skeletal anatomy. ‘Center’ was defined as the spinal coordinate right underneath the rib cage, that functions as a hinge as the mouse changes its direction of movement. ‘Spine1’ and ‘Spine2’ were defined as the place of the spine where the limbs attach.

8. After the label refinement step, re-upload the project to your Google drive and change the project_path in config.yaml to: /content/drive/My Drive/PupRetrievalTest-CW

9. In the OSF depository, go to the folder ‘Notebook’ and run the notebook ‘DeepLabCut tracking: CW-expand model’ similarly to Step 18.

10. Run stages 1-3 as explained in Steps 19-21 of this protocol.

11. Create a new dataset by running the code in Stage 4 in Colab (this can take a few minutes).

12. Go to /PupRetrievalTest-CW/dlc-models/iteration-0/PupRetrievalTest-trainset95shuffle1/train and open the pose_cfg.yaml. Change the batch size to 4.

13. Run the cell under Stage 5 and the network training will start. As we are using a batch size of 4, the learning curve generally plateaus around 60k, i.e. your network will lose flexibility. Therefore, we advise to stop training at 65k iterations.

14. The last 40 snapshots will be stored in /PupRetrievalTest-CW/dlc-models/iteration-0/PupRetrievalTest-trainset95shuffle1/train. Check in that folder whether snapshots starting from 35k iterations are present. If they are not present, then they will be in the bin and need to be restored.

    ! CAUTION Every snapshot has 3 corresponding files (.meta, .index, .data-00000-of-00001). It is important that all three files are restored.

15. Evaluate the created models as described in Box 2.

16. After identifying the best ranked model, proceed with Step 24.

20 Run the cell underneath ‘Stage 1 - DeepLabCut installation in Colab’ by clicking the ➸-button. This will take a few minutes.

```
#Install all the dependences - should be ran twice to install everything correctly.
!pip install deeplabcut
!pip install deeplabcut

# Use TensorFlow 1.x:
%tensorflow version 1.x

#GUIs don't work on the cloud, so we supress them:
import os
os.environ["DLClight"]="True"
import deeplabcut
```

21 In stage 2, Colab will be linked to the Google Drive in which the ‘PupRetrievalTest-CW’ folder is stored. After running the cell, an authorization code is demanded. Click on the provided URL to obtain this code, which will lead to a browser window. Here, select the drive account where the PupRetrievalTest-CW’-folder was stored. After giving permission for access, an authorization code is received. Copy this code, paste it in the designated box and press enter.

```
from google.colab import drive
drive.mount('/content/drive')
```

22 CRITICAL STEP Modify the parameters to be set in Stage 3. By default only the video format should be modified by changing ‘VideoType = ‘ to the format of the videos you want to analyze.

```
# PLEASE EDIT THIS:
VideoType = 'mp4' #Choose ‘avi’ or ‘mp4’

#don't edit this:
ProjectFolderName = ' PupRetrievalTest-CW’

videofile path = ['/content/drive/My Drive/'+ProjectFolderName+'/videos/']
videofile_path

path config file = '/content/drive/My Drive/'+ProjectFolderName+'/config.yaml'
path_config_file
```

23 CRITICAL STEP The function in Stage 4 analyzes the novel video. The results are stored in a .hd5 and .csv files in the same directory as the video.

```
deeplabcut.analyze_videos(path_config_file,videofile_path, videotype=VideoType, shuffle=1, save_as_csv=True)
```

24 CRITICAL STEP The last function in Stage 5 is for visualization, which creates videos in .mp4 format with the predicted features. These are also stored in the same directory as the original video. The algorithm should be able to track the animals correctly when they are not covered by the nest.

```
deeplabcut.create_labeled_video(path_config_file,videofile_path,shuffle=1, videotype=VideoType, draw_skeleton=True, color_by='bodypart')
```

CRITICAL STEP Keep in mind that the following steps include outlier corrections. In the folder 'PupRetrievalTest-CW/videos', our labeled videos are included as an example of an example of good tracking with outliers.

CRITICAL STEP In case that the tracking is unsatisfactory, three solutions are possible:

- Features are being tracked on the subjects although they are not stable (jumps in tracking). We advise to use the jump algorithm in the label refinement option in DeepLabCut as explained in Box 1.
- Features are being tracked but not on the subjects (underfitting of the tracking model). We advise to use the fitting algorithm in the label refinement option in DeepLabCut as explained in Box 1.
- Features are not being tracked. We advise to create an entirely novel DeepLabCut tracking dataset. All parameters we used for our model are available on the OSF directory in the document 'DeepLabCut_parameters'. Annotate approximately 500-700 frames and resume this protocol with Step 6 of Box 1.

25 Download all .csv files onto your computer.

Box 2 | Evaluating a new or expanded neural network in DLC ● Timing 60 min
Procedure CRITICAL Pose-estimation can also be adapted to experiments involving paternal, nulliparous females, brown-coat mice or mice with head posts. 1. Evaluate the network by considering the pixel error for all body parts. First, change 'p-cutoff' in the config.yaml to 0.1. Here, choose a model that is relatively low before p-cutoff, especially in the test data set. Further, the discrepancy between with or without cutoff pixel errors should be relatively

stable since the model should be certain of its pose-estimations without high error rate.

2. Again, evaluate the networks for all body parts with a p-cutoff of 0.5. There needs to be a discrepancy between the with or without cutoff values, since 0.5 calculates pixel errors for relatively high probability pose estimations.

3. Evaluate the network by only taking the pixel error of the nose of the mother into account with a p-cutoff of 0.1.

4. Combine the results of these evaluations to choose the best model. Create labeled videos to see whether performance is satisfying.

Data processing: behavioral classification using SimBA ● Timing

26 Download the SimBA folder on the repository (https://osf.io/rwhtd/) in the folder 'Dropbox: automatedPRT/Behavioral_classification – SimBA'.

27 In SimBA, load the project_config.ini file in /PupRetrievalTest_CW_BehavioralClassification/project_folder. The 'Load project'- window in SimBA will now open.

28 In the 'Load project'- window continue as follows. Import the videos to analyze and import tracking data (.csv) files created with DLC.

29 Proceed to the tab 'Video parameters' and in the 'Known distance (mm)' box fill in '267', if you used a type II animal cage. We use the length dimension of the top since the bottom part is covered with bedding. Then click 'Autopopulate table' and 'Set video parameters'. On the right, you see the column 'Get coord' with every video in your folder. You can click on the video and the first frame of your video will pop up. Double-click (left button) on the cage tag hinges (Fig. 1c) and its parallel over the length dimension. Follow the instructions in SimBA.

! CAUTION SimBA provides the option to duplicate this distance over all included videos. Only do this if certain there are no differences in angle or cropping.

30 Proceed to the tab outlier correction. Go to settings and use for each animal the 'Nose – Spine1' body parts to perform outlier correction. We use 2.5 and 4 as movement and location criterions, respectively. Use the more robust median for eventual outliers.

31 In the tab 'ROI', define ROI shapes by adding 1 circle and 1 polygon and click 'Show Shape Definitions Table'. Name the circle 'corenest' and the polygon 'nest'. Do not enter a number in the radius box after circle. Click 'Set Shape Definitions' and a 'ROI table'-window with all videos will pop up. We advise to draw ROIs for each file individually since these are important for outcome variables. For the ROI 'corenest', double click in the center of where your pups are laying together in the nest, and secondly click the outer border of where pups are located. Define the ROI 'nest' by double clicking on the borders of the nest with approximately a 15 mm margin (Fig 1d, example).

32 Click 'Analyze ROI' data and go to the tab 'Extract features'. First, run the 'extract features' button and afterwards 'Append ROI data to features'. A window will pop up asking which body parts to use. For the mother choose 'Nose', and for the pup choose 'Spine1'.

33 Proceed to the tab 'Run machine model' and under the section Run Machine Model, then click on 'Model settings'. A 'Select model to run'-window will open. Click 'Browse file' and per classifier select the correct .sav-file in (…)/PupRetrievalTest_CW_BehavioralClassification/models/generated_models. The threshold for the behavioral classifiers approach, carry and dig are respectively 0.47, 0.47 and 0.24. Minimum bout lengths are 500 ms for approach and 1000 ms for dig, and 200 ms for carry. Click 'Set model(s)' and in the 'Run machine model' tab, click 'Run RF Model'.

34 Go to the tab 'Add-ons' and click 'Pup retrieval – Analysis Protocol 1. Use the default parameters as shown in Figure 2 and click run. In the folder (…)/PupRetrievalTest_CW_BehavioralClassification/project_folder/logs, a .csv-file named 'Pup_retrieval_date' is created. This file contains the test output including 'frame to retrieval', 'latency to retrieval', Boolean value whether pup was retrieved (0-1), Boolean value whether mother retrieved pup into core nest (0-1), and per behavioral classifier: the total time spent on the behavior before retrieval, latency to first event, number of events before retrieval, mean duration and mean interval duration.

Troubleshooting

Troubleshooting advice can be found in Table 1.

Table 1 | Troubleshooting table

Step	**Problem**	**Possible reason**	**Solution**
	Deviant pup/nest weight on P5 (mean-2SD)		Do not include in analysis

9	Interference during test procedure	Unplanned event (such as a loud noise) occurred while sampling data	Do not include trial in analysis
21	Pose-estimation model does not fit on the subjects in the video. In the labeled video, features are repeatedly plotted on the same spot in the margin of the cage	An object is present at the side of the cage	Crop the video to exclude the object, if possible.
21	Pose-estimation model does not fit on the subjects in the video	Video is recorded too far or too close inside the cage, or video is not sampled centered top-down	This cannot be solved for the current pose estimation model. You can expand the neural network using Box 1
21	Pose-estimation model does not fit on the subjects in the video	The test animal is looking different from a C57bl6j mother used to create the pose estimation model or the pup is smaller/larger than P5 C57bl6j pups	Expand the neural network using Box 1
27	Pose estimation appears unstable after outlier	The criterion value for location and/or movement	Re-run outlier correction with a higher

	correction	outlier correction is too small	criterion value

Timing

Step 1, Stage I, sexing pups and culling: ~5 min/nest

Steps 2-5, Stage II, preparation of test including transport and habituation: ~70 min

Steps 6-10, Stage III, pup retrieval test sampling: variable, ranging from 20 to 40 min per nest

Step 11, Stage IV, pre-processing of video files: variable, approximately 5 to 10 min per video

Steps 12-13, Stage V, preparation of videos for analysis in Colab: 5 min,

Steps 14-19, Stage VI, preparation of Colab: 5-10 min

Step 20, Stage VII, video analysis in Colab: variable, > 1 min per video

Step 21, Stage VIII, video analysis in Colab: variable, ~ 1 min per video

Steps 23-25, Stage IX, preparation SimBA analysis: ~ 5 min

Step 26, Stage X, distance definition in SimBA: variable, ~ 5 to 10 min

Step 27, Stage XI, Outlier correction in SimBA: variable, < 1 min per video

Step 28, Stage XII, ROI definition in SimBA: variable, ~ 2 min per video

Steps 29-30, Stage XIII, ROI definition in SimBA: variable, ~ 30-60 min

Step 28, Stage XIV, Extract parameters: variable, ~ 10 min

Materials

Biological materials

. Breeding pairs

Pairs of primiparous C57BL/6JRj mice were purchased from Janvier Labs (Le Genest-Saint-Isle, France), and maintained for time-controlled breeding in standard type II cages. Males were only present in the home cage the night of mating and females were housed individually from gestation through weaning (P28).

. Laboratory-bred mouse pups on P5

For the test, each nest should contain a fixed number of pups with a balanced male:female ratio. On day of birth (P0), sex differentiation is based on the anogenital dark spot in males. Nests should be reduced to 6 pups with a preferable 3:3 male:female ratio (47-48).

! CAUTION Experiments using rodents must conform to local and national regulations. All our experiments were reviewed by the animal ethics committee of the University of Leuven (Belgium), in accordance with European Community Council Directive 86/609/EEC.

Reagents

. 70% (vol/vol) Ethanol solution

Equipment

. Clean gloves.

. Standard Type II cage, bottom dimensions 225 x 167 mm, top dimensions 267 x 208 mm and height 140 mm.

. Styrofoam box: 370 x 300 mm and height 330 mm.

. Stable heat pad: we use an IPower reptile heat pad with digital thermostat for small animals.

. Glass cup to place the pup in on the heat pad: we use a borosilicate 3.3 glass crystallizing dish (50 mm diameter, 30 mm height, VWR, cat. no. 216-0063).

. Balance (to weigh pups after testing)

. Stopwatch (to time maximum trial time)

. Tape to attach the thermometer probe of your heat pad to the glass cup.

Software

. Operating system: Windows (10).

! CAUTION SimBA is not compatible with other systems.

. Google drive

. Anaconda: free and open-source distribution of the Python programming language (https://www.anaconda.com/). SimBA is written in Python 3.6 (https://www.python.org/) and is not compatible with Python 2.

. SimBA: free and open-source toolbox available at https://github.com/sgoldenlab/simba. The code is written for Python 3.6.

. Microsoft Office Excel

Hardware

. Computer, The SimBA toolbox can be used on modern desktop workstations as well as laptops.

. Video camera: we use Foscam C2 IP-camera and online interface (EUport, Wageningen), at a resolution of 1280x720px and 10-30 frames per second (fps). The camera was fixed in the center of the Styrofoam box, approximately 50 cm above the floor.

Biological materials setup

. Animal housing

Adult mice (8-10 weeks old) were group-housed under standard housing conditions. Mice were kept at 12/12 hour light-dark cycle (lights on at 7 AM), water and food *ad libitum*, conditioned rooms (22°C, humidity 30%). A 14-day acclimation period before mating avoided transportation distress to affect the dams. Nests were transported to the test room at 8 AM and testing took place between 9 AM and 10 AM.

Equipment setup

. Testing setup

For testing, the home cage without grid is placed inside the Styrofoam box without lid to create a visually isolated environment. A camera is fixated 50 cm above the center of the floor of the Styrofoam box to record top-down. The heat pad is positioned next to the Styrofoam box with the glass cup placed on it. The thermometer probe should be taped to the bottom of the glass cup. The camera is connected to a laptop distanced approximately 1.5 m from the Styrofoam box. The balance is positioned between the laptop and the Styrofoam box.

3.7 Declarations

Availability of data and material: All our annotated images and videos, tracking models and behavioral classifiers are available on: https://doi.org/ 10.17605/OSF.IO/RWHTD. For further inquiries, please contact the corresponding author.

Competing interests: The authors declare no competing interests.

Authors' contributions: C.W. and R.D. designed the experimental strategy. C.W. and V.A.O-S. optimized experimental procedures. C.W., W.G., S.N. and S.G. conceptualized and wrote the code. C.W. labeled the data. C.W. and R.D. wrote the manuscript with input from all authors.

Acknowledgements: We thank Annemone Mia Gorssen Winters for exemplary behavior. This study was funded by an SB PhD fellow- ship (1S05818N) of the Research Foundation Flanders (FWO) to Carmen Winters. The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

3.8 References

(1) Braungart-Rieker, J. M., Garwood, M. M., Powers, B. P., & Wang, X. (2001). Parental sensitivity, infant affect, and affect regulation: Predictors of later attachment. Child development, 72(1), 252-270.

(2) Shin, H., Park, Y. J., Ryu, H., & Seomun, G. A. (2008). Maternal sensitivity: A concept analysis. *Journal of Advanced Nursing*, *64*(3), 304-314.

(3) Johnson, K. (2013). Maternal-Infant Bonding: A Review of Literature. *International Journal of Childbirth Education*, *28*(3).

(4) Numan, M. (2020). *The parental brain: Mechanisms, development, and evolution*. Oxford University Press.

(5) Sullivan, R., Perry, R., Sloan, A., Kleinhaus, K., & Burtchen, N. (2011). Infant bonding and attachment to the caregiver: insights from basic and clinical science. *Clinics in perinatology*, *38*(4), 643-655.

(6) Tarantino, L. M., Sullivan, P. F., & Meltzer-Brody, S. (2011). Using animal models to disentangle the role of genetic, epigenetic, and environmental influences on behavioral outcomes associated with maternal anxiety and depression. *Frontiers in psychiatry*, *2*, 44.

(7) Rilling, J. K., & Young, L. J. (2014). The biology of mammalian parenting and its effect on offspring social development. *Science*, *345*(6198), 771-776.

(8) Hahn, M. E., & Lavooy, M. J. (2005). A review of the methods of studies on infant ultrasound production and maternal retrieval in small rodents. *Behavior genetics*, *35*(1), 31-52.

(9) Hill, R., & Flanagan, J. (2020). The Maternal–Infant Bond: Clarifying the Concept. *International journal of nursing knowledge*, *31*(1), 14-18.

(10) Larsen, C. M., & Grattan, D. R. (2010). Prolactin-induced mitogenesis in the subventricular zone of the maternal brain during early pregnancy is essential for normal postpartum behavioral responses in the mother. *Endocrinology*, *151*(8), 3805-3814.

(11) Cohen, L., Rothschild, G., & Mizrahi, A. (2011). Multisensory integration of natural odors and sounds in the auditory cortex. *Neuron*, *72*(2), 357-369.

(12) Marlin, B. J., Mitre, M., D'amour, J. A., Chao, M. V., & Froemke, R. C. (2015). Oxytocin enables maternal behaviour by balancing cortical inhibition. *Nature*, *520*(7548), 499-504.

(13) Tasaka, G. I., Feigin, L., Maor, I., Groysman, M., DeNardo, L. A., Schiavo, J. K.,... & Mizrahi, A. (2020). The temporal association cortex plays a key role in auditory-driven maternal plasticity. *Neuron*, *107*(3), 566-579.

(14) Smart, J. L., & Preece, J. (1973). Maternal behaviour of undernourished mother rats. *Animal Behaviour*, *21*(3), 613-619.

(15) Giordano, A. L., Johnson, A. E., & Rosenblatt, J. S. (1990). Haloperidol-induced disruption of retrieval behavior and reversal with apomorphine in lactating rats. *Physiology & behavior*, *48*(1), 211-214.

(16) Mann, P. E., Pasternak, G. W., & Bridges, R. S. (1990). Mu1 opioid receptor involvement in maternal behavior. *Physiology & behavior*, *47*(1), 133-138.
(17) Zhang, Z., Dawson, P. A., Piper, M., & Simmons, D. G. (2019). Postnatal N-acetylcysteine administration rescues impaired social behaviors and neurogenesis in Slc13a4 haploinsufficient mice. *EBioMedicine*, *43*, 435-446.
(18) Zhao, C., & Li, M. (2009). The receptor mechanisms underlying the disruptive effects of haloperidol and clozapine on rat maternal behavior: A double dissociation between dopamine D2 and 5-HT2A/2C receptors. *Pharmacology Biochemistry and Behavior*, *93*(4), 433-442.
(19) Young, D. M., Schenk, A. K., Yang, S. B., Jan, Y. N., & Jan, L. Y. (2010). Altered ultrasonic vocalizations in a tuberous sclerosis mouse model of autism. *Proceedings of the National Academy of Sciences*, *107*(24), 11074-11079.
(20) Grabrucker, S., Pagano, J., Schweizer, J., Urrutia-Ruiz, C., Schön, M., Thome, K.,... & Boeckers, T. M. (2021). Activation of the medial preoptic area (MPOA) ameliorates loss of maternal behavior in a Shank2 mouse model for autism. *The EMBO Journal*, e104267.
(21) Allan, A. M., Chynoweth, J., Tyler, L. A., & Caldwell, K. K. (2003). A mouse model of prenatal ethanol exposure using a voluntary drinking paradigm. *Alcoholism: Clinical and Experimental Research*, *27*(12), 2009-2016.
(22) Pardon, M. C., Gérardin, P., Joubert, C., Pérez-Diaz, F., & Cohen-Salmon, C. (2000). Influence of prepartum chronic ultramild stress on maternal pup care behavior in mice. *Biological psychiatry*, *47*(10), 858-863.
(23) Andrieux, A., Salin, P., Schweitzer, A., Bégou, M., Pachoud, B., Brun, P.,... & Job, D. (2006). Microtubule stabilizer ameliorates synaptic function and behavior in a mouse model for schizophrenia. *Biological psychiatry*, *60*(11), 1224-1230.
(24) Carini, L. M., Murgatroyd, C. A., & Nephew, B. C. (2013). Using chronic social stress to model postpartum depression in lactating rodents. *Journal of visualized experiments: JoVE*, (76).
(25) Wahlsten, D., Metten, P., Phillips, T. J., Boehm, S. L., Burkhart-Kasch, S., Dorow, J.,... & Crabbe, J. C. (2003). Different data from different labs: lessons from studies of gene–environment interaction. *Journal of neurobiology*, *54*(1), 283-311.
(26) Worley, N. B., Djerdjaj, A., & Christianson, J. P. (2019). DeepLabCut Analysis of Social Novelty Preference. *bioRxiv*, 736983.
(27) von Ziegler, L., Sturman, O., & Bohacek, J. (2020). Big behavior: challenges and opportunities in a new era of deep behavior profiling. *Neuropsychopharmacology*, *46*(1), 33-44.
(28) Krishnan, K., Casenhiser, D. M., & Stevenson, P. K. (2020). Systematic analysis of goal-related movement sequences during maternal behavior in a female mouse model for Rett syndrome. *bioRxiv*.
(29) Hong, W., Kennedy, A., Burgos-Artizzu, X. P., Zelikowsky, M., Navonne, S. G., Perona, P., & Anderson, D. J. (2015). Automated measurement of mouse social behaviors using depth sensing, video tracking, and machine learning. *Proceedings of the National Academy of Sciences*, *112*(38), E5351-E5360.
(30) Leng, X., Wohl, M., Ishii, K., Nayak, P., & Asahina, K. (2020). Quantitative comparison of Drosophila behavior annotations by human observers and a machine learning algorithm. *bioRxiv*.
(31) Nilsson, S. R., Goodwin, N. L., Choong, J. J., Hwang, S., Wright, H. R., Norville, Z.,... & McLaughlin, R. J. (2020). Simple Behavioral Analysis (SimBA): an open source toolkit for computer classification of complex social behaviors in experimental animals. *BioRxiv*.
(32) Tuyttens, F. A. M., de Graaf, S., Heerkens, J. L., Jacobs, L., Nalon, E., Ott, S.,... & Ampe, B. (2014). Observer bias in animal behaviour research: can we believe what we score, if we score what we believe?. *Animal Behaviour*, *90*, 273-280.
(33) Goodwin, N. L., Nilsson, S. R., & Golden, S. A. (2020). Rage Against the Machine: Advancing the study of aggression ethology via machine learning. *Psychopharmacology*, 1-20.
(34) Fleming, A. S., & Kraemer, G. W. (2019). Molecular and genetic bases of mammalian maternal behavior. *Gender and the Genome*, *3*, 2470289719827306.
(35) Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature neuroscience*, *21*(9), 1281-1289.

(36) Sturman, O., von Ziegler, L., Schläppi, C., Akyol, F., Privitera, M., Slominski, D.,... & Bohacek, J. (2020). Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions. *Neuropsychopharmacology*, *45*(11), 1942-1952.
(37) Sturman, O. (2021). *Profiling stress-induced changes in approach avoidance behaviour in mice* (Doctoral dissertation, ETH Zurich).
(38) Kuhn, M. (2009). The caret package. *Journal of Statistical Software*, *28*(5).
(39) Lauer, J., Zhou, M., Ye, S., Menegas, W., Nath, T., Rahman, M. M.,... & Mathis, A. (2021). Multi-animal pose estimation and tracking with DeepLabCut. *bioRxiv*.
(40) van Dam, E. A., Noldus, L. P., & van Gerven, M. A. (2020). Deep learning improves automated rodent behavior recognition within a specific experimental setup. *Journal of neuroscience methods*, *332*, 108536.
(41) Brownlee, J. (2018). *Better deep learning: train faster, reduce overfitting, and make better predictions*. Machine Learning Mastery.
(42) Esposito, G., Hiroi, N., & Scattoni, M. L. (2017). Cry, baby, cry: expression of distress as a biomarker and modulator in autism spectrum disorder. *International Journal of Neuropsychopharmacology*, *20*(6), 498-503.
(43) Bridges, R., Zarrow, M. X., Gandelman, R., & Denenberg, V. H. (1972). Differences in maternal responsiveness between lactating and sensitized rats. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *5*(2), 123-127.
(44) Kobayashi, K., Matsushita, S., Shimizu, N., Masuko, S., Yamamoto, M., & Murata, T. (2021). Automated detection of mouse scratching behaviour using convolutional recurrent neural network. *Scientific reports*, *11*(1), 1-10.
(45) Heun-Johnson, H., & Levitt, P. (2016). Early-life stress paradigm transiently alters maternal behavior, dam-pup interactions, and offspring vocalizations in mice. Frontiers in Behavioral Neuroscience, 10, 142.
(46) Pennington, Z. T., Dong, Z., Feng, Y., Vetere, L. M., Page-Harley, L., Shuman, T., & Cai, D. J. (2019). ezTrack: An open-source video analysis pipeline for the investigation of animal behavior. *Scientific reports*, *9*(1), 1-11.
(47) Maestripieri, D., & Alleva, E. (1990). Maternal aggression and litter size in the female house mouse. *Ethology*, *84*(1), 27-34.
(48) Alleva, E., Caprioli, A., & Laviola, G. (1989). Litter gender composition affects maternal behavior of the primiparous mouse dam (Mus musculus). *Journal of comparative Psychology*, *103*(1), 83.
(49) Reis, A. R., De Azevedo, M. S., De Souza, M. A., Lutz, M. L., Alves, M. B., Izquierdo, I.,... & Lucion, A. B. (2014). Neonatal handling alters the structure of maternal behavior and affects mother–pup bonding. *Behavioural brain research*, *265*, 216-228.
(50) de Sá Couto-Pereira, N., Ferreira, C. F., Lampert, C., Arcego, D. M., Toniazzo, A. P., Bernardi, J. R.,... & Dalmaz, C. (2016). Neonatal interventions differently affect maternal care quality and have sexually dimorphic developmental effects on corticosterone secretion. *International Journal of Developmental Neuroscience*, *55*, 72-81.
(51) Champagne, F. A., Curley, J. P., Keverne, E. B., & Bateson, P. P. (2007). Natural variations in postpartum maternal care in inbred and outbred mice. *Physiology & Behavior*, 91(2-3), 325-334.

4 BAMBI: A new method for automated assessment of bidirectional early-life interaction between maternal behavior and pup vocalization in mouse dam-pup dyads*

* Adapted from Winters, C., Gorssen, W., Wöhr, M., & D'Hooge, R. (2023). BAMBI: A new method for automated assessment of bidirectional early-life interaction between maternal behavior and pup vocalization in mouse dam-pup dyads. Frontiers in Behavioral Neuroscience, 17. https://doi.org/10.3389%2Ffnbeh.2023.1139254

4.1 Abstract

Vital early-life dyadic interaction in mice requires a pup to signal its needs adequately, and a dam to recognize and respond to the pup's cues accurately and timely. Previous research might have missed important biological and/or environmental elements of this complex bidirectional interaction, because it often focused on one dyadic member only. In laboratory rodents, the Pup Retrieval Test (PRT) is the leading procedure to assess pup-directed maternal care. The present study describes BAMBI (Bidirectional Automated Mother-pup Behavioral Interaction test), a novel automated PRT methodology based on synchronous videorecording of maternal behavior and audiorecording of pup vocalizations, which allows to assess bidirectional dam-pup dyadic interaction. We were able to estimate pup retrieval and pup vocalization parameters accurately in 156 pups from 29 dams on postnatal days (PND) 5, 7, 9, 11 and 13. Moreover, we showed an association between number of emitted USVs and retrieval success, indicating dyadic interdependency and bidirectionality. BAMBI is a promising new automated homecage behavioral method that can be applied to both basic and preclinical studies investigating complex phenotypes related to early-life social development.

4.2 Introduction

Neonatal mouse pups depend on their dam for nutrition, thermoregulation, and protection (1). They produce acoustic signals to communicate their vital needs, and particularly ultrasonic vocalizations (USVs) are essential to evoke maternal care behaviors, such as retrieval in pups that have dangerously strayed from the nest (2, 3). Early-life USVs can be used to study the genetic and neural basis of early-life communication and to assess early-life communicative defects and their impact on social development (4-7). Moreover, early-life maternal care has been shown to affect the pup's physical and functional development in a very broad sense (8-12).

Establishing effective bidirectional communication does not only require that the pup is able to signal distress effectively, but also that the dam is able to perceive, process and respond accurately and timely to these cues (3, 12). The separation-induced USV test has been used to assess the quantity and quality of pup USVs after separation from its dam and litter (4), but it is essentially a unidirectional behavioral assay that focusses on the pup. On the other hand, assays such as the pup retrieval test (PRT) or USV playback tests center on maternal behaviors, such as search and retrieval (13-17). Some studies implemented both unidirectional procedures separately, but assessed statistical association afterwards (18-20). Combining both procedures in one behavioral assay has several advantages. First, the behavioral readouts can be sampled in a single assay, which reduces workload, and microenvironmental variability, originating from differences in animal transportation and handling, for example (21-22). Second, communication and social competence can be investigated as a bidirectional process in the same animals (23). Third, the complex interaction between deficits in dam and pup can be investigated (24). The latter is particularly important in rodent models of disorders with early-life communication deficits, such as autism or fetal alcohol syndrome (24-25). Therefore, we present BAMBI (Bidirectional Automated Mother-pup Behavioral Interaction test), a combined, automated approach to assess early-life communicative bidirectionality in laboratory mice. The automated PRT, as described in Winters et al. (26), was expanded with simultaneous recording and automated detection of pup USVs.

4.3 Materials and methods

4.3.1 Animal housing and breeding

C57BL/6J mice from Janvier Labs (Le Genest-Saint-Isle, France) and the KU Leuven Animal Facility (Leuven, Belgium) were time-specifically bred and kept at a 12/12-hour light-dark cycle (lights on at 7 AM), with ad libitum water and food in conditioned rooms (22°C, humidity 30%). The morning after mating was considered as gestational day (GD) 0.5. On GD0.5, dams were housed individually for the remainder of the pregnancy and pregnancies were confirmed between GD7.5 and GD10.5 by recording weight evolution based on Heyne et al (44). All experimental procedures were approved by the Animal Ethics Committee of KU Leuven (P028/2018), in accordance with European Community Council Directive 86/609/EEC, the ARRIVE guidelines and the ILAR Guide to the Care and Use of Experimental Animals.

4.3.2 Experimental groups

In compliance with the reduction principle, mice for the present methodological work were obtained from an independently designed pharmacological study, in which pregnant dams were injected with valproic acid sodium salt (VPA) in order to generate pups representing a neurodevelopmental model of autism. Pups were pharmacologically treated to attempt a rescue of the behavioral impairment. More specicifically, pregnant dams (N=44) received a single subcutaneous injection with 60 mg/ml VPA (Sigma Aldrich, Taufkirchen, Germany) dissolved in saline solution on GD12.5. The day of birth was considered as PND0. To standardize nest composition, nests were culled to six pups on PND0 and every nest needed to have at least four pups with both sexes present. These restrictions resulted in 29 dams with viable progeny and a total of 156 pups for further testing. Pups were subcutaneously injected daily from PND1-7 with a low (0.5 mg/kg) or a high (2 mg/kg) dose of THIQ (N-[(1R)-1-[(4-Chlorophenyl)methyl]-2-[4-cyclohexyl-4-(1H-1,2,4-trazol-1-ylmethyl)-1-piperidinyl]-2-oxoethyl]-1,2,3,4-tetrahydro-3-isoquinolinecarboxamide; Bio-techne, Abingdon, UK) or a PBS-DMSO control vehicle (doses adapted from Mastinu et al. (27)). THIQ was dissolved in PBS and 3,5% DMSO. In total, 9 litters (48 pups) were injected with low THIQ dose, 10 litters (50 pups) with a high THIQ dose and 10 litters (58 pups) with the PBS-DMSO control vehicle. The pharmacological effects are not the focus of the present study and will be described in a separate study. In the current study the aim is to present a proof of principle demonstration of the feasibility and validity of a new automated method for behavioral testing of early life mother-pup bidirectional interactions. Since pharmacological effects were not relevant for the present study, we employed a general linear model (GLM) in which drug effect was set as fixed factor, in order to correct for drug effects (see Calculation of parameters and statistics), and the three drug groups (low-dose THIQ, high-dose THIQ and vehicle) were pooled into one group. For the present study, animals were divided into the following groups: dams (n = 29) and

pups (n = 156). For the pup sex effect analysis, animals were divided into three groups: dams (n = 29), male pups (n = 72) and female pups (n = 84). For the subsequent analyses investigating the general behavioral interactions between mother and pups, we corrected pup sex effects through a GLM model in which pup sex was set as fixed factor and we pooled male and female pups together. Mice were tested at 5 time-points: pup postnatal day (PND) 5, 7, 9, 11 and 13.

Figure 1: Image of the BAMBI testing setup
BAMBI test was performed in the homecage and included a cup on PND7-13 to prevent the pups from crawling back into the nest. An ultrasound microphone was placed approximately 5 cm above the test pup's corner in order to minimize interference from USVs emitted by the pups in the nest (in the opposite corner). A videocamera was mounted above the homecage.

4.3.3 Pup retrieval test (PRT) protocol

The PRT was performed as described previously by Winters et al. (26). Briefly, the test is performed in the homecage which is placed inside a Styrofoam box (370 x 300 x 330 mm) to create a visually isolated environment. A single pup was removed from the nest and placed in a clean, glass cup pre-heated to 35°C using a heating pad. A trial was started by a beep when the dam was on the nesting site. Hereafter, the pup was placed in the furthest corner from the nest. Trials had a maximum duration of 100 seconds after the beep, and when the pup was not retrieved within this time, it was returned to the nest by the experimenter. On PND7-13, since pups had developed more mature motor skills, they were kept from crawling back into the nest by placing them in a cup (90 mm diameter and 55 mm height) as described by Esposito et al. (28). Per dam, the PRT was repeated six times on PND5, and due to practical limitations four times per dam on PND7-13. For all test ages, pup sex was counterbalanced per dam and pups were not marked during this test to avoid odor interference. Pups thus could not be identified, meaning that a pup might have been tested more than once. Maternal trial sequence was defined as the order of trials within a dam on a specific testing day. During each trial, PRT performance was scored by the experimenter performing the test (two experimenters in total) for latency to

retrieval (s) and retrieval success (0 = not retrieved, 1 = retrieved).

4.3.4 Ultrasonic recording and pre-processing

Pup USVs were recorded using an ultrasound microphone (Dodotronic Ultramic UM250K, Rome, Italy) connected to a personal computer with Avisoft SASLab Lite software (Avisoft, Bioacoustics, Berlin, Germany). The microphone was placed approximately 5 cm above the pups' corner or cup to minimize interference by USVs emitted by the pups in the nest. USVs were recorded for 100s after the beep that introduced the start of a trial, with a sampling rate of 250 kHz and 16 bits. Audacity® open-source software (Version 3.1.3, https://www.audacityteam.org/) was used to remove DC (direct current) offset and a equalization (EQ) curve filter was used to remove all signal below 40 kHz (**Figure 2**).

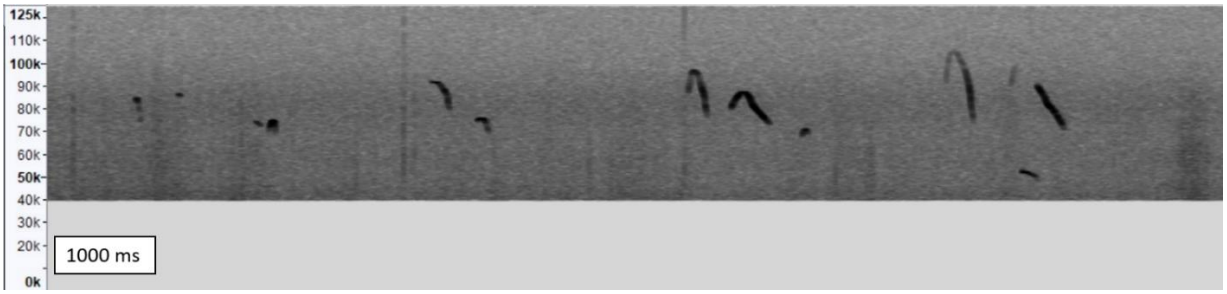


Figure 2: Exemplary spectrogram of ultrasonic vocalizations emitted by pups.
Frequency (Hz) is shown on the y-axis over a fixed time of 1000 ms on the x-axis.

4.3.5 Synchronization of USV and behavioral pup retrieval recording

A Foscam C2 IP-camera (EUport, Wageningen) was mounted over the homecage to record maternal behavior. Per dam, one video was recorded including six PRT trials on PND5 or four PRT trials on PND7-13. A PRT trial started after the dam was back on the nest, and a beep, manually played on a smartphone, introduced the placement of the pup in the furthest corner of the homecage. Synchronization of behavioral and audio data was done by identifying the beep using frame-precision Shotcut software (www.shotcut.org, version 22.10.25, Meltytech, LLC). This means video or audio signals can be split precisely per individual frame. Here, videos were recorded using a frame rate of 30 frames per second (fps) and videos were split before the first frame in which the beep was audible. The end of the video was defined 100 seconds after the first frame with an audible beep. Similarly, the audio recordings were recorded in Avisoft with a sampling rate of 250 kHz, whereas Shotcut was used to remove signal before the beep using a frame rate of 25 fps. Again, the start of the recording was defined as the sampling point before the first fragment in which the beep was audible. For USV recordings, the end was not defined as Avisoft automatically ended sampling after 100s.

4.3.6 Automated detection of pup USVs using DeepAudioSegmenter

USV detection was performed using a custom-build model with DeepAudioSegmenter (DAS, 29). DAS

0.26.7 was installed in an Anaconda environment with Python 3.9.13. and training was performed using the DAS notebook on Google Colaboratory (COLAB, (29)). Thirty-two audio recordings were pseudo-randomly selected across two different experiments at five ages (postnatal day 5, 7, 9, 11 and 13) and both sexes. Across all recordings, 2189 pup vocalizations were manually annotated as segments which is equivalent to 67.7 seconds of pup USVs. Per audio recording, 80% of the annotated USVs were assigned to the training dataset, 10% to the testing dataset and 10% to the validation dataset. A pup USV network was trained in Google COLAB and structural training parameters were chosen based on Steinfath et al. (29) and can be found in Supplementary File 1. DAS automatically stops training as the validation loss of the model has not improved in 20 epochs (29). The pup vocalization model did not improve after 44 epochs and performance of this detection model was assessed using precision, recall, F1 scores and overall accuracy. Precision is the percentage of "true cases" per "detected cases". Recall on the other hand is the percentage of "true cases" per "manually annotated cases". The F1 score is the harmonic mean of precision and recall. Data were post-processed for quality control using a custom-build R script to resolve inaccuracies. In the videos, the time was recorded between the first detectable beep segment and the first frame where the hand of the researcher was completely out of the setup after placing the pup in it. All detected USVs were removed from the recording during this time interval.

4.3.7 Expanded body part tracking

The resulting dataset included 212 frames and was used to re-train the original network from Winters et al. (26). DeepLabCut 2.2b8 (DLC, 30) was installed in an Anaconda environment with Python 3.7.7 on a laptop equipped with an Intel Core i5-8350U CPU and 8 GM RAM and Windows 10 64-bit operating system. Training, evaluation and analysis of the expanded model was performed using DLC in Google COLAB (https://colab.research.google.com/).

4.3.8 Learning strategy and performance evaluation of the PRT pose estimation model

The PRT dam–pup tracking model developed by Winters et al. (26) was trained to track only PND5 pups in a homecage without a cup. As C57BL/6J pup body shape changes significantly between PND5 and PND13, and the use of a cup is a significant context change that elicits different maternal poses, the model needed to learn these changes. A two-phase hybrid learning strategy was used similar to Gorssen et al. (31). In the first phase, fourteen extra single trail video recordings were selected because of their variability in pup age and/or modulated homecage environment. Fifteen frames per video were extracted using k-means clustering in DLC and labeled manually. Additionally, using the original model ten extra outlier frames per video were extracted using the DLC 'jump' algorithm. Labels in these outlier frames were manually refined and frames were only annotated if both dam and pup were visible. The resulting dataset included 212 frames and the original PRT model was retrained with a 95:5

train/test ratio using the same features as Winters et al. (26). The model was trained for 47 000 iterations and had a mean pixel error over all body parts of 4.29 px for the training dataset and 14.35 px for the test dataset.

In the second training phase, all original labeled data was combined with the data from the first training phase. The output model from the first training phase was then re-trained using the entire dataset with a 95:5 train/test ratio. After 18 000 iterations, the model had a mean pixel error over all body parts of 6.34 px for the training dataset and 10.11 px for the test dataset. Applying a p cut-off (p = 0.10) improved mean pixel error on the training dataset to 5.54 pixels (or 1.96 mm), and 8.82 pixels (or 3.13 mm) for the test dataset. Average pixels per millimeter did differ between the original dataset and the data used to extend the dataset. Distance calculation, performed in Simple Behavioral Analysis (SimBA, 32) as described by Winters et al. (26), showed an average of 2.27 (SD=0.3) px/mm in the original dataset, and an average of 2.87 (SD=0.16) px/mm for the newly annotated data. A custom-build R script was used to post-process the data (quality control) and to estimate retrieval time. First, a time correction was applied to ensure tracking started at the precise moment the pup was placed in the nest. Hereafter, the rolling median (90 frames) of the distance of pup to the nest was calculated to correct for inaccurate tracking in the first seconds of the PRT. The first frame where the rolling median >85 mm was determined. If this was not the first frame, the distance of pup to nest for all previous frames was set to 85 mm, as pups started at least >85 mm from the nest at the start of PRT. Frames with a mean pup tracking probability over all bodyparts <0.01 were discarded, as these estimates were considered unreliable. Next, a smoothing algorithm was used to approximate the distance of a pup to the nest using the *stat_smooth* function in R (loess method) with a smoothing factor of 0.25. Observed values deviating more than 15 mm from the smoothing estimate were set to missing. After these quality control steps, retrieval time was estimated as the first frame a pup entered the nest.

4.3.9 Calculation of parameters and statistics

A custom-build R-script (RStudio, Inc., Boston, MA) was used to allow direct comparison between parameters of video and audio analysis. Mean USV duration was calculated as the mean of all USVs emitted by the same pup within one trial. USV rate before retrieval was calculated as shown below:

$$USV\ rate\ (\frac{USVs}{s}) = \frac{Number\ USVs\ before\ retrieval}{Retrieval\ time\ (s)}$$

Statistical analyses were performed using the GLM package in R for (binomial) regression models and survival package in R for survival analysis via multivariate Cox regression for the trait retrieval success. All models were corrected for USV rate or average USV duration (covariate), sex (fixed effect), day of testing (fixed effect), maternal trial (covariate) and experimental condition (fixed effect).

4.4 Results

4.4.1 Performance evaluation of DAS audiodetection

The USV detection algorithm accomplished an overall accuracy of 99.7%. Noise was predicted with a precision of 99.8 %, a recall of 99.9% and a F1 score of 99.9%. Pup USVs were predicted with a precision of 94.3%, a recall of 90.2% and a F1 score of 92.2%.

4.4.2 Validation of retrieval parameters

To validate the performance of the automated PRT, automatically estimated retrieval times were compared with manual recordings. Retrieval success was estimated with an accuracy of 90.4% (95% CI=87.9 – 92.5), a sensitivity of 81.0% and specificity of 94.4%. The confusion matrix (**Table 1**) showed inconsistencies in the prediction of retrieval success of 65 of 670 (9.7%) data entries. After visual inspection, 8 files (Manual: pup not retrieved, Automated: pup retrieved) involved pups walking themselves back into the nest, for 11 files the automated retrieval estimation was more accurate than manual scores, and for 46 files manual scores were more accurate than automated estimations due to tracking errors.

	Manual not retrieved	**Manual retrieved**
Automated not retrieved	162	27
Automated retrieved	38	451

Table 1. Confusion matrix raw data.

For estimated retrieval time, Pearson correlations between manual recordings and automated analysis were high (r = 0.86). However, estimates using video analysis were on average 2.4 (SD=17.8) seconds faster than manual recordings. Within test day (PND5-13), Pearson correlations ranged between r=0.80 and r=0.92 (PND5: r= 0.80, PND7: r= 0.80, PND9: r= 0.92, PND11: r= 0.92, PND13: r= 0.86). To establish translatability for the current methodology, manual and automated recordings with a difference larger than 30 seconds were flagged based on the distribution of differences (**Figure 3a**). 54 Records were flagged of which 31 previously inspected retrieval inconsistencies and the remaining 23 records were visually inspected (**Figure 3b**). To ensure methodological correctness, 41 automated pup retrieval time estimations were corrected to their manual estimation. Also, pups that walked themselves into the nest were removed from the dataset as bidirectional behavior might be affected. The final dataset is visualized in **Figure 3c** and the confusion matrix is shown in **Table 2**. This corrected dataset had an accuracy of 95.1% (95% CI=93.2 – 96.6), sensitivity of 89.6% and specificity of 97.28%.

	Manual not retrieved	**Manual retrieved**
Automated not retrieved	172	13
Automated retrieved	20	465

Table 2. Confusion matrix corrected data.

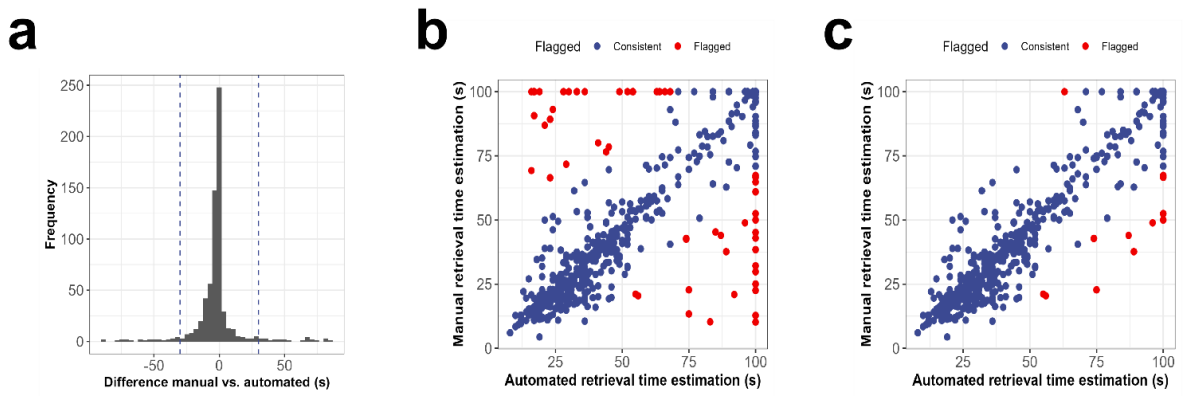


Figure 3: Post-processing quality control of retrieval time estimations.

a, Histogram representing the difference in seconds between manual and automated estimations of the retrieval time. Automated estimations of retrieval time were on average 2.4 seconds faster than manually registered estimations. **b**, Scatterplot displaying the relationship between raw manual and automated estimations. Differences smaller than 30 seconds were accepted and shown in blue, whereas differences larger than 30 seconds were flagged for visual inspection. **c**, Scatterplot displaying the relationship between corrected manual and automated estimations. After visual inspection of the flagged estimates of Figure 3b, the final estimate was either accepted (red) or corrected to the manual estimation.

4.4.3 Correlations between USV parameters

Correlational analysis showed that the total number of USVs emitted before retrieval was correlated with the USV rate before retrieval ($r=0.84$, $p<0.001$), mean USV duration ($r=0.44$, $p<0.001$) and first USV event ($r=-0.30$, $p<0.001$). The same pattern was observed for separate test days (Supplementary **Figure 2**).

4.4.4 Repeatability of traits over test days

Repeatability of traits was assessed by looking at the Pearson correlation matrix within a trait over time for the mean value of pups with the same sex within dams (**Supplementary Figure 3**). For maternal retrieval time, repeatability was generally moderate to high for consecutive test days, significant and consistently positive ($r=0.32$-0.63, $p<0.05$-0.001). The correlations suggest that dams who retrieve their pups faster on PND7 generally also will do so on the other days of testing. Pearson correlations between PND5 and the other days of testing were the lowest which might be due to the fact that this was the only day in which the cup paradigm was not used.

Repeatabilities for USV rate and mean USV duration were similarly assessed. Correlations were less pronounced, although most correlations were positive (**Supplementary Figures 3 and 4**). Particularly PND7 gave moderate correlations with the other test days for USV rate ($r=0.34$-0.50, $p<0.05$-0.001) and for mean USV duration ($r=0.39$-0.59, $p<0.01$-0.001) although not with PND13 data ($r=0.12$). For latency to first USV emission, no clear pattern was observed although most correlations were positive (**Supplementary Figure 5**).

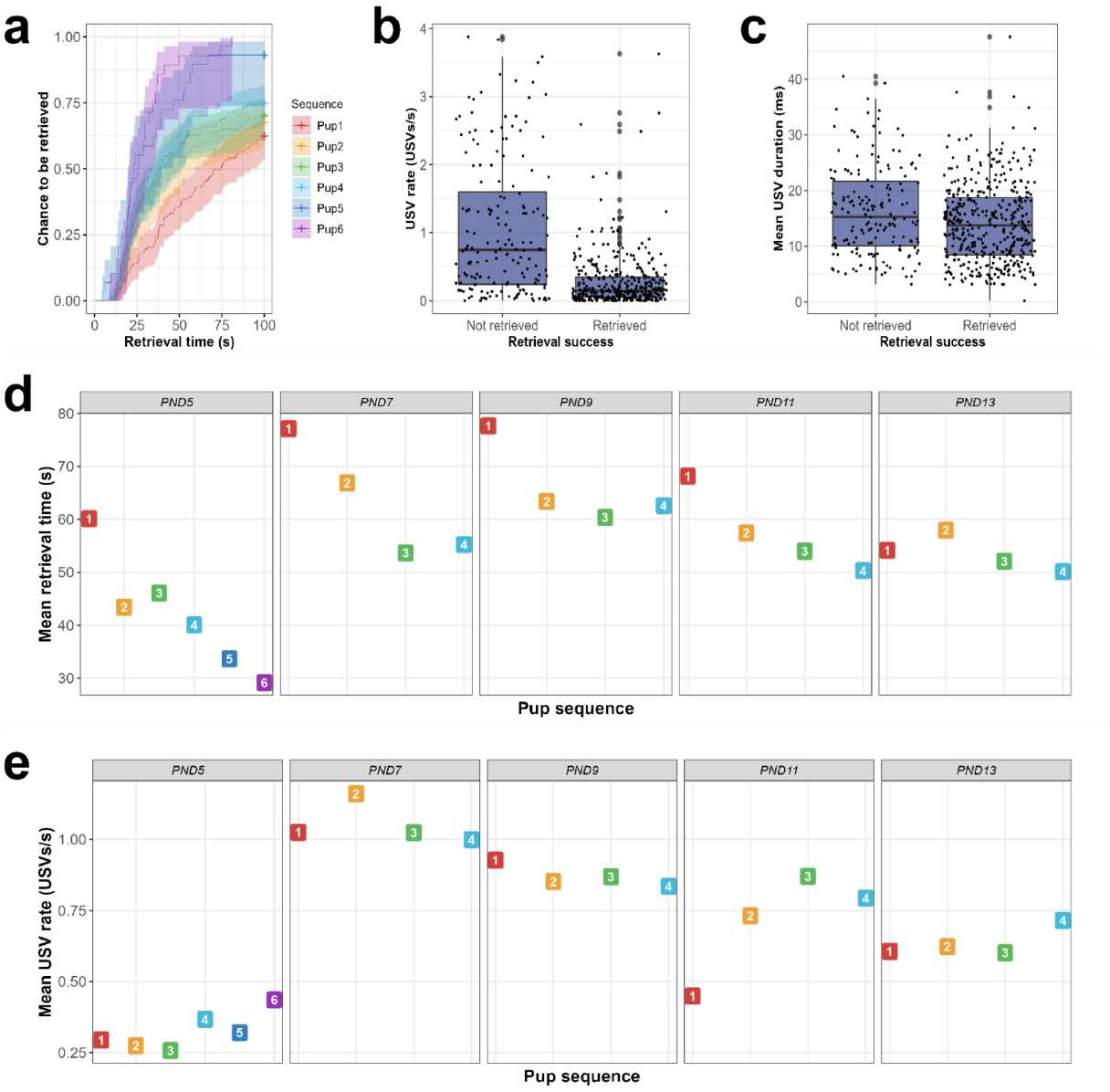


Figure 4: Results of bidirectionality analysis.

a, Survival plot representing the estimated probabilty to be retrieved over time in the PRT per maternal trial. Both retrieval time and chance to be retrieved increased as the maternal trial number increased, suggesting a maternal learning effect. **b**, Boxplot showing the number of USVs emitted per second when pups are either not retrieved or retrieved. Pups with a higher USV rate had a higher probability not to be retrieved ($p<0.001$). **c**, Boxplot showing the mean duration of USVs per pup when pups are either not retrieved or retrieved. Pups with a higher mean USV duration had an increased chance not to be retrieved ($p<0.001$). **d**, Mean plot showing the mean retrieval time per maternal trial per day. Although retrieval time decreases significantly for trials within days ($p<0.001$), the learning effect was not significant between days ($p=0.22$). **e**, Mean plot showing the mean USV rate per maternal trial sequence per day. USV rate was not affected by repeated trials ($p=0.59$), whereas test day significantly did ($p=0.02$).

4.4.5 Analysis of pup sex effect

No significant differences between pup sexes were found for USV rate before retrieval ($p=0.81$), indicating that the number of USVs, proportioned to the retrieval time, was comparable between pup sexes. However, USVs emitted by male pups had a significantly shorter duration compared to the USVs emitted by females ($p<0.001$). Nevertheless, this did not seem to affect maternal behavior. No significant effect of pup sex on maternal retrieval was observed ($p=0.07$).

4.4.6 Analysis of bidirectionality

Correlational analysis of PND5-13 data combined (**Supplementary Figure 6**), indicated a positive association between pup retrieval time and the amount of USVs the pup emitted (r=0.54, p<0.001), suggesting that pups that vocalized more were retrieved later. Hereafter, we looked at USV emission rate (number of USVs/retrieval time) and the number of USVs recorded during the first ten seconds of the test (USVs10sec), as most pups were retrieved after 10 seconds (5 pups <10s). This was done to correct for the fact that pups that are retrieved slower, also have more time to emit USVs. However, retrieval time was still positively correlated with USV emission rate (r=0.24, p<0.001). Interestingly, a significantly positive correlation was also found between retrieval time and $USVs_{10sec}$ (r=0.23, p<0.001).

Hereafter, we performed correlational analyses for each day separately, to exclude the use of the cup and/or age as cofounding variables for these results (**Supplementary Figure 6**). For total number of USVs emitted before retrieval, moderate, positive correlations were found with retrieval time for all testing days (r=0.45-0.61, p<0.001). This suggests that pups with a higher amount of vocalizations were generally retrieved later. Next, a correction for retrieval time was made by either looking at USV rate or $USVs_{10sec}$. Here, a significant positive correlation was only found on PND7-9 (r=0.31-0.33, p<0.01) for USV rate and on PND7 and PND13 (r= 0.21-0.29, p<0.05) for $USVs_{10sec}$. It should be noted that non-retrieved pups were assigned a retrieval time of 100 seconds, which might bias correlations.

The previous results query whether there is a difference in the number of vocalizations emitted by pups that are retrieved and those not retrieved. Binomial regression analysis of PND5-13 data combined, showed that the USV rate was a significant predictor of retrieval success (HR=0.58, p<0.001), which was also indicated by the boxplot (**Figure 4b**). The hazard ratio (HR) of 0.58 indicates that a USV rate increase of 1 USV/second reduces the probability of being retrieved by 42%. Hereafter, analyses were performed for each day separately, to exclude the use of the cup and/or age as cofounding variables for these results. **Figure 5** shows that median USV rate was higher in non-retrieved pups than in retrieved pups, although this difference was small on PND5 and PND13. Binomial regression analyses confirmed these results with negative estimated HR's on each test day (HR=0.46-.86) with only significant effects found on PND7, PND9 and PND11. The range of HR between 0.46-0.86 over separate test days indicates that a USV rate increase of 1 USV/second reduces the probability of being retrieved by 14-54%.

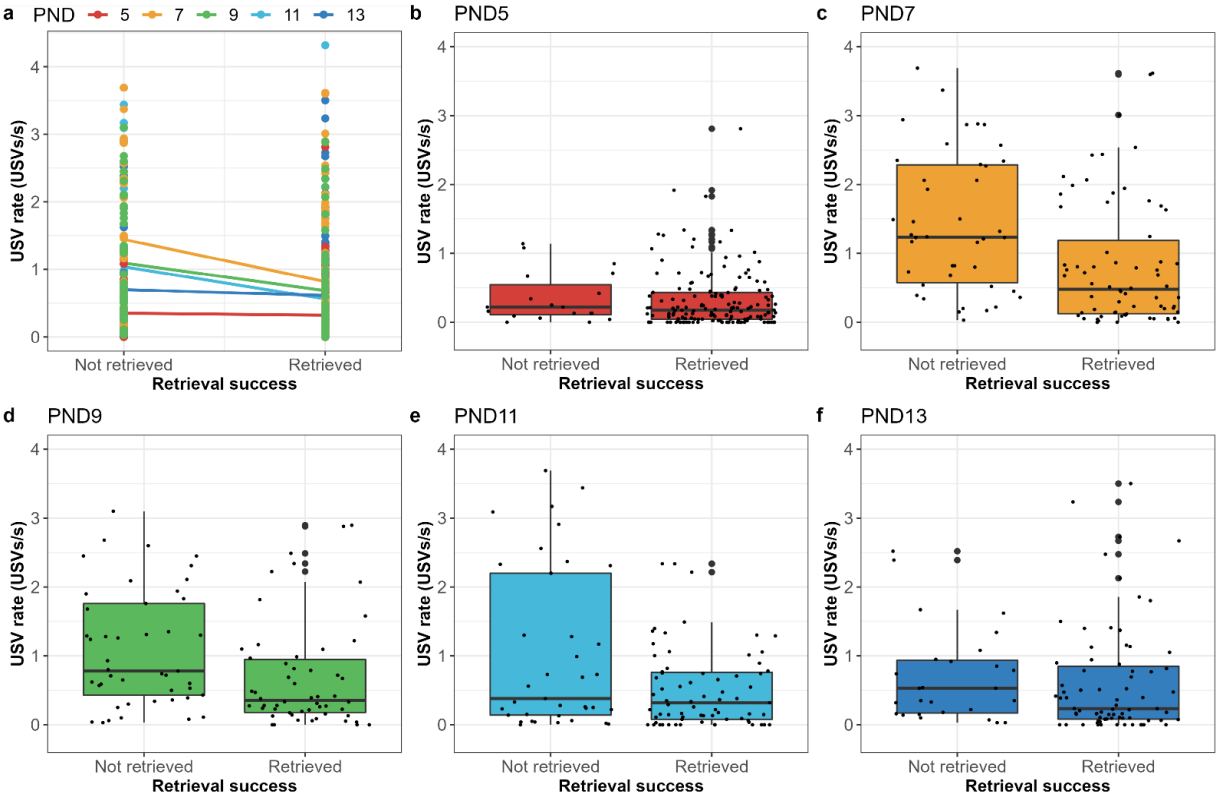


Figure 5: USV rate versus retrieval success for each test day separately.

a, Plot with linear regression of USV rate versus retrieval success scored as a binary variable for each test day separately. For all test days (PND5-13), USV rate was higher in non-retrieved pups than in retrieved pups, although regression estimates were close to 0 (horizontal regression line) for PND5 and PND13. **b**, Boxplot showing the USV rate (USVs per second) for pups which are either retrieved or not retrieved on PND5. **c**, Boxplot showing the USV rate (USVs per second) for pups which are either retrieved or not retrieved on PND7. **d**, Boxplot showing the USV rate (USVs per second) for pups which are either retrieved or not retrieved on PND9. **e**, Boxplot showing the USV rate (USVs per second) for pups which are either retrieved or not retrieved on PND11. **f**, Boxplot showing the USV rate (USVs per second) for pups which are either retrieved or not retrieved on PND13.

Furthermore, we wanted to see whether this could be explained by a few poorly retrieving dams (i.e., dams retrieving on fewer than 50% of the trials), such dams were removed from the dataset (n=7 dams). However, the effect of USV rate on retrieval success was still significant after removing poorly retrieving dams ($p<0.001$). As shown in **Supplementary Table 1**, some pups (n=67) did not vocalize before retrieval, although 64 of these pups were still retrieved by their dams. Of these 64 trials, 48% occurred on PND5, 20% on PND11 and 19% on PND13. Retrieval without pup vocalization is more common in pups with repeated maternal measurements, i.e. with a later position in maternal trial sequence within a litter (**Figure 4a**). Moreover, the sequence of maternal trial was found to influence retrieval success significantly (**Figure 4a**, $p<0.001$).

The significant effect of maternal trial suggests a learning effect, and as such, provides another possible explanation for the faster retrieval in pups that have a lower vocalization rate. That is, exposing a dam to multiple trials might affect her retrieval behavior and/or might affect pup vocalization rate. However, as shown in **Figure 4e**, USV rate was not significantly affected by maternal trial ($p=0.59$), although test day did ($p=0.02$). Over all days, a maternal learning effect was found to be statistically significant (Figure 4a, HR=1.19, $p<0.001$). The HR indicates that an increase in maternal trial by one

increases the probability of pup retrieval by 19%. As shown in **Figure 4d**, this maternal learning effect was manifest within repeated trials on the same day ($p<0.001$), but did not translate between days ($p=0.22$).

Lastly, the average duration of pup vocalizations was positively correlated with retrieval time ($r=0.14$, $p<0.001$), which was most pronounced on PND7-11 (**Supplementary Figure 7**). Pups emitting USVs with a longer average duration had a lower probability of being retrieved (**Figure 4c**). The estimated effect in a binomial model was -0.053 ($p<0.001$) which corresponds with a decreased hazard by a factor of 5% for one extra millisecond of USV.

4.5 Discussion

Bidirectional dam-pup dyad interactions are critical for pup survival. However, most studies investigated dyadic members and behaviors unilaterally (18, 20). In the current study, we describe BAMBI (Bidirectional Automated Mother-pup Behavioral Interaction test) to assess bidirectional dam-pup interaction in laboratory mice. This approach combines the automated PRT described by Winters et al. (26) with synchronous ultrasonic audio recording and subsequent automated USV detection. At first, we demonstrated the transferability of the previously established dam-pup model to a novel experiment with different traits. Further, a model was developed to detect simultaneously recorded pup USVs with high accuracy. Lastly, we applied this methodology on PRT data sampled on PND5, 7, 9, 11 and 13. Indeed, through synchronous videorecording of maternal behavior and audiorecording of pup vocalizations, BAMBI allowed to test bidirectional early-life mother-pup interactions in an unprecedented way.

We were able to expand the publicly available model (26), and optimized its performance for PRT data with different subject and environmental traits such as the inclusion of a cup. We used a hybrid learning strategy to increase variability relatively fast while minimizing bias. This hybrid learning strategy combined manual annotation of k-means selected frames and refinement of outlier frames selected by the DLC 'jump' algorithm. In our first attempts, these newly annotated data were added to the annotated dataset of Winters et al. (26) and retrained. However, pose estimation performance on videos with novel traits was insufficient (data not shown). We hypothesized this might be due to representation bias whereas the original dataset with robust PRT poses on PND5 outweighed the novel dataset with higher pose variability (33). Therefore, we used a two-step learning approach similar to Gorssen et al. (31). In a first step, the original model was retrained only with the newly annotated data, whereas in a second step, all annotated data were used to ensure the algorithm performed well on both the original and new data. The automated retrieval estimate can be seen as proof-of-concept and had a high accuracy of 90.4% over all test days. For future research, two remarks on this learning approach should be kept in mind. First, the train and test error after the second retraining step should be interpreted and reported with caution. That is, all data has been used in previous training phases and thus the test data might not be completely new anymore. Second, we found a difference in the average pixels per millimeter when comparing the original dataset and the dataset of the current study. Again, this indicates that the retraining pixel errors should be interpreted with caution.

Further, we were able to develop a model to detect ultrasonic vocalizations in the PRT accurately and automatically using DAS (29). Despite the wide range of available automated detection options, we chose to work with DAS based on a few selection criteria. First, both the toolbox and its basis software (i.e., Python) are completely open source. Second, the system is versatile which is necessary as this

PRT assay intends to investigate early-life communicative deficits, and thus, the emitted vocalizations might not be as expected (19, 34-36). The system therefore should be easily adaptable and relatively flexible. Third, the system should be able to handle background noise as the PRT is performed in freely moving animals, which are interacting with their environment. As argued in the work of Ey et al. (22), most available automated systems cannot (yet) handle background noise. However, the main limitation of DAS is that the output is limited to the temporal parameters start and end time of the vocalization. Although this was not a problem for the current study, it is a restriction when investigating communicative deficits. Additional spectrographic output parameters should be an integral part of communicative assessment to fully understand eventual deficits.

An obstacle in the current study was the synchronization of video and audio recordings. Both recording data were sampled using different software and could be synchronized by introduction of a beep at the start of the trial. Although we were able to precisely retrace this beep with frame accuracy, this required an intensive step of data processing. To find its way to standard operational practices, an integrated recording system would significantly reduce human involvement and workload. An exemplary integrated recording system was described in Ey, de Chaumont and Bourgeron (22). In this work, behavioral monitoring was done using the Live Mouse Tracker (LMT, 37) system in which synchronized USV sequences were recorded using the Avisoft UltraSoundGate Recording system's trigger function. The Avisoft burst recording yield an advantage when working with long-term recordings (22). However, in the PRT paradigm a maximum time of 100 seconds is defined and, as previously mentioned, intends to investigate abnormalities in early-life communicative behaviors. The use of burst recordings should be used with caution as it could miss deviant vocalizations and thus could lead to loss of data which cannot be corrected afterwards. Other options exist as most Avisoft UltraSoundGates have the possibility to connect a TTL cable, which can be used to start ultrasound recording together with another software, e.g. video recording.

Lastly, we demonstrated the effectiveness of our combined methodology by applying it on PRT data sampled on PND5, 7, 9, 11 and 13. It is important to add a note regarding the selection of the study subjects. In compliance with the reduction principle, mice of the present study were obtained from an independently designed pharmacological study. As a consequence, in the absence of controls for experimental disease models, subjects were exposed to VPA and pharmacological treatment, possibly affecting their behavior. Importantly, the aim of the present work was not to investigate pharmacological effects, but rather to present a proof of principle demonstration of the feasibility and validity of a new automated method for behavioral testing of early life mother-pup bidirectional interactions. Nevertheless, in order to address the issue of not being pharmacologically naive, statistical analyses performed in the current study employed a correction for pharmacological

treatments as a confounding variable, by using a GLM model in which drug effect was set as a fixed effect, which allowed to pool the different drug groups into a single group (see Experimental groups). Therefore, the general relationships between pup vocalizations and maternal retrieval found in our study can be considered relevant for future research.

We found an association between maternal retrieval success and pup calling behavior. Counterintuitively, we found that pups that were retrieved had a lower call rate during maternal separation than non-retrieved pups (Figure 4b), which was most pronounced on PND7-13. This effect was not caused by certain poorly retrieving mothers, nor testing day. Previous research (18, 38) reported a negative relationship between maternal caregiving behaviors and separation-induced pup calling. These studies found that high levels of maternal caregiving behavior in the first days of life lead to reduced numbers of USV later in life, probably because of reduced anxiety. In the same line, maternal carrying has been shown to have soothing effects on pup physiology including cardiac deceleration, immobility response and a reduction of emitted USVs, whereas the absence of this calming response has been reported to hinder maternal retrieval efficacy (39). Altogether, these findings seemingly go against a robust set of evidence from playback literature which show that pup USVs elicit retrieval behavior (13-17). Our hypothesis is that USVs do elicit retrieval behavior, but is dependent on a great number of factors (40) and an excessive amount of USV vocalizations might negatively influence maternal retrieval efficacy. This negative relationship might be due to a miscommunication in the mother-pup dyad. However, further research is necessary to test this hypothesis.

Studies that used maternal retrieval and separation-induced vocalizations separately suggested that these factors might be related. The present simultaneous registrations further confirm and detail this relationship. For example, we found that vocalizations during the first 10 seconds actually predicted retrieval success, notwithstanding corrections for age and maternal trial sequence. Still, this should not be taken as evidence that pup behavior tunes maternal behavior, as behavioral testing only started on PND5. In our results, we found a peak in USV rate at PND7-9 (Figure 4d), which corresponds with previous findings in literature (41). However, future research might consider earlier time points as communicative fitness might already be affected before PND5 in either quality and/or quantity of vocalizations.

Further, we show that dams subjected to repeated retrieval trials show a significant learning curve within the same test day, although this does not translate to an inter-day effect (Figure 4d). Between PND7 and 9 this might be explained by the introduction of a cup in the homecage. However, translation is still limited on the other four days that the cup is present. Research has shown that experience improves pup retrieval success (42-43). Mice tend to use a spatial memory-based strategy when

engaged repetitively in pup search and retrieval (43). Therefore, an overall decrease in retrieval time was to be expected as pups were always placed in the same corner. Additionally, Dunlap et al. (43) report that retrieval behavior further improves by sensory learning of associated cues. The beep at the start of the trial in the current experiment could have predicted the presence of an separated pup in the homecage. Our findings seem to contradict the findings of Dunlap et al. (43) although the number of retrieval repetitions is significantly higher than in our PRT procedure, and the test environment might play a role in the valence of pup stimuli (42). For the interpretation of USVs, this means that the functional relevance of USV emission is particularly high at the beginning. After repeated testing, USV emission seems to be less and less relevant, as evidenced by the fact that retrieval behavior even occurred in the absence of USV emission probably due to maternal learning. However, this maternal learning curve could also be used as a behavioral read-out.

In the present study, we adapted our previous automated homecage PRT (Winters et al., 2020) and we combined videorecording of maternal behavior with synchronous audiorecording of pup vocalizations in order to assess bidirectional dam-pup dyadic interaction. Our methodology expands the automated pup retrieval test with automated detection of pups' ultrasonic vocalizations. Moreover, we validated our results and showed that the number and rate of ultrasonic vocalizations are associated with retrieval success. BAMBI is a promising new automated homecage behavioral method that can be applied to both basic and preclinical studies on early-life social development.

4.6 Supplementary information

4.6.1 Supplementary File 1.

Structural parameters used to train the DAS automated detection network.

```
das.train.train(model_name='tcn_stft',
    data_dir=path_to_data,
    save_dir=path_to_data,
    pre_nb_conv=4,
    pre_nb_dft=33,
    pre_nb_filters=33,
    pre_kernel_size=64,
    nb_hist=8192,
    batch_size=32,
    batch_norm=True,
    kernel_size=16,
    nb_filters=32,
    ignore_boundaries=True,
    verbose=1,
    nb_conv=2,
    learning_rate=0.0005,
    use_separable=(True, True, False, False),
    nb_epoch=100)
```

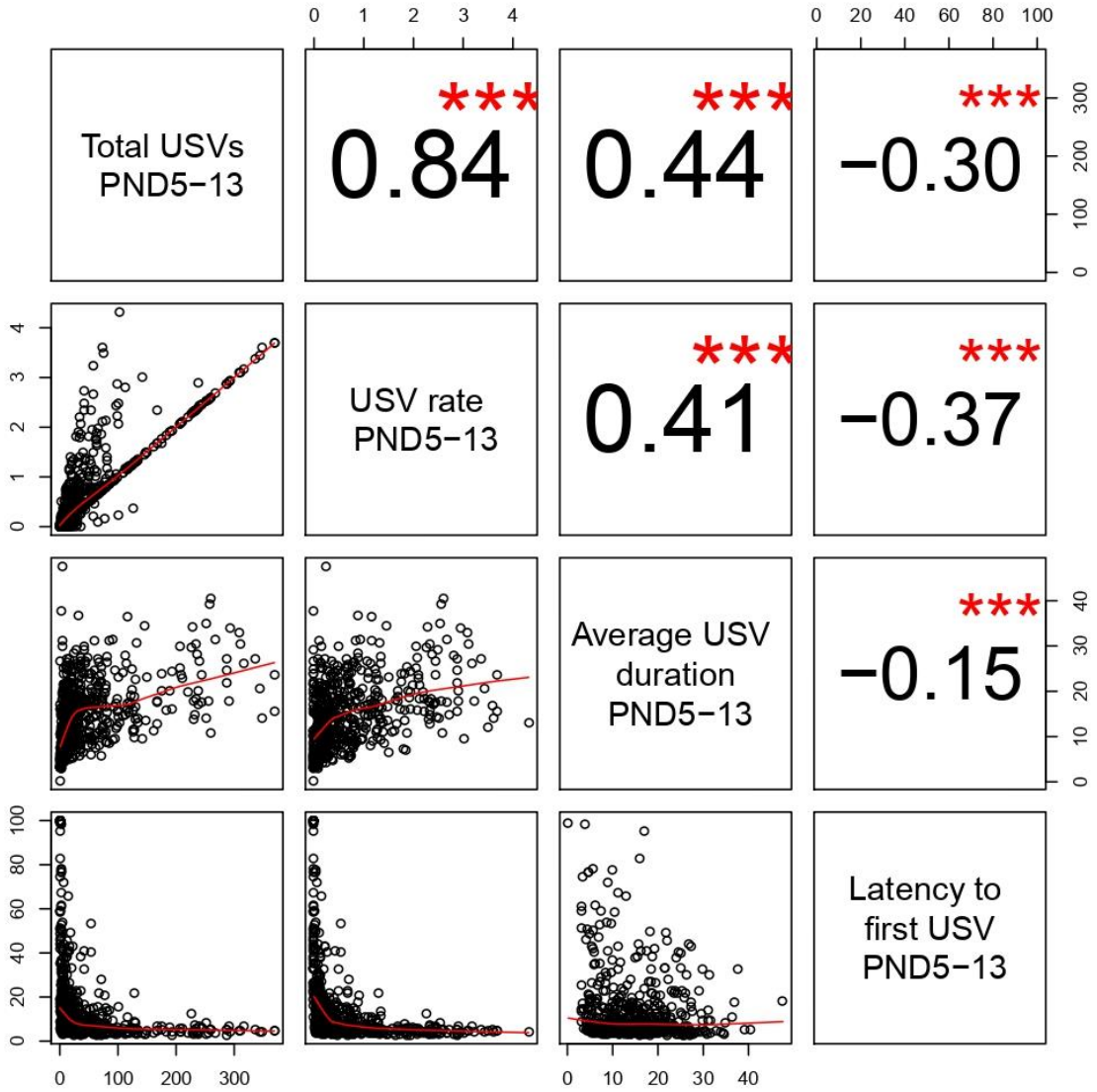
4.6.2 Supplementary Table 1.

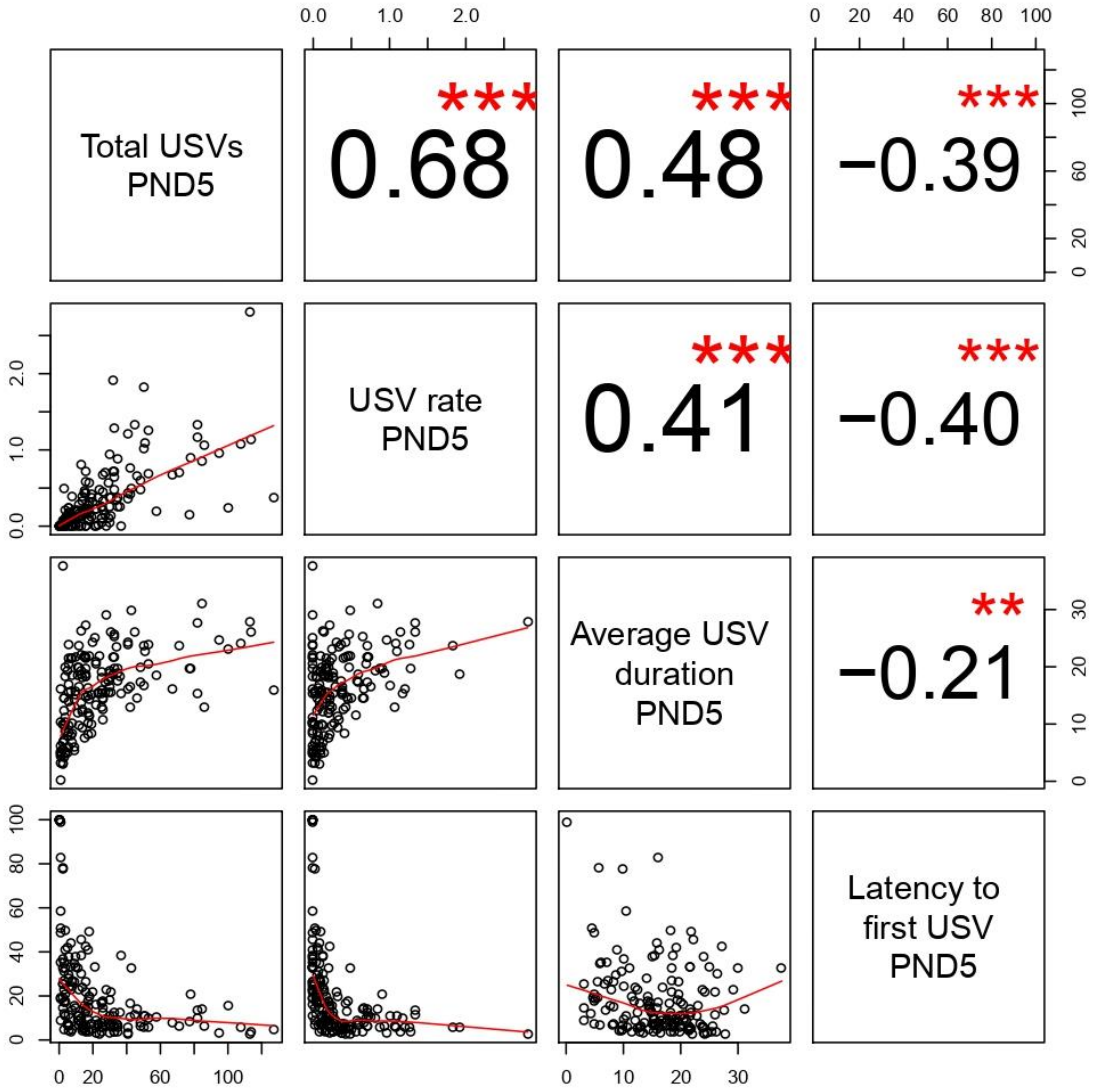
Table including pups that did not vocalize.

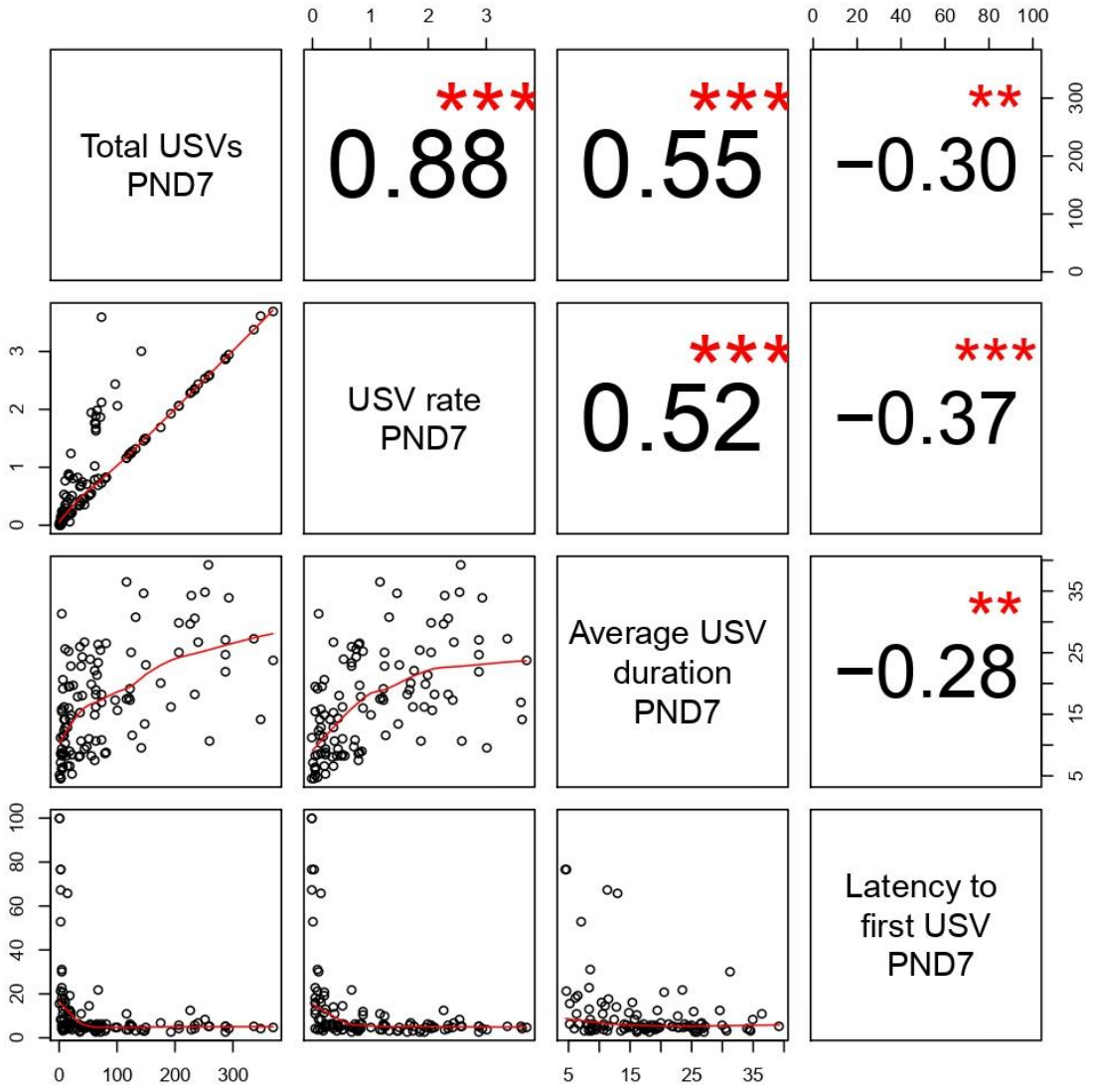
SEQUENCE	RETRIEVED	NOT RETRIEVED
PND5	31	2
Pup 1	*1*	*1*
Pup 2	*4*	*0*
Pup 3	*6*	*0*
Pup 4	*2*	*1*
Pup 5	*8*	*0*
Pup 6	*10*	*0*
PND7	4	0
Pup 1	*1*	*0*
Pup 2	*1*	*0*
Pup 3	*1*	*0*
Pup 4	*1*	*0*
PND9	4	0
Pup 1	*0*	*0*
Pup 2	*1*	*0*
Pup 3	*1*	*0*
Pup 4	2	*0*
PND11	13	1
Pup 1	*5*	*0*
Pup 2	*2*	*1*
Pup 3	*4*	*0*
Pup 4	2	*0*
PND13	12	0
Pup 1	*2*	*0*
Pup 2	*3*	*1*
Pup 3	*2*	*0*
Pup 4	5	*0*

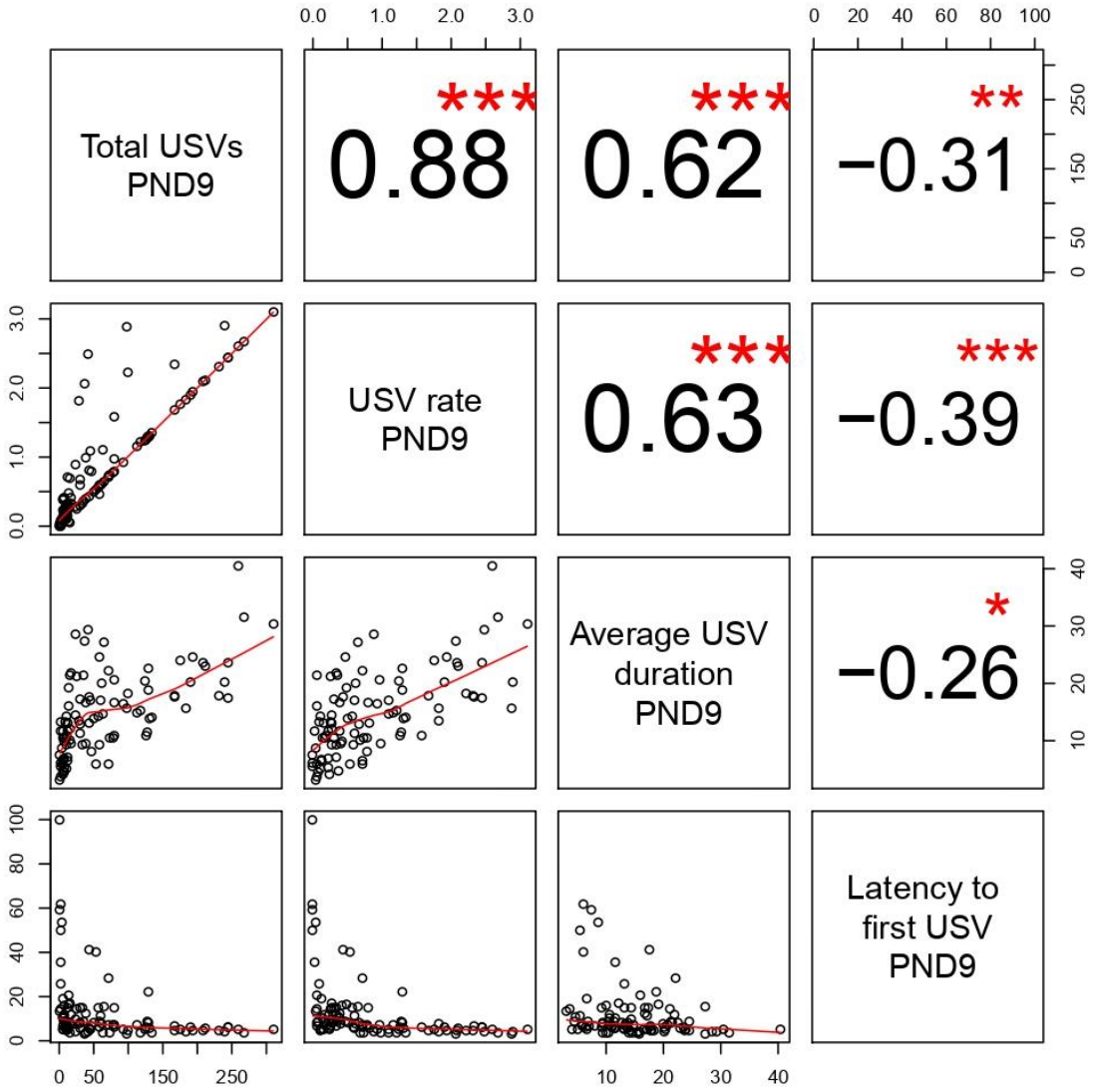
4.6.3 Supplementary Figure 1.

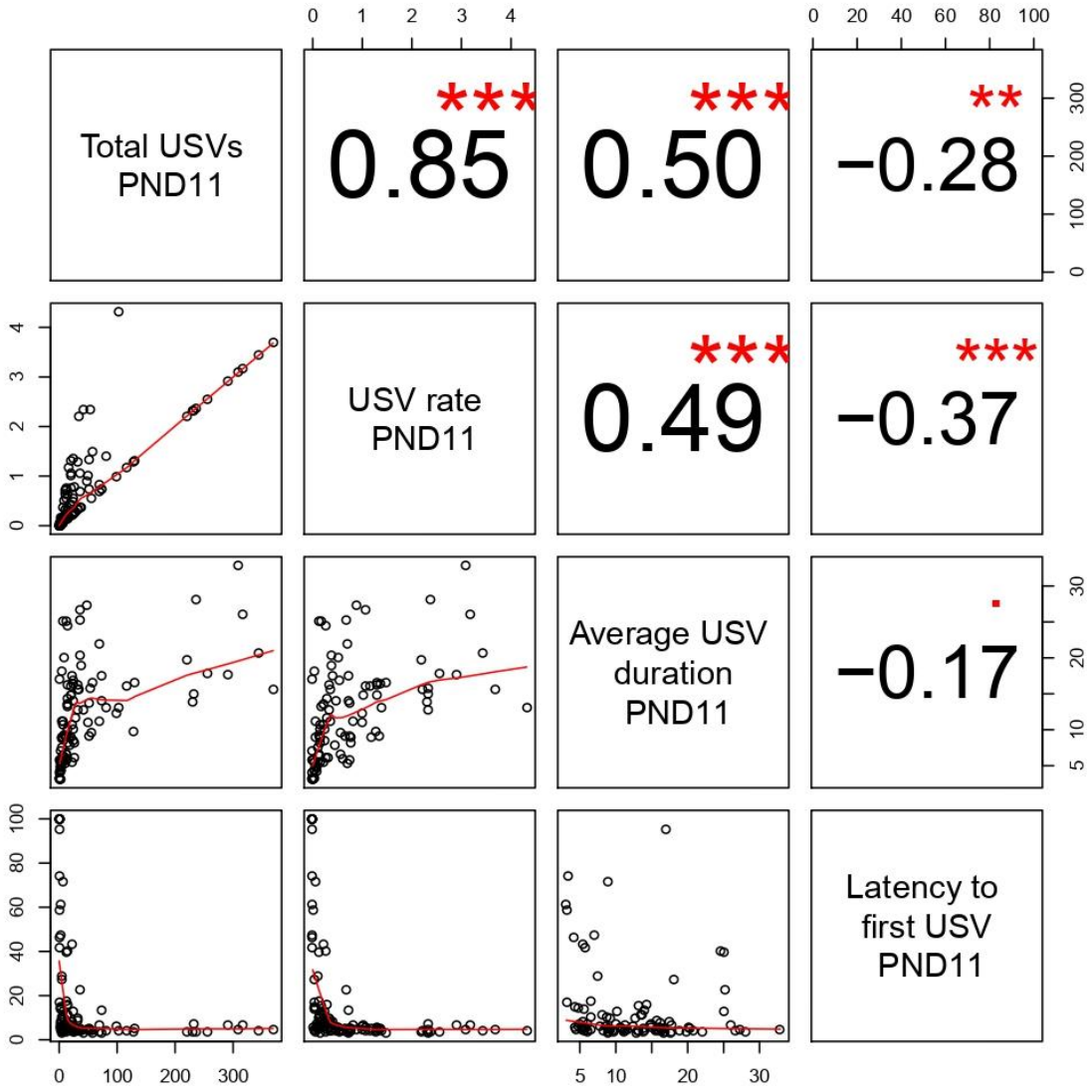
Pairwise correlational plots of total amount of emitted USVs (number USVs), USV rate (USVs/s), average USV duration (s) and latency to first USV (s). Pairwise correlation plots are first given for all test days combined (PND5-13) and then separately for PND5, 7, 9, 11 and 13. Below diagonal the pairwise correlation plot is shown. Above the diagonal the estimated Pearson correlation coefficient is given with significance value. (°p<0.10, *p<0.05, **p<0.01, ***p<0.001)

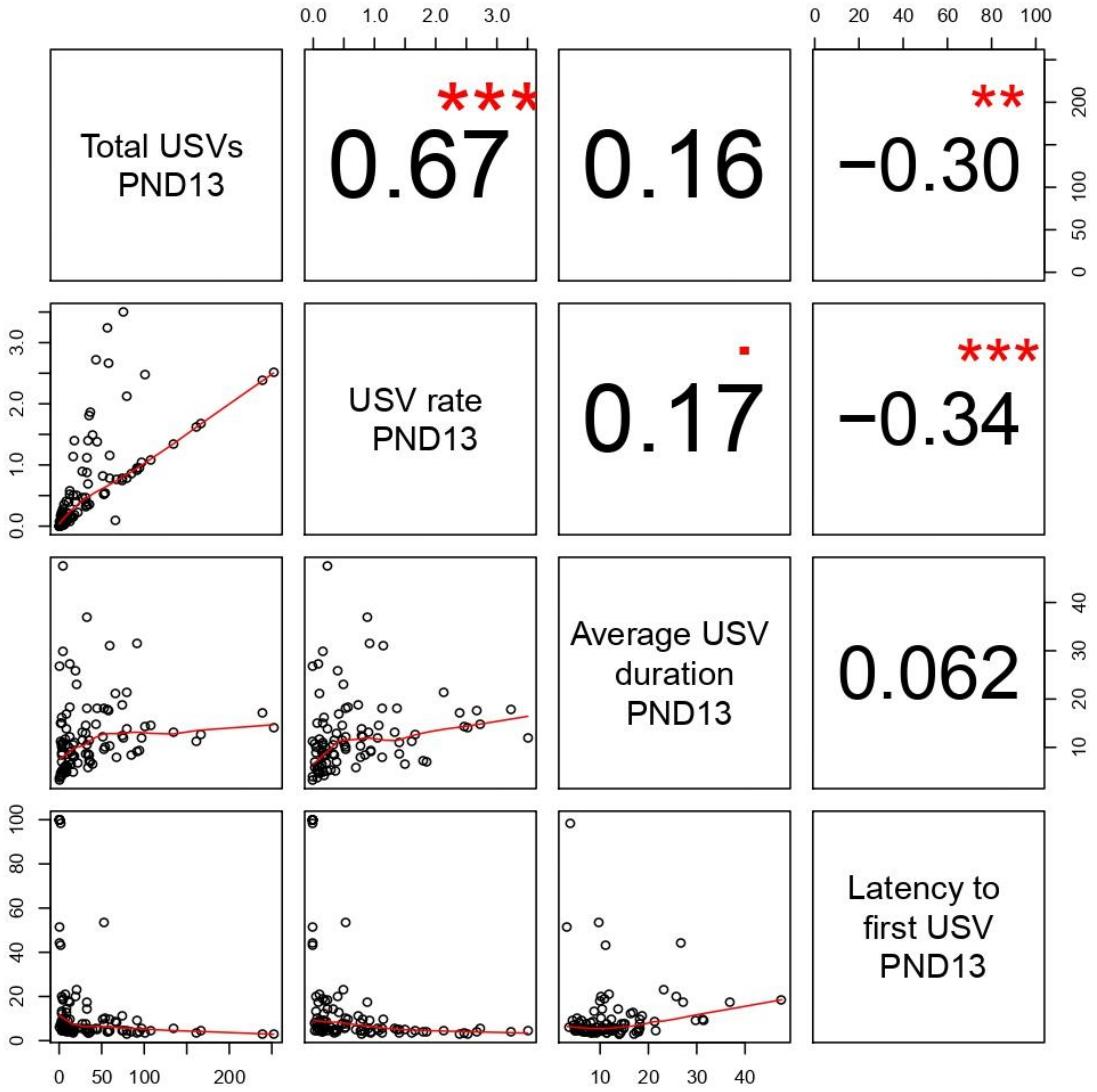



Total USVs PND5
0.68
***
0.48
***
−0.39
***
USV rate PND5
0.41
***
−0.40
***
Average USV duration PND5
−0.21
**
Latency to first USV PND5
0.0
1.0
2.0
0
20
40
60
80
100
0
10
20
30


Total USVs PND7
0.88
***
0.55
***
−0.30
**
USV rate PND7
0.52
***
−0.37
***
Average USV duration PND7
−0.28
**
Latency to first USV PND7
0 1 2 3
0 20 40 60 80 100
0 100 200 300
0 1 2 3
0 20 40 60 80 100
0 100 200 300
5 15 25 35

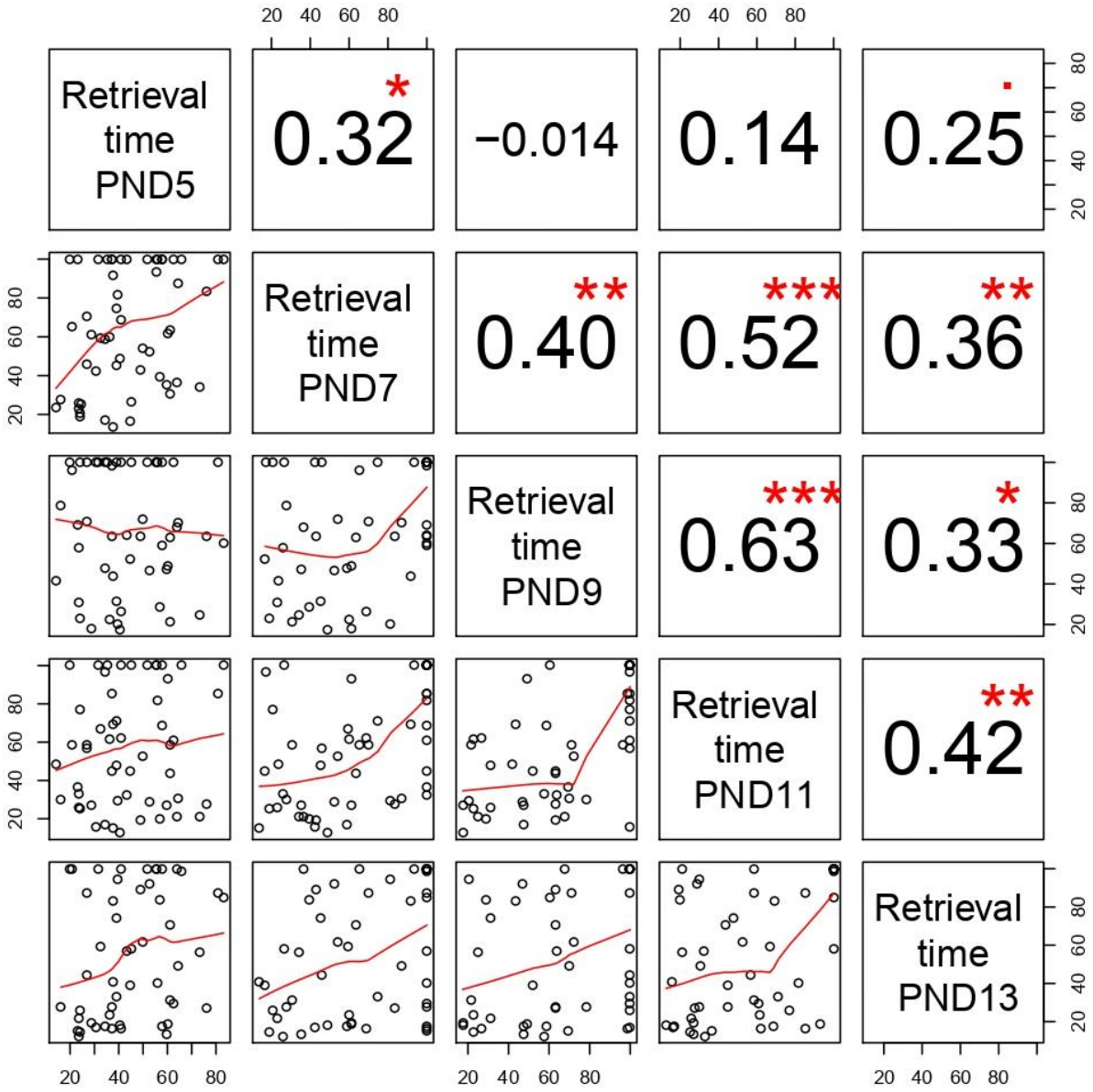

Total USVs PND9
0.88
***
0.62
***
−0.31
**
USV rate PND9
0.63
***
−0.39
***
Average USV duration PND9
−0.26
*
Latency to first USV PND9
0.0
1.0
2.0
3.0
0
20
40
60
80
100
0
50
150
250
10
20
30
40


Total USVs PND11
0.85
***
0.50
***
−0.28
**
USV rate PND11
0.49
***
−0.37
***
Average USV duration PND11
−0.17
.
Latency to first USV PND11
0 1 2 3 4
0 20 40 60 80 100
0 100 200 300
0 1 2 3 4
0 20 40 60 80 100
0 100 200 300
5 10 20 30
5 10 20 30


Total USVs PND13
0.67
***
0.16
−0.30
**
USV rate PND13
0.17
.
−0.34
***
Average USV duration PND13
0.062
Latency to first USV PND13
0.0 1.0 2.0 3.0
0 20 40 60 80 100
0 50 100 200
10 20 30 40

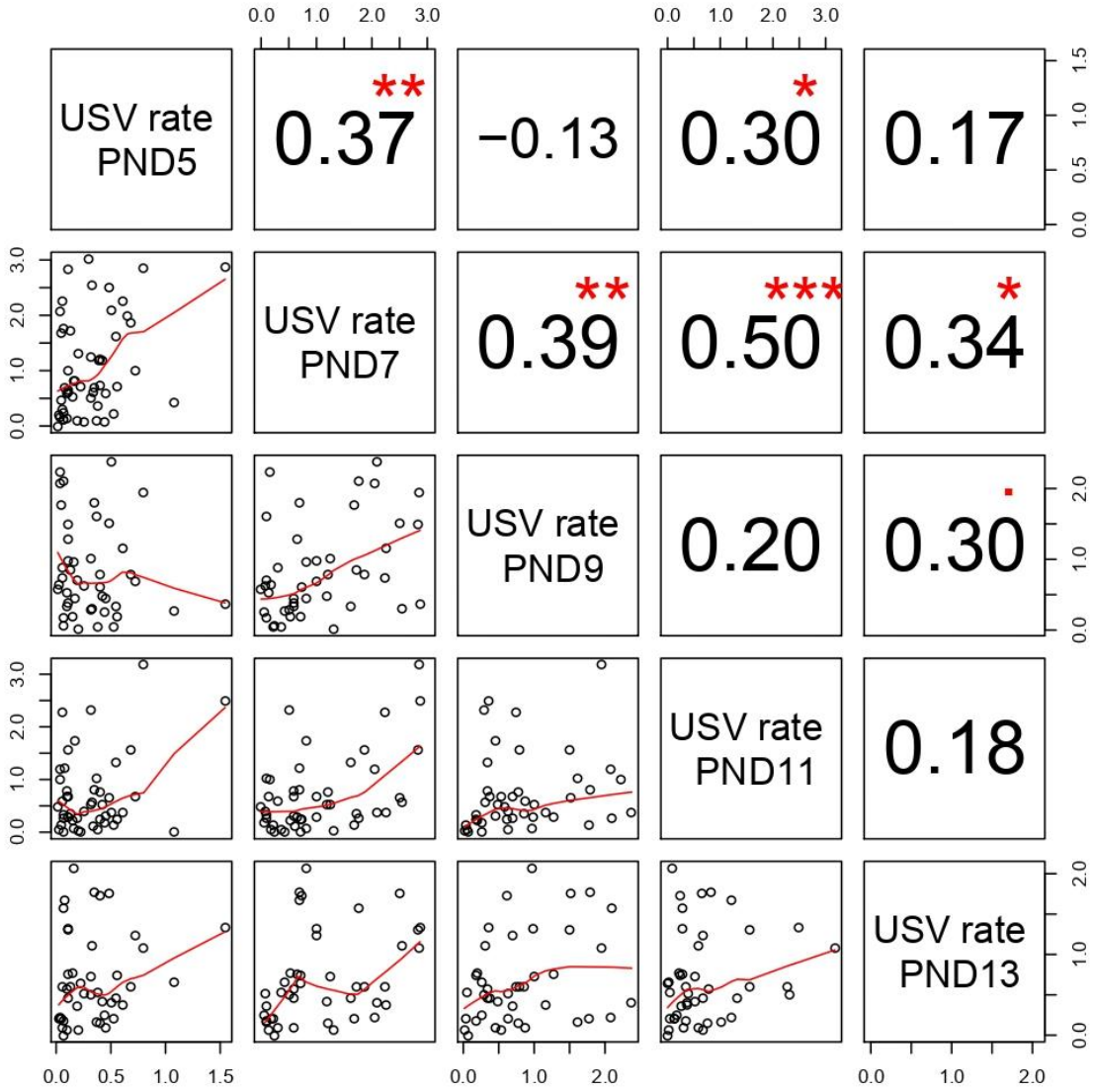
4.6.4 Supplementary Figure 2.

Repeatability of maternal retrieval time represented as a pairwise correlational plot of retrieval time (s) over PND5, 7, 9, 11 and 13. Below diagonal the pairwise correlation plot is shown. Above the diagonal the estimated Pearson correlation coefficient is given with significance value. (°p<0.10, *p<0.05, **p<0.01, ***p<0.001)



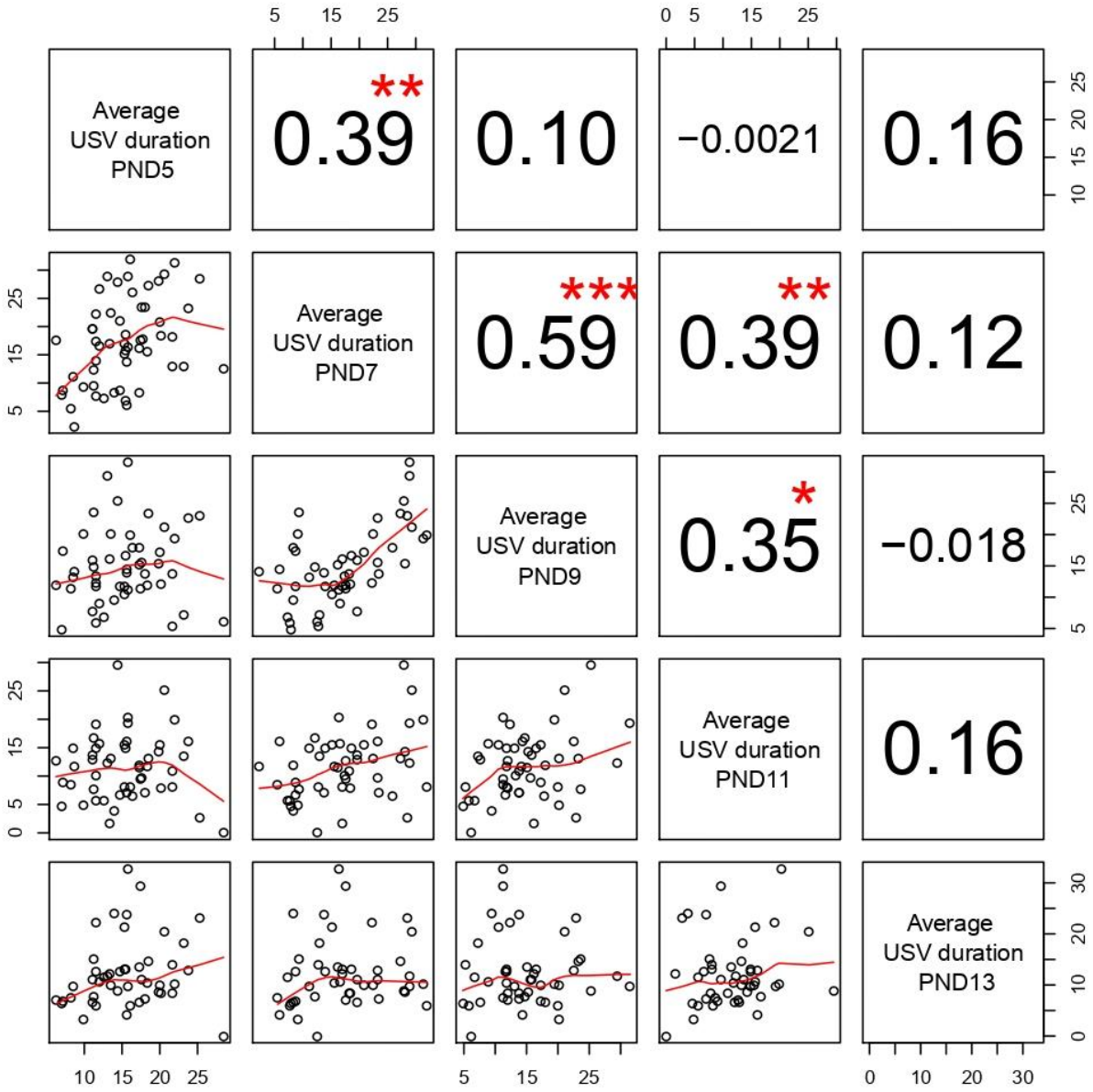
4.6.5 Supplementary Figure 3.

Repeatability of USV rate represented as a pairwise correlational plot of USV rate (USVs/s) over PND5, 7, 9, 11 and 13. Below diagonal the pairwise correlation plot is shown. Above the diagonal the estimated Pearson correlation coefficient is given with significance value. (°p<0.10, *p<0.05, **p<0.01, ***p<0.001)



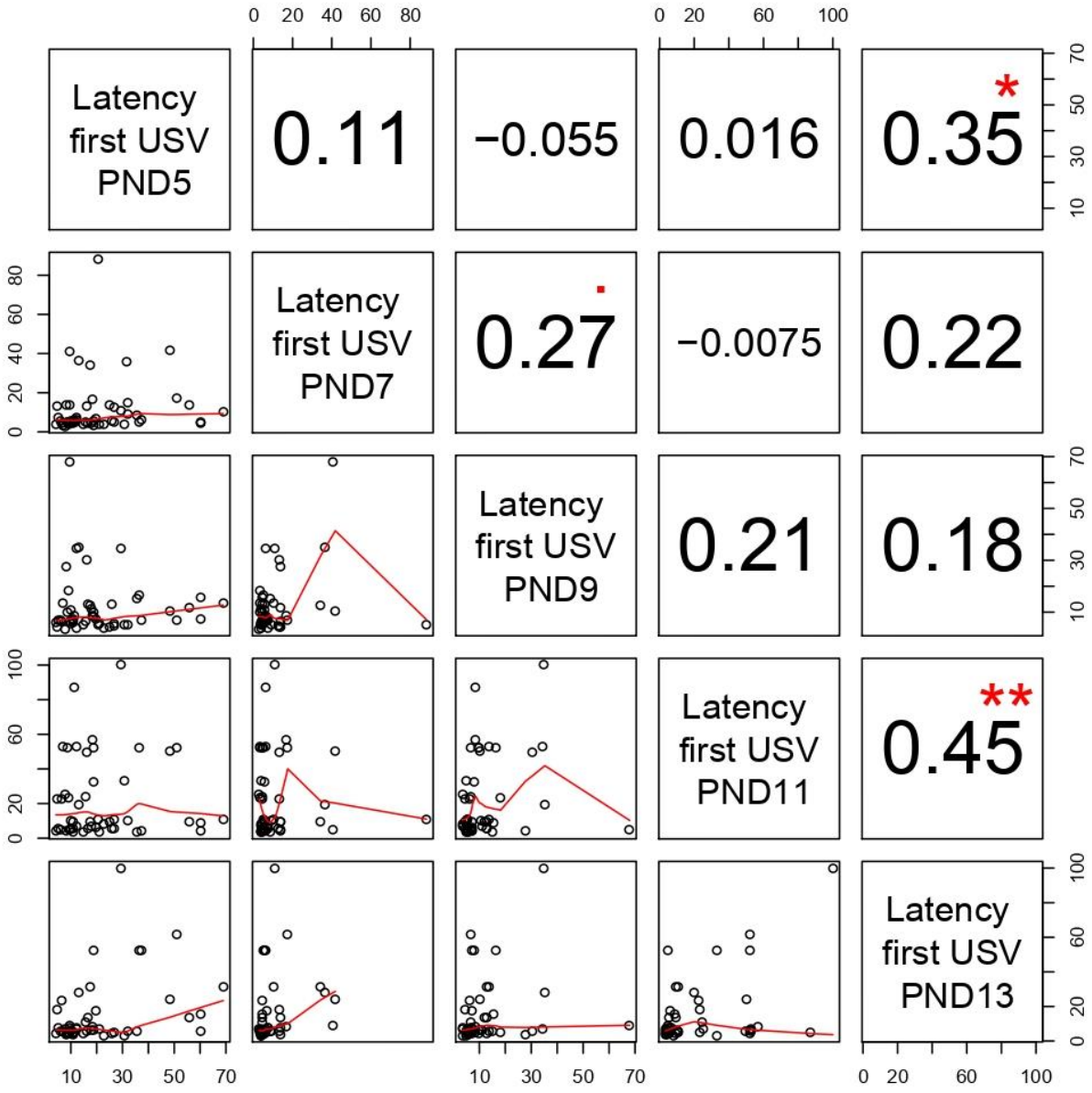
4.6.6 Supplementary Figure 4.

Repeatability of average USV duration represented as a pairwise correlational plot of average USV duration (s) over PND5, 7, 9, 11 and 13. Below diagonal the pairwise correlation plot is shown. Above the diagonal the estimated Pearson correlation coefficient is given with significance value. (°p<0.10, *p<0.05, **p<0.01, ***p<0.001)



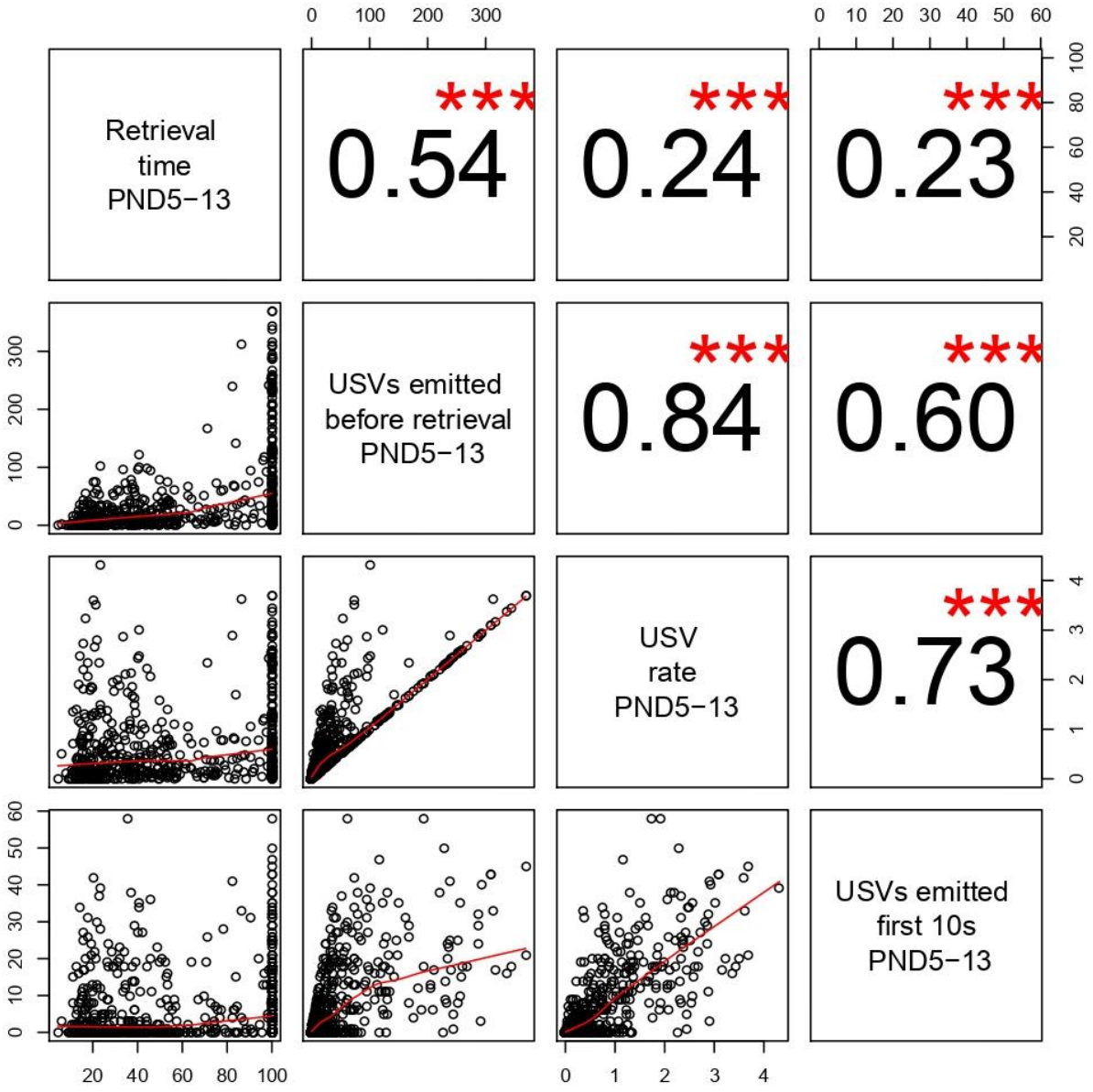
4.6.7 Supplementary Figure 5.

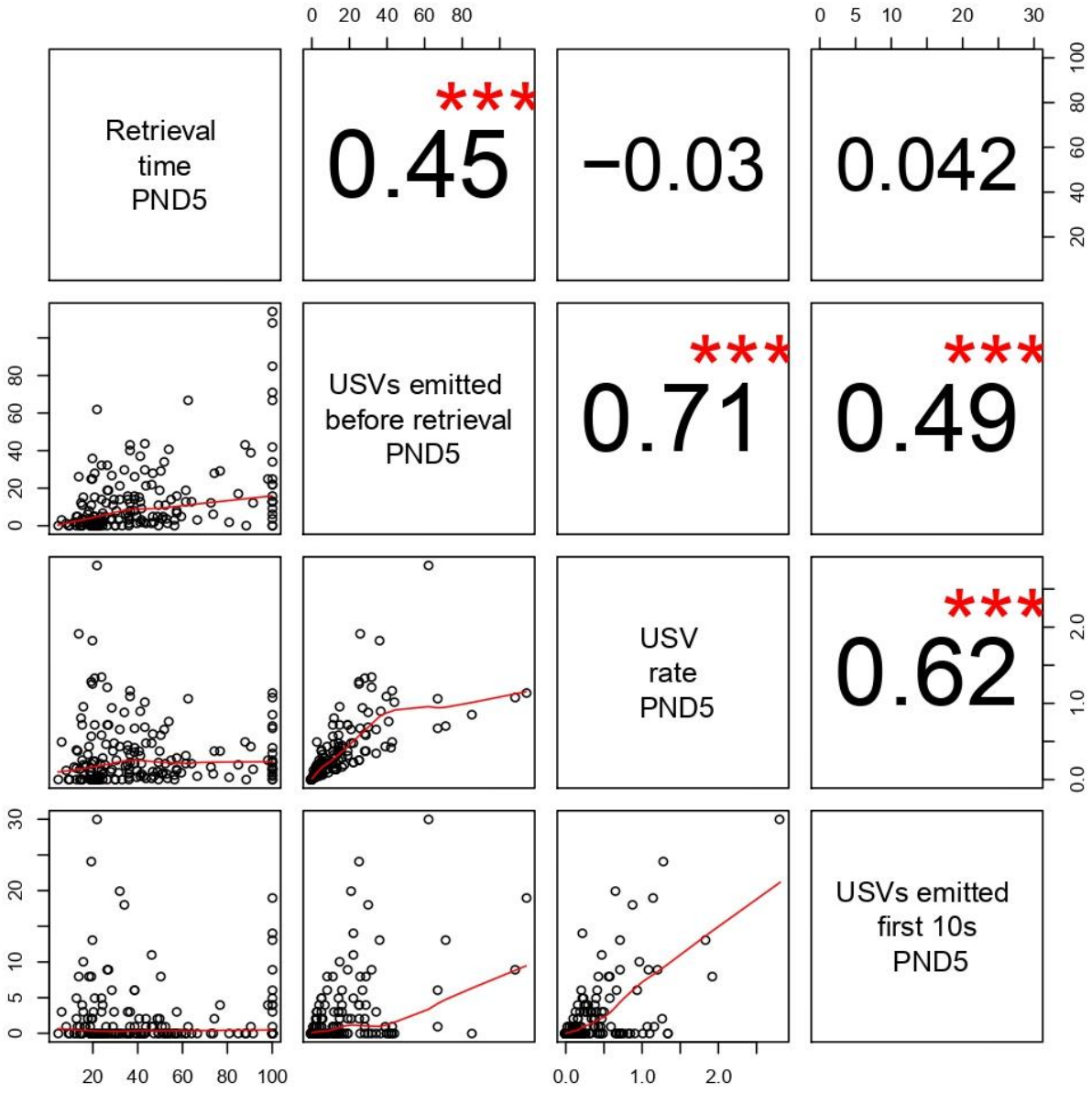
Repeatability of latency to first USV represented as a pairwise correlational plot of latency to first USV (s) over PND5, 7, 9, 11 and 13. Below diagonal the pairwise correlation plot is shown. Above the diagonal the estimated Pearson correlation coefficient is given with significance value. (°p<0.10, *p<0.05, **p<0.01, ***p<0.001)

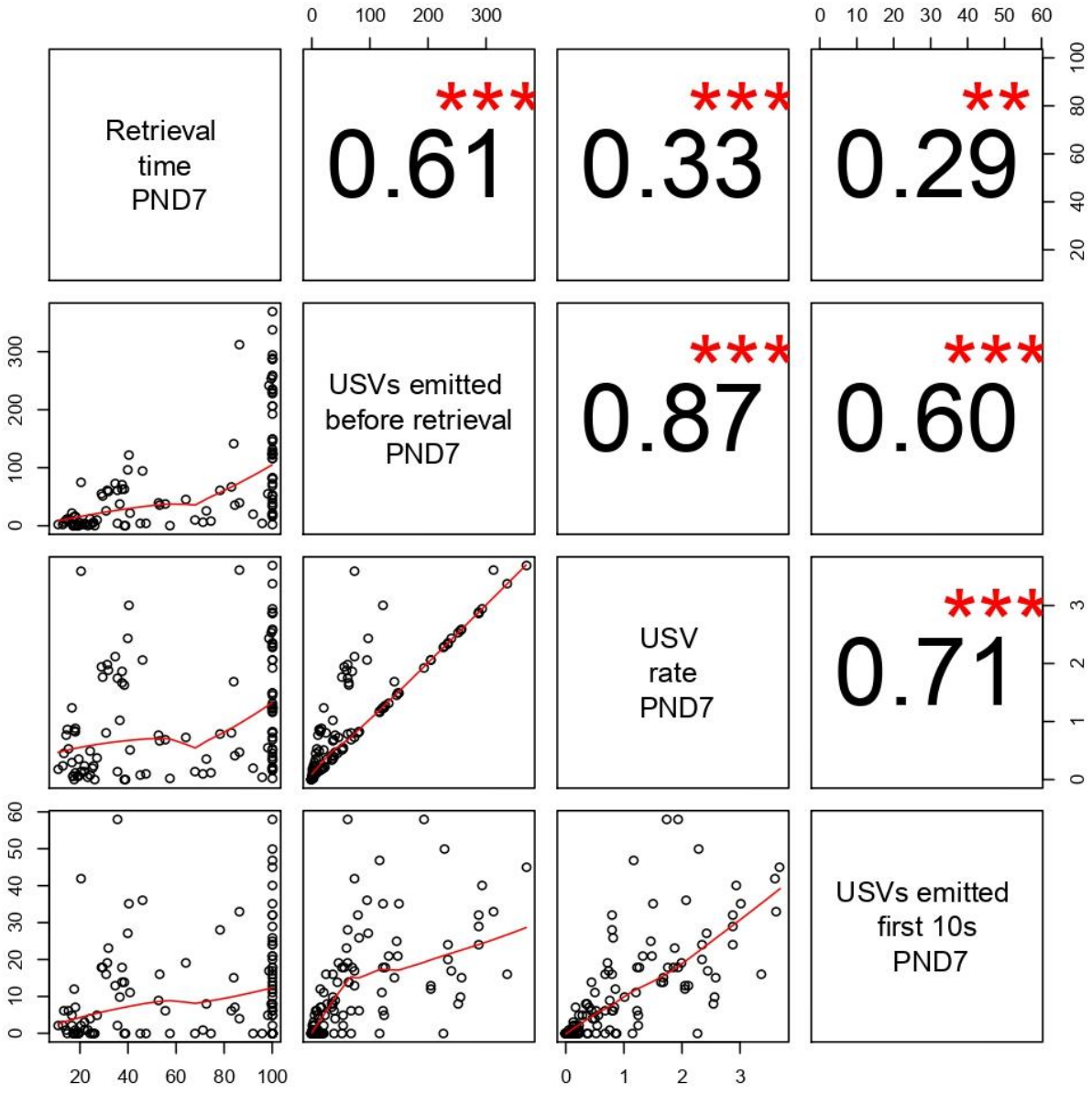


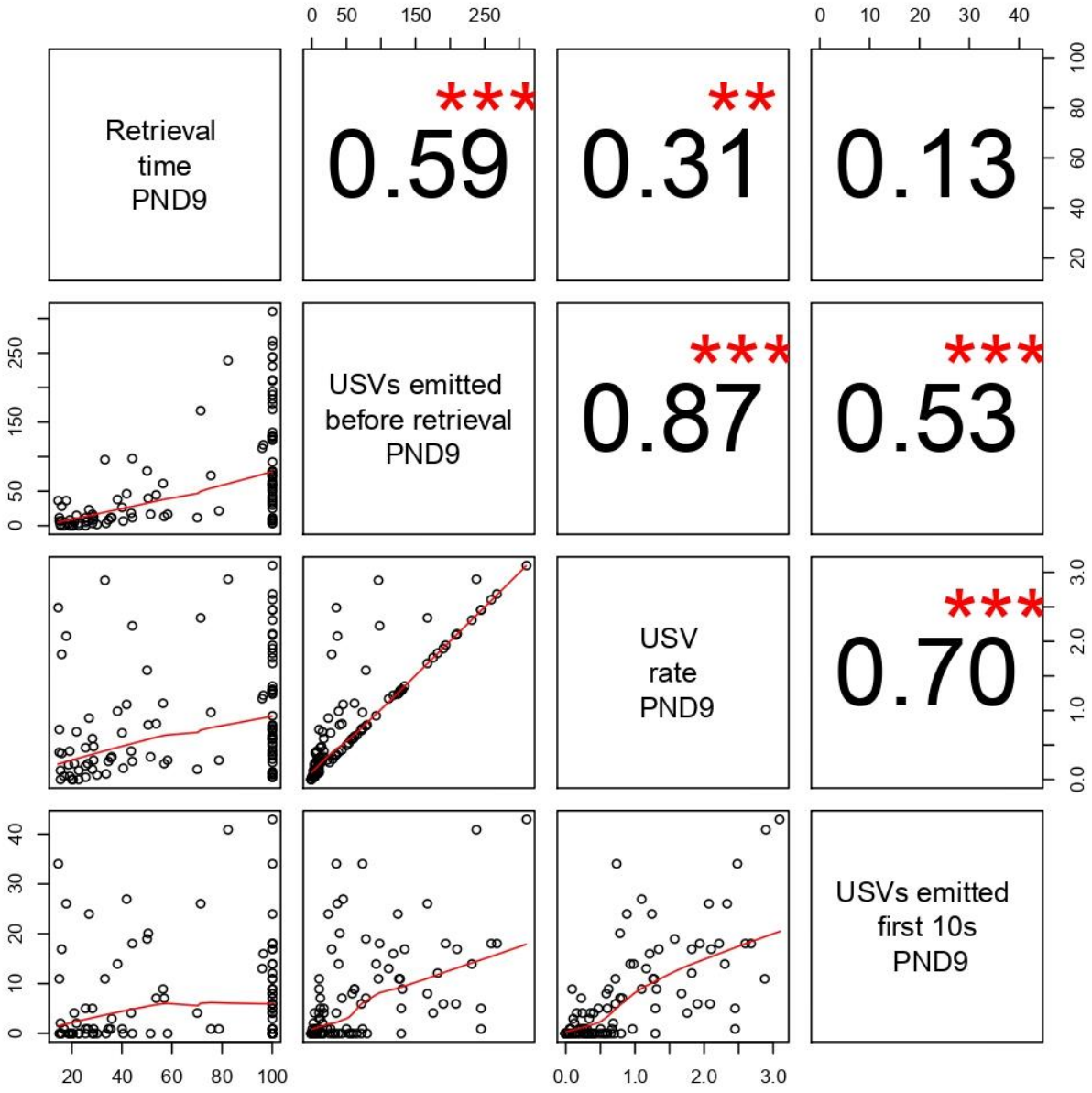
4.6.8 Supplementary Figure 6.

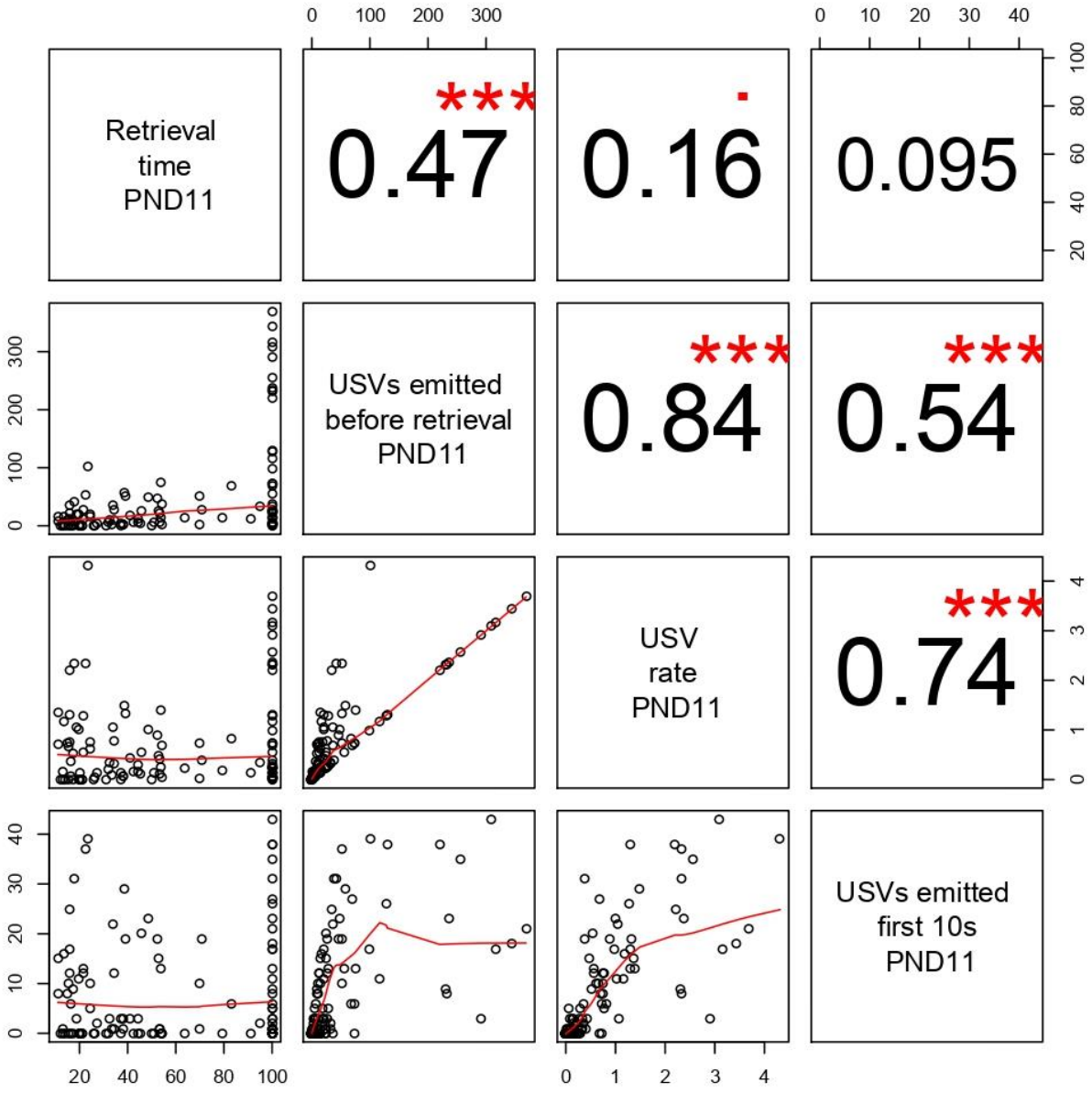
Pairwise correlational plots of retrieval time (s), USVs emitted before retrieval (number USVs), USV rate (USVs/s) and USVs emitted in the first 10 seconds of PRT trial (number USVs). Pairwise correlation plots are first given for all test days combined (PND5-13) and then separately for PND5, 7, 9, 11 and 13. Below diagonal the pairwise correlation plot is shown. Above the diagonal the estimated Pearson correlation coefficient is given with significance value. (°p<0.10, *p<0.05, **p<0.01, ***p<0.001)

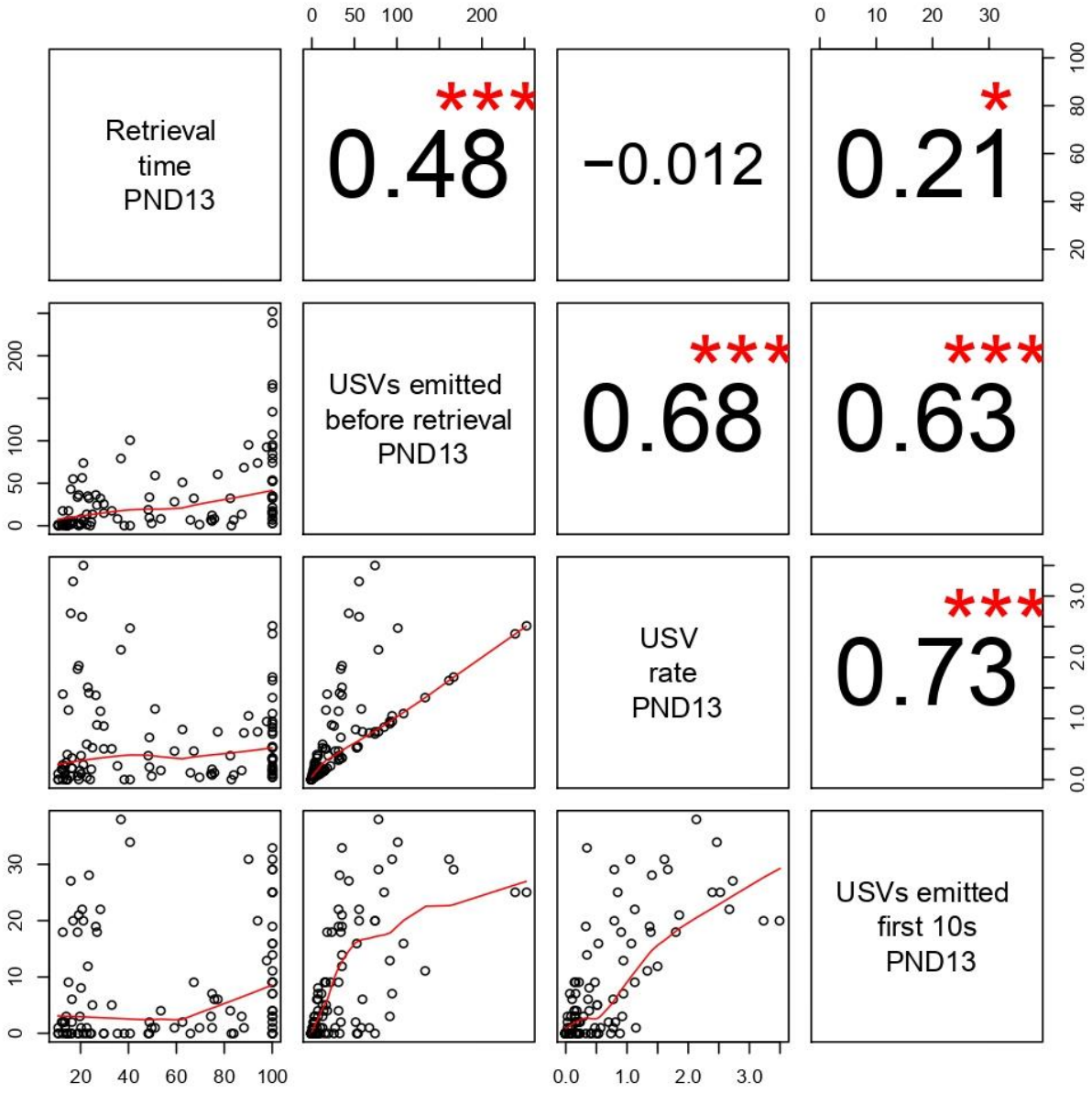



Retrieval time PND5
0.45
***
−0.03
0.042
USVs emitted before retrieval PND5
0.71
***
0.49
***
USV rate PND5
0.62
***
USVs emitted first 10s PND5
0 20 40 60 80
0 5 10 20 30
20 40 60 80 100
0.0 1.0 2.0


Retrieval time PND7
0.61
***
0.33
***
0.29
**
USVs emitted before retrieval PND7
0.87
***
0.60
***
USV rate PND7
0.71
***
USVs emitted first 10s PND7
0 100 200 300
0 10 20 30 40 50 60
0 100 200 300
20 40 60 80 100
0 1 2 3
0 10 20 30 40 50 60
20 40 60 80 100
0 1 2 3

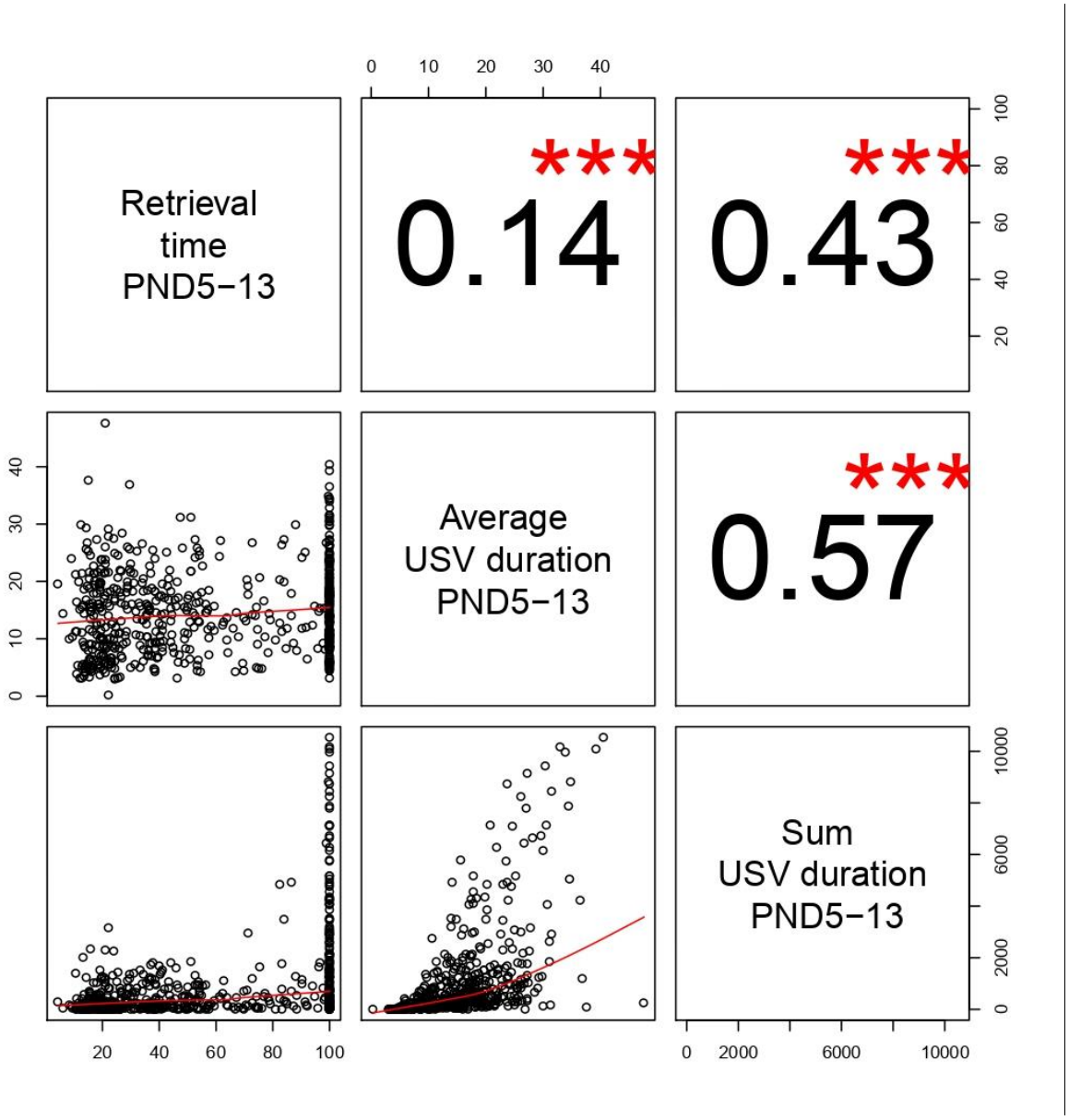

Retrieval time PND9
0.59
***
0.31
**
0.13
USVs emitted before retrieval PND9
0.87
***
0.53
***
USV rate PND9
0.70
***
USVs emitted first 10s PND9
0 50 150 250
0 10 20 30 40
20 40 60 80 100
0 50 150 250
0.0 1.0 2.0 3.0
0 10 20 30 40

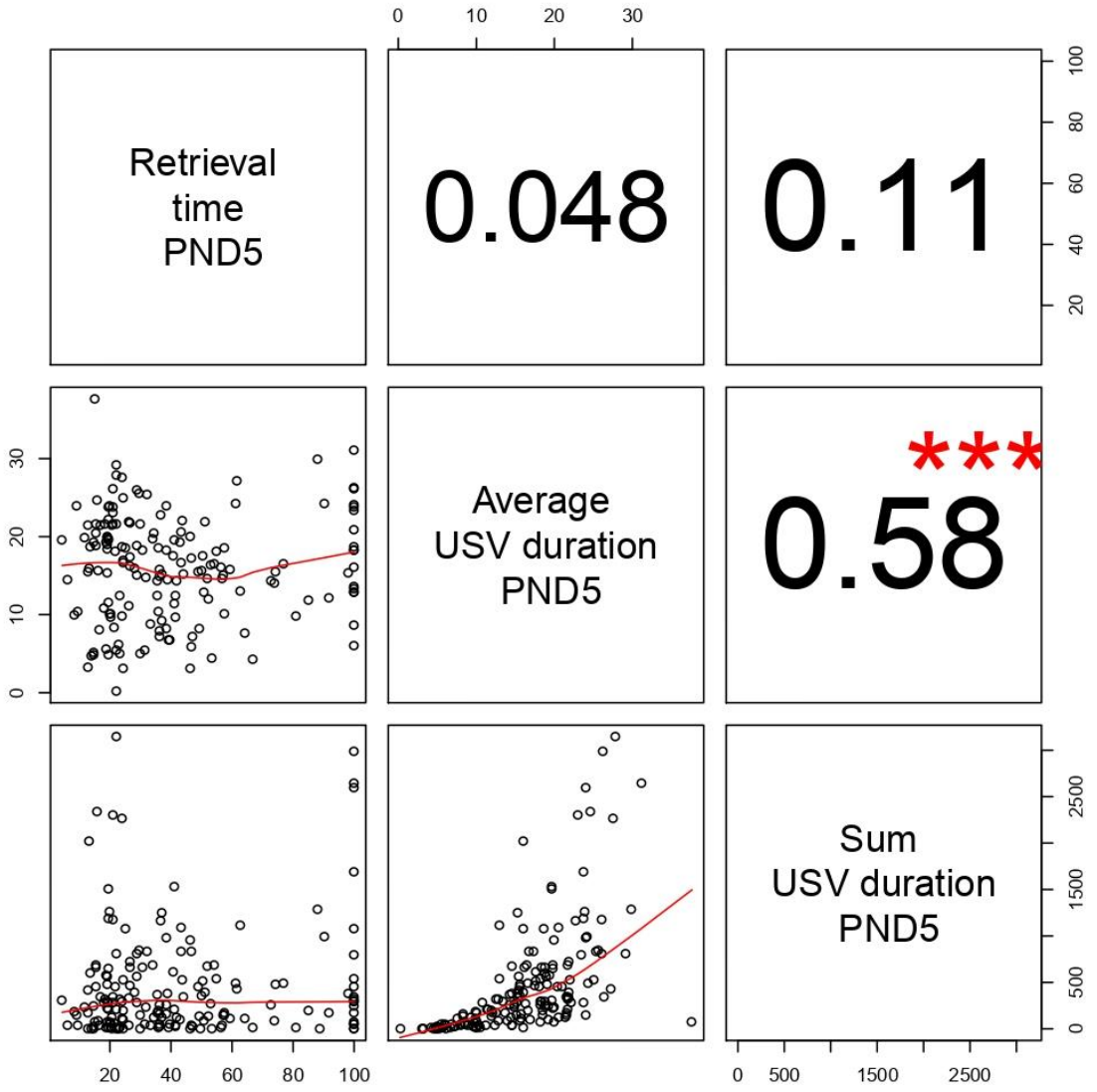

Retrieval time PND11
0.47
***
0.16
.
0.095
USVs emitted before retrieval PND11
***
0.84
***
0.54
USV rate PND11
***
0.74
USVs emitted first 10s PND11
0
100
200
300
0
10
20
30
40
0
100
200
300
20
40
60
80
100
0
1
2
3
4

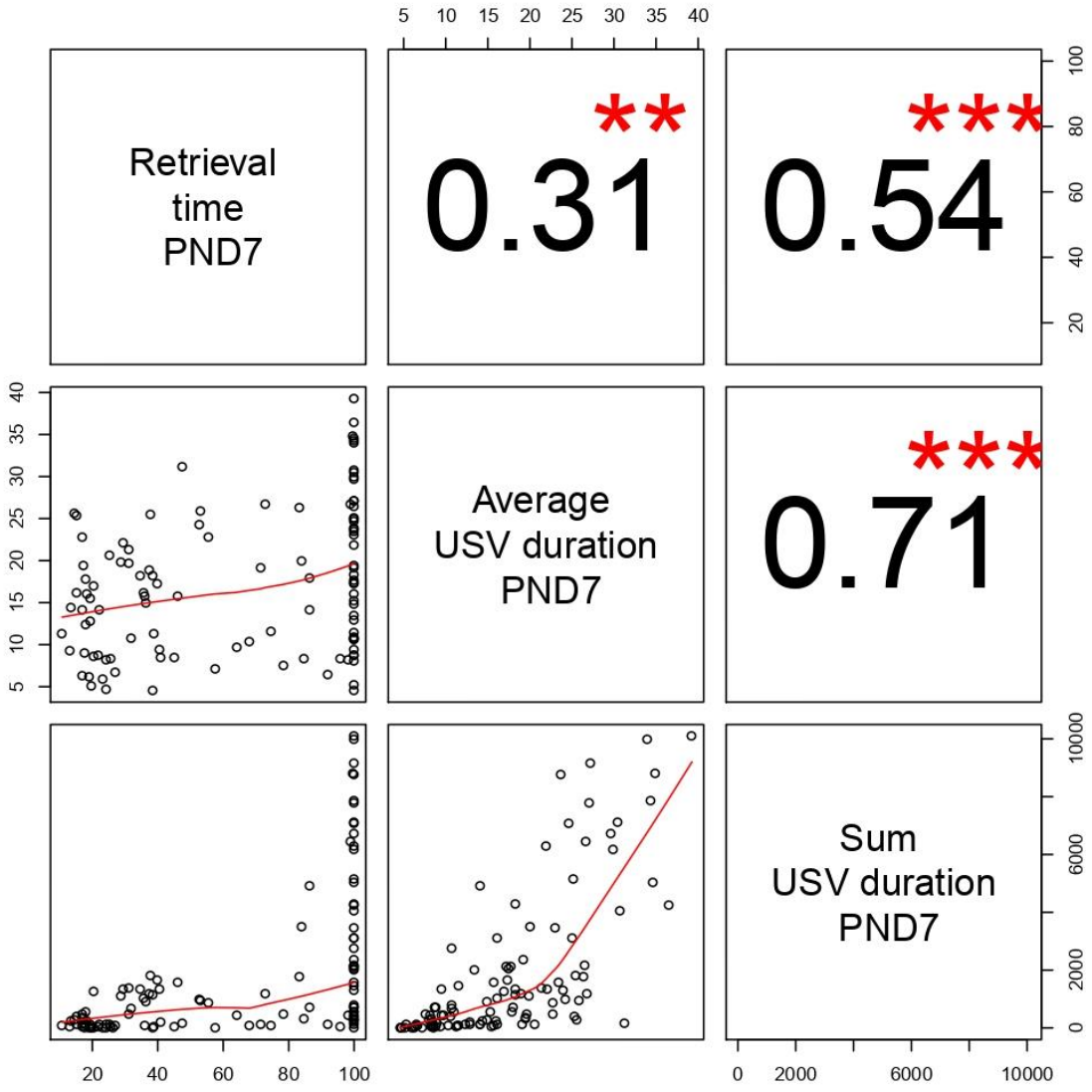

Retrieval time PND13
0.48
***
−0.012
0.21
*
USVs emitted before retrieval PND13
0.68
***
0.63
***
USV rate PND13
0.73
***
USVs emitted first 10s PND13
0 50 100 200
0 10 20 30
0 50 100 200
20 40 60 80 100
0.0 1.0 2.0 3.0
0 10 20 30
0.0 1.0 2.0 3.0

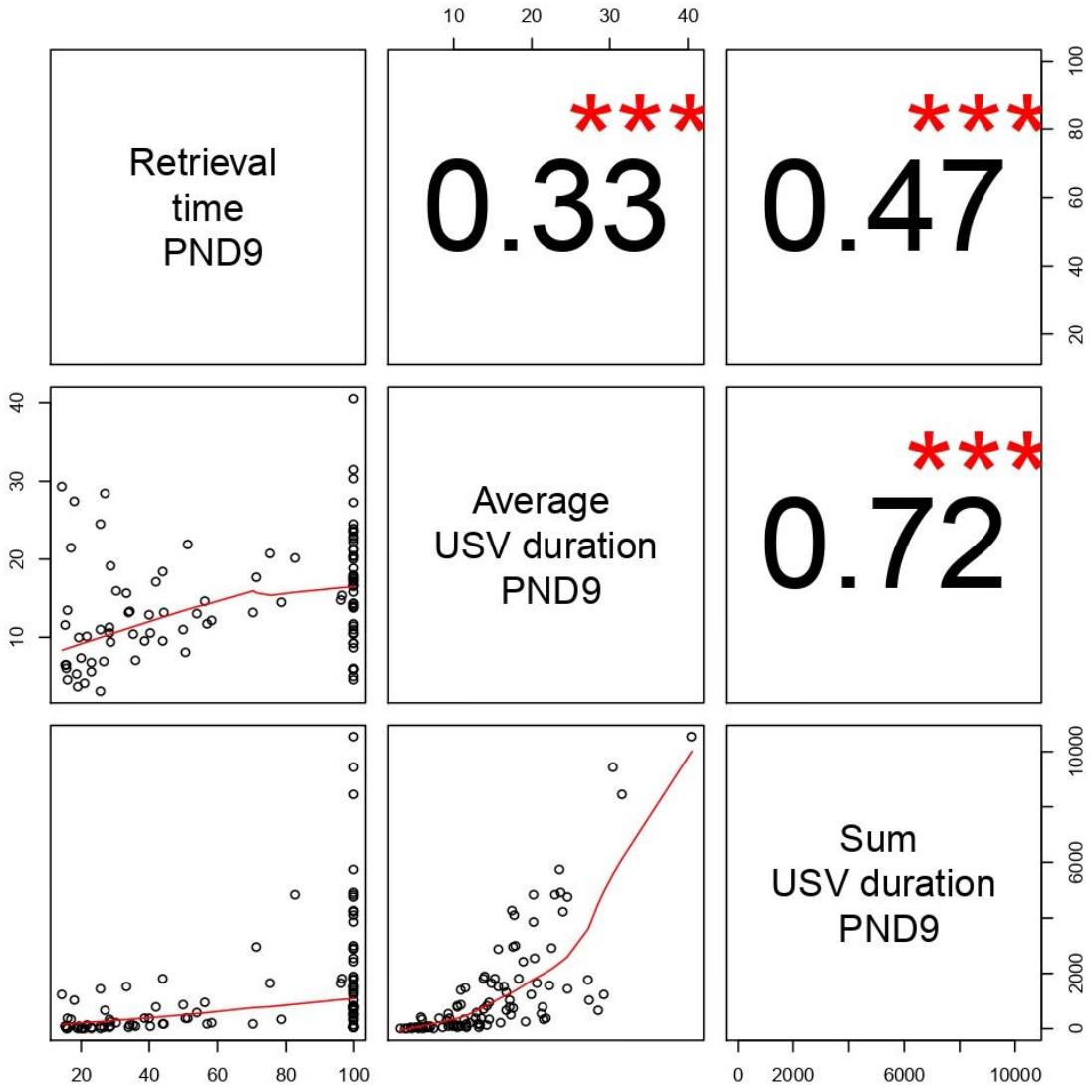
4.6.9 Supplementary Figure 7.

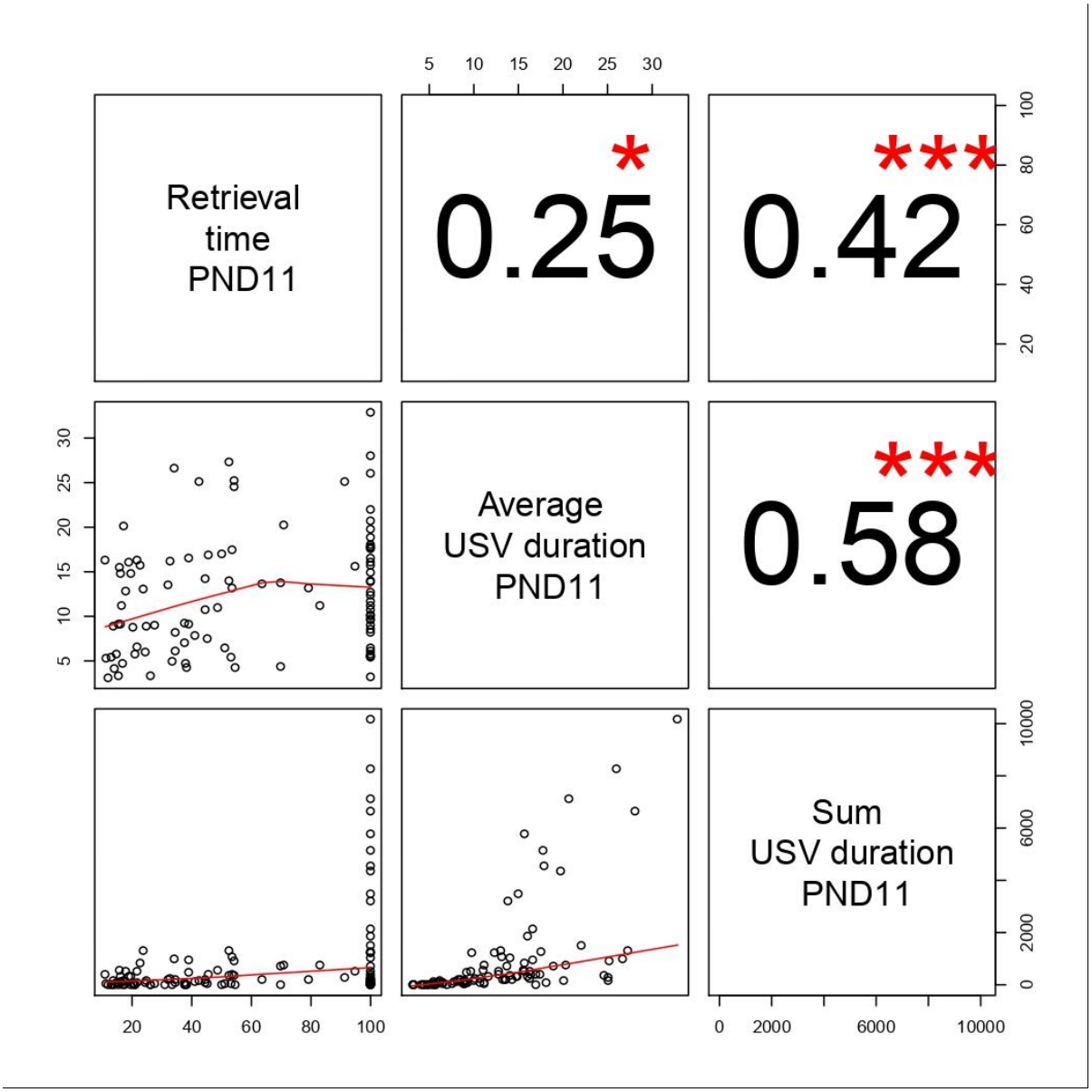
Pairwise correlational plots of retrieval time (s), average USV duration (s) and sum of USV duration (s). Pairwise correlation plots are first given for all test days combined (PND5-13) and then separately for PND5, 7, 9, 11 and 13. Below diagonal the pairwise correlation plot is shown. Above the diagonal the estimated Pearson correlation coefficient is given with significance value. (°p<0.10, *p<0.05, **p<0.01, ***p<0.001)

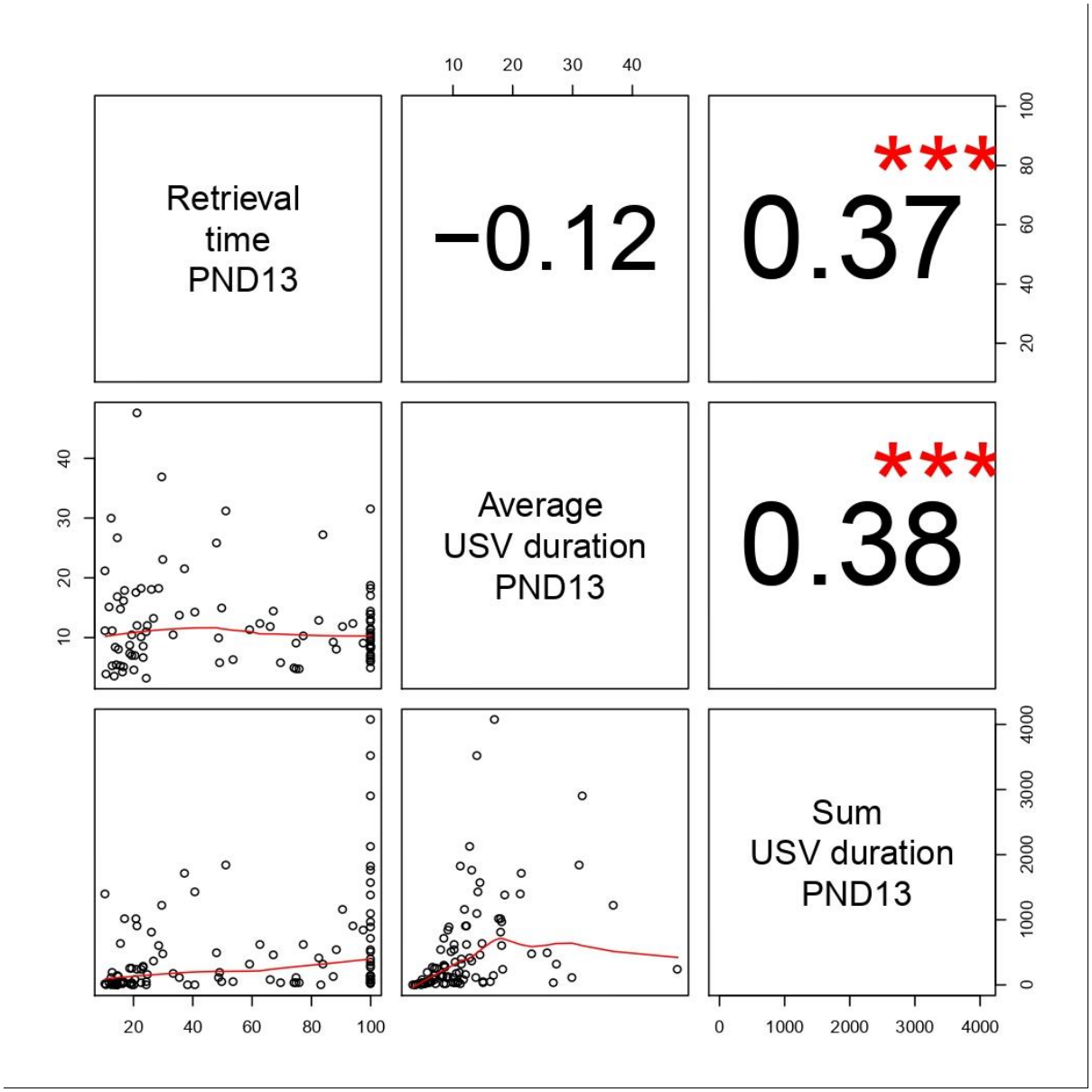



Retrieval time PND5
0.048
0.11
Average USV duration PND5
***
0.58
Sum USV duration PND5
0
10
20
30
20
40
60
80
100
0
500
1500
2500


Retrieval time PND7
0.31
**
0.54
***
Average USV duration PND7
0.71
***
Sum USV duration PND7
5 10 15 20 25 30 35 40
20 40 60 80 100
0 2000 6000 10000


Retrieval time PND9
0.33
***
0.47
***
Average USV duration PND9
0.72
***
Sum USV duration PND9
10
20
30
40
20
40
60
80
100
0
2000
6000
10000


Retrieval time PND11
0.25
*
0.42
***
Average USV duration PND11
0.58
***
Sum USV duration PND11
5 10 15 20 25 30
20 40 60 80 100
0 2000 6000 10000


Retrieval time PND13
−0.12
***
0.37
Average USV duration PND13
***
0.38
Sum USV duration PND13
10
20
30
40
10
20
30
40
20
40
60
80
100
20
40
60
80
100
0
1000
2000
3000
4000
0
1000
2000
3000
4000

4.7 Declarations

Availability of data and material: The expanded dam-infant tracking model and infant USV detection model for this study are available at: https://doi.org/10.17605/OSF.IO/VEJ4H. Further inquiries can be directed to the corresponding author (carmenwinters@hotmail.com)

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding: This study was funded by an SB PhD fellowship (1S05818N) of the Research Foundation Flanders (FWO) to Carmen Winters. Wim Gorssen was funded by an FR PhD fellowship (1104320N) of the Research Foundation Flanders. The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Authors' contributions: C.W. and R.D. designed the experimental strategy. C.W. optimized experimental procedures. C.W. and W.G., conceptualized and wrote the code. C.W. labeled the data. C.W. wrote the manuscript with input from all authors. All authors approved the final manuscript.

Acknowledgements: We thank Jan Clemens of DeepAudioSegmenter for helping to optimize the structural parameters in DAS, Anamarija Banjac for meticulous assistance in data sampling and Louise Moonen for animal facility support. Further, we would like to acknowledge the death of Bambi's mother as one of the most powerful moments in Disney cinemagraphic history by symbolizing the heartbreaking distress following mother-infant bond disruption.

4.8 References

1. Nowak, R., Porter, R., Levy, F., Orgeur, P., & Schaal, B. (2000). Role of mother-young interactions in the survival of offspring in domestic mammals. *Reviews of Reproduction*, *5*(3), 153–163. https://doi.org/10.1530/revreprod/5.3.153
2. Wöhr, M., Oddi, D., & D'Amato, F. R. (2010). Effect of altricial pup ultrasonic vocalization on maternal behavior. In *Handbook of Behavioral Neuroscience* (Vol. 19, Issue C, pp. 159–166). Elsevier Science & Technology. https://doi.org/10.1016/B978-0-12-374593-4.00016-4
3. Bornstein, M. H., Putnick, D. L., Rigo, P., Esposito, G., Swain, J. E., Suwalsky, J. T. D., Su, X., Du, X., Zhang, K., Cote, L. R., De Pisapia, N., & Venuti, P. (2017). Neurobiology of culturally common maternal responses to infant cry. *Proceedings of the National Academy of Sciences - PNAS*, *114*(45), E9465–E9473. https://doi.org/10.1073/pnas.1712022114
4. Hahn, M. E., & Lavooy, M. J. (2005). A review of the methods of studies on infant ultrasound production and maternal retrieval in small rodents. *Behavior Genetics*, *35*(1), 31–52. https://doi.org/10.1007/s10519-004-0854-7
5. Scattoni, M. L., Crawley, J., & Ricceri, L. (2009). Ultrasonic vocalizations: A tool for behavioural phenotyping of mouse models of neurodevelopmental disorders. *Neuroscience and Biobehavioral Reviews*, *33*(4), 508–515. https://doi.org/10.1016/j.neubiorev.2008.08.003
6. Potasiewicz, A., Holuj, M., Piotrowska, D., Zajda, K., Wojcik, M., Popik, P., & Nikiforuk, A. (2019). Evaluation of ultrasonic vocalizations in a neurodevelopmental model of schizophrenia during the early life stages of rats. *Neuropharmacology*, *146*, 28–38. https://doi.org/10.1016/j.neuropharm.2018.11.023
7. Reynolds, C. D., Nolan, S. O., Huebschman, J. L., Hodges, S. L., & Lugo, J. N. (2017). Early-life status epilepticus acutely impacts select quantitative and qualitative features of neonatal vocalization behavior: Spectrographic and temporal characterizations in C57BL/6 mice. *Epilepsy & Behavior*, *72*, 58–62. https://doi.org/10.1016/j.yebeh.2017.04.025

8. Caldji, C., Diorio, J., & Meaney, M. J. (2000). Variations in maternal care in infancy regulate the development of stress reactivity. *Biological Psychiatry*, *48*(12), 1164–1174. https://doi.org/10.1016/S0006-3223(00)01084-2
9. Meaney, M. J., Liu, D., Diorio, J., Day, J. C., & Francis, D. D. (2000). Maternal care, hippocampal synaptogenesis and cognitive development in rats. *Nature Neuroscience*, *3*(8), 799–806. https://doi.org/10.1038/77702
10. Braungart-Rieker, J. M., Garwood, M. M., Powers, B. P., & Wang, X. (2001). Parental Sensitivity, Infant Affect, and Affect Regulation: Predictors of Later Attachment. *Child Development*, *72*(1), 252–270. https://doi.org/10.1111/1467-8624.00277
11. Curley, J. P., & Champagne, F. A. (2015). Influence of maternal care on the developing brain: Mechanisms, temporal dynamics and sensitive periods. *Frontiers in Neuroendocrinology*, *40*, 52–66. https://doi.org/10.1016/j.yfrne.2015.11.001
12. Shin, H., Park, Y. J., Ryu, H., & Seomun, G. A. (2008). Maternal sensitivity: A concept analysis. Journal of advanced nursing, 64(3), 304-314.
13. Sewell, G. D. (1970). Ultrasonic Communication in Rodents. *Nature (London)*, *227*(5256), 410–410. https://doi.org/10.1038/227410a0
14. Smotherman, W. P., Bell, R. W., Starzec, J., Elias, J., & Zachman, T. A. (1974). Maternal responses to infant vocalizations and olfactory cues in rats and mice. *Behavioral Biology*, *12*(1), 55–66. https://doi.org/10.1016/S0091-6773(74)91026-8
15. Ehret, G. (1992). Categorical perception of mouse-pup ultrasounds in the temporal domain. *Animal Behaviour*, *43*(3), 409–416. https://doi.org/10.1016/S0003-3472(05)80101-0
16. Ehret, G. (2005). Infant rodent ultrasounds - A gate to the understanding of sound communication. *Behavior Genetics*, *35*(1), 19–29. https://doi.org/10.1007/s10519-004-0853-8
17. Ehret, G., & Haack, B. (1982). Ultrasound recognition in house mice: key-stimulus configuration and recognition mechanism. *Journal of comparative physiology*, *148*(2), 245-251. https://doi.org/10.1007/BF00619131
18. Wöhr, M., & Schwarting, R. K. W. (2008). Maternal Care, Isolation-Induced Infant Ultrasonic Calling, and Their Relations to Adult Anxiety-Related Behavior in the Rat. Behavioral Neuroscience, 122(2), 310–330. https://doi.org/10.1037/0735-7044.122.2.310
19. Bowers, J. M., Perez-Pouchoulen, M., Edwards, N. S., & McCarthy, M. M. (2013). Foxp2 mediates sex differences in ultrasonic vocalization by rat pups and Directs order of maternal retrieval. *The Journal of Neuroscience*, *33*(8), 3276–3283. https://doi.org/10.1523/JNEUROSCI.0425-12.2013
20. Abuaish, S., Tse, E. K., & McGowan, P. O. (2020). Perinatal high-fat diet impairs pup retrieval and induces sex-specific changes in ultrasonic vocalization characteristics of rat pups. *Developmental Psychobiology*, *62*(4), 436–445. https://doi.org/10.1002/dev.21923
21. Sukoff Rizzo, S. J., & Silverman, J. L. (2016). Methodological considerations for optimizing and validating behavioral assays. *Current protocols in mouse biology*, *6*(4), 364-379.
22. Ey, E., de Chaumont, F., & Bourgeron, T. (2020). Spontaneous social communication in laboratory mice - placing ultrasonic vocalizations in their behavioral context. *bioRxiv*. https://doi.org/10.1101/2020.07.09.195362
23. Vogel, A. P., Tsanas, A., & Scattoni, M. L. (2019). Quantifying ultrasonic mouse vocalizations using acoustic analysis in a supervised statistical machine learning framework. *Scientific Reports*, *9*(1), 8100–8100. https://doi.org/10.1038/s41598-019-44221-3
24. Kelly, S. J., Day, N., & Streissguth, A. P. (2000). Effects of prenatal alcohol exposure on social behavior in humans and other species. *Neurotoxicology and Teratology*, *22*(2), 143–149. https://doi.org/10.1016/S0892-0362(99)00073-2
25. Bosque Ortiz, G. M., Santana, G. M., & Dietrich, M. O. (2022). Deficiency of the paternally inherited gene Magel2 alters the development of separation-induced vocalization and maternal behavior in mice. *Genes, Brain and Behavior*, *21*(1), e12776–n/a. https://doi.org/10.1111/gbb.12776

26. Winters, C., Gorssen, W., Ossorio-Salazar, V. A., Nilsson, S., Golden, S., & D'Hooge, R. (2022). Automated procedure to assess pup retrieval in laboratory mice. *Scientific Reports*, *12*(1), 1663–1663. https://doi.org/10.1038/s41598-022-05641-w
27. Mastinu, A., Premoli, M., Maccarinelli, G., Grilli, M., Memo, M., & Bonini, S. A. (2018). Melanocortin 4 receptor stimulation improves social deficits in mice through oxytocin pathway. *Neuropharmacology*, *133*, 366–374. https://doi.org/10.1016/j.neuropharm.2018.02.007
28. Esposito, G., Truzzi, A., Yoshida, S., Ohnishi, R., Miyazawa, E., & Kuroda, K. O. (2019). Using maternal rescue of pups in a cup to investigate mother-infant interactions in mice/rodents. *Behavioural Brain Research*, *374*, 112081–112081. https://doi.org/10.1016/j.bbr.2019.112081
29. Steinfath, E., Palacios-Munoz, A., Rottschaefer, J. R., Yuezak, D., & Clemens, J. (2021). Fast and accurate annotation of acoustic signals with deep neural networks. *eLife*, *10*. https://doi.org/10.7554/eLife.68837.sa2
30. Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9), 1281–1289. https://doi.org/10.1038/s41593-018-0209-y
31. Gorssen, W., Winters, C., Meyermans, R., D'Hooge, R., Janssens, S., & Buys, N. (2022). Estimating genetics of body dimensions and activity levels in pigs using automated pose estimation. *Scientific Reports*, *12*(1), 15384–15384. https://doi.org/10.1038/s41598-022-19721-4
32. Nilsson, S. R., Goodwin, N. L., Choong, J. J., Hwang, S., Wright, H. R., Norville, Z. C.,... & Golden, S. A. (2020). Simple Behavioral Analysis (SimBA)–an open source toolkit for computer classification of complex social behaviors in experimental animals. *BioRxiv*. https://doi.org/10.1101/2020.04.19.049452
33. Krishnan, R., Sinha, A., Ahuja, N., Subedar, M., Tickoo, O., & Iyer, R. (2021). Mitigating Sampling Bias and Improving Robustness in Active Learning. https://doi.org/10.48550/arxiv.2109.06321
34. Scattoni, M. L., Gandhy, S. U., Ricceri, L., & Crawley, J. N. (2008). Unusual repertoire of vocalizations in the BTBR T+tf/J mouse model of autism. *PloS One*, *3*(8), e3067–e3067. https://doi.org/10.1371/journal.pone.0003067
35. Ey, E., Torquet, N., Le Sourd, A.-M., Leblond, C. S., Boeckers, T. M., Faure, P., & Bourgeron, T. (2013). The Autism ProSAP1/Shank2 mouse model displays quantitative and structural abnormalities in ultrasonic vocalisations. *Behavioural Brain Research*, *256*, 677–689. https://doi.org/10.1016/j.bbr.2013.08.031
36. Shahrier, M. A., & Wada, H. (2018). Effects of prenatal ethanol exposure on acoustic characteristics of ultrasonic vocalizations in rat pups. *Neurotoxicology (Park Forest South)*, *69*, 29–36. https://doi.org/10.1016/j.neuro.2018.08.006
37. de Chaumont, F., Ey, E., Torquet, N., Lagache, T., Dallongeville, S., Imbert, A., Legou, T., Le Sourd, A.-M., Faure, P., Bourgeron, T., & Olivo-Marin, J.-C. (2019). Real-time analysis of the behaviour of groups of mice via a depth-sensing camera and machine learning. *Nature Biomedical Engineering*, *3*(11), 930–942. https://doi.org/10.1038/s41551-019-0396-1
38. D'Amato, F. R., Scalera, E., Sarli, C., & Moles, A. (2005). Pups call, mothers rush: Does maternal responsiveness affect the amount of ultrasonic vocalizations in mouse pups? *Behavior Genetics*, *35*(1), 103–112. https://doi.org/10.1007/s10519-004-0860-9
39. Yoshida, S., Esposito, G., Ohnishi, R., Tsuneoka, Y., Okabe, S., Kikusui, T., Kato, T., & Kuroda, K. O. (2013). Transport Response is a filial-specific behavioral response to maternal carrying in C57BL/6 mice. *Frontiers in Zoology*, *10*(1), 1–11. https://doi.org/10.1186/1742-9994-10-50
40. Wöhr, M., Dahlhoff, M., Wolf, E., Holsboer, F., Schwarting, R. K. W., & Wotjak, C. T. (2008). Effects of Genetic Background, Gender, and Early Environmental Factors on Isolation-Induced Ultrasonic Calling in Mouse Pups: An Embryo-Transfer Study. *Behavior Genetics*, *38*(6), 579–595. https://doi.org/10.1007/s10519-008-9221-4
41. Sungur, A. Ö., Schwarting, R. K. W., & Wöhr, M. (2016). Early communication deficits in the Shank1 knockout mouse model for autism spectrum disorder: Developmental aspects and effects of social context. *Autism Research*, *9*(6), 696–709. https://doi.org/10.1002/aur.1564

42. Stolzenberg, D. S., Stevens, J. S., & Rissman, E. F. (2012). Experience-facilitated improvements in pup retrieval, evidence for an epigenetic effect. *Hormones and Behavior*, *62*(2), 128–135. https://doi.org/10.1016/j.yhbeh.2012.05.012
43. Dunlap, A. G., Besosa, C., Pascual, L. M., Chong, K. K., Walum, H., Kacsoh, D. B., Tankeu, B. B., Lu, K., & Liu, R. C. (2020). Becoming a better parent: Mice learn sounds that improve a stereotyped maternal behavior. *Hormones and Behavior*, *124*, 104779–104779. https://doi.org/10.1016/j.yhbeh.2020.104779
44. Heyne, G. W., Plisch, E. H., Melberg, C. G., Sandgren, E. P., Peter, J. A., & Lipinski, R. J. (2015). A Simple and Reliable Method for Early Pregnancy Detection in Inbred Mice. *Journal of the American Association for Laboratory Animal Science*, *54*(4), 368–371.

5 Beyond the Squeaks: understanding the interplay between deviant sociocommunicative behavior, maternal responses and the oxytocin system

C. Winters[1,3*], W. Gorssen[2], A. Banjac[4] & R. D'Hooge[1,3]

[1] Laboratory for Biological Psychology, KU Leuven, Tiensestraat 102 – PO Box 3714, 3000, Leuven, Belgium

[2] Department of Biosystems, Center for Animal Breeding and Genetics, KU Leuven, Leuven, Belgium

[3] KU Leuven, Leuven Brain Institute, Leuven B-3000, Belgium

[4] University of Ljubljana, Faculty of Medicine, Institute of Pathophysiology, Brain Research Lab

Keywords: early-life, communication, mouse, behavior, pup retrieval, ultrasonic vocalizations, oxytocin.

5.1 Abstract

The interactions between dams and their offspring during early life stages are pivotal for the development of social and emotional behaviors in rodents. Disruptions in these interactions during critical periods of development can exert enduring effects on neurodevelopment and behavior in offspring. To date, the impact of deviant infant sociocommunicative abilities on maternal behavior and dam-infant interactions remains largely unexplored. A key regulator of early-life dam-infant interactions is oxytocin, a neuropeptide involved in social bonding and maternal behavior.

This study aimed to provide novel insights into the impact of impaired early-life sociocommunicative abilities on maternal care throughout early-life development. Dams were prenatally injected with valproic acid (VPA) to establish offspring with a phenotype resembling autism spectrum disorder (ASD) with early-life sociocommunicative deficits. To provide evidence that this infant phenotype rather than the maternal VPA injection affects maternal behavior and mother-infant interactions, pups were treated with THIQ to stimulate endogenous oxytocin activity through the melanocortin 4 receptor agonist.

Our hypothesis posited that early postnatal MC4R stimulation in VPA-exposed pups would modulate maternal behavior and dam-infant interactions in a dose-dependent manner. To investigate this hypothesis, we conducted repeated assessments using the Bidirectional Automated Mother-pup Behavioral Interaction (BAMBI) test, while also regularly evaluating physical and developmental parameters.

Our findings indicated that early postnatal MC4R stimulation influenced weight gain, early developmental milestones, retrieval success, and ultrasonic vocalizations (USVs) characteristics in the pups. These findings shed light on the dynamic nature of social development and underscore the significance of the mother-infant dyad as potential therapeutic targets for individuals affected by neurodevelopmental disorders.

5.2 Introduction

Early-life interactions between dams and their offspring play a crucial role in the development of social and emotional behaviors in rodents. Disruptions in these interactions during critical developmental periods can have long-lasting effects on offspring neurodevelopment and behavior (1-4). Early-life social challenges are a core characteristic in autism spectrum disorder (ASD) and manifest early in development, typically leading to a diagnosis around the age of three years (5-6). The valproic acid (VPA) mouse model, which involves the prenatal exposure to VPA, has been widely used to study ASD (7-9). Prenatal VPA exposure in mice has been shown to induce behavioral abnormalities resembling those observed in individuals with ASD, including deficits in social interactions and communication already early in life (7-17). Early in development, prenatally VPA exposed pups exhibit behaviors such as reduced tendency to approach maternal bedding and increased startle response in presence of maternal odor, indicating disruptions in the mother-infant bond (14). Additionally, previous studies have extensively demonstrated early sociocommunicative deficits in prenatally VPA exposed pups in form of quantitative and qualitative alterations in ultrasonic vocalizations (USVs) including a decreased number of emitted calls, average call duration and peak amplitude as well as some inconsistent findings in peak frequency (13-14, 18-20). It has been suggested that these alterations in USVs imply a functional difference and as such elicit less effectively maternal caretaking behaviors hereby further exacerbating social impairments (14, 21).

Oxytocin (OT), a neuropeptide involved in social bonding and maternal behavior, has emerged as a key regulator of early-life dam-infant interactions. OT signaling promotes the activation of attraction and approach mechanisms in both dam and offspring (22). In dams, OT facilitates maternal care behaviors such as nursing, grooming and protection of offspring. In infants, OT plays a role in the emission of ultrasonic vocalizations (USVs) and altered OT signaling has been observed in context of ASD (23-24). Genetic knockout (KO) studies in mice for genes related to OT signaling have reported behavioral phenotypes reminiscent of ASD such as decreased social communication and impaired social cognitive memory (23-24). Moreover, impairments of the OT system have been reported in several animal models of ASD such as genetic *Fmr1, Cntnap2, Magel2, Oprm1, Shank3, Nlgn-3* models or the environmental VPA model (24). It has been suggested that such alterations in the infant OT system can affect the quality and intensity of early-life dam-infant interactions and as such contribute (further) to the sociocommunicative difficulties observed in psychopathologies such as ASD (21). Recently, studies have reported that OT treatment early in life in animal models of ASD improves adult behavioral impairments and this effect is blocked by administration of an OT antagonist (19, 24).

However, as a therapeutic drug OT has some major limitations including poor blood-brain barrier

penetrance after peripheral administration as well as short half-life and poor targeted release of exclusively central OT receptors in specific brain sites (25). Recent studies have revealed an alternative route for OT stimulation via the *melanocortin 4 receptor* (MC4R) agonism (24, 26). The MC4R colocalizes with OT neurons in the paraventricular nucleus (PVN) and activation of MC4R has been shown to activate PVN OT neurons in mice. Peñagarikano et al. (24) showed that both acute administration of OT or a selective MC4R agonist in *Cntnap2* mice, a monogenetic model for ASD, improved social deficits in adulthood. This rescue of social deficits using a selective MC4R agonist was blocked by an OT antagonist. Further, they reported that chronic early postnatal treatment with OT resulted in more lasting behavioral effects. Similarly, daily treatment during the first week of life with a MC3/4R agonist and selective MC4R agonist promotes social behavior in adulthood in prairie voles (27).

The primary aim of the present study was to provide novel insights into the dyadic impact of impaired infant sociocommunicative abilities in a mouse model of ASD. We used mice prenatally exposed to VPA as they exhibit sociocommunicative deficits early in development and the impact of these deficits on early-life dam-infant interactions remains largely unexplored. To exclude potential side-effects of VPA exposure on maternal care behaviors, all pups were prenatally exposed to VPA and were postnatally treated with different doses of the drug THIQ, a selective and potent MC4R agonist, to stimulate endogenous OT activity in a dose-dependent manner. We hypothesize that early postnatal MC4R stimulation in mice prenatally exposed to VPA will affect impaired sociocommunicative abilities observed in the VPA model which in turn will impact maternal behavior and dam-infant interactions. To address this hypothesis, pups were repeatedly tested using the Bidirectional Automated Mother-pup Behavioral Interaction (BAMBI) test (28). Additionally, physical and motor parameters were regularly assessed. The findings from this research provide insights into the dynamic nature of social development and the role of the mother-infant dyad in potential therapeutic targets in individuals affected by neurodevelopmental disorders.

5.3 Materials and methods

5.3.1 Animal housing and breeding

C57BL/6J mice obtained from Janvier Labs (Le Genest-Saint-Isle, France) and the KU Leuven Animal Facility (Leuven, Belgium) were bred with a specific timing and maintained in conditioned rooms with a 12/12-hour light-dark cycle (lights on at 7 AM). The mice had ad libitum access to water and food, and the rooms were maintained at a temperature of 22°C with a humidity level of 30%. The morning following successful mating was designated as gestational day (GD) 0.5. From GD0.5 onward, pregnant dams were housed individually, and the confirmation of pregnancy occurred between GD7.5 and

GD10.5 by monitoring weight changes following the approach described by Heyne et al. (29). All experimental procedures were conducted in compliance with the guidelines set forth by the Animal Ethics Committee of KU Leuven (P028/2018), adhering to the European Community Council Directive 86/609/EEC, the ARRIVE guidelines, and the ILAR Guide to the Care and Use of Experimental Animals. These animals were the same as described in Winters et al. (28).

5.3.2 Experimental groups

To induce a neurodevelopmental model of autism, pregnant dams were administered a subcutaneous injection of valproic acid sodium salt (VPA) on gestational day 12.5. The VPA solution (600 mg/kg) was obtained from Sigma Aldrich (Taufkirchen, Germany) and dissolved in saline solution. The day of birth was designated as postnatal day 0 (PND0). To maintain standardized nest composition, large nests were culled to six pups on PND0, while ensuring a minimum of four pups with both sexes present. This resulted in a total of 29 dams with viable progeny, comprising 156 pups for subsequent testing.

The pups were subjected to daily subcutaneous injections from PND1 to PND7. They received either a placebo (0.0 mg/kg), low (0.5 mg/kg) or a high (2.0 mg/kg) dose of THIQ (N-[(1R)-1-[(4-Chlorophenyl)methyl]-2-[4-cyclohexyl-4-(1H-1,2,4-trazol-1-ylmethyl)-1-piperidinyl]-2-oxoethyl]-1,2,3,4-tetrahydro-3-isoquinolinecarboxamide), which was dissolved in PBS and 3.5% DMSO. The doses used were adapted from Mastinu et al. (30) and within one litter all pups received the same treatment. The total number of pups injected was as follows: 48 pups from 9 litters received the low THIQ dose, 50 pups from 10 litters received the high THIQ dose, and 58 pups from 10 litters received the PBS-DMSO control vehicle.

5.3.3 Maternal weight evolution

Pregnant dams were weighed (a) before pregnancy, (b) on GD12.5, 13.5, 16.5 and 17.5, and (c) on PND8, 19 and 29 to monitor weight evolution during pregnancy, lactation and weaning. The used weighing scale had an accuracy of 0.001g.

5.3.4 Pre- and postnatal nestbuilding

The building of a brood nest starts during pregnancy and continues for the first two weeks of lactation (31). As such, assessment of nest quality has been used as a readout for maternal behavior in mice. On GD16 and PND8, the cardboard house as well as existing nest were removed from the homecage and dams were provided with approximately 7.15 gram paper snippets and 3 nesting pads in their food grid. Cages were returned to their shelf in the animalium and left undisturbed until the next morning. Nest quality was graded for nest presence (0-1), nest closed (0-1), percentage of the materials used (weight of materials used/total weight of provided materials), and a 1 to 5 nesting score based on Hess et al. (32). In this grading methodology, a grade-1 was scored for a flat and unfocused nest, grade-2 for

a shallow bowl, grade-3 indicates a half of a sphere, grade-4 indicates a high-walled bowl which is open above, whereas a grade-5 nest indicates a nest constructed as a full dome (as shown in 32). For each quarter wall that was heightened, a nest receives another 0.25 points.

5.3.5 Pup development

All pups were weighed and assessed for measurements of growth rate and development daily from PND1-18. After weaning, on PND29, pups were weighed once more. Daily assessment of the body weights and developmental milestones was conducted between 10.00 and 12.00h. The developmental milestones included tail pinch reflex, ear detachment, eye opening and surface righting. For ear detachment, each pup was given a score depending on the number of ears detached: '0' for no ears, '1' for either left or right ear, and '2' for both ears. Similarly, eye opening was scored with '0' for no eyes, '1' for either left or right eye, and '2' for both eyes. The righting reflex was performed by gently placing the pup on its back and assessing whether it is able to flip over and place all four paws on the ground within 1 s. Righting was graded as successful with a '1' score and unsuccessful with a '0' score. Order of the cages was fixed, whereas the pup order was randomized. Developmental milestones were assessed up to the day all pups reached the milestone. After injection, pup tails were marked to prevent injecting the same pup twice. As we observed difference in response to this 'tail pinch' the reaction was graded with '0' for no reaction, '1' for only a vocal squeak, and '2' if the pup squeaked and rolled over on its back.

5.3.6 Bidirectional Automated Mother-pup Behavioral Interaction (BAMBI) test

The Bidirectional Automated Mother-pup Behavioral Interaction (BAMBI) test, as described in Winters et al. (28), was conducted to assess early-life dam-infant interactions.

Setup. The BAMBI test was performed inside the homecage which was placed inside a Styrofoam box (dimensions: 370 mm × 300 mm × 330 mm) to minimize external distractions. On PND7-13, a cup (diameter: 90 mm, height: 55 mm) was placed inside the homecage to prevent the pups from crawling back into the nest during the test. Further, the setup was equipped with a Foscam C2 IP-camera (EUport, Wageningen) above the homecage and an ultrasound microphone (Dodotronic Ultramic UM250K, Rome, Italy). The microphone was positioned approximately 5 cm above the the furthest corner from the nest or cup to minimize interference by USVs emitted by the pups in the nest.

Acclimation. Prior to the test, the dam and pups were habituated for 60 minutes to the testing room inside the homecage with the grid on top to minimize stress. The cup was introduced in the homecage at the start of these 60 min. Before testing, the homecage was placed inside the Styrofoam box without the grid on top and animals were allowed habituation for another five minutes.

Test session. In the BAMBI test, dam and litter were tested together. The entire test session was video recorded and for each trial a separate audio was recorded using Avisoft SASLab Lite software (Avisoft, Bioacoustics, Berlin, Germany). To initiate a trial, a single pup was carefully taken out the nest and placed in a clean, glass cup that had been pre-warmed to 35°C using a heating pad. A manually activated beep announced the start of a trial and the isolated pup was gently positioned in the furthest corner from the nest under a ultrasound microphone (Dodotronic Ultramic UM250K, Rome, Italy). Trials had a maximum duration of 100 seconds after the beep. If the pup was not retrieved within this time, it was returned to the nest by the experimenter. For each dam, the PRT was repeated six times on PND5, and due to practical constraints, it was performed four times per dam on PND7-13. Pup sex was counterbalanced for each dam and the pups were not marked during the test to avoid any interference caused by odor cues. Therefore it should be noted that individual pups could not be identified and as such a pup may have been tested multiple times. The maternal trial sequence referred to the specific order of trials conducted within a dam on a given testing day.

Data analysis. The recorded video session and individual audio files were preprocessed and synchronized as described in Winters et al. (28). Using the elaborated dam-pup retrieval DeepLabCut tracking algorithm (28) relevant behavioral paramaters were extracted using a custom-build R-script (33). USV detections was performed using a custom-build Deep Audio Segmenter (DAS, 34) model. A custom-build R-script enabled to perform quality control as well as extract parameters including latency to first USV.

5.3.7 Statistical Analysis

Statistical analyses were performed using the GLM package in R for (binomial) regression models and survival package in R for survival analysis via multivariate Cox regression. For the maternal body weights, linear regression models were corrected for age (covariate) and experimental condition (fixed effect). For the pup body weights, linear regression models were corrected for age (covariate), maternal age at birth (covariate), sex (fixed effect) and experimental condition (fixed effect). For the developmental milestones, binomial regression models and/or Cox regression were corrected for sex (fixed effect), and experimental condition (fixed effect). For the BAMBI test, binomial regression models and Cox regression was used for the binary trait retrieval success, correcting for sex (fixed effect), day of testing (fixed effect), experimental condition (fixed effect) and maternal trial sequence (covariate). For the continuous phenotypes from BAMBI, linear regression was used, correcting for sex (fixed effect), day of testing (fixed effect), experimental condition (fixed effect) and maternal trial sequence (covariate).

5.4 Results

5.4.1 Maternal weight evolution

There were no significant differences in maternal weights between the experimental THIQ groups before and during pregnancy. However, the VPA injection at a dose of 600 mg/kg appeared to have a detrimental impact on dam body weight during pregnancy, as illustrated in **Figure 1a**. Whereas dams had an average daily gain of 1.25g (SD=0.38, min=0.24, max=1.84) between GD11-12, they lost 0.53g (SD=0.38, min=-1.96, max=1.70) on average after VPA injection. Subsequently, maternal weight increased again after E13.5, although our validation study indicated that their weight trajectory remained lower compared to the saline-injected control group[1]. During postnatal development, dams in the control group exhibited significantly higher weights compared to both the high THIQ group (p=0.04) and the low THIQ group (p<0.01).

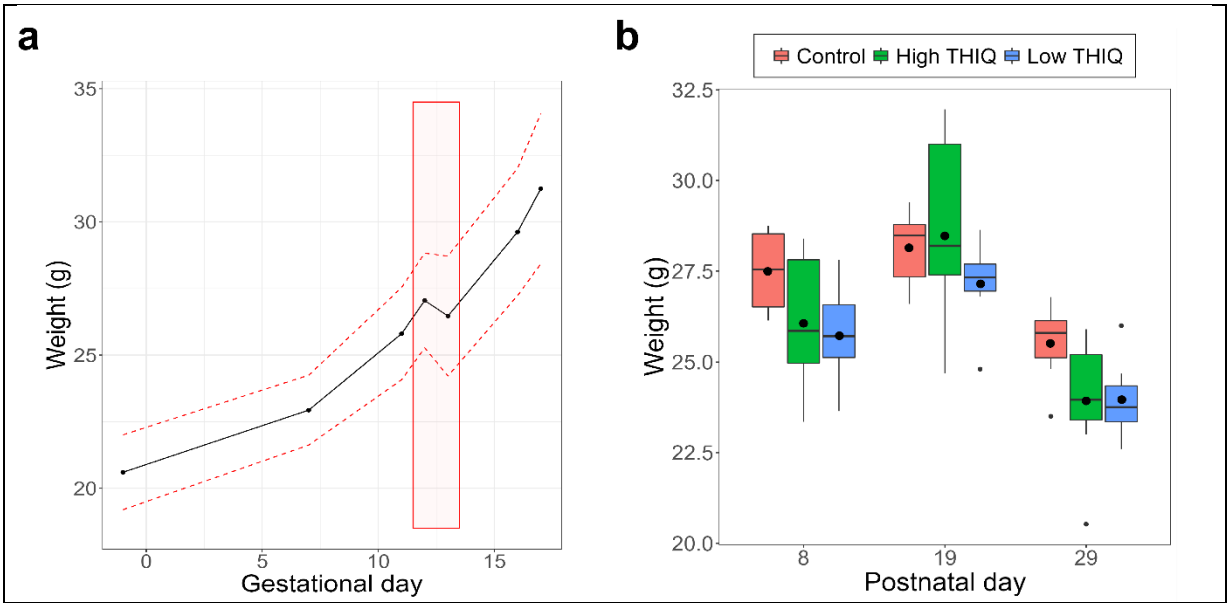


Figure 1. Maternal weight evolution. (**a**) Mean dam weight (g) during the gestational period (d). Maternal weight on average decreased after VPA injection. The solid black line indicates mean weight, whereas red dashed lines indicate ± 1SD and the red box highlights the impact of the VPA injection on dam weight. n=46 (**b**) Mean dam weight during the postnatal period. Dams of the control group weighed significantly more than dams from the high THIQ group (p=0.04) and the low THIQ group (p<0.01). Boxplots show median with upper and lower quartiles, whereas the black dot in the boxplot indicates the mean.

5.4.2 Gestational period and birth traits

Litters were pseudo-randomly assigned to one of three experimental THIQ groups (placebo, low THIQ or high THIQ) by a blinded experimenter. At birth, there were no significant differences observed between the experimental THIQ groups in terms of the number of pups born before culling, the number of alive males or females before culling, maternal age at birth, or maternal pre-pregnant weight. The dataset collected from postnatal day 1 (PND1) consisted of 156 pups derived from 29

[1] These data were collected in a validation experiment which is not part of current dissertation.

distinct mothers, comprising 85 females and 71 males. Throughout the period from PND1 to PND18, a total of eight pups died, and there was no apparent association between these deaths and the experimental treatment.

5.4.3 Maternal Nest building

The nesting behavior of the dams was assessed by recording their ability to construct a brooding nest overnight on GD16 and PND8. All dams, regardless of treatment, successfully built nests both prenatally and postnatally. The mean nest building score prenatally was 4.34 (SD=0.47, min=3.25, max=5.00) and postnatally 4.85 (SD=0.27, min=3.75, max=5.00). Statistical analysis revealed no significant differences in nest structure or quality between the THIQ-treated groups and the control mice.

5.4.4 Pup weight evolution

Daily weight measurements were obtained for pups from PND1-18. Notably, mice treated early in life with high THIQ ($p<0.001$) weighed significantly more compared to control animals, whereas mice treated with low THIQ did not differ from controls. Further, males weighed significantly more than females ($p<0.001$) and the weight effect of high THIQ significantly interacted with pup sex ($p<0.001$). That is, the effect of increased weight after high THIQ treatment was most prominent in females (**Figure 2a**). After weaning, on PND29, both high ($p=0.04$) and low ($p=0.06$) THIQ showed an increased weight compared to control animals. Moreover, males significantly weighed more than females ($p<0.001$) in all experimental groups.

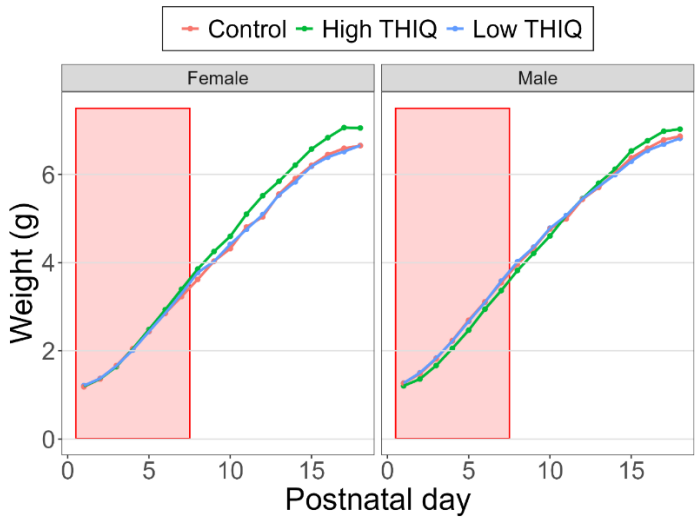


Figure 2: Mean pup weight evolution between PND1-18. Mean daily weight (g) of pups injected with sham, low THIQ or high THIQ between PND1-7. The pink box visualizes the injection period. The difference between the high THIQ group (green, n=43-51) and the control group (red, n=50-59) was significant ($p<0.001$). There was no difference between the low THIQ group (blue, n=36-43) and the control group. b. Sex differences for mean daily weight (g) of pups in three conditions. Whereas males have an increased weight compared to females in the controls and low THIQ conditions, males and females weigh similarly in the high THIQ group.

5.4.5 Developmental milestones

To assess the effect of THIQ treatment in early developmental stage, typical time-fixed behaviors and reflexes were tested. From PND6 to PND13 the pups' ability to surface right was tested. Pups treated with high or low THIQ showed a significant delay in the development of surface righting ability (High THIQ: log ratio=-46.31, p=0.03, Low THIQ: log ratio=-65.16, p<0.001). Although there was no significant effect of pup sex, THIQ treated females seemed to drive the effect of a delay in surface righting. The onset of eye opening occurred between PND13 and PND16. Compared to control pups, eye opening was delayed in both groups treated with THIQ (**Figure 3b**, low THIQ: log ratio=-63.18, p=0.01, high THIQ: log ratio=-65.24, p<0.01). Further, eye opening was developed faster in males compared to females (log ratio=166.50, p<0.01).

No other significant effects of THIQ treatment or sex were found ear detachment or tail pinch reflex.

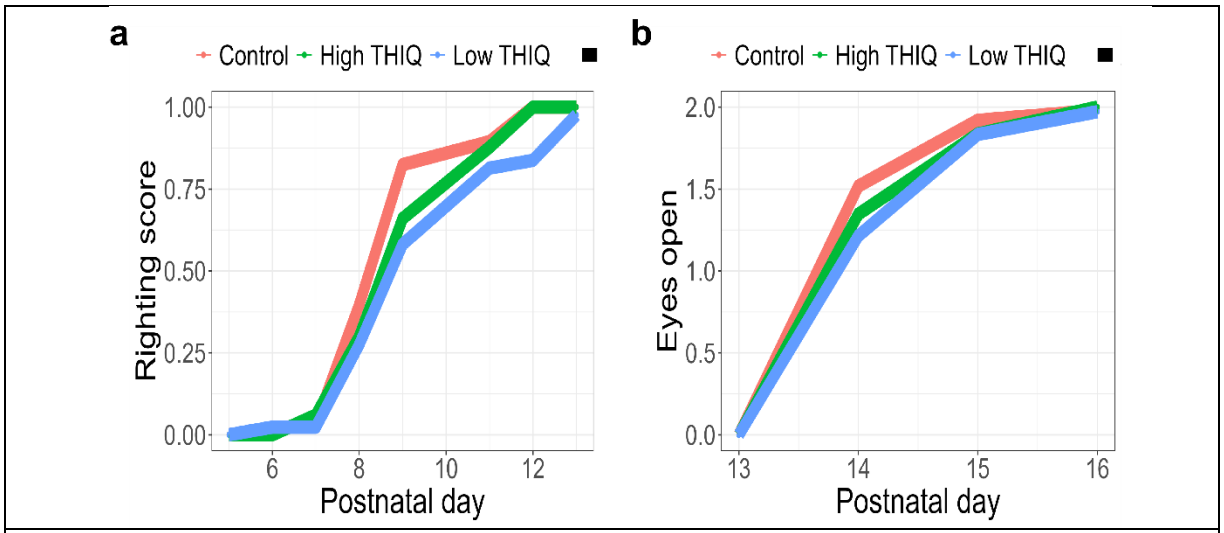


Figure 3. Evolution of developmental scores per condition over days. (**a**). Proportion of animals exhibiting mature surface righting (righting score). (**b**) mean number of opened eyes (eye opening score).

5.4.6 BAMBI: Maternal retrieval behavior

Day-by-day analysis showed different patterns in the THIQ conditions compared to controls in probability to be retrieved by the dam (**Table 1, Figure 4)**. On PND5, the probability to be retrieved is highest in control pups although differences were not significant. However, pups treated with high THIQ had a steadily increasing probability to be retrieved significantly faster than controls from PND7 to PND11 with a peak retrieval probability of +245.2% and a significant difference (p=0.02) on PND9. Hereafter, this pattern decreased and pups were retrieved later than control pups on PND13. Although pups treated with low THIQ did not show any significant differences compared to control pups, a similar pattern to the high THIQ condition was observed over all days. As retrieval success and latency to retrieval were highly, negatively correlated (r=-0.83) a similar pattern was observed for latency to retrieval. Dams of the high THIQ-treated pups retrieved their pups faster on PND7 (ES=-4.8s, p=0.57),

PND9 (ES=-11.8s, p=0.18), and PND11 (ES=-15.1s, p=0.07), but slower on PND13 (ES=+17.3s, p=0.05). Due to high internal correlations, similar patterns were also observed for latency to approach and latency to carry (**Supplemental Figure 1**).

PND	vs. control	Retrieval probability	p-value	vs. control	Retrieval probability	p-value
5	High THIQ	-59.1 %	0.17	Low THIQ	-44.2 %	0.40
7	High THIQ	+155.1 %	0.06	Low THIQ	+102.8 %	0.18
9	High THIQ	+245.2 %	0.02	Low THIQ	+11.0 %	0.84
11	High THIQ	+131.3 %	0.12	Low THIQ	+0.0 %	1.00
13	High THIQ	-65.6 %	0.08	Low THIQ	-64.2 %	0.10
5-13	High THIQ	+38.2 %	0.18	Low THIQ	-5.9 %	0.80

Table 1. Table showing day-by-day comparisons between THIQ conditions vs. controls in probability to be retrieved by the dam. PND5-13 indicates the overall results, combining all days. Results were obtained using binomial regression analysis.

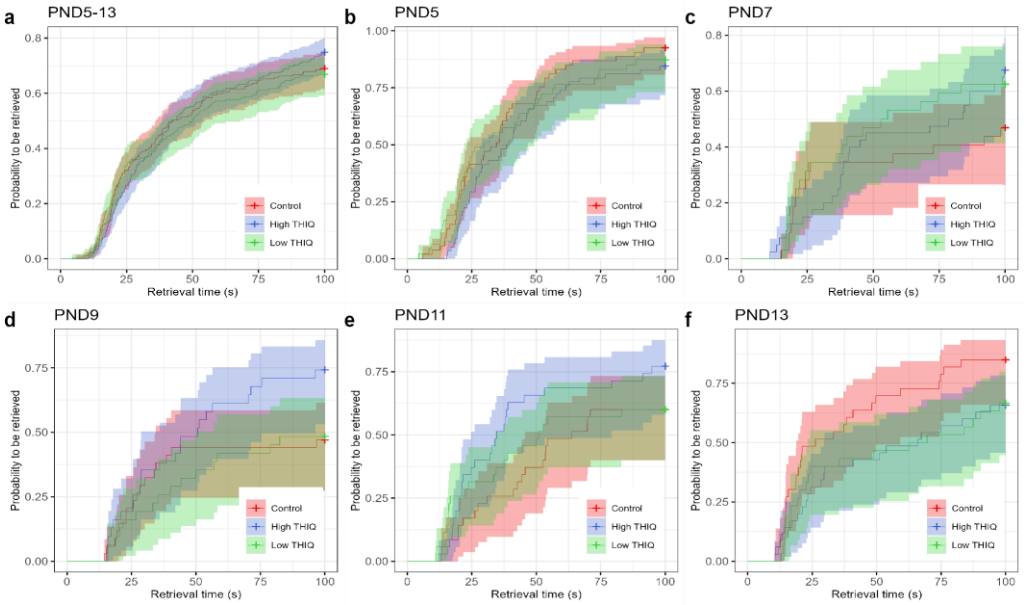


Figure 4. Survival curve for retrieval success in function of retrieval time (s). The y-axis shows the estimated probability of retrieval per experimental group in function of retrieval time (x-axis). As a consequence, retrieval probability increases over time and the end value at 100s represents the total estimated retrieval probability after the BAMBI test (marked with cross). Solid lines give the estimated retrieval probability, whereas colored areas indicate 95% confidence intervals. The titles represent at which postnatal day (PND) the test was performed. **a)** Retrieval probability for all tests combined (PND5-13). **b)** Retrieval probability on PND5. **c)** Retrieval probability on PND7. **d)** Retrieval probability on PND9. **e)** Retrieval probability on PND11. **f)** Retrieval probability on PND13.

5.4.7 BAMBI: Pup separation-induced vocalizations

Over all testing days, high THIQ-treated pups tended to emit a higher number of USVs during the first 10s of a trial, although the difference was not significant (ES=+1.4 USVs, p=0.09) and the largest differences were found on PND5-9 (**Table 2**). For pups treated with low THIQ no pronounced differnces

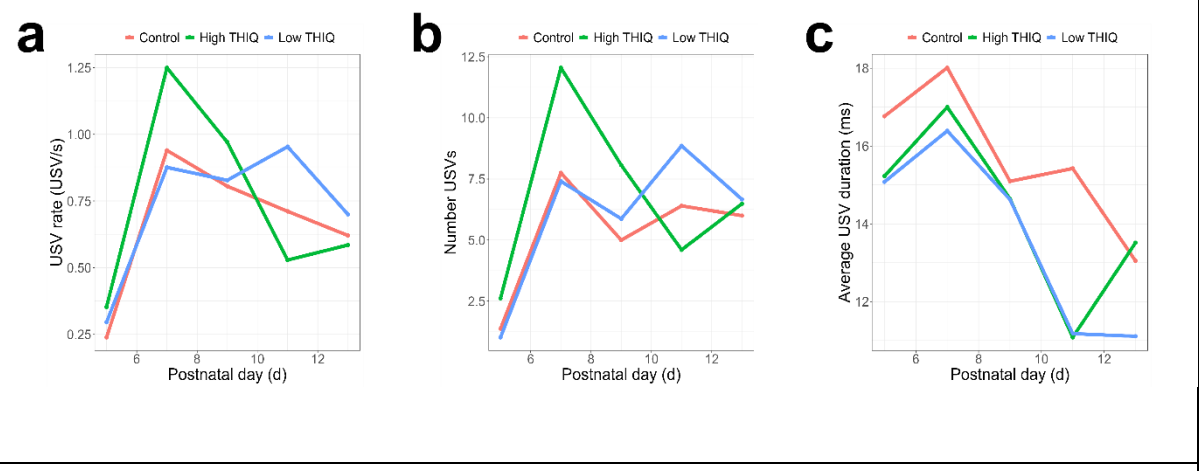


Figure 5. Evolution of mean values of USV traits per condition over days. (a) Mean value of USV rate (USVs/s) before retrieval. **(b)** Mean value of number of USVs emitted during the first 10s of a BAMBI trial. **(c)** Mean average duration of USVs (ms) emitted before retrieval.

were found. Moreover, THIQ treatment affected average duration of USVs. On average, vocalizations of pups treated with low THIQ were 1.8ms shorter (p=0.02), whereas vocalizations of high THIQ treated pups were 1.3ms shorter (p=0.09). The largest effect was seen on PND11 with highly significant differences (p<0.01) for both THIQ treatments. Moreover, over all testing days, males emitted vocalizations that were on average 2.0ms shorter compared to females (p<0.01). No differences were found of THIQ treatment, nor sex on the number of USVs emitted before retrieval.

Trait	PND	High THIQ vs. control ES (p-value)	Low THIQ vs. control ES (p-value)	Male vs. female ES (p-value)	Sequence ES (p-value)
USV_{10sec}	5	**+1.2 (0.06)**	-0.4 (0.61)	-0.10 (0.86)	0.12 (0.20)
USV_{10sec}	7	+4.1 (0.12)	-0.5 (0.85)	+3.2 (0.15)	-1.24 (0.21)
USV_{10sec}	9	+3.1 (0.14)	+1.0 (0.65)	-0.18 (0.92)	+0.74 (0.34)
USV_{10sec}	11	-1.7 (0.41)	+2.5 (0.25)	-1.12 (0.52)	**+2.03 (0.01*)**
USV_{10sec}	13	+0.3 (0.88)	+0.5 (0.80)	+0.87 (0.61)	+1.05 (0.17)
USV_{10sec}	5-13	**+1.4 (0.09)**	+0.5 (0.53)	+0.6 (0.39)	-0.38 (0.15)
Mean USV duration (ms)	5	-1.5 (0.24)	-1.8 (0.21)	**-2.05 (0.07)**	+0.40 (0.22)
Mean USV duration (ms)	7	-0.88 (0.67)	-1.5 (0.48)	-2.83 (0.10)	-0.07 (0.93)
Mean USV duration (ms)	9	-0.31 (0.86)	-0.4 (0.84)	**-1.32 (0.04*)**	-1.09 (0.10)

Mean USV duration (ms)	11	**-4.3 (0.006**)**	-4.2 (0.007)	-1.75 (0.17)	+0.92 (0.11)
Mean USV duration (ms)	13	+1.1 (0.58)	-1.3 (0.52)	-2.23 (0.16)	**-2.19 (0.003**)**
Mean USV duration (ms)	5-13	**-1.3 (0.09)**	**-1.8 (0.02*)**	**2.01 (0.002**)**	-0.09 (0.70)

Table 2. Effect sizes and p-values for the regression models on the number of ultrasonic vocalizations emitted in the first 10s of the BAMBI test (USV_{10sec}) and the mean USV duration (ms) per postnatal day (PND) of testing. Differences with p-value <0.10 are shown in bold. USV_{10sec} was chosen over the total number of USVs, as some pups were retrieved very fast by their dam. By only taking the first 10s into account, the number of USVs do not need to be corrected for the effect of retrieval time.

5.5 Discussion

In this study, we investigated the impact of impaired infant sociocommunicative abilities in the prenatal valproic acid (VPA) exposure model of autism spectrum disorders (ASD) on maternal behavior and early-life interactions. We further manipulated these impairments by stimulating endogenous oxytocin activity through the melanocortin 4 receptor (MC4R) agonist THIQ in a dose dependent way. We have shown that early postnatal MC4R stimulation impacts pups' weight gain, early developmental milestones, USV characteristics as well as maternal retrieval behavior. These findings provide insights into the dynamic nature of social development and the role of the mother-infant dyad in individuals affected by neurodevelopmental disorders.

Previous studies have demonstrated early developmental and sociocommunicative deficits in VPA pups (13-14, 18-20). Therefore, we chose this rodent model to explore the impact of early sociocommunicative deficits on dam-infant interactions. The hypothetical chain of events of atypical mother-infant communication proposed by Esposito et al. (21) suggests that these deficits imply a functional difference that can affect the mother's perception, motivation, and ability to respond promptly and appropriately to the infant's needs. As such, disruption in mother-infant dynamics may occur, further exacerbating social impairments.

Initially, we assessed the impact of infant THIQ administration on the weight development of both the dam and the pups. MC4R is a key regulator of energy homeostasis and plays a significant role in the control of food intake and energy expenditure (35). Mutations in the MC4R gene are recognized as the most common cause of monogenic early-onset obesity which is further supported by the early onset of obesity observed in MC4R-knockout mice (35-36). Conversely, studies have indicated that MC4R activation leads to reduced food intake and body weight, while MC4R antagonism results in hyperphagia and weight gain (37-38).

First, we investigated the effect of THIQ treatments of pups on the maternal weight. During the prenatal period, all dams received equal treatment, and as expected, no significant differences in maternal weight were observed. Interestingly, the maternal daily weight gain generally decreased after VPA injection at GD12-13, indicating an effect of this VPA treatment (**Figure 1a**). In an earlier pilot experiment, control dams injected with saline did not show a decrease in daily weight gain[2]. Additionally, early-life THIQ treatments had a significant effect on postnatal dam weight, resulting in decreased body weight compared to the control group (**Figure 1b**). This decrease in weight may be attributed to the increased appetite and daily weight gain observed in THIQ-treated pups.

[2] These data were collected in a validation experiment which is not part of current dissertation.

Consequently, their dams would be stimulated to allocate more energy to producing milk, leading to a decreased maternal weight.

Further, we investigated the effect of THIQ treatments on pup weight. Previous studies investigating the use of MC4R agonism in neonates reported a decreased weight (27, 39). However, in the current study, we found that chronic early-life treatment with a low or high THIQ in mice prenatally exposed to VPA increased weight early in life and this increase persisted at the time of weaning. This difference in findings may be attributed to the timing of weight measurements. Glavas et al. (39) and Barrett et al. (27) recorded pup weight during the period of MC4R agonist injections. As mentioned in Glavas et al. (39), the effects of MC4R agonism on weight are primarily prominent during the first two days of injections. After this initial period, weight steadily increases and becomes comparable to the control group, with the effects of MC4R agonism diminishing over time. similar pattern was also observed by Barrett et al. (27, 40) using a MC3/4R agonist, and it was less pronounced when using a selective MC4R agonist, although they reported normalized weights in prairie voles at weaning on PND21 for both sexes. Interestingly, in the current study the decrease in weight after two first day of injections was found in males but not females. Additionally, whereas males generally weigh more than females, our study found that after early-life treatment with high THIQ, males and females weights were comparable (**Figure 2**). This indicates that sex may be another important factor contributing to the observed differences. Glavas et al. (39) only used male rats and Barrett et al. (27) did not include sex as an effect in their weight analyses whereas males were overrepresented compared to females. However, one study did report a significant decrease in weight in neonatally isolated female prairie voles, although this lower weight gain appeared to be associated with the isolation treatment (40). However, the post-treatment weight curve for these neonatally isolated, prairie voles did resemble the weight curve in our females (40). The authors argued that the MC3/4R agonist buffered against the negative impact of early isolation in females. In this line, prenatal exposure to VPA has been linked to increased risk of teratogenic effects and specifically decreased weight gain has been reported before weaning, during adolescence or adulthood (9, 41-43). Future research is needed to investigate the long term effects of early-life MC4R agonism on body weight in both males and females.

Next, we showed that THIQ administration in VPA mice had an effect on the developmental milestones eye opening and surface righting. Eye opening is a developmental milestone in mice that reflects the maturation of sensory system, particularly visual processing, and typically occurs between PND12 and PND14 in C57BL6J mice (44-47). In the VPA model for ASD, delayed eye opening and Kim et al. (48) reported that all control C57BL6J pups had opened both eyes on PND17, whereas not all VPA-injected pups had opened their eyes on PND18 (9, 15, 41, 48). In the current study, however, all pups were prenatally exposed to VPA and all animals regardless of the experimental condition had their eyes open

on PND16. Eye opening is a dynamic process influenced by genetic factors, environmental cues and overall maturation of the nervous system. One possible explanation for this time difference is that in our study, all pups received significant early-life tactile stimulation from the researcher, as they were handled and gently massaged daily to promote the spreading of the injected fluid. It has been previously reported that rat pups receiving artificial tactile stimulation from the researcher exhibit accelerated eye opening (49). Additionally, Barnett & Burn (50) suggested a causal association between accelerated eye opening and elevated levels of maternal care in pups that underwent ear-marking. They found that dams consistently provided more attention to ear-clipped pups than to control pups. Since all pups in our current experiment received an injection during the first week of life, it is possible that maternal care levels, in general, were elevated.

Further, we found that chronic early-life treatment of VPA pups with THIQ resulted in a delay in eye opening (**Figure 3b**). Altough there is limited evidence directly linking OT to the timing of eye opening in rodents, a recent study did show that OT plays a role in the developing visual cortex prior to eye opening in mice (44). Additionaly, evidence has shown that early-life tactile stimulation such as maternal licking promotes OT release in the infant brain (51). As such, the previously mentioned results in studies of Smart et al. (49) and Barnett & Burn (50) may be modulated by OT release in the infant brain. Future research should be conducted to further understand the role of maternal behavior and OT in the timing of eye opening.

In our study, we observed a significant delay in surface righting in pups treated with high or low THIQ (**Figure 3a**). Surface righting serves as a behavioral indicator of motor development and sensory-motor integration in rodent pups, reflecting neurological and physical progress. The emergence of surface righting requires the concurrent development of postural control, muscle strength, and coordination (52). A delay in surface righting does not necessarily indicate impaired nervous system development but rather suggests muscle weakness or delayed muscular development (52). In normal C57BL6J mice, the mature surface righting reflex is typically observed at PND6-8 (47). However, in our study, less than 25% of the animals exhibited the mature reflex on PND9, and all animals showed the mature reflex by PND13. Previous studies have shown that prenatal exposure to VPA delays surface righting, and Kim et al. (48) reported a similar delay in the appearance of the mature reflex on PND13 in VPA-exposed animals compared to vehicle-injected animals, which exhibited it around PND10. Furthermore, maternal behavior also plays a crucial role in facilitating the development of motor skills, including surface righting, in pups. Maternal care provides essential sensory stimulation, support, and guidance for the pups' motor learning and coordination. Maternal deprivation studies have demonstrated that maternally deprived pups exhibit slower development of the surface righting reflex and poorer performance in other motor coordination tests (53-54). Interestingly, this effect of maternal

deprivation appears to be sex-dependent, as it was only significant in females (53). In our study, we also found that the delay in surface righting was sex-dependent, but our results indicated that treatment with THIQ in both doses delayed the development of a mature surface righting response. This finding contradicts the evidence suggesting that the surface righting reflex is unaffected in OT-knockout or OTR-knockout mice (55-56). While direct evidence in rodents is lacking, studies have reported delayed or even lost ability to surface right after treatment with OT in chicks and rabbits (57-58). On the other hand, surface righting has been linked to the maturation of cerebellar and/or vestibular function (59-60), and both areas exhibit MC4R expression in mice (61). Therefore, the interpretation of the observed delay in surface righting in our study remains unclear, and further research is needed to elucidate the underlying mechanisms.

In this study, we observed that chronic early-life treatment with THIQ had a time-dependent effect on maternal retrieval (**Figure 4**). Isolated pups treated with high THIQ tended to be retrieved the fastest between P7-11, while pups treated with low THIQ showed only slight improvement compared to control animals. These findings align with a study by Da Prato et al. (62), which demonstrated that acute intranasal administration of OT in a neonatal model for ASD (Magel2 KO) rescued atypical distress vocalizations and delayed maternal pup retrieval response. It should be noted that studies investigating maternal care behaviors in the context of ASD are relatively scarce, although a few studies have reported that maternal behavior remains largely unaffected in dams prenatally injected with VPA (14, 63). However, this is in contrast with findings of our validation experiment (data not shown) and Morel et al. (64) reporting alterations in retrieval behavior of dams prenatally injected with VPA. Possibly, these conflicting findings can be attributed to differences in behavioral test assays as no differences were found in studies assessing maternal behaviors through observation whereas our study and Morel et al. (64) used the PRT.

Human studies have reported associations between increased numbers and shorter durations of infant vocalizations, reduced fitness, and increased risk for ASD (56). A vast amount of literature has consistently reported reminiscent deficits in VPA pups (13-14, 18-20). It has been suggested that alterations in acoustic structure of emitted USVs may disrupt mother-infant communication due to reduced likelihood to attract maternal attention and consequent retrieval behavior (21). In our experiment, all pups were prenatally treated with VPA and, as previously reported, we found that retrieved pups in general have a lower call rate and lower mean USV duration during maternal separation compared to non-retrieved pups (28). Our results indicated that the number of vocalizations emitted before retrieval by pups treated with THIQ was not different compared to control animals. This finding was unexpected as previous research has suggested that OT signaling plays an important role in typical development of early-life USV emission in mice (66-67). Moreover, a previous

study in rats indicated that chronic treatment with OT from PND0-6 resulted in an increased number of emitted USVs on PND7 in VPA-exposed rat pups (19). Conversely, studies using oxytocin receptor antagonists or oxytocin-related knockout models have demonstrated a decrease in USV production (67-69, 70). However, we did find that high THIQ treated animals showed a tendency to emit more vocalizations during the first 10s of the trial, whereas no differences were found in the low THIQ group. Additionally, for high THIQ-treated pups we found that both the USV rate before retrieval and number of USVs in the first 10s of the trial followed an inverted U-shaped pattern from PND5 to PND13 (**Figure 5a-b**). This pattern is typical for the ontogenetic profile of USVs in mouse pups (71). In contrast, this pattern was less pronounced in control animals and in the low THIQ group.

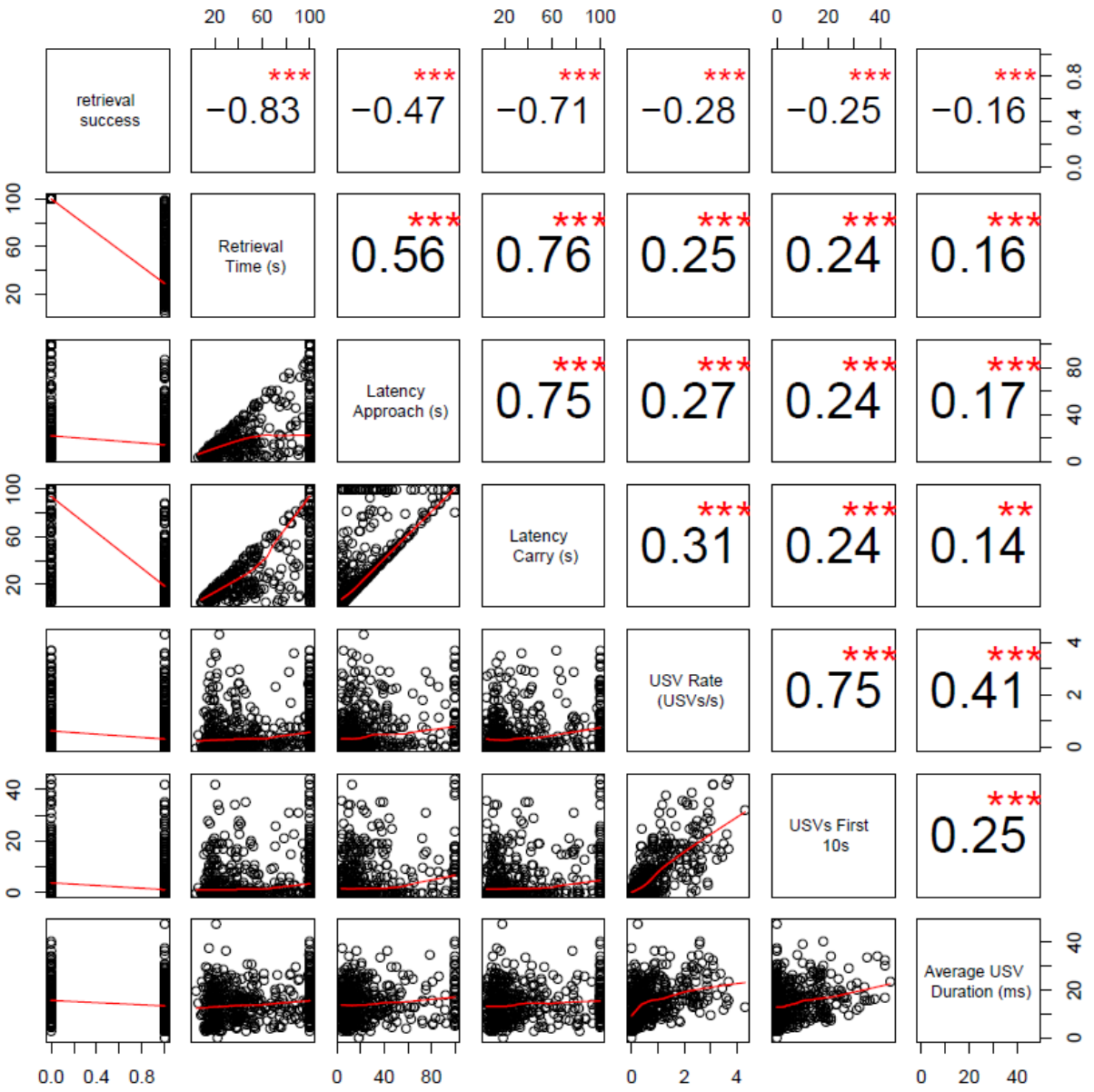
Moreover, results showed that treatment with THIQ affected the average duration of emitted USVs. More specifically both THIQ groups emitted shorter infant vocalizations compared to control VPA pups and this effect was more pronounced in the animals treated with the low dose THIQ. This finding is counter-intuitive, especially since a preference paradigm has shown that mouse dams prefer longer over shorter calls and do not respond to calls lasting under a certain duration (72-73). To the authors knowledge, there is no direct evidence for oxytocins' effect on mean call duration except one study reporting no differences in call duration of OTR mutant prairie voles on PND5 (66). However, Barrett et al. (14) argued that decreased USV duration and rate may indicate reduced distress from maternal separation as anxiolytics and antidepressants also reduce call duration in pups during isolation. Multiple studies have indicated that OT treatment is associated with anxiolytic and anti-stress effects on the brain (70, 74-75). Moreover, it should be noted that our BAMBI procedure used a short isolation of the pup in a heated, clean cup and a reintroduction to the home cage where the test is performed. This is different from most USV studies, which are performed in a clean environment. It has become evident that the presence of odor from dam and littermates leads to a calming effect on the pup which behaviorally translates in a reduction of USVs (76-77). OT has been suggested to play a critical role in odor recognition as well as a pups' fundamental attraction towards the dam (22). White et al. (78) have reported that prenatally VPA exposed rats show a impaired social recognition and/or reduced motivation to approach homecage-related odors on PND6-7 which normalized on PND13. However, they also report that maternal presence does regulate the stress response in VPA-exposed pups suggesting that the reduced tendency to approach is not due to olfactory impairment.

To conclude, our findings revealed that early postnatal MC4R stimulation directly impacted infant weight gain, early developmental milestones and ultrasonic vocalizations (USVs) characteristics as well as dam physiology and retrieval behavior. These results indicate the dynamic nature and complexity of early-life interactions between infant and dam. Further, they urge caution when therapeutically targeting neural pathways for intervention in neurodevelopmental disorders.

5.6 Supplementary information

5.6.1 Supplementary Figure 1.

Pairwise correlational plots of main BAMBI output parameters. Pairwise correlation plots are given for all test days combined (PND5-13). Below diagonal the pairwise correlation plot is shown. Above the diagonal the estimated Pearson correlation coefficient is given with significance value. (°p<0.10, *p<0.05, **p<0.01, ***p<0.001)



5.7 References

1. Weininger, O., McClelland, W. J., & Arima, R. K. (1954). Gentling and weight gain in the albino rat. *Canadian Journal of Psychology/Revue canadienne de psychologie, 8*(3), 147.
2. Levine, S., Alpert, M., & Lewis, G. W. (1957). Infantile experience and the maturation of the pituitary adrenal axis. *Science*, *126*(3287), 1347-1347.
3. Caldji, C., Tannenbaum, B., Sharma, S., Francis, D., Plotsky, P. M., & Meaney, M. J. (1998). Maternal Care during Infancy Regulates the Development of Neural Systems Mediating the Expression of Fearfulness in the Rat. *Proceedings of the National Academy of Sciences - PNAS*, *95*(9), 5335–5340. https://doi.org/10.1073/pnas.95.9.5335
4. Liu, D., Diorio, J., Tannenbaum, B., Caldji, C., Francis, D., Freedman, A.,... & Meaney, M. J. (1997). Maternal care, hippocampal glucocorticoid receptors, and hypothalamic-pituitary-adrenal responses to stress. *Science*, *277*(5332), 1659-1662.
5. Mäntymaa, M., Puura, K., Luoma, I., Vihtonen, V., Salmelin, R. K., & Tamminen, T. (2009). Child's behaviour in mother-child interaction predicts later emotional and behavioural problems. *Infant and Child Development*, *18*(5), 455–467. https://doi.org/10.1002/icd.633
6. Yamasue, H., & Domes, G. (2017). Oxytocin and Autism Spectrum Disorders. *Behavioral Pharmacology of Neuropeptides: Oxytocin*, 449–465. https://doi.org/10.1007/7854_2017_24
7. Mabunga, D. F. N., Gonzales, E. L. T., Kim, J., Kim, K. C., & Shin, C. Y. (2015). Exploring the Validity of Valproic Acid Animal Model of Autism. *Experimental Neurobiology*, *24*(4), 285–300. https://doi.org/10.5607/en.2015.24.4.285
8. Favre, M. R., Barkat, T. R., Lamendola, D., Khazen, G., Markram, H., & Markram, K. (2013). General developmental health in the VPA-rat model of autism. *Frontiers in Behavioral Neuroscience*, *7*, 88–88. https://doi.org/10.3389/fnbeh.2013.00088
9. Schneider, T., & Przewlocki, R. (2005). Behavioral Alterations in Rats Prenatally Exposed to Valproic Acid: Animal Model of Autism. *Neuropsychopharmacology (New York, N.Y.)*, *30*(1), 80–89. https://doi.org/10.1038/sj.npp.1300518
10. Gandal, M. J., Edgar, J. C., Ehrlichman, R. S., Mehta, M., Roberts, T. P., & Siegel, S. J. (2010). Validating γ Oscillations and Delayed Auditory Responses as Translational Biomarkers of Autism. *Biological Psychiatry (1969)*, *68*(12), 1100–1106. https://doi.org/10.1016/j.biopsych.2010.09.031
11. Moldrich, R. X., Leanage, G., She, D., Dolan-Evans, E., Nelson, M., Reza, N., & Reutens, D. C. (2013). Inhibition of histone deacetylase in utero causes sociability deficits in postnatal mice. *Behavioural Brain Research*, *257*, 253–264. https://doi.org/10.1016/j.bbr.2013.09.049
12. Servadio, M., Melancia, F., Manduca, A., Di Masi, A., Schiavi, S., Cartocci, V., Pallottini, V., Campolongo, P., Ascenzi, P., & Trezza, V. (2016). Targeting anandamide metabolism rescues core and associated autistic-like symptoms in rats prenatally exposed to valproic acid. *Translational Psychiatry*, *6*(9), e902–e902. https://doi.org/10.1038/tp.2016.182
13. Tyzio, R., Nardou, R., Ferrari, D. C., Tsintsadze, T., Shahrokhi, A., Eftekhari, S., Khalilov, I., Tsintsadze, V., Brouchoud, C., Chazal, G., Lemonnier, E., Lozovaya, N., Burnashev, N., & Ben-Ari, Y. (2014). Oxytocin-Mediated GABA Inhibition During Delivery Attenuates Autism Pathogenesis in Rodent Offspring. *Science (American Association for the Advancement of Science)*, *343*(6171), 675–679. https://doi.org/10.1126/science.1247190
14. Barrett, C. E., Hennessey, T. M., Gordon, K. M., Ryan, S. J., McNair, M. L., Ressler, K. J., & Rainnie, D. G. (2017). Developmental disruption of amygdala transcriptome and socioemotional behavior in rats exposed to valproic acid prenatally. *Molecular Autism*, *8*(1), 42–42. https://doi.org/10.1186/s13229-017-0160-x
15. Roullet, F. I., Wollaston, L., Decatanzaro, D., & Foster, J. A. (2010). Behavioral and molecular changes in the mouse in response to prenatal exposure to the anti-epileptic drug valproic acid. *Neuroscience*, *170*(2), 514–522. https://doi.org/10.1016/j.neuroscience.2010.06.069

16. Kataoka, S., Takuma, K., Hara, Y., Maeda, Y., Ago, Y., & Matsuda, T. (2013). Autism-like behaviours with transient histone hyperacetylation in mice treated prenatally with valproic acid. *The International Journal of Neuropsychopharmacology*, *16*(1), 91–103. https://doi.org/10.1017/S1461145711001714
17. Nicolini, C., & Fahnestock, M. (2018). The valproic acid-induced rodent model of autism. *Experimental Neurology*, *299*(Pt A), 217–227. https://doi.org/10.1016/j.expneurol.2017.04.017
18. Kuo, H. Y., & Liu, F. C. (2017). Valproic acid induces aberrant development of striatal compartments and corticostriatal pathways in a mouse model of autism spectrum disorder. *The FASEB Journal*, *31*(10), 4458-4471.
19. Dai, Y. C., Zhang, H. F., Schön, M., Böckers, T. M., Han, S. P., Han, J. S., & Zhang, R. (2018). Neonatal oxytocin treatment ameliorates autistic-like behaviors and oxytocin deficiency in valproic acid-induced rat model of autism. Frontiers in cellular neuroscience, 12, 355.
20. Gzielo, K., Potasiewicz, A., Hołuj, M., Litwa, E., Popik, P., & Nikiforuk, A. (2020). Valproic acid exposure impairs ultrasonic communication in infant, adolescent and adult rats. *European neuropsychopharmacology*, *41*, 52-62.
21. Esposito, G., Hiroi, N., & Scattoni, M. L. (2017). Cry, Baby, Cry: Expression of Distress As a Biomarker and Modulator in Autism Spectrum Disorder. *The International Journal of Neuropsychopharmacology*, *20*(6), 498–503. https://doi.org/10.1093/ijnp/pyx014
22. Nagasawa, M., Okabe, S., Mogi, K., & Kikusui, T. (2012). Oxytocin and mutual communication in mother-infant bonding. *Frontiers in Human Neuroscience*, *6*, 31–31. https://doi.org/10.3389/fnhum.2012.00031
23. Muscatelli, F., Desarménien, M. G., Matarazzo, V., & Grinevich, V. (2018). Oxytocin signaling in the early life of mammals: link to neurodevelopmental disorders associated with ASD. *Behavioral Pharmacology of Neuropeptides: Oxytocin*, 239-268.
24. Peñagarikano, O., Lázaro, M. T., Lu, X. H., Gordon, A., Dong, H., Lam, H. A.,... & Geschwind, D. H. (2015). Exogenous and evoked oxytocin restores social behavior in the Cntnap2 mouse model of autism. *Science translational medicine*, *7*(271), 271ra8-271ra8.
25. Sabatier, N. (2006). α-Melanocyte-stimulating hormone and oxytocin: a peptide signalling cascade in the hypothalamus. *Journal of neuroendocrinology*, *18*(9), 703-710.
26. Modi, M. E., Inoue, K., Barrett, C. E., Kittelberger, K. A., Smith, D. G., Landgraf, R., & Young, L. J. (2015). Melanocortin receptor agonists facilitate oxytocin-dependent partner preference formation in the prairie vole. *Neuropsychopharmacology*, *40*(8), 1856-1865.
27. Barrett, C. E., Modi, M. E., Zhang, B. C., Walum, H., Inoue, K., & Young, L. J. (2014). Neonatal melanocortin receptor agonist treatment reduces play fighting and promotes adult attachment in prairie voles in a sex-dependent manner. *Neuropharmacology*, *85*, 357-366.
28. Winters, C., Gorssen, W., Wöhr, M., & D'Hooge, R. (2023). BAMBI: A new method for automated assessment of bidirectional early-life interaction between maternal behavior and pup vocalization in mouse dam-pup dyads. *Frontiers in Behavioral Neuroscience*, *17*.
29. Heyne, G. W., Plisch, E. H., Melberg, C. G., Sandgren, E. P., Peter, J. A., & Lipinski, R. J. (2015). A Simple and Reliable Method for Early Pregnancy Detection in Inbred Mice. *Journal of the American Association for Laboratory Animal Science, 54(4)*, 368–371.
30. Mastinu, A., Premoli, M., Maccarinelli, G., Grilli, M., Memo, M., & Bonini, S. A. (2018). Melanocortin 4 receptor stimulation improves social deficits in mice through oxytocin pathway. *Neuropharmacology, 133*, 366–374. https://doi.org/10.1016/j.neuropharm.2018.02.007
31. Kuroda, K. O., Tachikawa, K., Yoshida, S., Tsuneoka, Y., & Numan, M. (2011). Neuromolecular basis of parental behavior in laboratory mice and rats: With special emphasis on technical issues of using mouse genetics. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *35*(5), 1205–1231. https://doi.org/10.1016/j.pnpbp.2011.02.008
32. Hess, S. E., Rohr, S., Dufour, B. D., Gaskill, B. N., Pajor, E. A., & Garner, J. P. (2008). Home improvement: C57BL/6J mice given more naturalistic nesting materials build better nests. *Journal of the American Association for Laboratory Animal Science*, *47*(6), 25-31.

33. R Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. *http://www. R-project. org/*.
34. Steinfath, E., Palacios-Muñoz, A., Rottschäfer, J. R., Yuezak, D., & Clemens, J. (2021). Fast and accurate annotation of acoustic signals with deep neural networks. *Elife*, *10*, e68837.
35. Tao, Y. X. (2005). Molecular mechanisms of the neural melanocortin receptor dysfunction in severe early onset obesity. *Molecular and cellular endocrinology*, *239*(1-2), 1-14.
36. Rouault, A. A., Srinivasan, D. K., Yin, T. C., Lee, A. A., & Sebag, J. A. (2017). Melanocortin receptor accessory proteins (MRAPs): functions in the melanocortin system and beyond. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 1863*(10), 2462-2467.
37. Fan, W., Boston, B. A., Kesterson, R. A., Hruby, V. J., & Cone, R. D. (1997). Role of melanocortinergic neurons in feeding and the agouti obesity syndrome. *Nature*, *385*(6612), 165-168.
38. Kask, A., Rägo, L., & Harro, J. (1998). Evidence for involvement of neuropeptide Y receptors in the regulation of food intake: studies with Y1-selective antagonist BIBP3226. British journal of pharmacology, 124(7), 1507-1515.
39. Glavas, M. M., Joachim, S. E., Draper, S. J., Smith, M. S., & Grove, K. L. (2007). Melanocortinergic activation by melanotan II inhibits feeding and increases uncoupling protein 1 messenger ribonucleic acid in the developing rat. *Endocrinology*, *148*(7), 3279-3287.
40. Barrett, C. E., Arambula, S. E., & Young, L. J. (2015). The oxytocin system promotes resilience to the effects of neonatal isolation on adult social attachment in female prairie voles. *Translational psychiatry*, *5*(7), e606-e606.
41. Yang, E. J., Ahn, S., Lee, K., Mahmood, U., & Kim, H. S. (2016). Early behavioral abnormalities and perinatal alterations of PTEN/AKT pathway in valproic acid autism model mice. *PloS one*, *11*(4), e0153298.
42. Ruhela, R. K., Soni, S., Sarma, P., Prakash, A., & Medhi, B. (2019). Negative geotaxis: An early age behavioral hallmark to VPA rat model of autism. *Annals of Neurosciences*, *26*(1), 25–31. https://doi.org/10.5214/ans.0972.7531.260106
43. Khalaj, R., Moghaddam, A. H., & Zare, M. (2018). Hesperetin and it nanocrystals ameliorate social behavior deficits and oxido-inflammatory stress in rat model of autism. *International Journal of Developmental Neuroscience*, *69*, 80-87.
44. Maldonado, P. P., Nuno-Perez, A., Kirchner, J. H., Hammock, E., Gjorgjieva, J., & Lohmann, C. (2021). Oxytocin shapes spontaneous activity patterns in the developing visual cortex by activating somatostatin interneurons. *Current Biology*, *31*(2), 322-333.
45. Rice, D. C., Reeve, E. A., Herlihy, A., Zoeller, R. T., Thompson, W. D., & Markowski, V. P. (2007). Developmental delays and locomotor activity in the C57BL6/J mouse following neonatal exposure to the fully-brominated PBDE, decabromodiphenyl ether. *Neurotoxicology and teratology*, *29*(4), 511-520.
46. Scattoni, M. L., Gandhy, S. U., Ricceri, L., & Crawley, J. N. (2008). Unusual repertoire of vocalizations in the BTBR T+tf/J mouse model of autism. *PloS One*, *3*(8), e3067–e3067. https://doi.org/10.1371/journal.pone.0003067
47. Castelhano-Carlos, M. J., Sousa, N., Ohl, F., & Baumans, V. (2010). Identification methods in newborn C57BL/6 mice: a developmental and behavioural evaluation. *Laboratory Animals*, *44*(2), 88-103.
48. Kim, U. J., Hong, N., & Ahn, J. C. (2022). Photobiomodulation Attenuated Cognitive Dysfunction and Neuroinflammation in a Prenatal Valproic Acid-Induced Autism Spectrum Disorder Mouse Model. *International Journal of Molecular Sciences*, *23*(24), 16099.
49. Smart, J. L., McMahon, A. C., Massey, R. F., Akbar, G. N., & Warren, M. A. (1990). Evidence of non-maternally mediated acceleration of eye-opening in 'enriched'artificially reared rat pups. *Developmental Brain Research*, *56*(1), 141-143.
50. Barnett, S. A., & Burn, J. (1967). Early stimulation and maternal behaviour. *Nature*, *213*, 150-152.
51. Lenz, K. M., & Sengelaub, D. R. (2010). Maternal care effects on the development of a sexually dimorphic motor system: the role of spinal oxytocin. *Hormones and behavior*, *58*(4), 575-581.

52. Hill, J. M., Lim, M. A., & Stone, M. M. (2008). Developmental milestones in the newborn mouse. *Neuropeptide techniques*, 131-149.
53. Horvath, G., Reglődi, D., Farkas, J., Vadasz, G., Mammel, B., Kvarik, T.,... & Kiss, P. (2015). Perinatal positive and negative influences on the early neurobehavioral reflex and motor development. *Perinatal programming of neurodevelopment*, 149-167.
54. Farkas, J., Reglodi, D., Gaszner, B., Szogyi, D., Horvath, G., Lubics, A.,... & Kiss, P. (2009). Effects of maternal separation on the neurobehavioral development of newborn Wistar rats. *Brain research bulletin*, *79*(3-4), 208-214.
55. Crawley, J. N., Chen, T., Puri, A., Washburn, R., Sullivan, T. L., Hill, J. M.,... & Young, W. S. (2007). Social approach behaviors in oxytocin knockout mice: comparison of two independent lines tested in different laboratory environments. *Neuropeptides*, *41*(3), 145-163.
56. Lee, H. J., Caldwell, H. K., Macbeth, A. H., Tolu, S. G., & Young 3rd, W. S. (2008). A conditional knockout mouse line of the oxytocin receptor. *Endocrinology*, *149*(7), 3256-3263.
57. Hess, L., Votava, M., Málek, J., Kurzová, A., & Slíva, J. (2016). Sedative effects of intranasal oxytocin in rabbits and rhesus monkeys. *Physiological research*, *65*.
58. Buchholz, R., Dodd, C. E., & Heppard, J. M. (2019). Discovery of a synergistic effect of oxytocin on the efficacy of Alpha-Chloralose, a drug used for capturing and controlling birds. *International Journal of Pest Management*, *65*(2), 130-136.
59. Wagner, G. C., Reuhl, K. R., Cheh, M., McRae, P., & Halladay, A. K. (2006). A new neurobehavioral model of autism in mice: pre-and postnatal exposure to sodium valproate. *Journal of autism and developmental disorders*, *36*, 779-793.
60. Mesquita, A. R., Pêgo, J. M., Summavielle, T., Maciel, P., Almeida, O. F. X., & Sousa, N. (2007). Neurodevelopment milestone abnormalities in rats exposed to stress in early life. *Neuroscience*, *147*(4), 1022-1033.
61. Shang, D., Xiong, J., Xiang, H. B., Hao, Y., & Liu, J. H. (2015). Melanocortinergic circuits from medial vestibular nuclei to the kidney defined by transneuronal transport of pseudorabies virus. *International Journal of Clinical and Experimental Pathology*, *8*(2), 1996.
62. Da Prato, L. C., Zayan, U., Abdallah, D., Point, V., Schaller, F., Pallesi-Pocachard, E.,... & Matarazzo, V. (2022). Early life oxytocin treatment improves thermo-sensory reactivity and maternal behavior in neonates lacking the autism-associated gene Magel2. *Neuropsychopharmacology*, *47*(11), 1901-1912.
63. Kazlauskas, N., Campolongo, M., Lucchina, L., Zappala, C., & Depino, A. M. (2016). Postnatal behavioral and inflammatory alterations in female pups prenatally exposed to valproic acid. *Psychoneuroendocrinology*, *72*, 11–21. https://doi.org/10.1016/j.psyneuen.2016.06.001
64. Morel, C., Christophe, A., Maguin-Gaté, K., Paoli, J., Turner, J. D., Schroeder, H., & Grova, N. (2022). Head-to-Head Study of Developmental Neurotoxicity and Resultant Phenotype in Rats: α-Hexabromocyclododecane versus Valproic Acid, a Recognized Model of Reference for Autism Spectrum Disorders. *Toxics (Basel)*, *10*(4), 180–. https://doi.org/10.3390/toxics10040180
65. Esposito, G., del Carmen Rostagno, M., Venuti, P., Haltigan, J. D., & Messinger, D. S. (2014). Brief Report: Atypical expression of distress during the separation phase of the strange situation procedure in infant siblings at high risk for ASD. *Journal of Autism and Developmental Disorders*, *44*, 975-980.
66. Miyazaki, S., Hiraoka, Y., Hidema, S., & Nishimori, K. (2016). Prenatal minocycline treatment alters synaptic protein expression, and rescues reduced mother call rate in oxytocin receptor-knockout mice. *Biochemical and biophysical research communications*, *472*(2), 319-323.
67. Takayanagi, Y., Yoshida, M., Bielsky, I. F., Ross, H. E., Kawamata, M., Onaka, T.,... & Nishimori, K. (2005). Pervasive social deficits, but normal parturition, in oxytocin receptor-deficient mice. *Proceedings of the National Academy of Sciences*, *102*(44), 16096-16101.
68. Liu, H. X., Lopatina, O., Higashida, C., Tsuji, T., Kato, I., Takasawa, S.,... & Higashida, H. (2008). Locomotor activity, ultrasonic vocalization and oxytocin levels in infant CD38 knockout mice. *Neuroscience letters*, *448*(1), 67-70.

69. Winslow, J. T., Hearn, E. F., Ferguson, J., Young, L. J., Matzuk, M. M., & Insel, T. R. (2000). Infant vocalization, adult aggression, and fear behavior of an oxytocin null mutant mouse. *Hormones and behavior*, *37*(2), 145-155.
70. Kramer, K. M., Cushing, B. S., & Carter, C. S. (2003). Developmental effects of oxytocin on stress response: single versus repeated exposure. *Physiology & behavior*, *79*(4-5), 775-782.
71. Branchi, I., Santucci, D., & Alleva, E. (2001). Ultrasonic vocalisation emitted by infant rodents: a tool for assessment of neurobehavioural development. *Behavioural brain research*, *125*(1-2), 49-56.
72. Smith, J. C. (1976). Responses of adult mice to models of infant calls. *Journal of Comparative and Physiological Psychology*, *90*(12), 1105.
73. Ehret, G. (1992). Categorical perception of mouse-pup ultrasounds in the temporal domain. *Animal Behaviour*, *43*(3), 409-416.
74. Morais, L. H., Golubeva, A. V., Casey, S., Scott, K. A., Ramos Costa, A. P., Moloney, G. M.,... & Cryan, J. F. (2021). Early-life oxytocin attenuates the social deficits induced by caesarean-section delivery in the mouse. *Neuropsychopharmacology*, *46*(11), 1958-1968.
75. Yoshida, M., Takayanagi, Y., Inoue, K., Kimura, T., Young, L. J., Onaka, T., & Nishimori, K. (2009). Evidence that oxytocin exerts anxiolytic effects via oxytocin receptor expressed in serotonergic neurons in mice. *Journal of Neuroscience*, *29*(7), 2259-2271.
76. Potasiewicz, A., Holuj, M., Piotrowska, D., Zajda, K., Wojcik, M., Popik, P., & Nikiforuk, A. (2019). Evaluation of ultrasonic vocalizations in a neurodevelopmental model of schizophrenia during the early life stages of rats. *Neuropharmacology*, *146*, 28-38.
77. Wöhr, M. (2015). Effect of social odor context on the emission of isolation-induced ultrasonic vocalizations in the BTBR T+ tf/J mouse model for autism. *Frontiers in Neuroscience*, *9*, 73.
78. White, A. M. M., An, X., & Debiec, J. (2022). Intact maternal buffering of stress response in infant rats despite altered responsivity towards maternal olfactory cues in the valproic acid model of autism-like behavior. *Frontiers in Behavioral Neuroscience*, *16*, 959485–959485. https://doi.org/10.3389/fnbeh.2022.959485

6 General discussion

The study of early-life mother-infant interactions is of paramount importance due to its profound impact on the development and wellbeing of both infant and the mother. Understanding the dynamics of these early-life mother-infant interactions, we can identify risk factors that may hinder healthy bonding and attachment, promote healthy attachment relationships, support optimal child development, and foster the wellbeing of both mothers and infants. Accurate profiling of these early-life behaviors are crucial to help understand the effects of sociocommunicative deficits. Sociocommunicative deficits are complex to phenotype and the neurobiology underlying this behavior can be even more variable. The current thesis focused on effects of early-life sociocommunicative deficits on mother-infant interactions and subsequent development in mice. We have successfully developed and benchmarked novel behavioral test and analysis procedures known as BAMBI (Bidirectional Automated Mother-pup Behavioral Interaction test). BAMBI is an automated PRT methodology based on synchronous videorecording of maternal behavior and audiorecording of pup vocalizations, enabling a comprehensive assessment of the bidirectional interaction between mother and pup. This innovative approach enhances the accuracy of phenotyping this complex dyadic relationship. To facilitate further research, we have made our methodology, annotated images, videos, tracking models, and behavioral classifiers are openly accessible. By sharing these resources, we aim to encourage the scientific community to expand upon our findings. Moreover, we applied BAMBI in an experiment involving dose-dependent melanocortin 4 receptor (MC4R) stimulation in a autism spectrum disorder (ASD) mouse model with impaired sociocommunicative abilities early in development. Using BAMBI, we could accurately phenotype the impact of this intervention on ethologically relevant maternal and infant behaviors. While further research is needed to fully understand the precise role and mechanisms of MC4R agonism in this dyadic relationship, BAMBI represents a crucial step towards comprehensive phenotyping and unraveling this intricate interaction.

6.1 Towards computational ethology

In neuroscience research, animal behavioral experiments are essential as they allow to observe effects of pharmacological manipulation, rehabilitation protocols and neurological diseases on behavioral patterns. Nevertheless, translating animal behaviors to humans remains challenging, and the phenotyping and analysis of these behaviors are still relatively complex. Many behavioral test assays rely on manual quantification and classification, which are subject to human judgment, introducing the risk of bias, high costs, and reduced reproducibility (1-4). To address these limitations, the emergence of computational ethology tools for automating behavioral analysis represents a logical progression in improving conventional scoring practices. Additionally, these new tools highlight the need for well-defined operationalization of ethologically-relevant behaviors and unlike commercial test environment

allow capturing animal behavior in more enriched and dynamic environments such as a home cage with nesting site. By leveraging these advancements, researchers can enhance the accuracy, efficiency, and ecological validity of behavioral analysis, ultimately advancing our understanding of the intricate relationships between neurological processes and behavior.

While the advancements in computational ethology are promising, it is crucial to recognize the significance of thorough quality control measures. Firstly, even though these tools allow for behavior registration in enriched and dynamic environments, it is important to acknowledge that even the slightest interference in environmental or experimental conditions can impact the behavioral outcomes of both the dam and the pups. Therefore, behavioral researchers need to be mindful of their presence and actions during data collection. Comprehensive early-life behavioral assessment entails not only standardizing the test procedures but also standardizing practices during the dam's pregnancy and postnatal period leading up to the test. Chapter 2 of this thesis encompasses a detailed test protocol that goes beyond the behavioral test assay and analysis. It includes specific recommendations for standardizing the prenatal and early postnatal environment, recognizing the critical influence of these factors on the behavioral outcomes under investigation. By incorporating these guidelines, we aim to ensure the integrity, consistency and replicability of studies, ultimately contributing to a more comprehensive understanding of early-life behaviors and their underlying mechanisms.

Secondly, it is important to note that computational ethology techniques, whether commercial or open-source, are not airtight tools. In fact, a significant amount of quality control measures should be implemented prior to data analysis to ensure the production of reliable and reproducible results, adhering to the "garbage in, garbage out"- principle. By exercising diligent quality control, researchers can enhance the validity and robustness of their findings, fostering a more accurate understanding of the intricate behaviors and interactions under investigation. While our BAMBI setup and analysis have demonstrated successful performance in the specific settings for which they were developed, it is crucial to exercise caution when applying them to new environments or contexts. Whenever using any test or methodology in a different setting, it is essential to conduct thorough performance checks. Specifically, it is recommended assessing the accuracy and reliability of the automated mice tracking and ultrasonic vocalization (USV) detection before proceeding with data analysis. BAMBI includes various steps of quality control and these are included in Chapter 3. By taking these precautionary measures, researchers can ensure the validity and robustness of the results obtained, avoiding potential misinterpretations or inaccuracies that may arise from using the BAMBI system in unfamiliar conditions.

Finally, minimizing potential sources of error, variability, or bias that may otherwise lead to

inconclusive or unreliable results is directly beneficial for animal welfare as it minimizes the need to repeat experiments or conduct additional animal studies. Moreover, rigorous quality control ensures that collected data is meaningful and of high quality, which can lead to more robust scientific conclusions. Reliable and reproducible data enable researchers to draw accurate inferences and make informed decisions based on solid evidence. This, in turn, promotes the responsible use of animals in research by maximizing the scientific value derived from each individual animal.

The objective of this thesis was to capture subtle behavioral patterns in mother-infant interactions. This was accomplished through the application of machine learning techniques in Chapter 3 and the introduction of a novel behavioral test assay in Chapter 4. The central idea was to identify changes in behavioral patterns that might be too subtle for human scoring or prone to biases associated with anthropomorphizing animal behavior. Computational ethology allows researchers to break down complex behaviors, such as maternal pup retrieval, into smaller behavioral units. In Chapter 3, maternal pup retrieval was deconstructed into specific behavioral units, including maternal approach, carry, and digging, providing a deeper understanding of the factors influencing aberrations in maternal retrieval behavior. However, it is essential to exercise caution in handling and interpreting these results due to the current limited research in this area. Further investigation is needed to explore the significance of these behavioral units and their etho-neurological implications, with the goal of establishing a reliable and well-defined reference framework.

Gaining a comprehensive understanding of the intricate dynamics of mother-infant interactions and their impact on offspring development necessitates a thorough assessment encompassing various aspects. This includes measuring maternal behavior, infant behavior, bidirectional interactions, developmental changes, as well as physical and sensory development. The significance of a well-designed test battery that captures ethologically-relevant dynamic changes throughout the postnatal period becomes evident, particularly in the context of this thesis focusing on the prenatal exposure to valproic acid as a rodent model for autism spectrum disorder (ASD). For many years, there have been reports indicating dam behavior is not affected in this rodent model based on manual registrations of maternal nursing, grooming, and other behaviors (5-7). However, it is important to recognize that maternal behavior is more nuanced and dynamic than mere counts of such events, and thus, relying solely on these reports may overlook subtle changes. More recent research focusing on specific maternal behaviors in different tasks and environments, such as retrieval and nest building, has revealed differences in the context of prenatal valproic acid exposure (8). Similary, we did find differences in our VPA validation study indicating that maternal retrieval performance was affected by prenatal VPA exposure (data not shown). Moreover, unlike other studies, we did find an impact of VPA on maternal behavior as reported in Chapter 5. From this perspective, the use of different research

methods such as cross-fostering is recommended to better understand different aspects and potential confounding factors in the VPA model (9). Finally, the development of similar methodologies such as AMBER (Automated Maternal Behavior during Early Life in Rodents) of Lapp (10) facilitate parallel use of different behavioral assays without the burden of manual scoring.

While the use of various testing procedures to understand the bigger picture of infant phenotypes is recommended when studying early life, it's important to note that, during this period, the range of tests available for investigating social behavior is limited due to the restricted behavioral repertoire of postnatal mouse pups. Frequently employed readouts include developmental milestones such as ear detachment, surface righting, and eye opening. However, despite their frequent use as measures of development, the underlying neurobiology of these milestones remains unclear. Therefore, the significance and translational value of reaching these milestones early or late is unclear in the context of development. To provide a more comprehensive perspective and a deeper understanding of the early developmental period, further research is needed.

6.2 Complex behavioral phenotypes and translation to human research

Developing a mouse model to replicate complex human phenotypes is inherently challenging. In addition, mimicking disrupted social functions in mice, especially during the early postnatal period, presents a formidable task. However, understanding the complexities of early-life mother-infant interactions in such models is vital as behaviors generate experiences and these experiences again generate behaviors. In this project, early-life sociocommunicative challenges were mimicked using the valproic acid (VPA) mouse model of autism. The VPA mouse model is a laboratory-based approach which simulates autism-like behaviors in mice through exposure to an environmental variable, in this case valproic acid. As the aim of this project was to investigate the effect of maternal interactions on social development, such an environmental model of ASD has the advantage that maternal behavior is assumed not altered by the intervention (9). However, some studies, including this project, reported alterations in maternal behavior in response to prenatal VPA exposure. Although, we provided evidence that maternal behavior is affected by pup behavior, our study did not exclude the possibility that prenatal VPA also has a direct effect on maternal behavior. Future research should apply cross-fostering to avoid confounding effects of alterations in maternal care behavior.

Additionally, in the context of conditions like ASD that have a genetic basis, sociocommunicative impairments may be inherited from one or both parents leading to a complex interplay of genetic and environmental factors. This intergenerational transmission of traits adds another layer of complexity to studying mother-infant interactions in the context of sociocommunicative impairments. For future studies, it might be relevant to use different mouse models of autism to study the dam-infant

interactions using our methods. For example, when using genetic mouse models of ASD, it can be argued that sociocommunicative impairments in both dam and infant may contribute to an aggrevated ASD-phenotype. It also will be necessary to better understand the neurobiology of, and the relationship between maternal attraction, motivation and responsiveness to infant social stimuli. As such, efforts can be made to improve the translatability of findings from animal research to human research.

Another limitation of this thesis is that only C57BL6j mice were used. The C57BL6j are generally spontaneously parental and exhibit a relatively strong maternal response (11-13). Therefore, other strains of mice (e.g. BALB/c) or other species (e.g. rats) might be more suitable to study the bidirectional impact of altered infant communication on maternal behavior. It also should be taken into account that the frequency range of components of maternal behavior can considerably differ between laboratories. Pedersen et al. (14) have argued that this might be partly explained by methodological differences as well as natural variations among C57BL6j dams reared and bred in different animal colonies.

Previous findings indicate that incorporating measurements of maternal behavior received during the early-life period may be important covariates. Phenotyping these early-life maternal behaviors may enhance the assessment of the validity of mouse models for human behavioral and emotional disorders (14). However, precise measurements of such abstract and complex behavioral concepts still remains challenging and often results in cumbersome quantification. Therefore, it is advisable to use different behavioral tests that measure various aspects of these abstract behaviors. Additionally, subdivision of a complex behavioral phenotype into underlying behaviors holds the potential to provide valuable insights into these abstract behavioral concepts such as maternal attraction, motivation and responsiveness. Moreover, this would allow to study their impact on infant development better. While this approach has shown promise in Chapter 5, it is crucial to emphasize that further research is necessary to fully understand the complexities of these behaviors. First, it will be valuable to understand the normal variability in these maternal care behaviors and which factors contribute to these differences. Pedersen and colleagues (14), for example, demonstrated that adult offspring of C57BL6j dams that exhibited high or low frequencies of pup licking within natural variability exhibited distinct behavioral outcomes. Future research should aim to investigate whether natural occurring variations in maternal retrieval, approach, carry or digging can predict distinct infant outcomes. Moreover, it would be interesting to understand the neurobiological mechanisms underlying these behaviors to obtain a better delineation of abstract concepts such as maternal attraction, motivation, and responsiveness.

Although BAMBI includes an important advance to develop a more ecological paradigm, it should be

noted that this procedure still includes an artificial intervention. Such an experimental task differs from the specific naturalistic operation that it is intended to mimic. Investigating early-life bidirectionality comes from isolating one pup and reintroducing it to the homecage to understand how both individuals process complex, often unpredictable, real-world sequences (15). However, the homecage environment is a relatively safe environment and the test procedure is always a controlled, passive, and repeated presentation of exact sequences of events that reoccur. In a naturalistic environment, it virtually never occurs that pups are gently removed and reintroduced right outside the nest. The species-specific relevance of this test methodology limits the extrapolation from laboratory mice to humans. However, previous research has suggested that, when methods are similar across species, many analogies can be noted in animal and rodent studies. For this purpose, the recent development of novel, more naturalistic perspectives to study mother-infant interactions in both animals and humans hold great promise to establish more translational paragdims and increase inter-species comparability (15).

Lastly, we offer new tools for studying mother-infant dyads, while emphasizing the need for researchers, educators, policymakers, and others to consider how results of such a behavioral assay might be interpreted in popular discussions (16). It is important to recognize that previous scientific findings or theories, such as the notion of the "refrigerator mother" proposed by Bettelheim, have unjustly held mothers solely accountable for infant psychopathology (17). While mothers undeniably play a significant role in infant development, it is crucial to acknowledge the influence of other factors, such as species-specific social structures and the distinctions between natural and laboratory environments. As suggested in Richardson et al. (16), responsible reporting of research outcomes should provide a contextual understanding by considering factors such as overall observed risk, treatability of the investigated pathology, other potential risk factors, and the establishment of causal inferences from research results. Animal researchers particularly should be cautious about making one-on-one translations from animal studies to humans and convey complexity as even in laboratory animals a plethora of environmental and genetic factors can affect behavioral outcomes.

6.3 Pharmacological intervention during early-development

As extensively discussed in the present thesis, the early developmental period plays a crucial role in growth and maturation, making it an attractive target for interventions aimed at addressing various psychopathologies. Early intervention has been advocated for conditions like ASD, as clinical signs often emerge early in life and core symptoms are not significantly changed throughout adulthood (18). Therefore it has been suggested that treatment should start at an early stage of life before the symptoms become fixed (6, 19). Due to the vulnerability of early development, interventions during this early-life period require careful consideration. Pharmacological studies, designed to target specific

molecular pathways, may have unintended off-target effects, leading to unforeseen complications and adverse outcomes. The developing brain undergoes rapid changes, necessitating the establishment of precise, appropriate timeframes, doses, and administration routes, taking into account individual variability.

In Chapter 5 of this thesis, we explored the administration of a selective and potent MC4R agonist at different doses to stimulate endogenous oxytocin signaling. The oxytocin system is an interesting target due to its suggested role in the development of the social neural circuitry, as well as its suggested role in ASD etiology (6, 19-20). However, caution must be exercised when manipulating the oxytocin system during the early-life period. Previous research by Bales et al. (21) demonstrated that chronic intranasal oxytocin administration during adolescence could have unintended, long-term negative consequences on social behavior in prairie voles. In our pharmacological intervention experiment, we observed effects on both maternal behavior and pup behavior and development following administration of a melanocortin-4 receptor agonist and these effects were dose-dependent. However, further research is required to fully interpret these specific effects, although it highlights the complex and extensive nature of early-life treatments, underscoring the need for careful consideration and ongoing investigation in this field.

6.4 References

1. Krishnan, K., Casenhiser, D. M., & Stevenson, P. K. (2020). Systematic analysis of goal-related movement sequences during maternal behavior in a female mouse model for Rett syndrome. *bioRxiv*.
2. Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature neuroscience*, *21*(9), 1281-1289.
3. Nilsson, S. R., Goodwin, N. L., Choong, J. J., Hwang, S., Wright, H. R., Norville, Z.,... & McLaughlin, R. J. (2020). Simple Behavioral Analysis (SimBA): an open source toolkit for computer classification of complex social behaviors in experimental animals. *BioRxiv*.
4. Steinfath, E., Palacios-Muñoz, A., Rottschäfer, J. R., Yuezak, D., & Clemens, J. (2021). Fast and accurate annotation of acoustic signals with deep neural networks. *Elife*, *10*, e68837.
5. Kazlauskas, N., Campolongo, M., Lucchina, L., Zappala, C., & Depino, A. M. (2016). Postnatal behavioral and inflammatory alterations in female pups prenatally exposed to valproic acid. *Psychoneuroendocrinology*, *72*, 11–21. https://doi.org/10.1016/j.psyneuen.2016.06.001
6. Jaber, M., Turpin, V., Schaffhauser, M., Thabault, M., Aubert, A., Joffre, C.,... & Laye, S. (2022). A balanced omega-6/omega-3 polyunsaturated fatty acid diet suffices to prevent autism spectrum disorder symptoms in an environmental mouse model.
7. Barrett, C. E., Hennessey, T. M., Gordon, K. M., Ryan, S. J., McNair, M. L., Ressler, K. J., & Rainnie, D. G. (2017). Developmental disruption of amygdala transcriptome and socioemotional behavior in rats exposed to valproic acid prenatally. *Molecular Autism*, *8*(1), 42–42. https://doi.org/10.1186/s13229-017-0160-x
8. Morel, C., Christophe, A., Maguin-Gaté, K., Paoli, J., Turner, J. D., Schroeder, H., & Grova, N. (2022). Head-to-Head Study of Developmental Neurotoxicity and Resultant Phenotype in Rats: α-Hexabromocyclododecane versus Valproic Acid, a Recognized Model of Reference for Autism Spectrum Disorders. *Toxics (Basel)*, *10*(4), 180–. https://doi.org/10.3390/toxics10040180
9. Chaliha, D., Albrecht, M., Vaccarezza, M., Takechi, R., Lam, V., Al-Salami, H., & Mamo, J. (2020). A systematic review of the valproic-acid-induced rodent model of autism. *Developmental neuroscience*, *42*(1), 12-48.
10. Lapp, H. (2023). AMBER: automated maternal behavior during early life in rodents. COST-TEATIME Webinar Series 2: e0002. Available at: https://www.youtube.com/watch?v=P4Fm52L434E&ab_channel=COST_TEATIME
11. Kuroda, K. O., Tachikawa, K., Yoshida, S., Tsuneoka, Y., & Numan, M. (2011). Neuromolecular basis of parental behavior in laboratory mice and rats: with special emphasis on technical issues of using mouse genetics. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *35*(5), 1205-1231.
12. Weber, E. M., & Olsson, I. A. S. (2008). Maternal behaviour in Mus musculus sp.: an ethological review. *Applied Animal Behaviour Science*, *114*(1-2), 1-22.
13. Francis, D. D., Szegda, K., Campbell, G., Martin, W. D., & Insel, T. R. (2003). Epigenetic sources of behavioral differences in mice. *Nature neuroscience*, *6*(5), 445-446.
14. Pedersen, C. A., Vadlamudi, S., Boccia, M. L., & Moy, S. S. (2011). Variations in maternal behavior in C57BL/6J mice: behavioral comparisons between adult offspring of high and low pup-licking mothers. *Frontiers in psychiatry*, *2*, 42.
15. Wass, S. V., & Goupil, L. (2022). Studying the developing brain in real-world contexts: moving from castles in the air to castles on the ground. *Frontiers in Integrative Neuroscience*, *16*, 896919.
16. Richardson, Sarah S., Cynthia R. Daniels, Matthew W. Gillman, Janet Golden, Rebecca Kukla, Christopher Kuzawa, and Janet Rich-Edwards. "Society: don't blame the mothers." *Nature* 512, no. 7513 (2014): 131-132.
17. Bennett, M., Webster, A. A., Goodall, E., Rowland, S., Bennett, M., Webster, A. A.,... & Rowland, S. (2018). Establishing contexts for support: Undoing the legacy of the "refrigerator mother" myth. *Life on the autism spectrum: Translating myths and misconceptions into positive futures*, 61-80.
18. Yamasue, H., & Domes, G. (2017). Oxytocin and Autism Spectrum Disorders. *Behavioral Pharmacology of Neuropeptides: Oxytocin*, 449–465. https://doi.org/10.1007/7854_2017_24
19. Klin, A., Micheletti, M., Klaiman, C., Shultz, S., Constantino, J. N., & Jones, W. (2020). Affording autism an

early brain development re-definition. *Development and Psychopathology*, *32*(4), 1175-1189.

20. Mogi, K., Ooyama, R., Nagasawa, M., & Kikusui, T. (2014). Effects of neonatal oxytocin manipulation on development of social behaviors in mice. *Physiology & Behavior*, *133*, 68–75. https://doi.org/10.1016/j.physbeh.2014.05.010
21. Bales, K. L., Perkeybile, A. M., Conley, O. G., Lee, M. H., Guoynes, C. D., Downing, G. M.,... & Mendoza, S. P. (2013). Chronic intranasal oxytocin causes long-term impairments in partner preference formation in male prairie voles. *Biological psychiatry*, *74*(3), 180-188.

7 Summary

Early-life social behaviors impact an organisms further physical, socioemotional and cognitive development profoundly. In all mammalian species the mother is a central interaction partner. Disruptions in early-life mother-infant interactions have been associated with an increased risk for physical, socioemotional and cognitive development. Therefore, preclinical research is necessary to better understand underlying neural mechanisms involved in healthy and abnormal sociocommunicative development. However, establishing valid and reproducible research findings requires precise and accurate assessment tools.

The Pup Retrieval test (PRT) is the leading behavioral assay to assess the maternal behavioral response to infant isolation distress in laboratory rodents. The PRT quantifies maternal behavior of a dam by the time it needs to approach, find and carry back an isolated pup. Despite the effectiveness and usefulness of the PRT, sampling procedures are currently lab and even researcher dependent, and test output is rather limited. In this project, we developed novel phenotypes for assessment of early-life bidirectional development based on the PRT.

First, we focused on extending the level of behavioral analysis and improving test accuracy, reliability and reproducibility. Using open-access deep neural networks and machine learning technology we developed an automated procedure which is able to accurately and reliably estimate the classic PRT parameters automatically. Additionally, we extended this to quantify ethologically relevant components of maternal retrieval behavior such as maternal approach and carrying.

Second, the automated PRT was used as basis for a novel behavioral test assay: BAMBI. BAMBI is the first behavioral test assay to assess early-life bidirectionality in mother-pup dyads in mice. In contrast to existing behavioral paradigms, BAMBI simultaneously records pup isolation vocalizations and maternal behaviors to investigate the action-reaction dynamic between them. This is particularly important in preclinical research of rodent models of disorders with early-life communication deficits such as autism spectrum disorders since infant and/or maternal factors might underlie and ameliorate infant social deficits. In our experiments, for example, we showed a significant association between number of ultrasonic vocalizations of an infant during PRT and retrieval success.

Finally, we applied this early-life methodology to the prenatal valproate exposure mouse model, an established model for autism spectrum disorder (ASD). Here, we administered THIQ, a potent and selective melanocortin-4 receptor agonist, in both male and female neonates to profile its impact on early-life and adult ASD-like behaviors. We were able to show differences between treatment groups in early-life social interactions with their mother as well as long lasting behavioral alterations.

8 Samenvatting

Vroege sociale gedragingen hebben een diepgaande invloed op de verdere fysieke, sociaal-emotionele en cognitieve ontwikkeling van een organisme. In alle zoogdiersoorten is de moeder een centrale interactiepartner. Verstoringen in vroege moeder-infant interacties worden in verband gebracht met een verhoogd risico op fysieke, sociaal-emotionele en cognitieve ontwikkelingsproblemen. Daarom is preklinisch onderzoek nodig om de onderliggende neurale mechanismen die betrokken zijn bij typische en atypische sociocommunicatieve ontwikkeling beter te begrijpen. Het verkrijgen van geldige en reproduceerbare onderzoeksresultaten vereist echter nauwkeurige en betrouwbare meetinstrumenten.

De Pup Retrieval-test (PRT) is de meest gebruikte gedragsproef om de moederlijke gedragsreactie op de isolatiedistress van een pup te beoordelen bij laboratoriummuizen. De PRT kwantificeert het gedrag van een moederdier aan de hand van de tijd die het kost om een geïsoleerde pup te benaderen, te vinden en terug te dragen naar het nest. Ondanks de effectiviteit en bruikbaarheid van de PRT, zijn de testprocedures momenteel afhankelijk van het laboratorium en zelfs van de onderzoeker, en is de output van de test nogal beperkt. In dit project hebben we op basis van de PRT nieuwe kenmerken ontwikkeld voor de beoordeling van vroegtijdige bidirectionele ontwikkeling.

Ten eerste hebben we ons gericht op het uitbreiden van het niveau van gedragsanalyse en het verbeteren van de nauwkeurigheid, betrouwbaarheid en reproduceerbaarheid van de PRT. Met behulp van open-access deep neural networks en machine learning-technologie hebben we een geautomatiseerde procedure ontwikkeld waarmee de klassieke PRT-parameters nauwkeurig en betrouwbaar automatisch kunnen worden geschat. Bovendien hebben we dit uitgebreid om ethologisch relevante componenten van moederlijk retrieval-gedrag, zoals moederlijke benadering en dragen, te kwantificeren.

In een tweede studie werd de geautomatiseerde PRT gebruikt als basis voor een nieuwe gedragstest: BAMBI. BAMBI is de eerste gedragstest die vroege bidirectionaliteit in moeder-pup dyades bij muizen beoordeelt. In tegenstelling tot bestaande gedragsparadigma's registreert BAMBI gelijktijdig pup vocalisaties en moederlijk gedrag om de wisselwerking tussen hen te onderzoeken. Dit is met name belangrijk in preklinisch onderzoek in knaagdiermodellen van stoornissen met vroege communicatieve beperkingen, zoals autismespectrumstoornissen, omdat zowel kinder- als moederfactoren ten grondslag kunnen liggen aan en de tekorten in sociale interactie van de pups kunnen verergeren of verzachten. In onze experimenten hebben we bijvoorbeeld een significante associatie aangetoond tussen het aantal ultrasone vocalisaties van een pup tijdens de PRT en retrieval succes.

Tot slot hebben we deze methodologie toegepast op het muismodel van prenatale blootstelling aan

valproïnezuur, een geaccepteerd model voor autismespectrumstoornis (ASS). Hier hebben we THIQ, een potente en selectieve agonist van de melanocortine-4 receptor, toegediend aan zowel mannelijke als vrouwelijke pasgeboren pups om de impact ervan op vroegtijdig ASS-gerelateerd gedrag te onderzoeken. We waren in staat om verschillen tussen behandelingsgroepen te laten zien in vroegtijdige sociocommunicatieve gedragingen in pup alsook veranderingen in moederlijk gedrag.

9 Popularized summary

How we interact with others when we're young has a big impact on our physical health, emotions, and thinking abilities as we grow up. For most mammals, including humans, mothers play an central role in these interactions. When there are problems in how infants and their mothers interact early in life, it can increase the chances of having difficulties in various areas later on. That's why scientists need to study these early-life interactions in animals before they can understand how it works in humans.

To study this, scientists use a test called the Pup Retrieval test (PRT). This test measures how a mother animal responds to her infants' distress when they are separated. The test looks at things like how long it takes for the mother to find and bring back her infant. Although the PRT is helpful, the way it's done can be different between labs and researchers, and the results are limited.

In this research project, we wanted to improve the PRT and develop new ways to understand the relationship between infants and their mothers. We used advanced technology like deep neural networks and machine learning to automate the test and make it more accurate and reliable. We also looked at other behaviors, like how the mother approaches and carries the pup, to get a better picture of their interactions.

In a second step, we used this improved PRT as a basis to develop a new behavioral test called BAMBI. BAMBI is the first test that looks at both the infants' sounds and the mother's behaviors at the same time. This helps scientists understand how mother and infant react to each other. This is important when studying animal models of conditions like autism, where infants and/or mothers might have difficulties communicating. In this experiment, we found that the number of sounds the infant makes during the PRT was linked to how well the mother brought them back to the nest.

Finally, we used this new approach to study a mouse model of autism. We gave certain drugs to newborn mice and looked at how it affected their behavior as they grew up. We found that the drug caused differences in how the mice interacted with their mothers early in life and also had effects on both maternal and infant development.

Overall, this research helps us understand the importance of early interactions between mothers and infants and how it can affect their development. By studying animal models, scientists can learn more about the biology of conditions like autism and find ways to help infants with communication difficulties.

10 Populaire samenvatting

Hoe we omgaan met anderen als we jong zijn heeft een serieuze impact op onze fysieke gezondheid, emoties en denkvermogen naarmate we opgroeien. Voor de meeste zoogdieren, inclusief mensen, spelen moeders een centrale rol in deze interacties. Wanneer er problemen zijn in de vroege interacties tussen kinderen en hun moeders, kan dit de kans op latere moeilijkheden vergroten. Daarom is het belangrijk dat wetenschappers deze vroege interacties in dieren te bestuderen om dit beter te kunnen begrijpen hoe dit werkt in mensen.

Om dit te bestuderen, gebruiken wetenschappers de Pup Retrieval test (PRT). Deze test meet hoe een moederdier reageert op de noodsignalen van haar jongen wanneer ze van elkaar gescheiden zijn. De test kijkt naar zaken zoals hoelang het duurt voordat de moeder haar jongen vindt en terugbrengt. Hoewel de PRT nuttig is, kan de uitvoering verschillen tussen laboratoria en onderzoekers, en de resultaten zijn beperkt.

In dit onderzoeksproject wilden we de PRT verbeteren en nieuwe manieren ontwikkelen om de relatie tussen jonge dieren en hun moeders te begrijpen. We gebruikten geavanceerde technologie zoals diepe neurale netwerken en machine learning om de test te automatiseren en nauwkeuriger en betrouwbaarder te maken. We keken ook naar andere gedragingen, zoals hoe de moeder haar jongen benadert en draagt, om een beter beeld te krijgen van hun interacties.

In een tweede stap gebruikten we deze verbeterde PRT als basis voor het ontwikkelen van een nieuwe gedragstest genaamd BAMBI. BAMBI is de eerste test die zowel de geluiden van de jongen als het gedrag van de moeder tegelijkertijd onderzoekt. Dit helpt wetenschappers te begrijpen hoe moeder en jong op elkaar reageren. Dit is belangrijk bij het bestuderen van diermodellen van aandoeningen zoals autisme, waarbij jongen en/of moeders moeite kunnen hebben met communicatie. In dit experiment ontdekten we dat het aantal geluiden dat de jongen maken tijdens de PRT verband hield met hoe goed de moeder ze terugbracht naar het nest.

Tot slot gebruikten we deze nieuwe aanpak om een muismodel van autisme te bestuderen. We gaven bepaalde medicatie aan pasgeboren muizen en keken hoe dit van invloed was op hun gedrag naarmate ze opgroeiden. We ontdekten dat de medicatie verschillen veroorzaakte in hoe de muizen in de vroege levensfase met hun moeders omgingen en ook effecten had op zowel de ontwikkeling van de moeder als de jongen.

Over het algemeen helpt dit onderzoek ons te begrijpen hoe belangrijk vroege interacties tussen moeders en jongen zijn en hoe dit hun ontwikkeling kan beïnvloeden. Door diermodellen te bestuderen, kunnen wetenschappers meer leren over de biologie van aandoeningen zoals autisme en manieren vinden om jongen met communicatiemoeilijkheden te helpen.

11 Acknowledgements, Personal Contribution and Conflict of Interest Statements

Scientific Acknowledgement

In addition to the acknowledgements mentioned at the end of chapters 3-5, I would like to express my gratitude to Prof. Dr. Rudi D'Hooge, Prof. Dr. Guy Bosmans, and Wim Gorssen for proofreading this thesis. I would also like to acknowledge the financial support of FWO.

Absolutely Not Scientific Acknowledgements

Dit manuscript is een testament van slapeloze nachten, existentiële crisissen en een borderline ongezonde obsessie met mijn onderzoek. Voor ik dit werk afsluit – en dat plan ik al dansend met een Duvel in de hand te doen – wil ik deze ruimte toch nog even gebruiken om een aantal kleppers van exceptioneel kaliber te erkennen voor hun rol in dit werk.

Ik wil een welgemeende dank aan mijn Wim uitdrukken voor zijn eindeloze steun tijndens mijn doctoraatsavontuur. Ik zou zeggen dat uw kaliber van “geweldige man zijn” statistisch significant is, maar daarvoor ben ik met te weinig mannen getrouwd geweest. Wel kan ik met zekerheid zeggen dat uw aanmoediging de perfecte, statistische outlier was, een anomalie die los stond van de rest en mij enorm naar deze eindstreep heeft getrokken. Zoals een p-waarde die nul benaderd, was uw geloof in mij statistisch significant en uw liefdevolle kritieken waren de standaardafwijking die mij gegrond hielden te midden van alle onzekerheid. Zonder Wim Gorssen zou mijn academisch pad even verwarrend geweest zijn als een polynomiale regressie proberen te fitten in een wolk van random gescatterde data punten. Bedankt lieve Wim, om mijn constante en mijn outlier te zijn.

Ik wil ook een diepe appreciatie uitdrukken aan mijn ouders. Uit mijn onderzoek heb ik geleerd dat jullie een spectaculaire bijdrage hebben geleverd aan mijn ontwikkeling op fysiek, sociaal en cognitief vlak. Bedankt om mij niet te mishandelen en om mij niet te verwaarlozen. Ik weet dat ik het even spannend heb gehouden door met meer buizen thuis te komen dan nodig zijn om het volledige Baltisch Pijpleidingsysteem te vervangen, maar nu sta ik hier met een doctoraatsmanuscript! Jullie zijn schitterende ouders, bedankt voor alles! Bedankt aan mijn schoonouders om hetzelfde met Wimmeke te doen en bedankt aan zijn broers, voornamelijk Jef, die hem zo in nederigheid gedrild hebben. Ook wil ik Jef er enorm voor bedanken voor zijn vele relativeringen waarbij hij mij altijd op het hart drukte dat de universiteit “ook maar een beschutte werkplaats is voor hoogopgeleiden”.

Voor het schrijven had ik mijn eigen bron van inspiratie in mijn heilige drievuldigheid: Jetje, Moontje en Haasje. Jullie elke dag meer en meer zien ontwikkelen in de enorme wappen die jullie zijn, is één van mij dichtste bronnen van geluk. Elk van jullie drie mist wel een hoek en dat maakt mij enorm trots op jullie. Naar aanleiding van dit werk zal ik jullie ook altijd steunen in het niet in de rij lopen, al stel ik

wel voor geen drugs te doen en naar school te gaan of op zijn minst geen drugs op school te doen.

Voor mijn lieve Amber, die naast één van mijn dichtste vriendinnen ook één van de beste mama's is die ik ken. Ik kan u niet zeggen hoeveel deugd het mij heeft gedaan om af en toe tot bij u te komen en en te horen "Maar eerst, wat drinkt ge, wichtje?" Ik ben enorm dankbaar dat ik in u iemand gevonden heb die vaak een touwladder heeft gegooid naar de bodem van de put, maar die eveneens op de eerste rij staat als goed nieuws gevierd wordt.

In dezelfde lijn, Paulien, ik heb altijd bij u uit mogen razen, ge hebt mij altijd uw eerlijke mening gegeven en ik ben u daar enorm dankbaar voor. Ge zijt mijn fashion guru en moreel kompas. Bovendien zijt ge een enorm belangrijk persoon in het leven van mijn dochters en ik kan amper uitdrukken hoe speciaal dat voor mij voelt. Lars, ik ben u zo dankbaar dat ge Paulien in mijn leven hebt gebracht en ik kan niet wachten op het moment dat ik ooit jullie kindje mag ontmoeten (Voor de onbekende lezers: het gaat om een hypothetisch kind, niet om een kind dat ze al jarenlang voor mij verbergen. Hoop ik toch). Ook ben ik u altijd enorm dankbaar geweest voor de momenten van: "ik heb nog een leuk nieuw spel uit dat we eens kunnen spelen", gezien ze mij altijd enorm uit mijn warrige hoofd hebben gehaald.

Tom en Hanne, mijn tweekoppige peetmonster en ouders van mijn twee favoriete jongens. Voor al die keren dat jullie ons uit de volwassen wereld hebben gehaald om te spelen. Kolo-wie? Zo nu en dan een welgemikte tackle en eens heerlijk uitrazen in de speeltuin. Ik hoop dat jullie nog heel lang onze speelkameraden blijven, want ik voel me enorm gelukkig als ik omringd ben door jullie speelse vreugde.

Toon, ik wil ook uw rol in het leven van Moontje heel even in de verf zetten. Uw rol in het leven van Moontje illustreert zo mooi hoe belangrijk ook andere zorgfiguren in het leven van een kind zijn en ge vervult die rol met glans.

Mijn lieve, lieve Catia. Je bent een mama naar wie ik enorm opkijk en ik hoop vaak dat ik kan zijn voor mijn kinderen wat jij voor de jouwe bent. Maar dat terzijde ben je een ongelofelijk lief mens en heb je een enorme invloed gehad op mijn academisch werk. Als ik een idee voor onderzoek heb, leg ik het meestal eerst aan jou voorleg omdat je zo kritisch en goed beredeneerd bent en zoveel kennis van zake hebt. Daarbij is er niets gezelliger dan een fles rode wijn met jou te delen en eens heerlijk te genieten van gaan dansen. Dat er nog velen mogen volgen! Daarbij wil ik toch ook zeker de rest van mijn oud-collega's van Autisme Limburg bedanken om dat zij ook een vat van inspiratie en wijn voor mij blijven!

Lieve Nele en Lore, je stonden langs mij op mijn trouwdag omdat jullie er altijd voor mij geweest zijn en daar heeft dit chaotische doctoraatsparcours niets aan veranderd. Ik ben enorm gelukkig dat ik vriendinnen zoals jullie aan mijn zijde heb! Al moet ik zeggen: dit was het tweede feest van mijn kant op rij… Tijd om te trouwen, dames. De doctor schrijft het voor.

Personal Contribution

Prof. Dr. Guy Bosmans, Prof. Dr. Rudi D'Hooge, Dr. Zsuzsanna Callaerts-Végh, and Prof. Dr. Patrick Callaerts contributed to the design of the FWO project. Prof. Dr. Rudi D'Hooge and Carmen Winters designed the studies. Carmen Winters conducted the majority of the experiments, with assistance from Anamarija Banjac for the behavioral pup battery in the THIQ experiment. Occasional assistance with behavioral experiments was provided by Leen Van Aerschot, Emilie Vanempten, and Louise Moonen. Prof. Dr. Rudi D'Hooge provided the behavioral setups for conducting all analyses. Carmen Winters performed all the analyses with highly-appreciated input from Wim Gorssen. The development of automated algorithms was performed with assistance from Prof. Dr. Jan Clemens, Prof. Dr. Sam Golden, and Dr. Simon Nilsson. Carmen Winters wrote the thesis, and this manuscript was reviewed and corrected by Prof. Dr. Rudi D'Hooge, Prof. Dr. Guy Bosmans, and Wim Gorssen.

Conflict of Interest

No competing interests to declare

12 Curriculum vitae

CARMEN WINTERS

Cell phone: +32 470 44 19 21
Email: carmenwinters@hotmail.com

OrcID:
https://orcid.org/0000-0002-8077-561X
ResearchGate:
https://www.researchgate.net/profile/Carmen-Winters-2

EDUCATION

PhD Katholieke Universiteit Leuven, Laboratory for Biological Psychology 2023
Dissertation: "Profiling early-life social development in laboratory mice"

MS Katholieke Universiteit Leuven, Educational Studies Sep, 2017
Extracurricular courses: Organische chemie, klinische chemie en medische beeldvorming, Advances in Neurobiology of Psychiatric Disorders, Neuropsychologie, Anatomie en histologie
Thesis: "The effects of chronic oxytocin treatment in a mouse model of delayed brain development"
Promotor: prof. dr. Rudi D'Hooge

BS Katholieke Universiteit Leuven, Educational Studies Sep, 2015

GRANTS AND AWARDS

Grant Research Foundation - Flanders (Brussels, BE) Jan, 2018 – Dec, 2019
"A good start is half the battle "The role of the oxytocin / dopamine system on parent young attachment and social interaction in mice.

Grant Research Foundation - Flanders (Brussels, BE) Jan, 2020 – Dec, 2022
"A good start is half the battle "The role of the oxytocin / dopamine system on parent young attachment and social interaction in mice.
- 2 year grant extension
- 1 year administrative extension (maternity leave)

Grant Research Foundation - Flanders (Brussels, BE) Jan, 2023 – May, 2023
"A good start is half the battle "The role of the oxytocin / dopamine system on parent young attachment and social interaction in mice.
- 5 month Covid19 extension

Award Wageningen Publishers Award for most innovative talk, 73rd EAAP Annual Meeting, Porto, 2022/09/05.

RESEARCH EXPERIENCE

PhD Dissertation, Katholieke Universiteit Leuven, Leuven 2017-2023
Supervisor: prof. dr. Rudi D'Hooge
Co-supervisor: prof. dr. Guy Bosmans

- Designed a validation experiment of the valproic acid (VPA) mouse model of autism spectrum disorder
- Conducted behavioral research of VPA validation experiment using early-life behavioral assays and adult behavioral battery including social, cognitive and anxiety-related test assays
- Analyzed data from experiments to determine significant behavioral changes
- Performing animal care tasks and required safety procedures
- Maintained accurate records of daily activities and test interferences
- Automated and optimized the pup retrieval test using convolutional neural models and random forest behavioral classifiers
- Automated a synchronous behavioral assay to assess early-life bidirectionality
- Automated infant ultrasonic vocalization detection using convolutional neural network
- Designed and conducted a drug intervention experiment
- Visualization and analysis of large behavioral/physiological datasets
- Communicated results of experiments to colleagues and general public

Master Dissertation, Katholieke Universiteit Leuven, Leuven 2016-2017
Supervisor: prof. dr. Rudi D'Hooge

- Conducted behavioral research on the acute effects of oxytocin administration to a mouse model of delayed development using an adult behavioral battery including social, cognitive and anxiety-related test assays
- Analyzed data from experiments to determine significant behavioral changes
- Performing animal care tasks and required safety procedures
- Maintained accurate records of daily activities and test interferences
- Analyzed and summarized data in a thesis dissertation
- Presented findings to supervisor and jury

Master internship, Katholieke Universiteit Leuven, Leuven 2016-2017
Supervisor: prof. dr. Kaat Alaerts

- Recruited participants
- Served as live model for (un)reciprocated gaze
- Conducted a pilot study on the association between variations in attachment style and differential expression of sympathetic autonomic arousal upon live dyadic gaze interactions in human participants using skin conductance responses, EEG, eye tracking, respiration rate and heart rate.

TEACHING EXPERIENCE

Katholieke Universiteit Leuven, Leuven Sep, 2019 to Sep, 2022
Teaching Assistant, B-KUL-P0M03A - Ontwikkelingspsychologie, deel 1: kindertijd

- Assess student essays
- Give students feedback

Katholieke Universiteit Leuven, Leuven Sep, 2019 to Sep, 2022
Teaching Assistant, B-KUL-P0M04A - Ontwikkelingspsychologie, deel 2: adolescentie tot late volwassenheid

- Monitoring the quality of course material
- Assess student essays
- Give students feedback

Katholieke Universiteit Leuven, Leuven Sep, 2018 to Jan, 2019
Teaching Assistant, B-KUL-P0T68 - Psychologen aan het werk

- Supervise weekly practices
- Facilitate discussions
- Organize exam feedback sessions
- Give students exam feedback

Doctoral Students Supervised
Banjac Anamarija, "Melanocortin-4 Receptor Stimulation in Neonate Mice to Improve Autism-Related Behaviors", Sep 2022. Currently PhD student at Institute of Pathophysiology, University of Ljublijana.

Vanden Eynde Laura, "The Validation of Mouse Models in Research of Autism Treatments: Communication Differences in Pups' an Overview, Scientific and Ethical Discussion.", Sep 2021

Pauwels Joba, "The Behavioural Effects of Melanocortin-4 Receptor Activation", Sep 2021

PUBLICATIONS

Journal Publications

Winters, C., Gorssen, W., Ossorio-Salazar, V. A., Nilsson, S., Golden, S., & D'Hooge, R. (2022). Automated procedure to assess pup retrieval in laboratory mice. *Scientific Reports*, *12*(1), 1663–1663. https://doi.org/10.1038/s41598-022-05641-w

Gorssen, W., Winters, C., Meyermans, R., D'Hooge, R., Janssens, S., & Buys, N. (2022). Estimating genetics of body dimensions and activity levels in pigs using automated pose estimation. *Scientific Reports*, *12*(1), 15384–15384. https://doi.org/10.1038/s41598-022-19721-4

Winters, C., Gorssen, W., Wöhr, M., & D'Hooge, R. (2023). BAMBI: A new method for automated assessment of bidirectional early-life interaction between maternal behavior and pup vocalization in mouse dam-pup dyads. Frontiers in Behavioral Neuroscience, 17.

Gorssen, W., Winters, C., Meyermans, R., Chapard, L., Hooyberghs, K., Janssens, S., Huisman, A., Peeters, K., Mulder, H., & Buys, N., (2023). A promising resilience parameter for breeding: the use of weight and feed trajectories in growing pigs. Journal of Animal Science and Biotechnology [Accepted 1/06/2023].

Journal Papers in preparation

Winters, C., Gorssen, W., Banjac, A., & D'Hooge, R. "Melanocortin 4 receptor stimulation in neonate mice modulates autism-related development and behaviors,".

Datasets and software

Data, tracking model and behavioral classifiers for pup retrieval test: https://doi.org/10.17605/OSF.IO/RWHTD

Data and tracking algorithm for fattening pigs: https://doi.org/10.17605/OSF.IO/QKW5Y

'RESENTATIONS AND INVITED LECTURES

Presentation, "BAMBI: profiling early-life bidirectional interactions in mother-infant dyads in laboratory mice", TEATIME Webinar, [Online], 2023/06/19.

Presentation, "An open source pose estimation model for fattening pigs during weighing", 73rd EAAP Annual Meeting, Porto, 2022/09/05.

Lecture, "Early-life bidirectionality between mother and offspring in laboratory mice", Seminars on Theory and Research, Faculty of Psychology and Educational Sciences, [Online], 2021/12/15.

Lecture, "Early-life bidirectionality between mother and offspring in laboratory mice", Brain & Cognition seminar, Faculty of Psychology and Educational Sciences, [Online], 2021/12/21.

Presentation, "Automated assessment of early-life bidirectionality between mouse dams and their offspring", SIRG SoNeAt Virtual Mini-Conference 2021, [Online], 2021/09/23.

Lecture, "Is it over yet? Managing emotions in the pandemic world. Patient and parent perspectives," European Cystic Fibrosis Conference, [Online], 2021/06/09.

Presentation, "Bonding is a two-mice job: Quantification of mother-child interactions in mice", 8th International Meeting of the FWO Research Community "A multiple levels of analysis approach to typical and atypical development", Leuven, 2021/09/23.

Lecture, "Early-life bonding between mother and offspring: humans and mice", Seminars on Theory and Research, Faculty of Psychology and Educational Sciences, Leuven, 2019/12/11.

SCIENCE COMMUNICATION AND COMMUNICATION TO GENERAL PUBLIC

General public presentation, "Autisme spectrum stoornis: van muis naar mens", Autisme Limburg, Neerpelt, 2019/07/16.

Technical journal: Pig brother is watching you; Genetica van activiteit en lichaamsdimensies via automatische beeldanalyse. *Varkensbedrijf*, *31*(10), 32–33.

Technical journal: Gedrag van varkens sturen via genetica. *Pig Business, 16*(10), 18-21.

Online magazine: Interview FWO kennismakers: Lockdown zorgde voor kruisbestuiving tussen onderzoekskoppel. [December 2022]

Technical catalogue: TEATIME COST HCM catalogue [2023]

LANGUAGES

Dutch: Native Language

English: Advanced Listener, Speaker, Reader and Writer

COMPUTER SKILLS

Statistics and AI: Multivariate statistics, Machine learning, convolutional neural networks, random forests

Programming languages and other technologies: R, Python, DeepLabCut, SimBA, DAS, Excel, PowerPoint

MEMBERSHIP AND VOLUNTEERING

Autisme Limburg, Overpelt: Individual supervision (2010-2013)

Autisme Limburg, Overpelt: Group supervision (2014-2017)

Autisme Limburg, Overpelt: Board member Overpelt (2014-2017)

Membership, International Society for Developmental Psychobiology

Workgroup, COST TEATIME