# The behavioral phenotype of early life adversity: A 3-level meta-analysis of rodent studies

V. Bonapersona[a,*], J. Kentrop[a], C.J. Van Lissa[b], R. van der Veen[c], M. Joëls[a,d], R.A. Sarabdjitsingh[a]

[a] Department of Translational Neuroscience, UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht University, the Netherlands
[b] Department of Methodology and Statistics, Utrecht University, the Netherlands
[c] Centre for Child and Family Studies, Leiden University, the Netherlands
[d] University Medical Center Groningen, Groningen University, the Netherlands

## ARTICLE INFO

## ABSTRACT

Altered cognitive performance is considered an intermediate phenotype mediating early life adversity (ELA) effects on later-life development of mental disorders, e.g. depression. Whereas most human studies are limited to correlational conclusions, rodent studies can prospectively investigate how ELA alters cognitive performance in several domains. Despite the volume of reports, there is no consensus on i) the behavioral domains being affected by ELA and ii) the extent of these effects. To test how ELA (here: aberrant maternal care) affects specific behavioral domains, we used a 3-level mixed-effect meta-analysis, and thoroughly explored heterogeneity with MetaForest, a novel machine-learning approach. Our results are based on > 400 independent experiments, involving ~8600 animals. Especially in males, ELA promotes memory formation during stressful learning but impairs non-stressful learning. Furthermore, ELA increases anxiety-like and decreases social behavior. The ELA phenotype was strongest when i) combined with other negative experiences ("hits"); ii) in rats; iii) in ELA models of ~10days duration. All data is easily accessible with MaBapp (https://osf.io/ra947/), allowing researchers to run tailor-made meta-analyses, thereby revealing the optimal choice of experimental protocols and study power.

## 1. Introduction

Early life adversity (ELA) is a consistent risk factor of psychiatric disorders (Kessler et al., 2010; Teicher et al., 2016), and it is regularly associated with poorer cognitive outcomes later in life (Masson et al., 2016; Nelson et al., 2007; Vargas et al., 2018). Indeed, impaired cognitive processing is a prominent feature of psychopathologies (Masson et al., 2016; Millan et al., 2012; Monfils and Holmes, 2018), e.g. dysregulated contextual memory in post-traumatic stress disorder (Liberzon and Abelson, 2016) or social cognition in schizophrenia (Green et al., 2015). ELA may therefore alter cognitive development, thereby resulting in behavioral abnormalities that may render individuals more vulnerable to psychiatric disorders (Ammerman et al., 1986).

To investigate exactly how ELA affects cognitive processing, rodent models are a valuable resource: they complement human studies by in-depth and thorough investigations of otherwise hard-to-study mechanisms. In animal experiments, genetic and environmental influences can be more precisely controlled and experimentally varied than in humans (Knop et al., 2017). Furthermore, prospective designs are more feasible. For example, rodent studies have disentangled the different components of mother-pup interaction, a critical factor of early development across mammalian species (Bowlby, 1951; Harlow and Harlow, 1965; Meaney, 2001). This has helped uncover links between disturbed maternal care and disturbed emotional and cognitive functioning later in life, implicating the stress system (Levine, 2005) and "hidden regulators" (Hofer, 1978).

Rodent studies have also highlighted paradoxical ELA effects on cognitive abilities. For instance, Benetti et al. (Benetti et al., 2009) reported that rats with a history of ELA had impaired memory in the object recognition task. Conversely, Champagne et al. (Champagne et al., 2008) reported that ELA mice display increased memory in a fear conditioning paradigm. Both tests have historically been used as memory tasks, albeit in a non-stressful and stressful context respectively. Possibly, the equivocal results are due to different underlying biological mechanisms (e.g. learning in stressful versus non-stressful situations) or pertain to the divergent methodology used (e.g. type of test or ELA model, species, experimenters, labs). A few studies have investigated the latter by testing the same animals in different memory tasks (Bredy et al., 2003; Ivy et al., 2010; Kanatsou et al., 2017; Mello

---

* Corresponding author.
  *E-mail address:* v.bonapersona-2@umcutrecht.nl (V. Bonapersona).

et al., 2009). Although these studies favor the former explanation, the limited amount of animals used (Button et al., 2013) – alongside the heterogeneous methodology – prohibits firm conclusions.

To address this conundrum, we here carried out a large-scale 3-level meta-analysis of all peer-reviewed preclinical literature on the subject, and tested the hypothesis that ELA (here defined as aberrant maternal care, i.e. differing from care seen in undisturbed, standard housed laboratory mice and rats) differentially affects specific behavioral domains in adulthood. We focused on memory formation after stressful or non-stressful learning, anxiety-like and social behavior, given their relevance for psychopathologies. We addressed (potential) sex-differences by investigating males and females separately. Furthermore, we tested whether the presence of multiple hits (e.g. other negative life experiences, independent of the developmental stage, see S1.4) (Daskalakis et al., 2013) amplified ELA effects. Finally, we applied the novel, machine-learning based analysis *MetaForest* (van Lissa, 2018a) to identify the most important moderators of ELA effects on behavior.

Based on this comprehensive analysis, we evaluate the translational potential of ELA rodent models. With the aid of a specially developed web-based tool MaBapp (*Meta-Analysis of Behavior application*) (https://osf.io/ra947/), interested researchers can perform their own meta-analysis and retrieve valuable *ad hoc* information for experimental design and power calculations.

## 2. Methods

We adhered to SYRCLE's guidelines (De Vries et al., 2015; Leenaars et al., 2012), and to the PRISMA (Moher et al., 2009) reporting checklist. To ease reading of the methodology, definitions of technical terms are provided in Supplemental Methods (S1.1). A summary of the general approach can be found in Fig. 1.

### 2.1. Search strategy

The electronic databases PubMed and Web of Science (Medline) were used to conduct a comprehensive literature search on *the effects of ELA on behavior* on December 6th 2017. The search string was constructed with the terms "behavioral tests", "ELA" (as aberrant postnatal maternal care) and "rodents" (S1.2).

Prior to the beginning of the study, four experts were consulted. After elaborate discussions they agreed upon *i)* the selection of tests and related outcomes (S1.3), *ii)* their classification into behavioral domains (S1.3) and *iii)* the definition of multiple hits (S1.4). The results of each individual test, independent of categorization, are available for consultation on MaBapp (Section 5.1). Studies' titles and abstracts were screened independently by two researchers (VB & JK) and selected if the inclusion criteria were met (S1.5). Studies' inclusion was performed blinded to the studies' results. In case of doubt, the full text was inspected. Any disagreement was resolved by greater scrutiny and discussion.

To limit subjectivity in the data gathering and entry process, data from eligible studies were extracted in a standardized dataset alongside its explanatory codebook (https://osf.io/ra947/).

For each individual comparison, we calculated Hedge's G (Viechtbauer, 2010), a standardized mean difference with a correction for small samples (Vesterinen et al., 2014). S1.6 details the extraction of statistical information as well as handling of missing values. We estimated the summary statistics of data presented only graphically with *Ruler for Windows* (Latour, 2006), of which we previously validated the accuracy (Bonapersona et al., 2018). If the data was not reported in any format (or other crucial information was missing e.g. sex), we contacted two authors per manuscript published after 2008 (response rate 52.6%). If no answer was received within two months and after a reminder, the authors were considered not reachable, and the comparison was excluded.

### 2.2. Meta-analysis: research questions and statistical approach

To avoid possible biases, the experimenter (VB) was blinded to the ELA effects while coding the analysis. This was achieved by randomly multiplying half of the effect sizes by -1.

#### 2.2.1. Hypothesis-testing

We built a 3-level mixed effect meta-analysis with restricted maximum likelihood estimation. In our experimental design, the 3 levels correspond to variance of effect size between 1) animals, 2) outcomes and 3) experiments. This approach accounts for the violation of the assumption of independency when the data is collected from the same animals (Aarts et al., 2014; Bonapersona et al., 2018; Cheung, 2014), thereby improving the robustness of the conclusions drawn. We included "domains" and "hits" as moderators in order to address the following two research question: 1) *what are the effects of ELA on each behavioral domain?*; 2) *are the effects enhanced if the animals experienced multiple hits?*. Since both questions were answered with the same model, effect sizes were estimated only once.

We ensured that all behavioral measurements were in the same theoretical direction by multiplying – whenever necessary – the effect sizes by -1 (S1.3) (Vesterinen et al., 2014). Although this was essential for the model estimation, we here report effect sizes in a more interpretable manner: an increase in Hedge's G signifies an enhancement of the behavioral domain under study (e.g. more anxiety-like behavior, better memory).

We conducted several sensitivity analyses (S1.7) to assess the robustness and consistency of our conclusions. We examined whether the quality of the studies affected the estimation of the results by dissecting the influence of reporting bias, blinding, randomization and study power. Furthermore, we thoroughly investigated influential and outlying cases (Viechtbauer and Cheung, 2010) according to multiple definitions (S1.7).

To compensate for methodological limitations, we tested the presence of publication bias with various qualitative/quantitative methods (S1.8), and quantified its influence with fail-safe N (Rosenthal, 1979) and trim-and-fill analyses (Duval and Tweedie, 2000) (S1.8).

Risk of bias was evaluated with SYRCLE's assessment tool (Hooijmans et al., 2014), where we distinguished between study-level and outcome-level biases (Moher et al., 2009). Lack of reporting of experimental details was scored as an *unclear* risk of bias.

Heterogeneity was assessed with Cochrane Q-test (Cheung, 2014) and $I^2$, which was estimated at each of the 3-levels of the model to determine how much variance could be attributed to differences within (level 2) or between experiments (level 3) (Assink and Wibbelink, 2016). Estimates of explained variance can be positively biased when based on the data used to estimate the model (Hastie et al., 2009). For this reason, we used 10-fold cross-validation to obtain an estimate of how much variance our model might explain in new data. This cross-validated estimate of $R^2$ ($R_{cv}^2$) is robust to overfitting and provides evidence for the results' generalizability.

P-values were corrected with Bonferroni for family-wise error rate (each research question considered as a separate family of tests) to limit capitalization on chance. Since we expected the amplitude of effect sizes to differ between sexes (Loi et al., 2017; Walker et al., 2017), we considered males and females as two separate datasets.

#### 2.2.2. Exploratory analysis

We used MetaForest (van Lissa, 2018a), a novel exploratory approach to identify the most important moderators of the ELA effects on behavioral domains. This innovative, data-driven technique adapts random forests (a machine learning algorithm) for meta-analysis, by means of bootstrap sampling. MetaForest ranks moderators based on their influence on the effect size.

Preclinical experiments often adopt diverse protocols. Although this can be an advantage (Karp, 2018), in a meta-analysis it induces
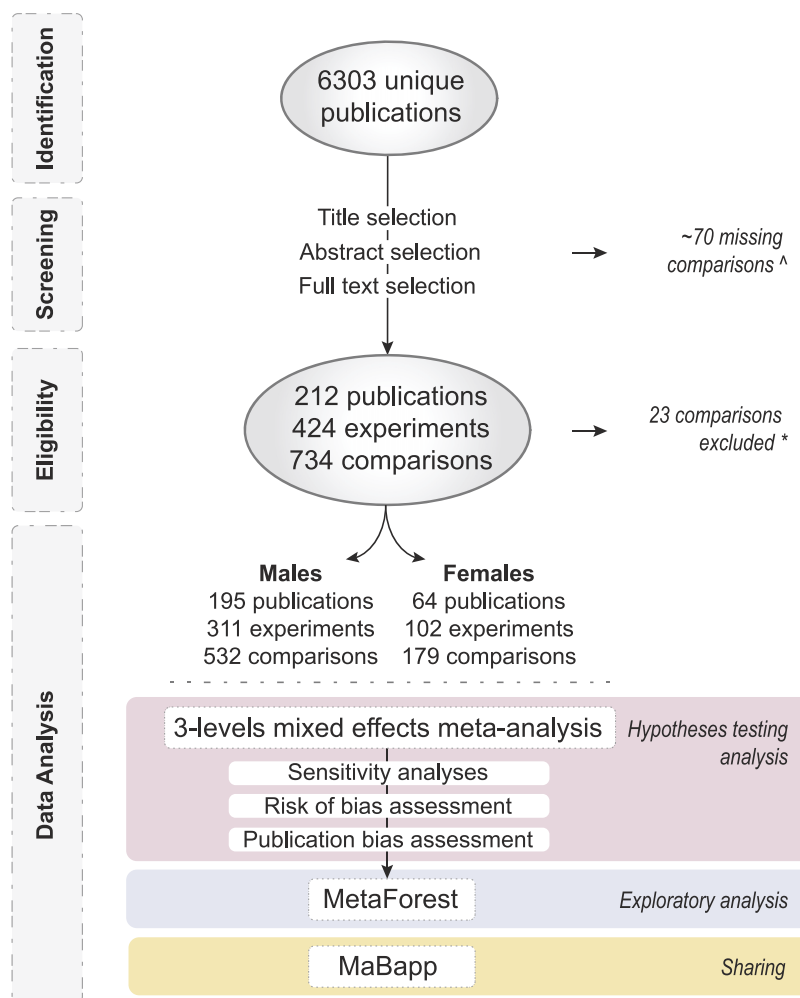
**Fig. 1.** Flow chart of study selection and analysis. Of note: in 47 publications, both males and females were tested. ^ = estimation of missing comparisons (S2.1); * = comparisons excluded from the meta-analysis due to controversial behavioral domain categorization (S2.2).

substantial heterogeneity. Therefore, we classified the published experimental protocols in > 30 standardized variables with the intent to identify potential methodological sources of heterogeneity. Based on theoretical importance, we selected 26 of these moderators for inclusion in the MetaForest analysis. We used 10-fold cross-validation (S1.9) to determine the optimal tuning parameters that minimized RMSE: uniform weighting, 4 candidate moderators at each split, and a minimum node size of 2. The marginal bivariate relationship of each moderator with effect size was averaged over the values of all other moderators (S1.9). Residual heterogeneity was estimated with tau$^2$ (S1.9).

Lastly, we created MaBapp (https://osf.io/ra947/) for anyone to perform their own meta-analysis on the topic by selecting their favorite characteristics (Section 5.1).

Analyses were conducted in R (version 3.5.1) (R Core Team, 2015), using the following packages: 1) *metafor* (Viechtbauer, 2010) for conducting the analysis, 2) *metaforest* (van Lissa, 2018b) for data exploration, 3) *shiny* (Chang et al., 2017) to create MaBapp, and 4) *dplyr* (Wickham et al., 2018) for general data handling. For further specifications about the analysis, the R script and the data are available (https://osf.io/ra947/).

## 3. Results

### 3.1. Studies selection and characteristics

In total ~8600 animals (age$_{weeks}$ median[IQR] = 12[4]; proportion rats = 68%) were included in the analysis, 77.7% of which were males. Anxiety-like behavior was the domain most investigated (48.8%), elevated plus maze the most popular test (14.3%), and maternal separation the ELA paradigm most often used (48.9%). For additional descriptive information on study characteristics, see S2.3.

Although no publication reported on all SYRCLE's potential bias items, 41 publications (19.3%) were blinded as well as randomized, and overall we estimated a risk of bias of 3[1] (median[IQR]) on a 10 points scale (S2.4). Lastly, at a systematic review level (S2.5), 68.5% of comparisons were either not-significant (n$_{comp}$ = 386) or the result could not be directly interpreted from the information provided (n$_{comp}$ = 117).

### 3.2. ELA effects are pronounced in males and with "multiple hits"

The effect sizes included ranged between -6.4 and 6.1 (mean [SD] = 0.29[1.06]), with 95% of comparisons between -2 and 2. Sample size ranged between 6 and 59 animals (mean[SD] = 22[7.8]), and differed < 20% between control and ELA groups in 90% of the cases (estimation).

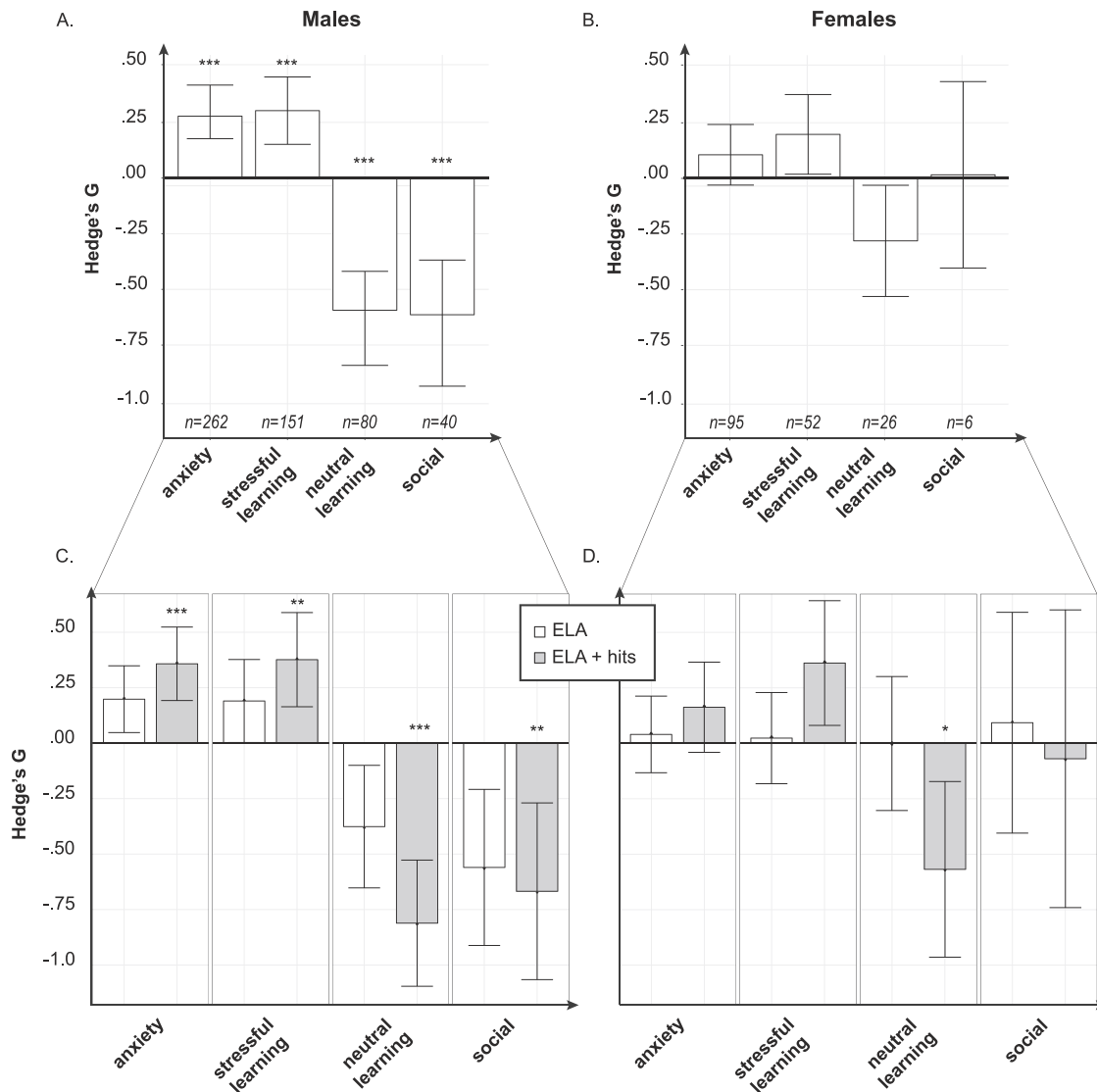When qualitatively comparing sexes, the effects of ELA were more

**Fig. 2.** Effects of ELA on behavioral domains in males (A) and females (B), and the role of multiple hits (in addition to ELA, grey bars) compared to only ELA (white bar) in mediating these effects (males: C, females: D). Each bar represents the size of the effect (G, standardized mean difference) of the ELA manipulation when comparing a control and an experimental group. * = p < 0.05, ** = p < 0.01, *** = p < 0.001.

evident in males than in females. **Male** rodents with a history of ELA displayed increased anxiety-like (Hedges'G[95%CI] = 0.278[0.165, 0.39], z = 4.819, p < 0.000), improved memory after stressful learning (Hedge'sG[95%CI] = 0.283[0.141, 0.425], z = 3.9, p < 0.000), impaired memory after non-stressful learning (Hedge'sG[95%CI] = -0.594[-0.792, -0.395], z = -5.86, p < 0.000) and decreased social behavior (Hedge'sG[95%CI] = -0.614[-0.88, -0.348], z = -4.521, p < .000, Fig. 2A, S2.6). We were unable to confirm any effect of ELA on behavior in **females**, although directionality was generally comparable in both sexes (Fig. 2B, S2.7).

Overall, the presence of multiple hits (for our definition, see S1.4) intensified the effects of ELA in **males** (Hedge'sG[95%CI] = 0.222[0.018, 0.426], z = 2.131, p = 0.033) yet marginally in **females** (Hedges'G[95%CI] = 0.297[-0.003, 0.596], z = 1.939, p = 0.052). Although these enhancing effects were not significant at a single-domain level (posthoc analysis, Fig. 2C-D, S2.6/S2.7), memory after non-stressful learning was the most impacted domain in **males** (difference in Hedge's G = 0.435, z = 2.156, p = 0.124) as well as in **females** (difference in Hedge's G = 0.565, z = 2.234, p = 0.102).

### 3.2.1. Sensitivity analyses and publication bias

Qualitative evaluation of funnel plot asymmetry suggested the presence of publication bias, which was confirmed by Egger's regression and Begg's test (S2.8). Nonetheless, fail-safe N as well as trim-and-fill analyses confirmed that – albeit present – publication bias is unlikely to distort the interpretation of the results (S2.8). Furthermore, the robustness of the male and female models was confirmed by several sensitivity analyses (S2.9).

### 3.3. Exploration of moderators

Although the models of the hypotheses-testing analysis described a significant proportion of variance ($R^2_{cvmales}$ = 0.026, $R^2_{cvfemales}$ = 0.03), substantial heterogeneity was recorded in both models (**males**: Q (524) = 1763.118, p < 0.000; **females**: Q(171) = 326.93, p < 0.000, S2.10). This was not surprising due to the diversity of publications included in the meta-analysis.

To investigate the source of the heterogeneity, we used MetaForest, a new statistical technique that ranks moderators (Fig. 3A) based on their predictive value. These can roughly be divided in 4 groups, describing: *i)* characteristics of the animals (e.g. origin of the breeding
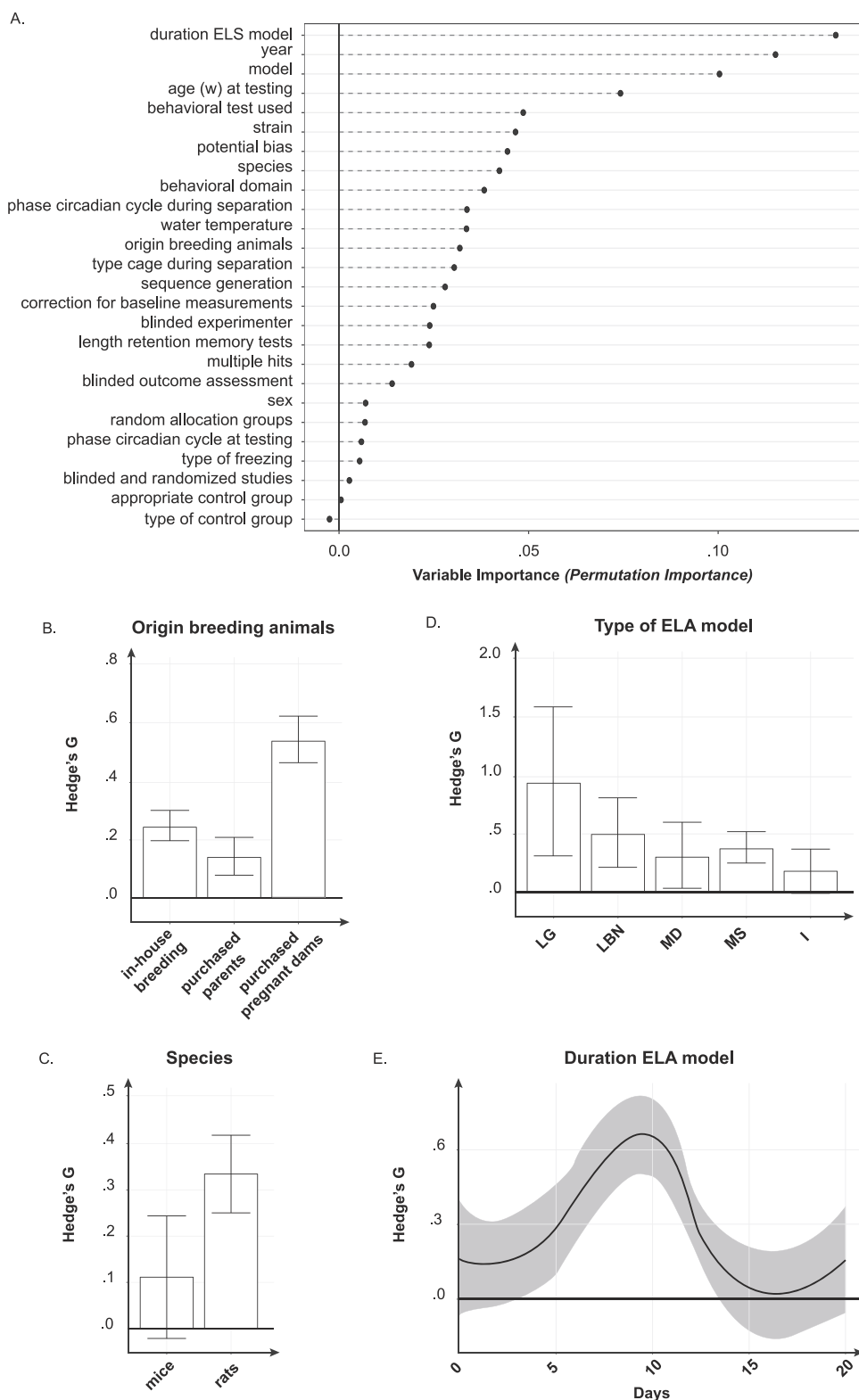
A.



**Fig. 3.** Exploratory MetaForest analysis. (A) Rank moderators' importance. Variable/permutation importance is a measure of how strongly each moderator explains differences in effect size, capturing (non-)linear relationships as well as higher order interactions. For information about MetaForest's partial dependence plots, see S2.11. Effect sizes distinguished by origin of the breeding animals (B), species (C), type of ELA model (D) and duration of ELA (E). Results are expressed as Hedge's G[95%CI]. The usefulness of this exploration can be best appreciated with the aid of MaBapp. For example, the overall estimate of the effects of ELA on anxiety-like behavior is Hedge's G = 0.24. However, if we select only the LBN model, the effect size rises to 0.37. If we combine LBN and rats, the effect size further rises to 0.60. If we then select only elevated plus maze as respectively behavioral test, the effect size rises to 0.81. LG = licking-and-grooming, LBN = limited bedding and nesting, MD = maternal deprivation, MS = maternal separation, I = isolation.

B.



C.



D.



E.



animals (Fig. 3B) and species investigated (Fig. 3C)), *ii)* ELA model used (e.g. type of model and duration of ELA (Fig. 3D-E)), *iii)* outcome measures (e.g. domain and test used), and iv) potential bias (e.g. blinding and randomization). MetaForest confirmed that the selected moderators account for a substantial portion of the variance ($R_{cv}^2$[SD] = 0.12[0.09]).

Offspring of dams purchased pregnant had larger effect sizes than offspring bred in the own facility (Fig. 3B). Rats had overall larger effect

sizes than mice (Fig. 3C). Concerning ELA models (Fig. 3D), selecting the extremes of natural variation (licking-and-grooming model) yielded the strongest phenotype. Lastly, effect sizes appeared to be maximal with a 10 days' ELA duration (Fig. 3E).

## 4. Discussion

In this study, we substantiate that adversities early in life

profoundly and lastingly change rodent behavior. Due to low power (Button et al., 2013) and heterogeneous methodologies, results at a single-study level are often inconclusive and difficult to interpret. Here, by adopting a meta-analytic approach, we provide extensive evidence that ELA (due to maternal care that differs from that provided by un-disturbed, standard-housed dams) has differential effects on memory: it enhances memory if learning occurs in a stressful situation, but it hampers learning under non-stressful circumstances. Furthermore, ELA increases anxiety-like and decreases social behavior, particularly in males. In line with the multiple-hits hypotheses (Daskalakis et al., 2013; Walker et al., 2009), the effects are amplified if the animals experience other stressful life events (e.g. prenatal stress due to transport of pregnant females), independent of the developmental period during which these occur (S1.4). These results are independent of the type of ELA or behavioral test used, and are remarkably similar to what has been reported at a correlational level in humans (Pechtel and Pizzagalli, 2011; Suor et al., 2015). Altogether, our results point to a clear and robust phenotype of ELA in four behavioral domains and complement the human literature by supporting a causative role of ELA in altering behavior, which may predispose individuals to precipitate symptoms of psychiatric disorders.

### 4.1. Methodological considerations

The lack of sufficient power to detect experimental effects is an emerging issue in preclinical literature (Bonapersona et al., 2018) that seriously hampers research interpretation (Button et al., 2013). As a consequence, results from single-studies are useful for hypotheses generation but do require replication. The ability to recreate experiments (replication) and/or to reach similar conclusions via different methods (reproducibility) are fundamental aspects of scientific inquiry. Underpowered research undermines both aspects, as the conclusions drawn are likely to be uncertain (Ioannidis, 2005).

Indeed, in our study the majority of comparisons (68.5%) was not-significant at a systematic review level, but the effects were significant when analyzed meta-analytically. In addition to study preregistration, realistic power calculations, and testing by several independent teams (Button et al., 2013; Ioannidis, 2005), statistical tools such as meta-analyses can therefore be very useful to substantiate conclusions from animal data and translate them more reliably to patients (Hooijmans and Ritskes-Hoitinga, 2013). Furthermore, our study showcases how "negative" research is also fruitful, and reminds how (lack of) formal statistical significance (typically p-*value* < 0.05) must not be a decisive requirement to publish research.

In this project, we intertwine these concepts with state-of the-art statistical methodology, adopting an approach never used in preclinical studies. Firstly, our meta-analysis was built with a 3-level model (Cheung, 2014), which allows for a more robust estimation of the effects by accounting for the dependency of same-animal's data (Bonapersona et al., 2018; Rosenthal, 1991). Secondly, a leading strength of preclinical meta-analyses is the systematic exploration of heterogeneity (Hooijmans and Ritskes-Hoitinga, 2013). Instead of the standard subgroup/meta-regression approach, we opted for an ex-ploratory analysis using MetaForest (van Lissa, 2018a), a newly de-veloped technique that ranks moderators' importance by adapting the machine learning algorithm *random forests* to summary-statistics' data. A major strength of MetaForest is its robustness to overfitting, and its ability to accommodate non-linear effects (van Lissa, 2018a), as shown by the impact of ELA duration on effect sizes.

Thirdly, we extensively coded potential (biological and experi-mental) moderators. Although possibly relevant moderators were not included due to insufficient reporting (e.g. temperature during se-paration (Pryce et al., 2003), cross-fostering (Penke et al., 2001), cul-ling (Veenema et al., 2007)), this dataset treasures relevant information for future experimental designs. To facilitate others to exploit this da-taset, we created **MaBapp** (https://osf.io/ra947/), a web-based app

with a user-friendly interface through which anyone can perform his/ her own meta-analysis on the topic of ELA and behavioral domains. Within the app, a wide variety of features can be selected, such as ELA models and their components (e.g. type, timing, predictability), beha-vioral tests used, age and sex of the animals, etc. Based on the char-acteristics indicated, the app reports forest, funnel and cumulative plots, as well as a list of relevant publications. The app is a useful re-source, which can be used to *i)* comprehensively retrieve relevant publications, *ii)* explore the literature at an individual researcher's needs' level, *iii)* define new hypotheses, *iv)* evaluate publication bias and replicability of findings, and *v)* estimate realistic effect sizes on which to ground future research.

The validity of our conclusions is not limited to the robustness of the models used but grounded on the vast primary evidence included (> 200 publications). As a consequence, accidental findings have little weight. Although the methods and approach we adopt are rigorous and reasonably conservative, the quality of the conclusions critically de-pends on the quality of the studies and data included. From our qua-litative bias assessment, the risk for potential bias was lower than previously reported in Neuroscience (Antonic et al., 2013; Bonapersona et al., 2018; Egan et al., 2011); yet, only ~20% of studies stated being blinded as well as randomized. Furthermore, any meta-analytic dataset is burdened with missing data, due to publication bias or to the pre-ferred investigation of certain factors over others (Cooper et al., 2009). Our models did display evidence of publication bias, yet they were robust to several corrections and sensitivity analyses. Although we cannot fully exclude that the above-mentioned limitations may affect the outcome, it is unlikely that the conclusions drawn would be sub-stantially impacted. Nevertheless, we have attempted to address these methodological issues as comprehensively as possible in our analysis.

### 4.2. Considerations on ELA models

ELA encompasses a wide range of pre- and post-natal experiences, but we here focused on altered maternal care (relative to care provided by undisturbed, standard-housed dams). Although this definition limits the generalizability of the conclusions, it is essential to enable the comparability (thus meaningful quantitative synthesis) of the studies incorporated in our meta-analysis.

The behavioral changes we report are presumably a convergent phenotype of distinct, model-dependent, underlying biological me-chanisms. An organism's development is not linear nor simultaneous for every component, but it occurs in critical periods (Hensch, 2005). For example, postnatal day (P)2-P5 is a sensitive period for the maturation of the adrenal glands (Levine and Lewis, 1959), P9 for prepulse in-hibition (Ellenbroek and Cools, 2002), and ~P10 for adrenal respon-siveness (Witek-Janusek, 1988). Furthermore, higher cognitive func-tions develop as multistage processes of sequential nature (Hensch, 2005). Accordingly, ELA may particularly disrupt the development of competences whose critical period is active during the time of stress, thereby heightening the variability of the ELA phenotype.

Evidence supporting these notions derives from studies using a single 24 h maternal deprivation paradigm, which show a persistent yet paradoxical hypo- and hyper- responsiveness of juvenile ACTH if de-privation occurred at P3 or P11 respectively (Van Oers et al., 1998). Thus, while meta-analyses may serve to discern patterns among vast amounts of studies, exploratory studies experimentally dissecting components of ELA in rodents remain indispensable for addressing the underlying mechanisms of action of ELA to the brain (for example: (Peña et al., 2017; Singh-Taylor et al., 2018)).

#### 4.2.1. Suggestions for future ELA research

Given that the criteria for construct and face validity of ELA models have been met (Suchecki, 2018), our results provide a practical fra-mework where researchers can anticipate the ELA effect on cognitive outcomes and/or build their own ELA model accordingly. Our

exploratory analysis gives insights in the suitability of the models and tests to choose, depending on the question.

Based on this analysis, we tentatively conclude that *i)* rats seem overall more sensitive to ELA-induced changes than mice. Moreover, *ii)* elements such as transporting pregnant dams appear to amplify the effects of ELA. Such stressful life events may have substantial impact on the system, in line with the multiple-hit theory (Daskalakis et al., 2013). As evident from Fig. 3, *iii)* a duration of ~10 days ELA produced the most robust phenotype. Finally, *iv)* the limited bedding and nesting (LBN) model produced the largest effect sizes when compared to separation/deprivation models. Given this reliability, in combination with the feasibility and translational validity, LBN seems an influential paradigm to investigate the mechanisms of chronic stress early in life (Rice et al., 2008; Walker et al., 2017).

According to the rank of moderators by MetaForest, publication year, age of testing, strain and behavioral test used account for a substantial portion of the variance. The impact of publication year has previously been reported in several areas of biology (Jennions and Møller, 2002), and could be the result of the *Winner's curse* (Button et al., 2013). In brief, the first published studies on any topic are likely to be biased towards extreme effect sizes. This bias tends to disappear as evidence accumulates, thereby providing an explanation for the influence of publication year in our dataset.

Conversely, age of testing, strain and behavioral test used did not show any theory-interpretable pattern. One explanation could be that there is no preferable age/strain/test, but that the different elements of the study design interact with one another. For example, the open field (OF) and the elevated plus maze (EPM) are behavioral tests used to assess anxiety-like behavior. Conceptually, they both aim to create a conflict between the rodents' exploratory drive and their fear of exposed spaces (Wigger and Neumann, 1999). With MaBapp, we can explore the confidence interval of these two tests following the LBN model (*OF*: Hedge'sG[95%CI] = 0.12[-0.21, 0.44]; *EPM*: Hedge'sG[95%CI] = 0.49[0.22, 0.75]) or maternal separation (*OF*: Hedge'sG[95%CI] = 0.32[0.14, 0.5]; *EPM*: Hedge'sG[95%CI] = 0.4[0.15, 0.65]). Tentatively, the EPM appears more sensitive than the OF to represent the effects of the LBN model, while rather similar when investigating the effects of maternal separation. Similarly, we can explore the interaction between these tests and any specific strain. For example, C57Bl/6 mice appear more sensitive to the EPM (Hedge'sG[95%CI] = 0.38[0.07,0.68]) than to the OF (Hedge'sG[95%CI] = 0.00[-0.27, 0.28]), independent of the ELA model used. These examples illustrate the complexity of these interactions. Unfortunately, the information so far available is insufficient to conduct meaningful quantitative analyses. Nonetheless, researchers can now make more informed decision on experimental designs by exploring with MaBapp (feasible) possibilities that fit their needs. Alternatively, we refer researchers to primary publications in which the effects of age (Oitzl et al., 2000) or strain (Millstein and Holmes, 2007) were experimentally investigated.

To reduce variability and improve comparability across studies, ELA should be preferably applied with consistent protocols (S1.5), unless manipulation of particular aspect(s) of the model is under investigation. Clearly, the importance of individual variation is a factor that should not be overlooked. In our analysis, the paradigm of licking-and-grooming – which is not experimentally induced but based on natural variation in maternal care – consistently evoked the largest effect sizes, although these were based on fewer publications than the other models.

### 4.3. Translational potential

ELA is one of the most consistent environmental risk factors for the development of psychopathology (Teicher et al., 2016). Although the effects of ELA on the brain can be adaptive, they may evolve into dysfunctional elements in genetically predisposed individuals (Teicher et al., 2016). Behavioral performance in specific cognitive domains seems to be a relevant intermediate phenotype (Ammerman et al.,

1986), as it may mediate the effects of ELA on psychopathology. For example, in post-traumatic stress disorder, enhanced memory of stressful events becomes pathological after a later-life trauma (Liberzon and Abelson, 2016).

In humans, the concept of ELA is extremely varied. Even when considering solely maltreatment, this can be characterized by repeated or sustained episodes of various forms of neglect and abuse (Teicher and Samson, 2013). Furthermore, the environmental variation is intertwined with socio-economic status, complex relations (e.g. family, neighborhoods, peers, school), and intergenerational transmissions (Teicher and Samson, 2013). Rodent paradigms do not capture the complexity of human ELA, but they can model specific *aspects* of the human variability in a well-controlled setting. For example, LBN is based on the erratic and unpredictability of maternal care (Baram et al., 2012; Rice et al., 2008; Walker et al., 2017) that has been established as a hallmark in childhood abuse situations (Whipple and Webster-Stratton, 1991). Similarly, cognitive performance (e.g. memory after stressful learning) can be modelled in rodents, albeit with clear restraints: the tasks are obviously different, should be interpreted in relation to the animal's normal behavior, and cannot investigate a range of outcomes such as verbal abilities, critical for social interaction and psychopathology (Cohen, 2001), also in relation to ELA (Miller et al., 2018).

Explaining how ELA increases psychopathology risk requires the understanding of its complex interplay with other susceptibility/resilience factors, such as genetic background and later life stressors (Bale et al., 2010). This mechanistic investigation is difficult to achieve in humans, where limited material, difficulty of prospective and longitudinal designs, complexity and lack of control over the environment and genetic variation hamper causal inferences of ELA to later life cognitive performance. To this end, animal studies can be of considerable added value (Walker et al., 2017).

An interesting issue in evaluating the translational potential of ELA rodent models is sex differences. In our analysis, males showed larger effect sizes (albeit in the same direction) than females to the effects of ELA on all outcomes, thereby confirming previous preclinical literature (Loi et al., 2017). Conversely, in clinical populations, females appear more sensitive to childhood trauma as well as to the development of stress-related psychopathologies (Walker et al., 2017), although sex differences depend on the type of disorder (Riecher-Rössler, 2017). A plausible biological explanation for this discrepancy is the developmental timing during which stress occurs. Although humans and rodents are altricial species, the brain of newborn rats corresponds roughly to 23/24-week old human fetuses (Plotsky et al., 2000). Interestingly, the sensitivity to adversities in the last trimester of gestation in humans has been suggested to affect males more than females (Bale and Epperson, 2015). Experimentally manipulating the timing of ELA exposure may further elucidate 'female' stress-sensitive periods. It therefore remains to be established whether the effects of ELA on cognitive domains are truly different between sexes. Our analyses suggest that the effects may not be sexually dysmorphic in nature but may result from the experimental designs used. For example, ELA models and behavioral tests were originally developed for males: maternal care shows clear sex-specific differences (Oomen et al., 2009; van Hasselt et al., 2012), and females perform poorly in behavior tests such as object recognition and object-in-location (Loi et al., 2017; Walker et al., 2017). Consequently these paradigms may not be sensitive enough for a female's phenotype. Indeed, the recorded effects were in the same direction across sexes, and MetaForest attributed to sex a relatively modest variable importance. Our results showcase the necessity to study sex as a biological variable (Bale and Epperson, 2015; McCarthy, 2016), which requires the development of tests and models that are female-specific. This step is required for a more meaningful comparison between rodent and humans, and a delineation of the underlying sex-dependent mechanisms of ELA.

Despite these drawbacks, our meta-analysis confirms and

importantly extends standing hypotheses on ELA based on exploratory studies. To aid future investigations in this field, we provide a online tool to evaluate existing literature and direct the experimental design of new studies.

## Acknowledgements

## Appendix A. Supplementary methods and results

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.neubiorev.2019.04.021.

## References

Aarts, E., Verhage, M., Veenvliet, J.V., Dolan, C.V., van der Sluis, S., 2014. A solution to dependency: using multilevel analysis to accommodate nested data. Nat. Neurosci. 17, 491–496. https://doi.org/10.1038/nn.3648.

Ammerman, R.T., Cassisi, J.E., Hersen, M., Van Hasselt, V.B., 1986. Consequences of physical abuse and neglect in children. Clin. Psychol. Rev. 6, 291–310. https://doi.org/10.1016/0272-7358(86)90003-6.

Antonic, A., Sena, E.S., Lees, J.S., Wills, T.E., Skeers, P., Batchelor, P.E., Macleod, M.R., Howells, D.W., 2013. Stem cell transplantation in traumatic spinal cord injury: a systematic review and meta-analysis of animal studies. PLoS Biol. 11. https://doi.org/10.1371/journal.pbio.1001738.

Assink, M., Wibbelink, C.J.M., 2016. Fitting three-level meta-analytic models in R: a step-by-step tutorial. Quant. Methods Psychol. 12, 154–174. https://doi.org/10.20982/tqmp.12.3.p154.

Bale, T.L., Epperson, C.N., 2015. Sex differences and stress across the lifespan. Nat. Neurosci. 18, 1413–1420. https://doi.org/10.1038/nn.4112.

Bale, T.L., Baram, T.Z., Brown, A.S., Goldstein, J.M., Insel, T.R., McCarthy, M.M., Nemeroff, C.B., Reyes, T.M., Simerly, R.B., Susser, E.S., Nestler, E.J., 2010. Early life programming and neurodevelopmental disorders. Biol. Psychiatry 68, 314–319. https://doi.org/10.1016/j.biopsych.2010.05.028.

Baram, T.Z., Davis, E.P., Obenaus, A., Sandman, C.A., Small, S.L., Solodkin, A., Stern, H., 2012. Fragmentation and unpredictability of early-life experience in mental disorders. Am. J. Psychiatry 169, 907–915. https://doi.org/10.1176/appi.ajp.2012.11091347.

Benetti, F., Mello, P.B., Bonini, J.S., Monteiro, S., Cammarota, M., Izquierdo, I., 2009. Early postnatal maternal deprivation in rats induces memory deficits in adult life that can be reversed by donepezil and galantamine. Int. J. Dev. Neurosci. 27, 59–64. https://doi.org/10.1016/j.ijdevneu.2008.09.200.

Bonapersona, V., Joels, M., Sarabdjitsingh, R.A., 2018. Effects of early life stress on biochemical indicators of the dopaminergic system: a 3 level meta-analysis of rodent studies. Neurosci. Biobehav. Rev. 95, 1–16. https://doi.org/10.1016/j.neubiorev.2018.09.003.

Bowlby, J., 1951. Maternal Care and Mental Health. World Heal. Organ. Monogr. Ser.

Bredy, T.W., Humpartzoomian, R.A., Cain, D.P., Meaney, M.J., 2003. Partial reversal of the effect of maternal care on cognitive function through environmental enrichment. Neuroscience 118, 571–576. https://doi.org/10.1016/S0306-4522(02)00918-1.

Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S.J., Munafò, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nat. Rev. Neurosci. 14, 365–376. https://doi.org/10.1038/nrn3475.

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. J. Neurosci. 28, 6037–6045. https://doi.org/10.1523/JNEUROSCI.0526-08.2008.

Chang, W., Cheng, J., Allaire, J., Xie, Y., McPherson, J., 2017. Shiny: Web Application Framework for R.Shiny: Web Application Framework for R.

Cheung, M.W.L., 2014. Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. Psychol. Methods 19, 211–229. https://doi.org/10.1037/a0032968.

Cohen, N., 2001. Language Impairment and Psychopathology in Infants, Children, and Adolescnets. Sage Publications.

Cooper, H., Hedges, L.V., Valentine, J.C., 2009. The Handbook of Research Synthesis and Meta-Analysis, 2nd edition. Russel Sage Foundation 2nd ed.

Daskalakis, N.P., Bagot, R.C., Parker, K.J., Vinkers, C.H., de Kloet, E.R.R., N.P., D., R.C., B., K.J., P., C.H., V., E.R., de K, 2013. The three-hit concept of vulnerability and resilience: toward understanding adaptation to early-life adversity outcome. Psychoneuroendocrinology 38, 1858–1873. LK- http://sfx.library.uu.nl/utrecht?sid=EMBASE&issn=03064530& id=doi:10.1016%2Fj.psyneuen.2013.06.008&atitle=The+three-hit+concept+of +vulnerability+and+resilience%3A+Toward+understanding+adaptation+to +early-life+adversity+outcome&stitle=Psychoneuroendocrinology& title=Psychoneuroendocrinology&volume=38&issue=9&spage=1858& epage=1873&aulast=Daskalakis&aufirst=Nikolaos+P.&auinit=N.P.& aufull=Daskalakis+N.P.&coden=PSYCD&isbn=&pages=1858-1873&date=2013 &auinit1=N&auinitm. https://doi.org/10.1016/j.psyneuen.2013.06.008.

De Vries, R.B.M., Hooijmans, C.R., Langendam, M.W., Leenaars, M., Ritskes-Hoitinga, M., Wever, K.E., 2015. A protocol format for the preparation, registration and publication of systematic reviews of animal intervention studies. Evid.-Based Preclin. Med. 1, 1–9. https://doi.org/10.1002/ebm2.7.

Duval, S., Tweedie, R., 2000. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 56, 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x.

Egan, K.J., Sena, E.S., Vesterinen, H.M., 2011. Making the most of animal data - improving the prospect of success in pragmatic trials in the neurosciences. Trials 12https://doi.org/10.1186/1745-6215-12-S1-A102. no pagination.

Ellenbroek, B.A., Cools, A.R., 2002. Early maternal deprivation and prepulse inhibition: the role of the postdeprivation environment. Pharmacol. Biochem. Behav. 73, 177–184. https://doi.org/10.1016/S0091-3057(02)00794-3.

Green, M.F., Horan, W.P., Lee, J., 2015. Social cognition in schizophrenia. Nat. Rev. Neurosci. 16, 620–631. https://doi.org/10.1038/nrn4005.

Harlow, H., Harlow, M., 1965. The affectional systems. In: Schrier, A., Harlow, H., Stollnitz, F. (Eds.), Behavior of Nonhuman Primates. Academic Press, New York, pp. 287–334.

Hastie, T., Friedman, J., Tibshirani, R., 2009. The Elements of Statistical Learning: Data Mining, Inference and Prediction, Second Edi. Ed. Springer Series in Statistics.

Hensch, T.K., 2005. Critical period plasticity in local cortical circuits. Nat. Rev. Neurosci. 6, 877–888. https://doi.org/10.1038/nrn1787.

Hofer, M.A., 1978. Hidden regulatory processes in early social relationships. In: Bateson, P., Klopfer, P. (Eds.), Perspectives in Ethology: Social Behavior. Springer, Boston, pp. 135–201.

Hooijmans, C.R., Ritskes-Hoitinga, M., 2013. Progress in using systematic reviews of animal studies to improve translational research. PLoS Med. 10, 1–4. https://doi.org/10.1371/journal.pmed.1001482.

Hooijmans, C.R., Rovers, M.M., Vries, R.B.M.De, Leenaars, M., Ritskes-hoitinga, M., Langendam, M.W., 2014. SYRCLE' s risk of bias tool for animal studies. BMC Med. Res. Methodol. 14, 1–9. https://doi.org/10.1186/1471-2288-14-43.

Ioannidis, J.P.A., 2005. Why most published research findings are false. PLoS Med. 2, e124. https://doi.org/10.1371/journal.pmed.0020124.

Ivy, A.S., Rex, C.S., Chen, Y., Dube, C., Maras, P.M., Grigoriadis, D.E., Gall, C.M., Lynch, G., Baram, T.Z., 2010. Hippocampal dysfunction and cognitive impairments provoked by chronic early-life stress involve excessive activation of CRH receptors. J. Neurosci. 30, 13005–13015. https://doi.org/10.1523/JNEUROSCI.1784-10.2010.

Jennions, M.D., Møller, A.P., 2002. Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution. Proc. R. Soc. B Biol. Sci. 269, 43–48. https://doi.org/10.1098/rspb.2001.1832.

Kanatsou, S., Karst, H., Kortesidou, D., van den Akker, R.A., den Blaauwen, J., Harris, A.P., Seckl, J.R., Krugers, H.J., Joels, M., S., K, H., K, D., K., R.A, V.D.A., J, den B., A.P, H., J.R, S., H.J, K., M., J., 2017. Overexpression of mineralocorticoid receptors in the mouse forebrain partly alleviates the effects of chronic early life stress on spatial memory, neurogenesis and synaptic function in the dentate gyrus. Front. Cell. Neurosci. 11, 1–13. https://doi.org/10.3389/fncel.2017.00132.

Karp, N.A., 2018. Reproducible preclinical research—Is embracing variability the answer? PLoS Biol. 16, 1–5. https://doi.org/10.1371/journal.pbio.2005413.

Kessler, R.C., McLaughlin, K.A., Green, J.G., Gruber, M.J., Sampson, N.A., Zaslavsky, A.M., Aguilar-Gaxiola, S., Alhamzawi, A.O., Alonso, J., Angermeyer, M., Benjet, C., Bromet, E., Chatterji, S., de Girolamo, G., Demyttenaere, K., Fayyad, J., Florescu, S., Gal, G., Gureje, O., Haro, J.M., Hu, C.-Y., Karam, E.G., Kawakami, N., Lee, S., Lépine, J.-P., Ormel, J., Posada-Villa, J., Sagar, R., Tsang, A., Ustün, T.B., Vassilev, S., Viana, M.C., Williams, D.R., 2010. Childhood adversities and adult psychopathology in the WHO world mental health surveys. Br. J. Psychiatry 197, 378–385. https://doi.org/10.1192/bjp.bp.110.080499.

Knop, J., Joëls, M., van der Veen, R., 2017. The added value of rodent models in studying parental influence on offspring development: opportunities, limitations and future perspectives. Curr. Opin. Psychol. 15, 174–181. https://doi.org/10.1016/j.copsyc.2017.02.030.

Latour, R., 2006. A Ruler for Windows [WWW Document]. URL https://a-ruler-for-windows.en.softonic.com/.

Leenaars, M., Hooijmans, C.R., van Veggel, N., ter Riet, G., Leeflang, M., Hooft, L., van der Wilt, G.J., Tillema, A., Ritskes-Hoitinga, M., 2012. A step-by-step guide to systematically identify all relevant animal studies. Lab. Anim. 46, 24–31. https://doi.org/10.1258/la.2011.011087.

Levine, S., 2005. Developmental determinants of sensitivity and resistance to stress. Psychoneuroendocrinology 30, 939–946. https://doi.org/10.1016/j.psyneuen.2005.03.013.

Levine, S., Lewis, G.W., 1959. Critical period for effects of infantile experience on maturation of stress response. Science 80 (129), 42–43. https://doi.org/10.1126/science.129.3340.42.

Liberzon, I., Abelson, J.L., 2016. Context processing and the neurobiology of post-traumatic stress disorder. Neuron 92, 14–30. https://doi.org/10.1016/j.neuron.2016.09.039.

Loi, M., Mossink, J.C.L., Meerhoff, G.F., Den Blaauwen, J.L., Lucassen, P.J., Joëls, M., 2017. Effects of early-life stress on cognitive function and hippocampal structure in female rodents. Neuroscience 342, 101–119. https://doi.org/10.1016/j.neuroscience.2015.08.024.

Masson, M., East-Richard, C., Cellard, C., 2016. A meta-analysis on the impact of psychiatric disorders and maltreatment on cognition. Neuropsychology 30, 143–156. https://doi.org/10.1037/neu0000228.

McCarthy, M.M., 2016. Multifaceted origins of sex differences in the brain. Philos. Trans. R. Soc. B Biol. Sci. 371. https://doi.org/10.1098/rstb.2015.0106.

Meaney, M.J., 2001. Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations. Annu. Rev. Neurosci. 24, 1161–1192. https://doi.org/10.1146/annurev.neuro.24.1.1161.

Mello, P.B., Benetti, F., Cammarota, M., Izquierdo, I., 2009. Physical exercise can reverse the deficit in fear memory induced by maternal deprivation. Neurobiol. Learn. Mem. 92, 364–369. https://doi.org/10.1016/j.nlm.2009.04.004.

Millan, M.J., Agid, Y., Brüne, M., Bullmore, E.T., Carter, C.S., Clayton, N.S., Connor, R., Davis, S., Deakin, B., Derubeis, R.J., Dubois, B., Geyer, M.A., Goodwin, G.M., Gorwood, P., Jay, T.M., Joëls, M., Mansuy, I.M., Meyer-Lindenberg, A., Murphy, D., Rolls, E., Saletu, B., Spedding, M., Sweeney, J., Whittington, M., Young, L.J., 2012. Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. Nat. Rev. Drug Discov. 11, 141–168. https://doi.org/10.1038/nrd3628.

Miller, A.B., Sheridan, M.A., Hanson, J.L., McLaughlin, K.A., Bates, J.E., Lansford, J.E., Pettit, G.S., Dodge, K.A., 2018. Dimensions of deprivation and threat, psychopathology, and potential mediators: a multi-year longitudinal analysis. J. Abnorm. Psychol. 127, 160–170. https://doi.org/10.1037/abn0000331.

Millstein, R.A., Holmes, A., 2007. Effects of repeated maternal separation on anxiety- and depression-related phenotypes in different mouse strains. Neurosci. Biobehav. Rev. 31, 3–17. https://doi.org/10.1016/j.neubiorev.2006.05.003.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Altman, D., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J.A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J.J., Devereaux, P.J., Dickersin, K., Egger, M., Ernst, E., Gøtzsche, P.C., Grimshaw, J., Guyatt, G., Higgins, J., Ioannidis, J.P.A., Kleijnen, J., Lang, T., Magrini, N., McNamee, D., Moja, L., Mulrow, C., Napoli, M., Oxman, A., Pham, B., Rennie, D., Sampson, M., Schulz, K.F., Shekelle, P.G., Tovey, D., Tugwell, P., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 6. https://doi.org/10.1371/journal.pmed.1000097.

Monfils, M.H., Holmes, E.A., 2018. Memory boundaries: opening a window inspired by reconsolidation to treat anxiety, trauma-related, and addiction disorders. The Lancet Psychiatry 0366. https://doi.org/10.1016/S2215-0366(18)30270-0.

Nelson, C.A., Zeanah, C.H., Fox, N.A., Marshall, P.J., Smyke, A.T., Guthrie, D., 2007. Cognitive recovery in socially deprived Young children: the Bucharest early intervention project. Science 318 (80), 1937–1940. https://doi.org/10.1126/science.1143921.

Oitzl, M., Workel, J., Fluttert, M., Frösch, F., Ronde Kloet, E., 2000. Maternal deprivation affects behaviour from youth to senescence: amplication of individual differences in spatial learning and memory in senescent Brown Norway rats. Eur. J. Neurosci. 12, 3771–3780. https://doi.org/10.1046/j.1460-9568.2000.00231.x.

Oomen, C.A., Girardi, C.E.N., Cahyadi, R., Verbeek, E.C., Krugers, H., Joëls, M., Lucassen, P.J., 2009. Opposite effects of early maternal deprivation on neurogenesis in male versus female rats. PLoS One 4. https://doi.org/10.1371/journal.pone.0003675.

Pechtel, P., Pizzagalli, D.A., 2011. Effects of early life stress on cognitive and affective function: an integrated review of human literature. Psychopharmacology (Berl.) 214, 55–70. https://doi.org/10.1007/s00213-010-2009-2.

Peña, C.J., Kronman, H.G., Walker, D.M., Cates, H.M., Rosemary, C., Purushothaman, I., Issler, O., Loh, Y.E., Leong, T., Kiraly, D., Goodman, E., Neve, R.L., Shen, L., Nestler, E.J., 2017. Early life stress confers lifelong stress scsceptibility in mice via ventral tegmental area OTX2. Science 1 (80), 1185–1188. https://doi.org/10.1126/science.aan4491.

Penke, Z., Felszeghy, K., Fernette, B., Sage, D., Nyakas, C., Burlet, A., 2001. Postnatal maternal deprivation produces long-lasting modifications of the stress response, feeding and stress-related behaviour in the rat. Eur. J. Neurosci. 14, 747–755. https://doi.org/10.1046/j.0953-816x.2001.01691.x.

Plotsky, P., Bradley, C., Anand, K., 2000. Behavioral and neuroendocrine consequences of neonatal stress. Pain Res. Clin. Manag. 10, 77–100.

Pryce, C.R., Bettschen, D., Nanz-Bahr, N.I., Feldon, J., 2003. Comparison of the effects of early handling and early deprivation on conditioned stimulus, context, and spatial learning and memory in adult rats. Behav. Neurosci. 117, 883–893. https://doi.org/10.1037/0735-7044.117.5.883.

R Core Team, 2015. R: a Language And Environment For Statistical Computing. R Found. Stat. Comput.

Rice, C.J., Sandman, C.A., Lenjavi, M.R., Baram, T.Z., 2008. A novel mouse model for acute and long-lasting consequences of early life stress. Endocrinology 149, 4892–4900. https://doi.org/10.1210/en.2008-0633.

Riecher-Rössler, A., 2017. Sex and gender differences in mental disorders. Lancet Psychiatry 4, 8–9. https://doi.org/10.1016/S2215-0366(16)30348-0.

Rosenthal, R., 1979. The file drawer problem and tolerance for null results. Psychol. Bull. 86, 638–641. https://doi.org/10.1037/0033-2909.86.3.638.

Rosenthal, R., 1991. Meta-Analytic Procedures for Social Research. Sage.

Singh-Taylor, A., Molet, J., Jiang, S., Korosi, A., Bolton, J.L., Noam, Y., Simeone, K., Cope, J., Chen, Y., Mortazavi, A., Baram, T.Z., 2018. NRSF-dependent epigenetic mechanisms contribute to programming of stress-sensitive neurons by neonatal experience, promoting resilience. Mol. Psychiatry 23, 648–657. https://doi.org/10.1038/mp.2016.240.

Suchecki, D., 2018. Maternal regulation of the infant's hypothalamic-pituitary-adrenal axis stress response: seymour 'Gig' levine's legacy to neuroendocrinology. J. Neuroendocrinol. 30, 1–17. https://doi.org/10.1111/jne.12610.

Suor, J.H., Sturge-Apple, M.L., Davies, P.T., Cicchetti, D., Manning, L.G., 2015. Tracing differential pathways of risk: associations among family adversity, cortisol, and cognitive functioning in childhood. Child Dev. 86, 1142–1158. https://doi.org/10.1111/cdev.12376.

Teicher, M.H., Samson, J.A., 2013. Childhood maltreatment and psychopathology: a case for ecophenotypic variants as clinically and neurobiologically distinct subtypes. Am. J. Psychiatry 170, 1114–1133. https://doi.org/10.1176/appi.ajp.2013.12070957.

Teicher, M.H., Samson, J.A., Anderson, C.M., Ohashi, K., 2016. The effects of childhood maltreatment on brain structure, function and connectivity. Nat. Rev. Neurosci. 17, 652–666. https://doi.org/10.1038/nrn.2016.111.

van Hasselt, F.N., Boudewijns, Z.S.R.M., Van Der Knaap, N.J.F., Krugers, H.J., Joëls, M., 2012. Maternal care received by individual pups correlates with adult CA1 dendritic morphology and synaptic plasticity in a sex-dependent manner. J. Neuroendocrinol. 24, 331–340. https://doi.org/10.1111/j.1365-2826.2011.02233.x.

van Lissa, C.J., 2018a. Metaforest: exploring Heterogeneity In Meta-Analysis Using Random Forests. PsyArXivhttps://doi.org/10.31234/osf.io/myg6s.

van Lissa, C.J., 2018b. Package 'Metaforest.' Cran.

Van Oers, H.Jj., De Kloet, E.R., Levine, S., 1998. Early vs. late maternal deprivation differentially alters the endocrine and hypothalamic responses to stress. Dev. Brain Res. 111, 245–252. https://doi.org/10.1016/S0165-3806(98)00143-6.

Vargas, T., Lam, P.H., Azis, M., Osborne, K.J., Lieberman, A., Mittal, V.A., 2018. Childhood trauma and neurocognition in adults with psychotic disorders: a systematic review and meta-analysis. Schizophr. Bull. 1–14. https://doi.org/10.1093/schbul/sby150.

Veenema, A.H., Bredewold, R., Neumann, I.D., 2007. Opposite effects of maternal separation on intermale and maternal aggression in C57BL/6 mice: link to hypothalamic vasopressin and oxytocin immunoreactivity. Psychoneuroendocrinology 32, 437–450. https://doi.org/10.1016/j.psyneuen.2007.02.008.

Vesterinen, H.M., Sena, E.S., Egan, K.J., Hirst, T.C., Churolov, L., Currie, G.L., Antonic, A., Howells, D.W., Macleod, M.R., 2014. Meta-analysis of data from animal studies: a practical guide. J. Neurosci. Methods 221, 92–102. https://doi.org/10.1016/j.jneumeth.2013.09.010.

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. J. Stat. Softw. 36, 1–48. https://doi.org/10.1103/PhysRevB.91.121108.

Viechtbauer, W., Cheung, M.W.-L., 2010. Outlier and influence diagnostics for meta-analysis. Res. Synth. Methods 1, 112–125. https://doi.org/10.1002/jrsm.11.

Walker, A.K., Nakamura, T., Byrne, R.J., Naicker, S., Tynan, R.J., Hunter, M., Hodgson, D.M., 2009. Neonatal lipopolysaccharide and adult stress exposure predisposes rats to anxiety-like behaviour and blunted corticosterone responses: implications for the double-hit hypothesis. Psychoneuroendocrinology 34, 1515–1525. https://doi.org/10.1016/j.psyneuen.2009.05.010.

Walker, C.-D., Bath, K.G., Joels, M., Korosi, A., Larauche, M., Lucassen, P.J., Morris, M.J., Raineki, C., Roth, T.L., Sullivan, R.M., Taché, Y., Baram, T.Z., 2017. Chronic early life stress induced by limited bedding and nesting (LBN) material in rodents: critical considerations of methodology, outcomes and translational potential. Stress 0, 1–28. https://doi.org/10.1080/10253890.2017.1343296.

Whipple, E.E., Webster-Stratton, C., 1991. The role of parental stress in physically abusive families. Child Abuse Negl. 15, 279–291. https://doi.org/10.1016/0145-2134(91)90072-L.

Wickham, H., François, R., Henry, L., Müller, K., RSudio, 2018. Package 'dplyr' Version 0.7.6.

Wigger, A., Neumann, I.D., 1999. Periodic maternal deprivation induces gender-dependent alterations in behavioral and neuroendocrine responses to emotional stress in adult rats. Physiol. Behav. 66, 293–302. https://doi.org/10.1016/S0031-9384(98)00300-X. LK = http://sfx.library.uu.nl/utrecht?sid = EMBASE&issn = 00319384&id = doi:10.1016%2FS0031-9384%2898%2900300-X&atitle = Periodic + maternal + deprivation + induces + gender-dependent + alterations + in + behavioral + and + neuroendocrine + responses + to + emotional + stress + in + adult + rats&stitle = Physiol. + Behav.&title = Physiology + and + Behavior&volume = 66&issue = 2&spage = 293&epage = 302&aulast = Wigger&aufirst = Alexandra&auinit = A.&aufull = Wigger + A.&coden = PHBHA&isbn = &pages = 293-302&date = 1999&auinit1 = A&auinitm = .

Witek-Janusek, L., 1988. Pituitary-adrenal response to bacterial endotoxin in developing rats. Am. J. Physiol. 255, E525–E530.