Tatum van Kampen
6444520
February 2, 2023
Utrecht University

ANALYSING ANXIETY

The Effects of SSRIs on Anxiety-like Behaviour: A Systematic Review & Meta-Analysis

Abstract

Anxiety disorders are among the largest contributors to the global health related burden. Despite its high occurrence, the cause of anxiety is poorly understood. The serotonergic system appears to play a big role in the pathophysiology, which has led to the development of Selective Serotonin Reuptake Inhibitors (SSRIs) and their implementation in treatment strategies. However, the specific aspects of anxiety that are modulated by SSRIs are not yet established. Since anxiety disorders can exist on the basis of unconditioned fear, assessing anxiety in ethological paradigms could be advantageous in the endeavour of obtaining a complete understanding of the mechanism behind SSRIs. A well-rounded overview of different aspects of unconditioned anxiety can be presented through the evaluation of four anxiety tests: the Elevated Plus Maze Test (EPM), Marble Burying Test (MB), Ultrasonic Vocalisations Test (USV) and Stress-Induced Hyperthermia test (SIH). The aim of this systematic review and meta-analysis is to determine the effects of the six clinically relevant SSRIs on anxiety-like behaviour in animal studies using the EPM, MB, USV and SIH. A systematic search was conducted in PubMed and EMBASE for each anxiety test, yielding a total of 186 publications meeting inclusion criteria, of which 178 were eligible for meta-analysis. Descriptive analysis showed a great majority studying male rats and mice, with either acute or chronic SSRI treatment. Quantitative analysis revealed a significant decrease in anxiety-like behaviour in the EPM, MB and USV after SSRI administration. Additional moderator analysis through Bayesian Penalized Meta-Regression showed significant intercepts, and in the case of the MB, significant moderating influences of Human Equivalent Dose and sex. Analysis of the SIH did not yield any significant results. It can be concluded that SSRIs are effective in the treatment of various aspects of anxiety, namely approach-avoidance imbalance, obsessive compulsive tendencies and stress-induced distress calls. Furthermore, this SRMA provides insights in the optimal experimental set-up for animal research of putative anxiolytic drugs. However, additional assessment of the SIH is required to better comprehend the relation between the serotonergic system and the autonomic nervous system. Additionally, further research is needed to evaluate anxiety as it is frequently presented in the clinical setting, to gain a better understanding of interpersonal treatment responses.

1 | Introduction

An overwhelming sensation of nervousness, a churning stomach, jitters throughout the body topped with a racing heart: anxiety. A perfectly healthy and normal emotion, provided that the severity and duration remain within appropriate levels. Exorbitant experiences of anxiety can be indicative of a mental health disorder. Within the mental health disorders, anxiety disorders are among the most common, which can present themselves in different forms, primarily General Anxiety Disorder (GAD), Post-Traumatic Stress Disorder (PTSD), Obsessive Compulsive Disorder (OCD), Panic Disorder and Social Phobia. [1] With an estimated 4800 out of 100.000 people being affected by anxiety disorders worldwide annually, it is vital to gain a complete understanding of their pathophysiology and pharmacological treatment, in pursuance of an improved global mental health status, especially when considering the consequences of the COVID-19 pandemic adding to the global health related burden. [2]

While the exact pathophysiology of anxiety disorders remains a mystery, several mechanisms have been proposed. The leading theory revolves around disrupted modulation in the central nervous system. [3] Two neurotransmitter systems are considered as the most probable to play a role in the modulatory steps ultimately resulting in anxiety: the serotonergic and noradrenergic systems. [4] The common ground within these systems is that they both project to limbic structures, dysregulation of which is deemed paramount in the development of anxiety. [5] In a malfunctioning state, both neurotransmitter systems have an opposite reaction: the noradrenergic system becomes hyper-responsive to stress and fear stimuli, whereas the serotonergic system becomes hypo-

responsive, resulting in decreased stress reactivity, as well as decreased tolerance to aversion. [6] Both these modified responses can result in anxiety, or anxiety-like symptoms. Therapeutic action alleviating anxiety can thus be achieved by intervening in either neurotransmitter system. This is where SSRIs enter the picture.

Originally created to treat depression, Selective Serotonin Reuptake Inhibitors (SSRIs) are a class of drugs that have since evolved to being effective treatment for various mental illnesses, including anxiety disorders, to the point that they have become standard pharmacological therapy for anxiety disorders. [7,8] Currently, there are six clinically relevant SSRIs on the market that share the common property of serotonin reuptake inhibition: fluoxetine, citalopram, escitalopram, sertraline, paroxetine and fluvoxamine. [9] This potent and selective inhibition is achieved through the serotonin transporter (SERT). As opposed to the classically emphasized role of presynaptic axon terminals, action at the somatodendritic end appears to explain the therapeutic effects of SSRIs. As explained by the monoamine hypothesis, risen serotonin (5-HT) levels in the somatodendritic area, as a result of SSRI administration causing SERT blockage, stimulate nearby 5-HT_{1A} autoreceptors. This results in desensitisation and downregulation of these receptors, causing an inability for 5-HT to inhibit its own release. The subsequent rise of 5-HT levels desensitizes postsynaptic 5-HT receptors as well, a key step in the mechanism of action. Without opportunity for reuptake, 5-HT levels keep on rising in the synapse. Keeping in mind the hyporesponsive state of the serotonergic system in anxiety, the heightened 5-HT levels are the key in its treatment, since SSRI usage compensates for this impaired system. [9] While the primary mechanism of action for each of the SSRIs is identical, secondary pharmacological characteristics differentiate these agents. The distinction can be made through these unique properties: fluoxetine has 5-HT_{2c} antagonistic properties, citalopram is comprised of two enantiomers, with the R enantiomer being responsible for antihistaminic properties, while escitalopram only contains the active S enantiomer, sertraline has dopamine transporter (DAT) antagonistic properties as well as the possibility to bind to the sigma-1 receptor, paroxetine exerts both muscarinic anticholinergic and noradrenergic inhibitory action, and fluvoxamine has sigma-1 receptor binding properties, more potent than sertraline. This diversity may be accountable for different selectivity, potency and tolerability profiles. [8–10]

Despite being theoretically and clinically effective, ambiguities surrounding SSRIs remain. The specific aspects of anxiety that are modulated by SSRIs are not yet established. Insights into these unknowns can be gained through animal models. [11] Such models can be used to resolve unanswered questions, some of which impossible to answer via human studies, for instance regarding behavioural mechanisms. [12,13] Generally, two types of models can be distinguished: models provoking conditioned responses to stressful and often painful stimuli, and ethologically based paradigms, where a natural response to milder stressful situations is observed. [13] Different paradigms evaluate different types of anxiety and can thus be translated to various anxiety disorders present in humans. The unconditioned ethological paradigms in particular are considered to be qualitative analogues of human anxiety-like behaviour. [14]

One well-established behavioural paradigm is the Elevated Plus Maze Test (EPM). [15] This test entails an apparatus in the shape of a plus sign (+), with the four arms radiated from an elevated central platform. Two opposing arms are open, the other two arms are enclosed by walls. [15,16] Observed

behaviour in this apparatus is based on internal conflict, namely natural aversion of rodents to open spaces, counterbalanced with a natural tendency to explore new spaces. This is also known as the approach-avoidance conflict. During the EPM, this will present itself in an innate preference for the closed arms. The higher the level of anxiety, the greater the aversion of the open arms, possibly indicative of a dysfunctioning approach-avoidance system. Hence, measuring the time and entries in either arm accurately represents an animals' anxiety related to the approach-avoidance conflict. Additionally, effects of drugs possibly exerting influence on this phenomenon can be evaluated through this paradigm. [12,13,17–19] Since human anxiety can present itself with an approach-avoidance component as well, for example in an aversion of leaving one's house in fear of the dangers of the outside world, conclusions from animal drug studies on the EPM can be translated into implications for human treatment. [20]

Still, dysregulated approach-avoidance behaviour is not the only form in which anxiety can present itself. Other systems can become defective as well, among which the normal behavioural routine of rodents to bury. Studies show that burying and digging are expressed under both non-anxiogenic circumstances, such as nesting, as well as in anxiogenic scenarios, for instance when confronted by predators. [21] However, when this healthy behaviour turns aberrant, it may represent compulsive tendencies. Compulsive behaviour is a characteristic of multiple behavioural disorders and commonly seen in Obsessive Compulsive Disorder. To study this type of anxiety, several animal models have been developed, including the Marble Burying Test (MB). [21–23] Like the EPM, the MB is an ethologically based animal model. [24] During this test, an animal, either with or without pharmacological treatment, is placed in a clear cage bedded with sawdust, upon which glass marbles are presented. After a certain amount of time in the cage, the animal is removed and the number of marbles buried (at least for two-thirds) is observed. The amount of marbles buried positively correlates with the level of compulsion-based anxiety. With the MB being a validated model for this type of anxiety, it can provide insights and implications for human therapy. [25,26]

In contrast to the EPM and MB, which are both paradigms to evaluate a state of general anxiety, other ethological paradigms have put their focus on stimulus-based anxiety. [27] When exploring anxiety, stimuli commonly used to induce anxiety are separation, stress and foot-shocks. After exposure to a specific stimulus, the level of anxiety can be gauged through the evaluation of ultrasonic vocalisations (22 kHz). This is known as the Ultrasonic Vocalisations Test (USV). Rodents produce ultrasonic vocalisations as a sign of distress and this can thus be used as an indicator of anxiety level. An advantage of the USV compared to the aforementioned tests is that it is not dependent on conflict, exploratory tendencies and motor skills. [27–30] Given that pups are completely dependent on their mothers, as they are deaf, blind, poikilothermic and without fur in the first weeks of their lives, USV as a distress call effectively examines unconditioned anxiety. [31]

In addition to physical outings, anxiety can induce physiological responses, by affecting the autonomic nervous system. [32] One process sensitive to anxiety is body temperature. Rise of body temperature is induced by stressful stimuli and is present in a variety of species, humans included. [33] This phenomenon became the subject of another animal model: the Stress-Induced Hyperthermia Test (SIH). A typical SIH test involves rectal temperature measurement, which simultaneously serves as stressor, to assess baseline body temperature. After some time, temperature is measured again, thereby evaluating the stressors' effects. Additionally, pharmacological components can be administered in advance, in order to review

their effects on the rising body temperature. Since this stress response is objective and similar in rodents and humans, findings can be easily translated into implications for human treatment. [32–34]

With a variety of animal models representative of different types of anxiety and six clinically relevant SSRIs influencing anxiety-like behaviour, it is inevitable to lose sight of the wood for the trees. The aim of this systematic review and meta-analysis is to determine the effects of the six clinically relevant SSRIs on anxiety-like behaviour in animal studies using the Elevated Plus Maze Test, Marble Burying Test, Ultrasonic Vocalisations Test and Stress-Induced Hyperthermia Test. This information can contribute to a better understanding of the similarities, but more importantly, the differences between the SSRIs. Integration of various anxiety tests into this overview add to a comprehension of the influence of SSRIs on the specific aspects of anxiety, as well as provide more insights into the role of the serotonergic system in anxiety and anxiety-like behaviour. Furthermore, by determining the sensitivity of different anxiety tests for the SSRIs, whose clinical effectiveness has been proven, these findings can contribute to the optimisation of test set-ups when determining the putative anxiolytic activity of other drugs, thereby reducing unnecessary animal testing. Together, these findings may result in more informed decisions regarding the effects of different SSRIs in a particular type of anxiety-like disorder.

2 | Methods

2.1 | Study Protocol

The review protocol was formulated in accordance with SYRCLE's Systematic Review Protocol for Animal Intervention Studies and has been registered in the International Prospective Register of Systematic Reviews (PROSPERO) under registration number CRD42022371871 as of December 16, 2022. Reporting of the review occurred on the basis of the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) statement. [35]

2.2 | Literature Search

A systematic search was conducted in PubMed and EMBASE on September 12, 2022. The search strategy is based on components regarding pharmacological treatment and the tested paradigm of anxiety, which comes down to SSRIs and the four anxiety tests (EPM, MB, USV and SIH). Where possible, both regulated terms (i.e. MeSH and Emtree terms) and text words were used. Since not all anxiety tests were regulated in controlled terms, terms with probable connection to the respective model were used in order to obtain as many potentially relevant articles as possible. Along with this lack of controlled terms comes the need to use an animal filter, as the used search terms were not restricted to animal models. [36,37] Ultimately, four searches were conducted, all of which including the following concepts: fluoxetine, citalopram, escitalopram, sertraline, paroxetine and fluvoxamine. For each of the searches, these concepts were combined with one of the anxiety tests of interest: Elevated Plus Maze, Marble Burying, Ultrasonic Vocalisations and Stress-Induced Hyperthermia. A complete list of the components of the search string can be found in supplementary file S1.

2.3 | Study Selection

During the selection process and the subsequent data extraction, search results remained separated by anxiety test, ultimately resulting in four independent data sets. Study selection consisted of two screening phases. After removal of duplicates, two reviewers independently screened all identified articles for eligibility. In both screening phases, discrepancies in decisions were resolved through discussion, or when needed, the consultation of a third reviewer. The first screening phase revolved around the evaluation of title and abstract. The second screening entailed full-text assessment. Eligibility was determined based on the

Table 1: Inclusion criteria

Category	Inclusion criteria
Type of study	Originally peer reviewed published studies
	Controlled studies
Type of animals	All non-human animals
Type of intervention	One of six clinically effective SSRIs
	All dosing schedules
Outcome measures	Any outcome parameter assessing the level of anxiety-like behaviour in the respective anxiety test (EPM, MB, USV or SIH)
Language	English

inclusion and exclusion criteria listed in Table 1 and Table 2. Of the exclusion criteria, criteria 6 through 15 were only utilized in the second phase of screening (full text). Animal studies often make use of behavioural test batteries, but these batteries can potentially exert a significant influence on test outcomes, especially when examining stress-induced behaviour. [38,39] Therefore, such articles were included on one condition: as the aim was for the animals to be as naïve as possible, only studies conducting one of the defined anxiety tests as their first or only assay were included. Subsequently, when the test procedure contained a pretest of said anxiety test, the study was only included, if this pretest was used as a selection process.

2.4 | Data Extraction

2.4.1 | Extraction of study characteristics

Eligible studies were subjected to data extraction. Study characteristics were extracted by one reviewer, with a second reviewer being available in the case of uncertainties, while also fulfilling a supervisory role by checking the extracted data. General characteristics regarding the test procedures were extracted and, dependent on the performed test, additional test-specific information was extracted, both of which can be found in Table 3. For all tests, the Human Equivalent Dose was calculated from the reported dose using conversion factors. [40]

Table 2: Exclusion criteria, sorted by priority. Criteria 1-5 determined eligibility in phase one (title abstract screening); all criteria were implemented in phase two (full text screening).

	Exclusion criteria
1	Not an original full publication
2	Not an in vivo animal study, but a human, in vitro or ex vivo study
3	No EPM, MB, USV or SIH used, or test not used to measure anxiety
4	No SSRI treatment used
5	SSRI not given directly to subject
6	No appropriate placebo or vehicle-controlled experiment
7	No information available / retrievable within one reference on the protocol of the anxiety test
8	No information available / retrievable on specific SSRI used
9	No results reported from the used anxiety test
10	Use of an additional pharmacological treatment before / during / after SSRI treatment and before the anxiety test
11	SSRI treatment not tested on one of the four defined anxiety tests
12	Studies on one of the four defined anxiety tests where the SSRI was not given up until or within 24 hours of measuring anxiety behaviour
13	Animals are tested in other behavioural tests before tested in the anxiety test of interest
14	Pretesting of the anxiety test is not part of selection within the test procedure
15	Full article text not retrievable or not in the English language

Table 3: Extracted characteristics

General Characteristics	
Study Identification	Title
	All authors
	Year
Animal model	Housing
	Time of test (active or passive)
	Day / night schedule
	Species
	Strain
	Age
	Bodyweight
	Sex
	Disease induction
Intervention	Type of SSRI
	Dose
	Frequency (acute, subchronic or chronic)*
	Route of administration
	Timing of administration relative to disease induction
Anxiety test procedure	Test duration
	Details of pretest (when applicable)
Test-specific characteristics	
Marble Burying	Total number of marbles
Ultrasonic Vocalisations	Type of USV
	Foot-shock delivery protocol (when applicable)
Stress-Induced Hyperthermia	Type of SIH (group or individual)
	Type of temperature measurement
	Time interval between first and second temperature measurement

**Acute treatment was defined as 1 day of administration, subchronic treatment as 2-6 administration days and chronic treatment as 7 or more administration days.*

2.4.2 | Extraction of outcome measures

With the four anxiety tests each examining another aspect of anxiety, outcome measures differ slightly between tests:

- Elevated Plus Maze: any type of behaviour assessing anxiety;
- Marble Burying: number of marbles buried;
- Ultrasonic Vocalisations: number and duration of vocalisations;
- Stress-Induced Hyperthermia: baseline body temperature, mean difference body temperature.

For each of the respective reported outcome measures, mean (mean difference for the SIH), standard error of the mean (SEM) or standard deviation (SD) and number of animals was extracted for the intervention group, as well as for the control group. Additionally, for the MB and USV, general test results regarding locomotion were extracted. Data extraction occurred from text, tables and graphs. When the data was only displayed graphically, a digital ruler was used to extract the data. In the case of missing data, authors were contacted requesting the needed information. If the information could not be retrieved, the corresponding study was excluded from quantitative analysis. When the number of animals was presented as a range, the largest value was used to calculate SD. Outcome data presented in different units than mean, SED or SD were converted accordingly using necessary formulas, assuming a normal distribution of data.

2.5 | Data Analysis

All code for the analysis was preregistered, with the Preregistration-As-Code using the Workflow for Open Reproducible Code in Science [41], and is available at *https://github.com/cjvanlissa/meta_anx_ssri.git*.

Meta-analysis was carried out in R (R Core Team) [42], using R-packages *metafor* [43] and *pema* [44]. The random effects model was

applied, since variation in true effect size was expected. Each anxiety test was analysed separately. Effect sizes were obtained through three-level meta-analysis, presented as Hedges' g with 95% Confidence Interval (CI) and visualised in forest plots. [45] In the case of a reported sample size range, the lowest value was used in the calculation. Individual effect sizes were pooled to realise an overall Hedges' g and 95% CI. Heterogeneity was assessed using I^2 values.

Biological duplicates were avoided by only incorporating one outcome variable per animal. This was only the case for the EPM and the USV, where multiple types of behaviour can be observed during the test period. Systematic selection of samples was achieved through sorting outcome measures by priority. Priority order was determined on the basis of perceived relevance in the manifestation of anxiety. Regarding the EPM, the following order was established: (1) entries in open arms as a percentage of total entries (%EOA), (2) time spent in open arms as a percentages of total time spent in any of the arms (%TOA), (3) absolute number of entries in open arms (EOA) and (4) absolute amount of time spent in open arms (TOA). Articles not reporting any of these outcomes were individually evaluated, to include their most relevant outcome measure in the analysis. Two types of outcomes can be reported in the USV: duration of vocalisations and number of vocalisations. The latter was deemed most relevant and included in the meta-analysis, in the case of both outcomes being reported. Multiple use of a control group was corrected for by dividing the number of animals in the control group by the number of experiments using these animals.

To allow for more detailed analysis, six categorical moderators were established: type of SSRI, treatment frequency, disease induction, species, sex and use of pretest. For both the USV and the SIH, one additional moderator was included: type of test. Within these categorical moderators, dummy variables were formulated. In order to be able to draw more well-rounded conclusions, dummy variables were classified into groups containing comparable dummies when necessary. For instance, this was the case for disease induction, where a great variety of stress models were implemented across study protocols, which were combined into one overarching dummy variable: 'stress'. The defined dummy variables are presented in Table 4. Supplementary to the categorical moderators, one continuous moderator was included in the analysis: human equivalent dose (HED). HED was chosen over the original reported dose to allow for a uniform comparison, as doses are equalised across species.

Table 4: Dummy variables within moderator categories

Categorical moderator	Dummy variables
Type of SSRI	Fluoxetine
	Citalopram
	Escitalopram
	Sertraline
	Paroxetine
	Fluvoxamine
Treatment frequency	Acute
	Subchronic
	Chronic
Disease induction	Healthy
	Stress
	Other
Species	Rat
	Mouse
	Gerbil
	Frog
	Hamster
Sex	Male
	Female
	Both
	NR
Use of pretest	Yes
	No
Type of USV	Separation-induced
	Physical stress induced
Type of SIH	Individual
	Group

Since the number of moderators is relatively high compared to the number of studies, there was a risk for overfitting and model non-identification. Overcoming this problem often present in a classical meta-regression requires a technique performing variable selection. The conduction of a Bayesian Penalized Meta-regression (BRMA) was preferred over a classical meta-regression, since it identifies which moderators are important in predicting the effect size. BRMA shrinks all regression coefficients to zero, causing the need for these coefficients to overwhelmingly differ from the prior before significantly differing from zero. This type of analysis thus aids empirical model identification and allows for a greater certainty in drawing conclusions on a population level.

The BRMA requires choosing a reference variable to examine the effect of a categorical variable. The other dummy variables then encode the difference between said dummy variable and the reference variable. Results of the BRMA are presented as an intercept and the effect of dummy variables. The intercept depicts the effect size of a study falling within the reference category in each categorical variable. Effects of dummy variables represent the differences between this variable and the reference category; a significant effect displays that the dummy variable's mean significantly differs from that of the reference category, while variables in all other categories remain equal. Reference categories for each categorical variable are presented in Table 5. The reference category was chosen based on the greatest expected effect size of the variable and the highest quantity of available information on the variable.

It is important to note that the direction of effect size differs between anxiety tests. For the MB, USV and SIH, an anxiolytic effect of SSRIs is manifested as a reduction in the behaviour of interest (e.g. less marbles buried in the MB). For these three tests, an anxiolytic effect is represented by a positive effect size. The opposite holds true for the EPM. It is expected for an anxiolytic drug to tip the approach-avoidance balance towards exploratory behaviour, resulting in an increase in the behaviour of interest. Therefore, the effect size will be negative in the case of an observed anxiolytic effect.

2.5.1 | Risk of Bias

Presence of bias was assessed through the use of SYRCLE's risk of bias tool [46], with the addition of two reporting questions related to randomisation and blinding, and ultimately visualised into graphs. Two reviewers determined the risk of bias. Discrepancies were resolved either through discussion or consultation with a third reviewer. Risk of bias was scored as either 'low' or 'high': 'low' when sufficient measures were taken to minimalize or prevent risk of bias, and 'high' when such measures were not taken. Items of the risk of bias tool were marked as 'unclear' when the

Table 5: Reference categories for categorical variables per anxiety test

	Reference category			
	EPM	**MB**	**USV**	**SIH**
Type of SSRI	Fluoxetine	Fluoxetine	Fluoxetine	Fluoxetine
Treatment frequency	Chronic	Chronic	Chronic	Chronic
Disease induction	Stress	Stress	Healthy	Stress
Species	Rat	Mouse	Rat	Mouse
Sex	Male	Male	Both	Male
Use of pretest	No	No	No	No
Type of USV	-	-	Separation-induced	-
Type of SIH	-	-	-	Individual

provided information was insufficient to determine risk of bias. Assessment of the item 'selective outcome reporting' (reporting bias) was twofold: checking whether reported outcome measures lined up between the method and result sections (1), and reviewing the Animal Study Registry (German Centre for the Protection of Laboratory Animals, 2019), as well as Preclinicaltrials (Preclinicaltrials.eu, 2019), to examine the preregistration of the included studies (2).

3 | Results

3.1 | Study Characteristics

3.1.1 | Study Selection

Systematic searches for all four anxiety tests were conducted in PubMed and EMBASE on September 12, 2022.

ELEVATED PLUS MAZE

The selection process for the EPM is depicted in Figure 1. The systematic search yielded a total of 1293 results, with 906 unique articles remaining after removal of duplicates. These articles were screened based on their title and abstract, resulting in 684 articles eligible for full-text assessment. Based on the reasons in Figure 1, 579 articles were excluded during this second screening phase. The final selection consisted of 105 articles included in the systematic review, adding up to a total of 241 experiments. Out of these 105 articles, 100 articles were eligible for quantitative analysis.

MARBLE BURYING

A total of 853 articles resulted from the systematic search involving the MB. As displayed in Figure 2, this came down to 627 unique articles. During title and abstract screening, 461 articles failed to meet inclusion criteria. Subsequent full-text evaluation led to the inclusion of 63 out of the remaining 166 articles. These 63 articles consisted of 189 experiments. Two articles were excluded from the meta-analysis, adding up to 61 included articles.

ULTRASONIC VOCALISATIONS

Figure 3 shows the selection process for the USV, which started with 400 search results and 251 unique articles after removal of duplicates. Title and abstract screening resulted in 61 articles eligible for the second screening phase (full text). Assessment of these articles led to exclusion of 50 articles not meeting the inclusion criteria. Eleven articles, with a total of 52 experiments, were included in the systematic review, all of which were eligible for quantitative analysis.

STRESS-INDUCED HYPERTHERMIA

Systematic search of the SIH generated 86 results, as presented in Figure 4. Removal of duplicates resulted in 53 unique articles, of which 31 were excluded for not meeting the inclusion criteria of the title and abstract screening. The remaining 22 articles were evaluated based on their full text. Following the exclusion criteria, 15 articles were excluded from the review, ultimately leading to the inclusion of 7 articles and 15 experiments. One of the 7 articles lacked information regarding the number of animals used and was thus excluded from the meta-analysis.

3.1.2 | Descriptive analysis & meta-analysis per anxiety test

All extracted characteristics of the included studies from each of the four anxiety tests are listed in Supplementary File S2. An overview of the pooled effect size of all four anxiety tests is presented in Table 6.

Table 6: Overall effect sizes for each anxiety test

	Overall effect size (Hedges' g (95% CI))	
EPM	- 0.45	(-0.66, -0.25)
MB	1.88	(1.64, 2.12)
USV	1.04	(0.58, 1.49)
SIH	0.68	(-0.08, 1.44)

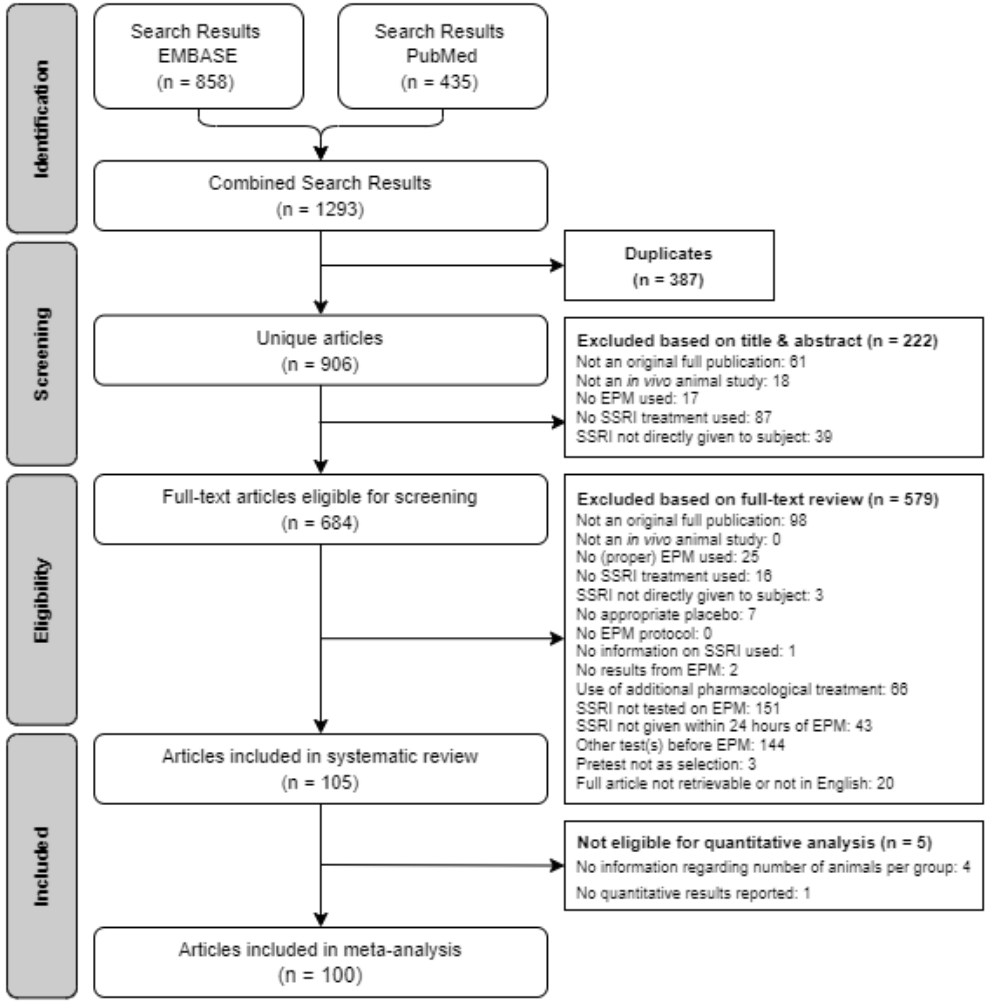

Figure 1: Flowchart of article selection, EPM

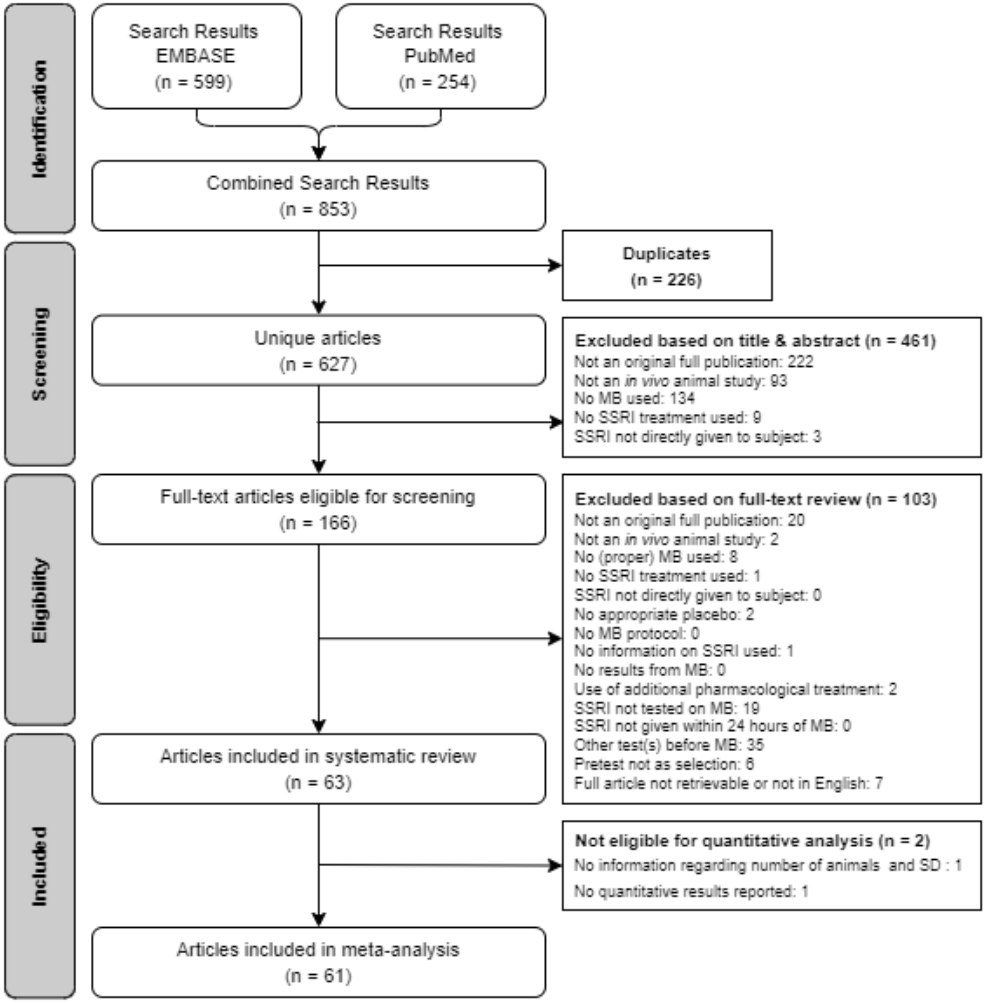

Figure 2: Flowchart of article selection, MB

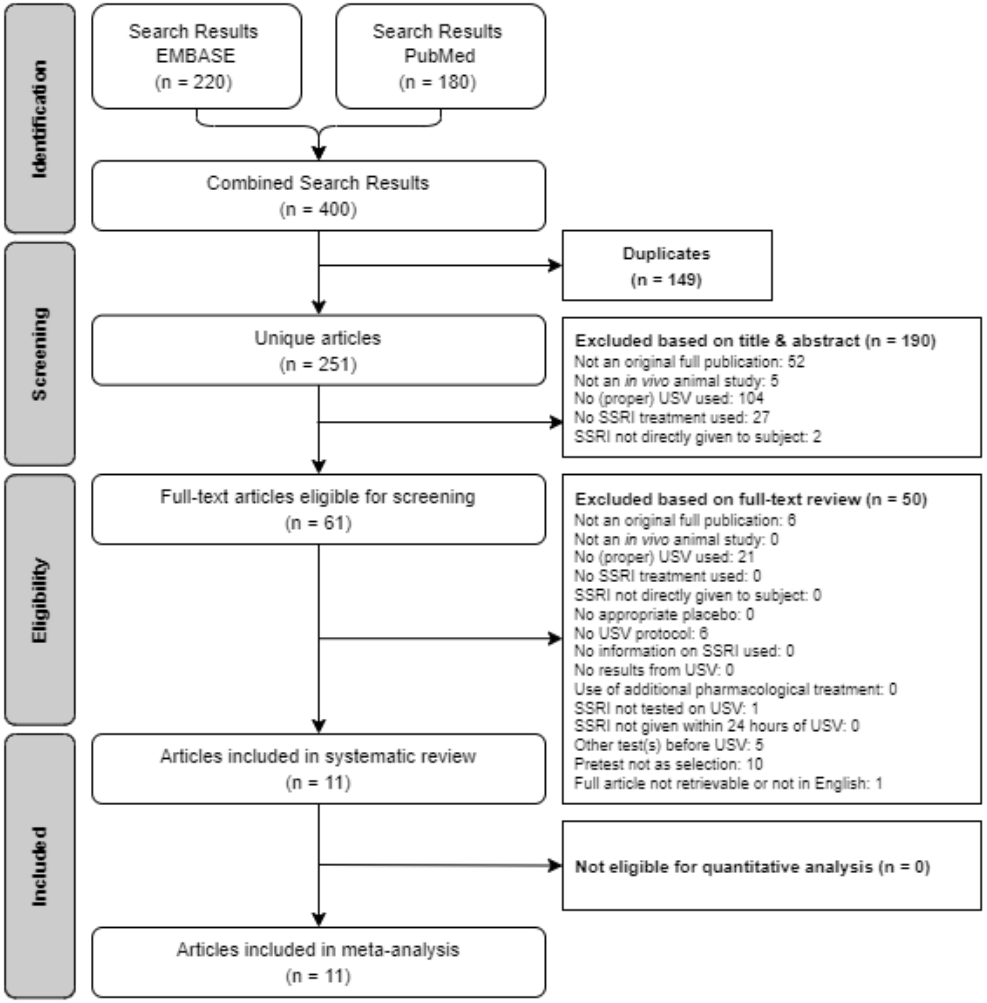

Figure 3: Flowchart of article selection, USV

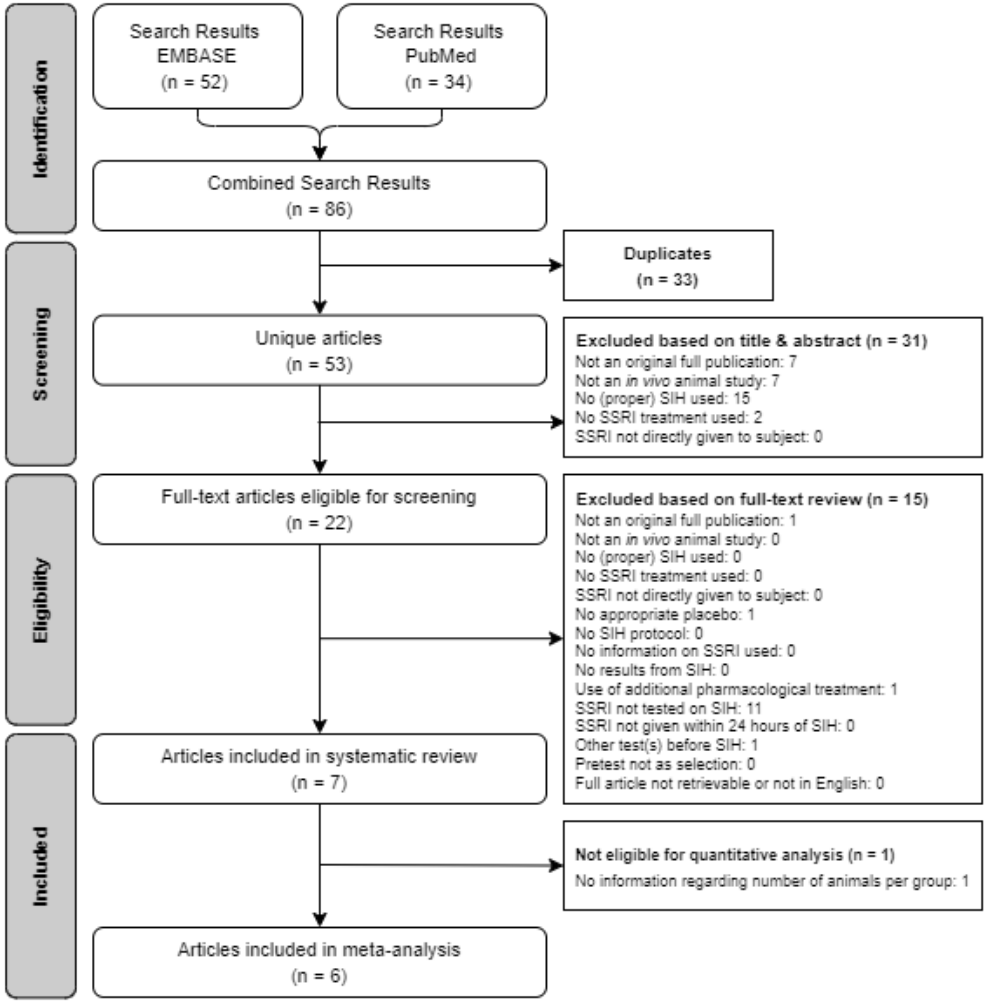

Figure 4: Flowchart of article selection, SIH

ELEVATED PLUS MAZE

Description of included studies

Rodents were the subject of all articles performing the EPM but one, which used frogs. Most articles used rats (n = 59) or mice (n = 43). A single article reported on hamsters and gerbils each. The sex of these animals was to the utmost extent male, namely 84 articles. Furthermore, 8 articles reported on females, 10 on mixed groups of males and females, and 4 failed to document the sex of their subjects. The state of these animals was largely healthy (n = 71). Fifty-four articles additionally or exclusively reported on diseased animals. Various methods of disease induction were applied, including stress models (n = 41), such as Chronic (Unpredictable) Mild Stress (n = 10), Restraint Stress (n = 5) and Predator (Scent) Stress (n = 6), but also genetic modification (n = 5) and STZ-induced diabetes (n = 3). All SSRIs were evaluated across the studies, fluoxetine being the most common with 60 articles (rats: 1-30 mg/kg, mice: 1-30 mg/kg, gerbils: 1-30 mg/kg, frogs: 5-20 mg/kg), followed by paroxetine (n = 16, rats: 0.1-17 mg/kg, mice: 5-20 mg/kg, gerbils: 0.3-30 mg/kg) and citalopram (n = 13, rats: 1-17 mg/kg, mice: 5-10 mg/kg, hamster: 10 mg/kg). Several studies conducted experiments using escitalopram (n = 9, rats: 0.5-27.5 mg/kg, mice: 10-20 mg/kg), fluvoxamine (n = 8, rats: 1-30 mg/kg, mice: 1-20 mg/kg) and sertraline (n = 7, rats: 5-40 mg/kg, mice: 5-10 mg/kg), and SSRI administration occurred mostly orally (n = 28) or through intraperitoneal injection (n = 56). Frequency of administration was primarily acute (n = 48) and chronic (n = 59). Just 7 articles studied subchronic administration.

Meta-analysis

Out of 105 articles included in the systematic review, 100 were eligible for quantitative analysis. The 5 articles excluded from the meta-analysis either lacked information regarding the number of animals in each group (n = 4), or reported no quantitative results at all (n = 1). The 100 included articles combined to a total of 218 experiments. The forest plot depicting overall effect size is shown in Supplementary File S3. SSRIs significantly reduced anxiety-like behaviour in the EPM (Hedges' g: -0.45; 95% CI: -0.66, -0.25; p=0.00; I^2 = 70%; k=218). This effect size was established as a result of 39 experiments showing an anxiolytic effect of SSRIs, 22 experiments showing an anxiogenic effect, and 157 experiments finding no significant effect. However, the substantial between-studies heterogeneity, represented by an I^2 value of 70%, indicates that these pooled results should be interpreted with caution. To evaluate the contribution of various moderators to the effect size and possibly explain the high heterogeneity, BRMA was conducted. The intercept of the BRMA model was significant (SMD: -0.42; 95% CI: -0.80, -0.01), which demonstrates that SSRIs indeed have anxiolytic effects, on the condition that the set-up of the experiment is in compliance with the reference variables (fluoxetine, chronic administration, stressed animals, rats, male, not pretested). No other dummy variables significantly moderated the effect size.

MARBLE BURYING

Description of included studies

The Marble Burying Test was exclusively conducted on mice, the vast majority of which being male, with 54 articles. Three articles studied females or mixed groups each. Two articles did not report on the animals' sex and one used gender-matched groups. Out of a total 63 articles, 60 were conducted on healthy subjects. Seven articles additionally of exclusively reported on diseased animals, using a variety of disease induction models. Over half of the articles studied fluoxetine (n = 37; 0.16-160 mg/kg). Fifteen articles reported on fluvoxamine (0.16-60 mg/kg), similarly to the 13 articles studying paroxetine (0.1-40 mg/kg). Citalopram was occasionally used (n = 10; 0.3-60 mg/kg), whereas sertraline (10 mg/kg) and

escitalopram (1-5 mg/kg) were only studied in 2 articles each. Route of administration was limited to oral, intraperitoneal and subcutaneous administration, intraperitoneal administration being the most common, with 42 articles. In 55 articles, acute administration of the SSRI was employed. Chronic and subchronic administration was utilized in 11 and 3 articles, respectively. Overall, study characteristics of MB studies were relatively homogenous, with the most variety being present in the type of SSRI used.

Meta-analysis

Sixty-one of the 63 included articles were eligible for meta-analysis. One of the two remaining articles failed to provide information regarding variability and the number of animals, the other did not report quantitative results at all, and thus both were excluded from quantitative analysis. Effect sizes of the 183 included experiments, as well as a pooled effect size, are displayed in the forest plot in Supplementary File S4. A single experiment found an anxiogenic effect of SSRIs, 126 experiments an anxiolytc effect, and 56 experiments observed no effect. Overall, SSRI treated animals showed significantly less anxiety-like behaviour compared to control animals (Hedges' g: 1.88; 95% CI: 1.62, 2.12; p=0.00, I^2=15%, k=183). As was anticipated following descriptive analysis, between-studies heterogeneity was low. Analysis through the BRMA model showed a significantly positive effect of SSRIs on the reduction of anxiety-like behaviour (SMD: 1.41; 95% CI: 0.43, 2.46), when all reference variables were present in the study design (fluoxetine, chronic administration, stressed animals, mice, male, not pretested). Two variables significantly moderated effect size. Firstly, the human equivalent dose was correlated positively to the effect size (SMD: 0.37; 95% CI: 0.22, 0.52). Secondly, in the categorical variable sex, articles not reporting the sex of their animals showed a greater anxiolytic effect than articles studying males, which was the reference variable (SMD: 4.25; 95% CI: 0.30, 7.76).

ULTRASONIC VOCALISATIONS

Description of included studies

In contrast to the MB, USV experiments were exclusively conducted on rats rather than mice, with the exception of 1 article. The sex of these animals leaned mostly towards mixed groups with a slight majority of 5 articles. Three articles reported on males. Additionally, 2 articles neglected to disclose the sex of the subjects, and the last article reported on females. All of the animals were healthy, except for one article also studying genetically modified organisms. The type of USV test was divided into two categories: separation-induced USV (n = 8) and physical stress induced USV (n = 3). The latter category was comprised of a shock-induced USV (n = 2) and a restraint stress-induced USV (n = 1). With several articles evaluating multiple SSRIs, treatment with every SSRI was covered. The SSRIs studied most were fluoxetine (n = 7, rats: 1-30 mg/kg) and paroxetine (n = 3, rats: 0.01-10 mg/kg, mice: 0.03 mg/kg). Two articles reported on citalopram (rats: 0.3-30 mg/kg). Fluvoxamine (rats: 0.3-3 mg/kg) and escitalopram (rats: 0.3-10 mg/kg) were evaluated in a single article each. SSRIs were administered either orally, intraperitoneally or subcutaneously, with intraperitoneal injection being the primary route (n = 6). SSRIs were acutely administered, with the exception of 2 articles, which evaluated subchronic administration.

Meta-analysis

All 11 articles included in the systematic review were eligible for inclusion in the meta-analysis. As shown in the forest plot in Supplementary File S5, SSRIs exerted a significant anxiolytic effect (Hedges' g: 1.04 , 95% CI: 0.58, 1.49, p=0.00, I^2=57%, k=52). All experiments reported either an anxiolytic effect (k=17), or no effect at

all (k=35). The intercept in the BRMA was significant (SMD: 1.03, 95% CI: 0.16, 1.89), demonstrating that SSRIs significantly reduce anxiety-like behaviour when evaluated in the USV according to the reference variables (fluoxetine, chronic administration, healthy animals, rats, both sexes, not pretested, separation-induced USV).

STRESS-INDUCED HYPERTHERMIA

Description of included studies

The study subjects in the SIH were solely mice, apart from a single article evaluating rats, and all animals were male. In each protocol, rectal temperature measurement functioned as the stressor, while one article exposed the subjects to an open field as an additional stressor. Five articles conducted an individual SIH; the remaining 2 articles carried out a group SIH. The low total number of experiments did not allow for the inclusion of all six SSRIs. Fluoxetine was administered in 3 articles (rats: 10 mg/kg, mice: 10-20 mg/kg) and 2 articles used paroxetine (mice: 0.3-10 mg/kg). One article was conducted using escitalopram (mice: 20 mg/kg) and the last article evaluated fluvoxamine (mice: 3-10 mg/kg). Citalopram and sertraline were not studied in any article. Four articles administered the SSRIs intraperitoneally and 3 articles used oral administration. A sole article chronically administered the SSRI, one subchronically, and 6 articles utilized acute administration.

Meta-analysis

As one article failed to provide information regarding the number of animals per group, only 6 out of the 7 articles could be included in the meta-analysis. The forest plot displayed in Supplementary File S6 shows that only 3 experiments found an anxiolytic effect of SSRIs, whereas the other 9 experiments reported no effect. The pooled effect size did not show a significant effect of SSRIs (Hedges' g: 0.68, 95% CI: -0.08, 1.44, p=0.08, I^2=73%, k=12). There was a high level of between-studies heterogeneity, as depicted by the I^2 value of 73%, which could not be explained through additional moderator analysis, since BRMA did not yield any significant results.

3.2 | Risk of Bias

ELEVATED PLUS MAZE

Figure 5 depicts the findings from the adaptation of SYRCLE's RoB Tool. Overall, the risk of bias was unclear across the articles

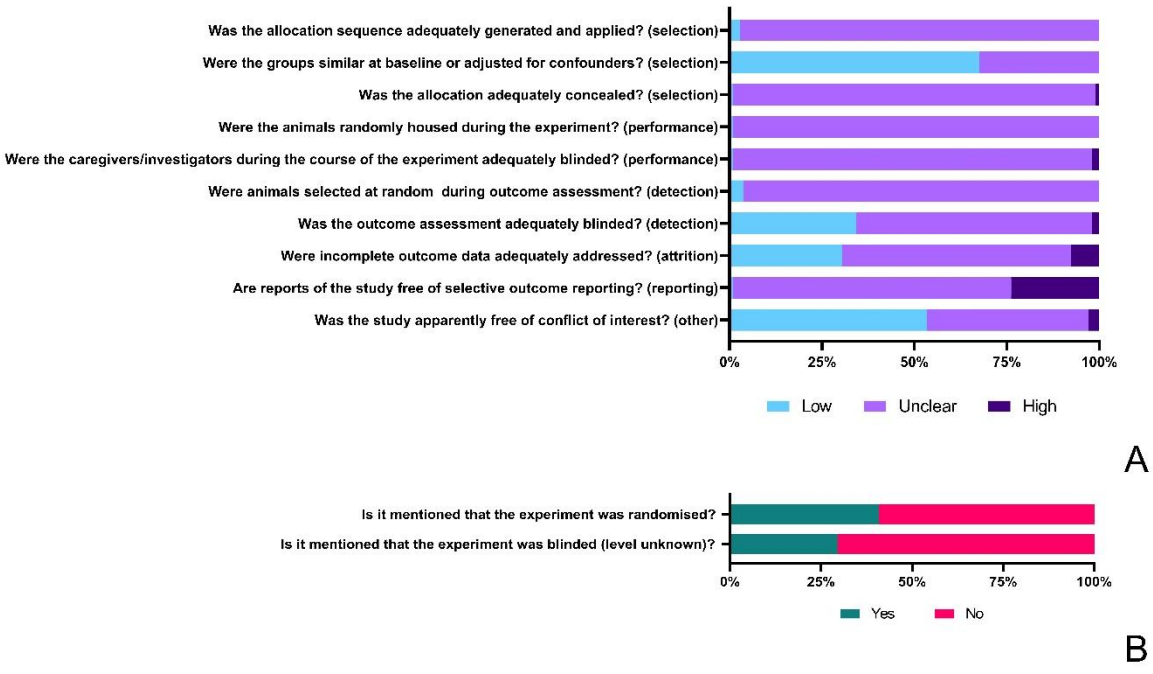

Figure 5: Risk of Bias evaluation for EPM articles, according to SYRCLE's RoB Tool (adapted)

(64%). There were two items in the RoB tool where the risk of bias could be determined for more than half of the articles: the similarity of the groups at baseline, and the presence of a conflict of interest, which were both established to be low in over half of the articles: 68% and 53%, respectively. The highest risk of bias involved reporting bias. Although in most cases (75%) this risk was unclear, almost a quarter of the articles, namely 24%, failed to align the outcome measures reported in the method section to those in the result section, causing a higher risk of reporting bias in those articles. The added reporting questions (Figure 5B) show that a quarter to half of the articles used some form of blinding (30%), randomisation (41%), or both.

MARBLE BURYING

Sixty-nine percent of the time, the risk of bias across MB articles was unclear (Figure 6A). An exception lies in selection bias, which 71% the articles sufficiently (partly) prevented through the creation of similar groups of animals. Experiments were only randomised in 6% of the articles, and a mere 19% used blinding (Figure 6B).

ULTRASONIC VOCALISATIONS

In every experiment, the outcome measure was assessed automatically, resulting in a low risk of detection bias. Almost half of the articles, 45%, established similar groups of animals at baseline. On the other hand, a few articles had a higher risk of reporting bias, by selectively reporting outcomes. This was the case for 18% of the articles. The level of bias remained unclear for 65% altogether (Figure 7A). While some articles applied randomisation in their protocol (36%), hardly any article blinded their experimental set-up (9%), as visualised in Figure 7B.

STRESS-INDUCED HYPERTHERMIA

As shown in Figure 8A, the risk of bias was again mostly unclear (65%). As was the case for the other anxiety tests, most articles studied groups of animals that were similar at baseline (71%),

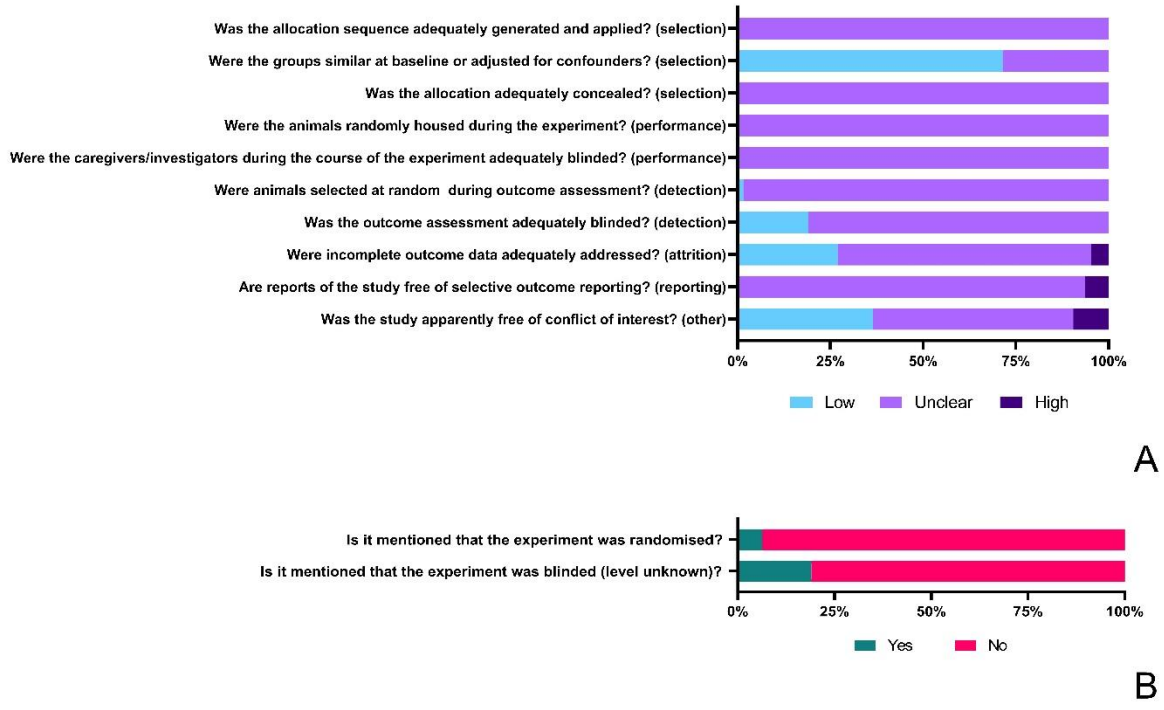

Figure 6: Risk of Bias evaluation for MB articles, according to SYRCLE's RoB Tool (adapted)

reducing the risk of selection bias. Information regarding the presence of attrition bias was reported in 57% of the articles, with all of those articles taking sufficient measures to prevent this type of bias. A potential conflict of interest was observed in a couple articles, namely 29%. The reporting questions depicted in Figure 8B revealed that the majority of articles randomised their experiments (57%), and 29% applied blinding.

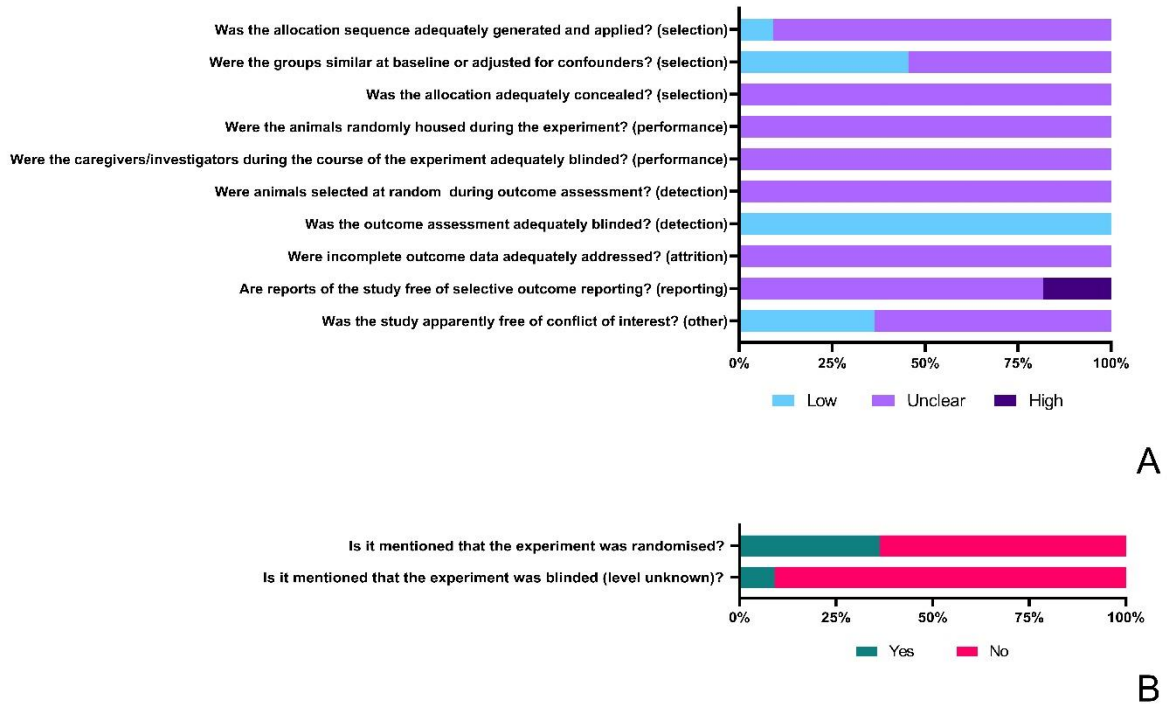

Figure 7: Risk of Bias evaluation for USV articles, according to SYRCLE's RoB Tool (adapted)

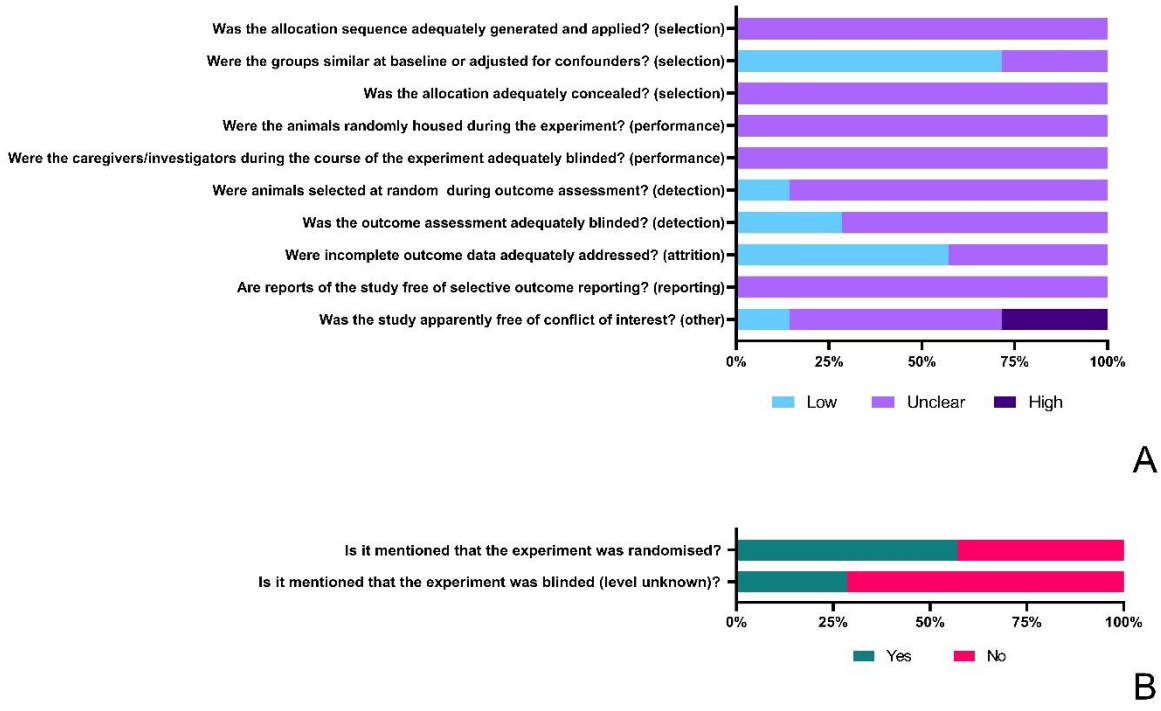

Figure 8: Risk of Bias evaluation for SIH articles, according to SYRCLE's RoB Tool (adapted)

4 | Discussion

This systematic review and meta-analysis is the first to evaluate the effects of all clinically effective SSRIs on multiple anxiety tests, each representing a different aspect of anxiety, and to integrate these results into a comprehensive overview of the aspects of anxiety influenced by these SSRIs. Results are based on a total of 186 articles, with over half reporting on the EPM and an estimated one third studying the MB. The performed meta-analyses suggest that SSRIs reduce the level of anxiety as measured in the EPM, MB and USV. The same cannot be said for the SIH, where no anxiolytic effect was observed after SSRI administration.

The anxiolytic effect observed in the EPM is proposed to be the responsibility of the neural circuitry of approach and avoidance. In a state of anxiety, dysfunctional approach-avoidance behaviour can either be due to disproportionate avoidance or insufficient approach. [47] The distinction between these processes can be made through the involved brain structures: the ventral striatum is associated with approach, while the amygdala is associated with avoidance. The latter is part of the limbic system, which has been established to be the primary projection site of the serotonergic pathways. [6] Combining this information suggests that the approach-avoidance imbalance as evaluated in the EPM stems from excessive avoidance rather than insufficient approach, as a result of a dysfunctioning amygdala. Administration of SSRIs could be hypothesised to partially restore amygdala function and thus the present imbalance in approach-avoidance, consequently reducing anxiety-like behaviour.

On the other hand, neurobiology underlying the effects of SSRIs on the MB has not been clear-cut. [48] This SRMA has proven SSRIs to be effective in reducing obsessive compulsive behaviour, but research on other antidepressants, specifically those that do not bind to the serotonergic transporter with an equivalent affinity, shows a general ineffectiveness of those drugs. [48] These findings strongly suggest that the serotonergic system plays a big role in the pathophysiology of compulsive disorders. However, evidence finding a dysfunction in this system is conflicting, causing an inability to pinpoint an underlying abnormality in the serotonergic system. It is thus hypothesized that SSRIs exert their influence through an intact serotonergic system to compensate for abnormalities in a different system, functionally coupled to the serotonergic system. Some findings point towards an imbalance between direct and indirect frontal-cortical pathways as the pathophysiological disturbance present in obsessive compulsive behaviour. Serotonergic neurons from the midbrain project to the orbitofrontal cortex, making this a plausible theory, worthy of more detailed research. [48,49]

While the reduction in obsessive compulsive behaviour as measured in the MB cannot be directly linked to the SERT blockage of SSRIs, the link between SERT blockage and the anxiolytic effects observed in the USV is relatively straightforward. It has been shown that SERT-knockout mice emit less ultrasonic vocalisations than wildtype mice. [50,51] This demonstrates the involvement of the serotonin transporter in anxiety-like distress calls. Given that SSRIs exert their effect through inhibiting specifically this transporter, the detected anxiolytic effect of SSRIs on the USV has a straightforward explanation, while also demonstrating the role of the serotonergic system in emotional regulation.

Despite this SRMA not finding any significant results concerning the SIH, it is interesting to look into the relation between the serotonergic system and the autonomic nervous system, particularly temperature regulation. While very few articles specifically target this relationship, the answer might lie in the serotonin receptors. Research shows that

activation of 5-HT_{1A}, 5-HT_3 and 5-HT_7 receptor subtypes lead to a decrease in body temperature. [52,53] Body temperature rises when experiencing stress [34], and since SSRI administration causes SERT blockage, which results in a higher level of serotonin present in the synapse, this could possibly counteract that increase in body temperature, by binding to the aforementioned receptors. Nevertheless, the lack of observed effect in the SIH suggests that this effect is rather subtle and might only be assessed with the right experimental set-up. Alternatively, the lack of conducted research and evidence could also contribute to an absence of power.

A common denominator in three of the paradigms (EPM, USV, SIH) was the high level of heterogeneity observed. The great variety in experimental set-ups across articles could partly account for this high level of heterogeneity. However, additional moderator analysis was performed in an effort to explain the heterogeneity even further. This BRMA found anxiolytic effects of SSRIs on the EPM and USV when the reference variables were implemented in the study protocol, but not for the SIH. Unfortunately, moderating effects of dummy variables were not found, therefore not providing an explanation for the high level of heterogeneity.

On the other hand, the MB showed a low level of heterogeneity, which makes sense considering the similarity of experimental protocol across articles. Still, BRMA was conducted to gain further insights into the influence of different moderators. Like the EPM and USV, an anxiolytic effect was found in studies applying all reference categories. Additionally, accompanying that finding were some other interesting results. One dummy variable showed a greater anxiolytic effect of SSRIs than the reference category: studies not reporting the sex of their animals. This finding will be further discussed in Section 4.3. The second and last finding was a moderating effect of the continuous moderator HED. It can be concluded that a higher dose of SSRI shows a greater anxiolytic effect on the MB. This is in line with clinical practice: OCD often sustains a higher maximum dosage of SSRIs than other anxiety-like disorders. [54] It has to be noted that a continuous increase of dosage does not lead to an ever-increasing anxiolytic effect. Toxicity is an important restriction. HEDs were low compared to the clinically used doses and in all cases lower than the minimal clinical dose. This would explain why a higher HED yields a more anxiolytic result; it is closer to the clinical dose effective in humans. However, the metabolism of rodents prevents the use of higher dosages to achieve clinical levels. Too high of a dose could be counter effective and induce serotonin syndrome rather than anxiolytic effects. A balance has to be found between effectiveness in animal research and translational validity in humans.

4.1 | Knowledge gaps

It goes without saying that most knowledge gaps are present around the SIH. Only seven articles met inclusion criteria, indicating that the effect of SSRIs on the SIH has not yet been widely studied. This in itself takes away the opportunity to draw conclusions with a high level of power. Additionally, when investigating the individual moderators, the amount of data lessens even more. Fluoxetine was the most studied SSRI, investigated in only three articles. Citalopram and sertraline were not studied at all. Moreover, no statements can be made on differences between species, since all but one experiments were conducted on mice. Overall, very little can be said about the true effect of SSRIs on the SIH, since not enough research has been conducted yet.

To a lesser extent, the same holds true for the USV. While eleven articles were enough to yield significant results, these still have to be interpreted with caution, since the power of this significance was not high. Especially in the

moderator analysis, information on different variables was scarce. For instance, only three articles induced ultrasonic vocalisations through physical stress, and both fluvoxamine and escitalopram were only studied in a single article each. Thus, although the USV is sensitive to SSRIs as a whole, distinctions regarding an optimal experimental set-up cannot be made, due to the quantity of performed research.

While knowledge gaps in the USV and SIH are both dependent on an overall lack of available research, those in the MB originate from another phenomenon. The MB is a well-established paradigm, but experimental set-ups are alike across articles. On the positive side, this allows for an overall effect size representative of the overall experimental population. On the flip side, distinctions between experimental set-ups cannot be assessed, since the experimental protocols are that similar. One major component missing in the MB data set is the presence of stressed animals. Only seven out of sixty-three articles applied stress protocols. Modifications in the serotonergic systems lie at the base of anxiety-like disorders and since healthy animals do not have these alterations, they may react differently to SSRI administration. [6] A second lack of variety is present in the studied species. All experiments were conducted on mice, leaving no room for interspecies evaluation.

The anxiety test with the most data available, and the most variety within the data was the EPM. Both rats and mice were studied, healthy as well as stressed animals were included, and all SSRIs were well-represented. However, one knowledge gap exists across this paradigm, as well as across the other three anxiety tests: representation of all sexes. When reviewing the data as a whole, there is a discrepancy between the sex of the animals studied and the sex more commonly affected by anxiety. A strong majority of the experiments was conducted on males, while females are known to suffer from anxiety disorders more frequently. [55,56] Although studies are limited, there is also some evidence that females respond differently to antidepressants. [55] Since females were only the subject of some articles, no sturdy conclusions can be drawn from moderator analyses about differences in treatment response between males and females.

4.2 | Limitations of included articles

Reliability of the findings depends on the quality of the included articles. Assessment of this quality through SYRCLE's Risk of Bias Tool yielded some interesting results. Most noticeable was the unclear level of reporting bias. Since protocols of preclinical animal studies are rarely preregistered, evaluation of reporting bias relied mostly on the comparison between reported outcome measures in the method section and the result section. Although these reports lined up in the majority of the articles, the lack of preregistration caused the risk of reporting bias to be unclear, as the originally intended outcome measures could not be verified. However, regardless of the preregistration, a number of articles failed to align the reported outcome measures, usually by not reporting results of all outcome measures stated in the method section. As the EPM allows for the evaluation of a wide variety of outcome measures, the risk of reporting bias was the highest in this paradigm. Reasons behind the discrepancies remain unknown; they could be due to experimenters not finding significant results and thus leaving those out, or worse, finding significant results, but contrary to those expected. This could undermine the quality of these articles and the advancement of this specific field of research.

A second limitation of the included articles involves the unit of outcome reporting. The desired unit of outcome reporting in this SRMA was the mean, accompanied with either a SEM or SD. In a handful of cases, a different unit was chosen to present the results. For an

equal analysis of all articles to be conducted, deviating values had to be converted. The downside that comes with recalculating those values into mean, SEM or SD is the assumption that the data points were normally distributed. It is likely that authors chose an alternative unit to report their result, to show that their data was in fact not normally distributed. For instance, this is common for time based outcome measures, as they are often right-skewed. Fortunately, the possible uncertainty induced by converting outcome measures remained low in terms of the meta-analysis as a whole, considering that no more than a tenth of all articles had to undergo this process.

Furthermore, one aspect to be mentioned revolves around the many possible (and reported) outcome measures of the EPM. Although several behavioural outcomes are valuable and relevant indicators of anxiety-like behaviour, prevention of biological duplicates as well as enlargement of homogeneity caused a need for prioritization of these outcomes. Observation of multiple behaviours during a test session allows for a lower number of animals needed for animal testing. The downside hereof is the limitation to evaluate a variety of outcome behaviours. Subsequently, the decision was made to rank outcome measures based on two principles: the perceived relevance in the assessment of approach-avoidance anxiety, but also the quantity of the outcome reported. Consequently, exploration of the open arms was determined to be most relevant. Despite time spent in open arms being reported more frequently, entries in open arms was deemed more representative of approach-avoidance anxiety. Individual differences in baseline willingness to explore were accounted for by prioritizing open arm entries or time spent in open arms presented as a percentage of total entries or total time spent in the arms.

4.3 | Limitations of study protocol

The results of this SRMA should always be considered in light of the protocol's limitations. A first limitation lies in the direction of effect size in the EPM, combined with the aforementioned variety in outcome measures. While the primary outcome measures (%EOA, %TOA, EOA and TOA) are expected to increase under anxiolytic treatment, not all outcome measures are. The variety in observable outcome behaviours caused not all articles to evaluate one of the primary outcome measures. As a result, a few articles were analysed based on a different outcome measure. Unfortunately, this resulted in the inclusion of nine experiments reporting a outcome measure with an inverted direction of effect. Eight of these experiments reported an anxiety index, and the ninth experiment reported the time spent in the closed arms; both of these outcomes are expected to decline under anxiolytic treatment. These effect sizes were not corrected for in the forest plots and BRMA and are thus opposite to the actual observed effect in the experiment. Fortunately, it is expected that since the number of experiments with inverted effect sizes is low, the overall effect size was not influenced by this limitation.

A second limitation stems from the article selection process. One exclusion criterion of the full-text assessment was the conduction of other behavioural tests prior to the anxiety test of interest. As not all articles specified a clear timeline, order of reporting was used to determine the test order. This could have resulted in the inclusion of animals which had already been tested in other tests in a behavioural test battery, or, on the other hand, exclusion of experiments testing naïve animals, but reporting it in a different order. However, in the majority of the articles, a clear timeline was presented, minimizing the chance

of faulty decisions in the selection process.

The last limitation is present in the results of the MB. Usually, an experiment deviating more than three standard deviations from the mean (Z-score > 3) is considered an outlier. In the case of the MB, four articles with the following article codes were determined to be outliers: Arzoo 2017A (Z-score: 10.5) [57], Arzoo 2017B (Z-score: 5.7) [58], Kaurav 2012 (Z-score: 10.2) [59], and Mesripour 2018 (Z-score: 3.5) [60]. Typically, outliers are eliminated from the meta-analysis, in order to obtain an overall effect size representative of the true population. However, in this SRMA these outliers were kept in the sample and treated as any other data point. The danger of this method is overvaluation of the outliers, consequently causing the overall effect size to potentially vary from the true effect size. [61] Furthermore, this limitation could be the cause of the observed significant moderating effect of sex in the BRMA. Articles not reporting the sex of their animals demonstrated a greater anxiolytic effect of SSRIs compared to articles reporting on males. However, this dummy variable consisted of three experiments, two of which were outliers (Arzoo 2017A and Arzoo 2017B). It is likely that these outliers have caused a false positive result in the BRMA, considering their impact on this dummy variable.

4.4 | Conclusion

This SRMA has analysed the effects of SSRIs on anxiety as measured in four anxiety tests: the Elevated Plus Maze Test, Marble Burying Test, Ultrasonic Vocalisations Test, and Stress-Induced Hyperthermia Test. Results showed a reduction in anxiety-like behaviour after SSRI administration in three of the paradigms (EPM, MB, USV). A limited level of research prevented the acquisition of significant results in the SIH. Heterogeneity between studies was high for all paradigms, except the MB, which is representative of the variety of protocols used in each paradigm. BRMA could not explain the high level of heterogeneity across anxiety tests, but did depict a positive correlation between HED and anxiolytic effect in the MB. BRMA additionally showed sensitivity of three anxiety tests (EPM, MB, USV) for SSRIs, when the study protocol entailed the reference variables as shown in Table 5. It can be concluded that SSRIs are effective in the treatment of various aspects of anxiety, namely an approach-avoidance imbalance, obsessive compulsive tendencies and stress-induced distress calls. These findings add to a better comprehension of the role of the serotonergic system in different forms of anxiety. Additionally, these findings provide insights in the optimal experimental set-up for evaluating putative anxiolytic effects of drugs and thereby reducing redundant animal testing. Variables showing the most sensitivity of the effects of SSRIs on the anxiety tests are fluoxetine, and chronic administration on naïve animals. These animals are preferred to be stressed, in the case of a USV through separation-induced stress. Rats show the greatest sensitivity of effects of SSRIs as measured on the EPM and the USV, while this holds true for mice in case of the MB. The latter is only due to the lack of data available on rats in the MB. As male animals were tested most frequently, this sex naturally was the most sensitive to effects of SSRIs on anxiety tests. The lack of moderating effects of dummy variables suggests that these were not able to explain the remaining level of heterogeneity. In order to fully understand the effects of SSRIs on anxiety disorders most commonly presented in the clinical setting, more research needs to be conducted to fill the knowledge gaps, for instance on female animals. Furthermore, more research on the SIH is needed to understand the effects of the serotonergic system on stress responses in the autonomic nervous system. In considering the implications of this study, aspects that always have to be kept in mind are the heterogeneous nature of anxiety and interpersonal differences in the clinical setting.

This review is the first step in the process to obtain a full comprehension of the effects of individual SSRIs on different anxiety disorders and opens up opportunities for further discoveries in the endeavour of improving the global mental health status.

5 | References

[1] Heimberg RG, Hofmann SG, Liebowitz MR, Schneier FR, Smits JAJ, Stein MB, et al. SOCIAL ANXIETY DISORDER IN DSM-5. Depress Anxiety 2014;31:472–9. https://doi.org/10.1002/da.22231.

[2] Santomauro DF, Mantilla Herrera AM, Shadid J, Zheng P, Ashbaugh C, Pigott DM, et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. Lancet 2021;398. https://doi.org/10.1016/S0140-6736(21)02143-7.

[3] Benjamin J. Sadock MD, Virginia A. Sadock MD, Pedro Ruiz MD. Kaplan & Sadock's COMPREHENSIVE TEXTBOOK OF PSYCHIATRY VOLUME I 9th ed. vol. 53. 2019.

[4] Adwas AA, Jbireal JM, Azab AE. Anxiety: Insights into Signs, Symptoms, Etiology, Pathophysiology, and Treatment. East African Sch J Med Sci 2019;2.

[5] Farb DH, Ratner MH. Targeting the modulation of neural circuitry for the treatment of anxiety disorders. Pharmacol Rev 2014. https://doi.org/10.1124/pr.114.009126.

[6] Ressler KJ, Nemeroff CB. Role of serotonergic and noradrenergic systems in the pathophysiology of depression and anxiety disorders. Depress Anxiety 2000. https://doi.org/10.1002/1520-6394(2000)12:1+<2::AID-DA2>3.0.CO;2-4.

[7] Baldwin DS, Anderson IM, Nutt DJ, Bandelow B, Bond A, Davidson JRT, et al. Evidence-based guidelines for the pharmacological treatment of anxiety disorders: Recommendations from the British Association for Psychopharmacology. J Psychopharmacol 2005. https://doi.org/10.1177/0269881105059253.

[8] Hiemke C, Härtter S. Pharmacokinetics of selective serotonin reuptake inhibitors. Pharmacol Ther 2000. https://doi.org/10.1016/S0163-7258(99)00048-0.

[9] Stahl SM. Stahl's essential psychopharmacology: Neuroscientific basis and practical applications. 4th ed. Cambridge University Press; 2013.

[10] Hyttel J. Pharmacological characterization of selective serotonin reuptake inhibitors (SSRIs). Int Clin Psychopharmacol 1994. https://doi.org/10.1097/00004850-199403001-00004.

[11] Cryan JF, Sweeney FF. The age of anxiety: Role of animal models of anxiolytic action in drug discovery. Br J Pharmacol 2011. https://doi.org/10.1111/j.1476-5381.2011.01362.x.

[12] Bourin M, Petit-Demoulière B, Nic Dhonnchadha B, Hascöet M. Animal models of anxiety in mice. Fundam Clin Pharmacol 2007. https://doi.org/10.1111/j.1472-8206.2007.00526.x.

[13] Kumar V, Bhat ZA, Kumar D. Animal models of anxiety: A comprehensive review. J Pharmacol Toxicol Methods 2013. https://doi.org/10.1016/j.vascn.2013.05.003.

[14] Rodgers RJ. Animal models of "anxiety": Where next? Behav Pharmacol 1997;8:477–96. https://doi.org/10.1097/00008877-199711000-00003 LK - https://utrechtuniversity.on.worldcat.org/atoztitles/link?sid=EMBASE&sid=EMBASE&issn=09558810&id=doi:10.1097%2F00008877-199711000-00003&atitle=Animal+models+of+%27anxiety%27%3A+Where+next%3F&stitle=BEHAV.+PHARMACOL.&title=Behavioural+Pharmacology&volume=8&issue=6-7&spage=477&epage=496&aulast=Rodgers&aufirst=R.J.&auinit=R.J.&aufull=Rodgers+R.J.&coden=BPHAE&isbn=&page

s=477-496&date=1997&auinit1=R&auinitm=J.

[15] Kraeuter AK, Guest PC, Sarnyai Z. The Elevated Plus Maze Test for Measuring Anxiety-Like Behavior in Rodents. Methods Mol. Biol., vol. 1916, 2019. https://doi.org/10.1007/978-1-4939-8994-2_4.

[16] Lister RG. Ethologically-based animal models of anxiety disorders. Pharmacol Ther 1990;46. https://doi.org/10.1016/0163-7258(90)90021-S.

[17] Rodgers RJ, Dalvi A. Anxiety, defence and the elevated plus-maze. Neurosci Biobehav Rev 1997;21:801–10. https://doi.org/https://doi.org/10.1016/S0149-7634(96)00058-9.

[18] Komada M, Takao K, Miyakawa T. Elevated plus maze for mice. J Vis Exp 2008. https://doi.org/10.3791/1088.

[19] La-Vu M, Tobias BC, Schuette PJ, Adhikari A. To Approach or Avoid: An Introductory Overview of the Study of Anxiety Using Rodent Assays. Front Behav Neurosci 2020;14. https://doi.org/10.3389/fnbeh.2020.00145 LK - https://utrechtuniversity.on.worldcat.org/atoztitles/link?sid=EMBASE&sid=EMBASE&issn=16625153&id=doi:10.3389%2Ffnbeh.2020.00145&atitle=To+Approach+or+Avoid%3A+An+Introductory+Overview+of+the+Study+of+Anxiety+Using+Rodent+Assays&stitle=Front.+Behav.+Neurosci.&title=Frontiers+in+Behavioral+Neuroscience&volume=14&issue=&spage=&epage=&aulast=La-Vu&aufirst=Mimi&auinit=M.&aufull=La-Vu+M.&coden=&isbn=&pages=-&date=2020&auinit1=M&auinitm=.

[20] Biedermann S V, Biedermann DG, Wenzlaff F, Kurjak T, Nouri S, Auer MK, et al. An elevated plus-maze in mixed reality for studying human anxiety-related behavior. BMC Biol 2017;15:125. https://doi.org/10.1186/s12915-017-0463-6.

[21] Wolmarans de W, Stein DJ, Harvey BH. Of mice and marbles: Novel perspectives on burying behavior as a screening test for psychiatric illness. Cogn Affect Behav Neurosci n.d.;16:551–60.

[22] Jimenez-Gomez C, Osentoski A, Woods JH. Pharmacological evaluation of the adequacy of marble burying as an animal model of compulsion and/or anxiety. Behav Pharmacol 2011;22. https://doi.org/10.1097/FBP.0b013e32834afebe.

[23] Taylor GT, Lerch S, Chourbaji S. Marble burying as compulsive behaviors in male and female mice. Acta Neurobiol Exp 2017;77:254–60.

[24] Alonso P, López-Solà C, Real E, Segalàs C, Menchón JM. Animal models of obsessive-compulsive disorder: utility and limitations. Neuropsychiatr Dis Treat 2015:11–1939. https://doi.org/10.2147/NDT.S62785.

[25] Albelda N, Joel D. Current Animal Models of Obsessive Compulsive Disorder: An Update 2012. https://doi.org/10.1016/j.neuroscience.2011.08.070.

[26] de Brouwer G, Fick A, Harvey BH, Wolmarans DW. A critical inquiry into marble-burying as a preclinical screening paradigm of relevance for anxiety and obsessive–compulsive disorder: Mapping the way forward. Cogn Affect Behav Neurosci 2018 191 2018;19:1–39. https://doi.org/10.3758/S13415-018-00653-4.

[27] Schwarting RKW, Wöhr M. On the relationships between ultrasonic calling and anxiety-related behavior in rats. Brazilian J Med Biol Res 2012;45:337–48. https://doi.org/10.1590/S0100-879X2012007500038.

[28] De Vry J, Benz U, Schreiber R, Traber J. Shock-induced ultrasonic vocalization in young adult rats: a model for testing putative anti-anxiety drugs. Eur J Pharmacol 1993;249. https://doi.org/10.1016/0014-2999(93)90530-U.

[29] Fendt M, Gonzalez-Guerrero CP, Kahl E. Observational fear learning in rats: Role of trait anxiety and ultrasonic vocalization. Brain Sci 2021;11. https://doi.org/10.3390/brainsci110404

23.

[30] Wöhr M, Schwarting RKW. Rodent ultrasonic communication and its relevance for models of neuropsychiatric disorders. E-Neuroforum 2010;16. https://doi.org/10.1007/s13295-010-0012-z.

[31] Groenink L, Verdouw PM, van Oorschot R, Olivier B. Models of anxiety: Ultrasonic vocalizations of isolated rat pups. Curr Protoc Pharmacol 2008. https://doi.org/10.1002/0471141755.ph0518s43.

[32] Vinkers CH, Van Oorschot R, Olivier B, Groenink L. Stress-induced hyperthermia in the mouse. Neuromethods 2009;42. https://doi.org/10.1007/978-1-60761-303-9_8.

[33] Adriaan Bouwknecht J, Olivier B, Paylor RE. The stress-induced hyperthermia paradigm as a physiological animal model for anxiety: A review of pharmacological and genetic studies in the mouse. Neurosci Biobehav Rev 2007;31. https://doi.org/10.1016/j.neubiorev.2006.02.002.

[34] Olivier B, Zethof T, Pattij T, Van Boogaert M, Van Oorschot R, Leahy C, et al. Stress-induced hyperthermia and anxiety: pharmacological validation. Eur J Pharmacol 2003;463:117–32. https://doi.org/10.1016/S0014-2999(03)01326-8.

[35] Moher D, Liberati A, Tetzlaff J, Altman DG, Altman D, Antes G, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med 2009;6. https://doi.org/10.1371/journal.pmed.1000097.

[36] de Vries RBM, Hooijmans CR, Tillema A, Leenaars M, Ritskes-Hoitinga M. A search filter for increasing the retrieval of animal studies in Embase. Lab Anim 2011;45:268–70. https://doi.org/10.1258/LA.2011.011056.

[37] de Vries RBM, Hooijmans CR, Tillema A, Leenaars M, Ritskes-Hoitinga M. Letter to the Editor. Lab Anim 2014;48:88. https://doi.org/10.1177/0023677213494374/SUPPL_FILE/LAN494374_SUPPLEMENT.DOC.

[38] Lad HV, Liu L, Paya-Cano JL, Parsons MJ, Kember R, Fernandes C, et al. Behavioural battery testing: Evaluation and behavioural outcomes in 8 inbred mouse strains. Physiol Behav 2010;99. https://doi.org/10.1016/j.physbeh.2009.11.007.

[39] McIlwain KL, Merriweather MY, Yuva-Paylor LA, Paylor R. The use of behavioral test batteries: Effects of training history. Physiol Behav 2001;73. https://doi.org/10.1016/S0031-9384(01)00528-5.

[40] van Gerven J, Cohen A. Integrating data from the Investigational Medicinal Product Dossier/investigator's brochure. A new tool for translational integration of preclinical effects. Br J Clin Pharmacol 2018;84. https://doi.org/10.1111/bcp.13529.

[41] Van Lissa CJ, Brandmaier AM, Brinkman L, Lamprecht A-L, Peikert A, Struiksma ME, et al. WORCS: A workflow for open reproducible code in science. Data Sci 2021;4. https://doi.org/10.3233/ds-210031.

[42] R Core Team. R core team (2021). R A Lang Environ Stat Comput R Found Stat Comput Vienna, Austria URL Http//Www R-Project Org 2021.

[43] Viechtbauer W. Conducting Meta-analysis in R with the metafor package. J Stat Softw 2010;36.

[44] Lissa CJ van, Erp S van, Clapper E-B. Selecting relevant moderators with Bayesian regularized meta-regression n.d. https://doi.org/10.31234/OSF.IO/6PHS5.

[45] Van den Noortgate W, López-López JA, Marín-Martínez F, Sánchez-Meca J. Meta-analysis of multiple outcomes: a multilevel approach. Behav Res Methods 2014;47. https://doi.org/10.3758/s13428-014-0527-2.

[46] Hooijmans CR, Rovers MM, De Vries RBM, Leenaars M, Ritskes-Hoitinga M,

Langendam MW. SYRCLE's risk of bias tool for animal studies. BMC Med Res Methodol 2014;14. https://doi.org/10.1186/1471-2288-14-43.

[47] Aupperle RL, Paulus MP. Neural systems underlying approach and avoidance in anxiety disorders. Dialogues Clin Neurosci 2010;12:517–31. https://doi.org/10.31887/dcns.2010.12.4/raupperle.

[48] Goodman WK, Storch EA, Sheth SA. Harmonizing the neurobiology and treatment of obsessive-compulsive disorder. Am J Psychiatry 2021;178:17–29. https://doi.org/10.1176/APPI.AJP.2020.20111601/ASSET/IMAGES/LARGE/APPI.AJP.2020.20111601F1.JPEG.

[49] Joel D, Doljansky J, Roz N, Rehavi M. Role of the orbital cortex and of the serotonergic system in a rat model of obsessive compulsive disorder. Neuroscience 2005;130:25–36.

[50] Willadsen M, Uengoer M, Schwarting RKW, Homberg JR, Wöhr M. Reduced emission of alarm 22-kHz ultrasonic vocalizations during fear conditioning in rats lacking the serotonin transporter. Prog Neuro-Psychopharmacology Biol Psychiatry 2021;108:110072. https://doi.org/10.1016/J.PNPBP.2020.110072.

[51] Hohmann CF, Blue ME. The Role of Serotonin in Cortical Development: Implications for Autism Spectrum Disorder. Handb. Behav. Neurosci., vol. 21, 2010. https://doi.org/10.1016/S1569-7339(10)70104-1.

[52] Ishiwata T. Role of serotonergic system in thermoregulation in rats. J Phys Fit Sport Med 2014;3:445–50. https://doi.org/10.7600/JPFSM.3.445.

[53] Voronova IP. 5-HT Receptors and Temperature Homeostasis. Biomol 2021, Vol 11, Page 1914 2021;11:1914. https://doi.org/10.3390/BIOM11121914.

[54] Informatorium Medicamentorum | KNMP Kennisbank n.d. https://kennisbank-knmp-nl.proxy.library.uu.nl/article/Informatorium_Medicamentorum/intro.html (accessed June 19, 2021).

[55] Palanza P. Animal models of anxiety and depression: how are females different? n.d.

[56] McLean CP, Asnaani A, Litz BT, Hofmann SG. Gender differences in anxiety disorders: Prevalence, course of illness, comorbidity and burden of illness. J Psychiatr Res 2011;45:1027–35. https://doi.org/10.1016/j.jpsychires.2011.03.006.

[57] Arzoo, Parle M. Anti-obsessive–compulsive activity of honey. Asian J Pharm Clin Res 2017;10:206–9.

[58] Arzoo, Parle M. Investigation of the anti-obsessive comulsive activity of Pyrus communis juice in mice. Int J Pharm Sci Rev Res 2017;42:16–20.

[59] Kaurav BP, Wanjari MM, Ch, ekar A, Chauhan NS, Upmanyu N. Influence of Withania somnifera on obsessive compulsive disorder in mice. Asian Pac J Trop Med n.d.;5:380–4.

[60] Mesripour A, Rezaei S, Hajhashemi V. Evaluating the effect of donepezil on depression and obsessive–compulsion disorder in mice models and proposing the mechanism involved. Thai J Pharm Sci 2018;42:1–5.

[61] Ghosh D, Vogt A. Outliers: An Evaluation of Methodologies n.d.