Table of Contents

1. Abstract ........................................................................................................................................ 3

2. Introduction .................................................................................................................................. 3
   2.1 Modelling depression .................................................................................................................. 4
   2.2 Evaluation of validity .................................................................................................................. 4
      2.2.1 Face validity ....................................................................................................................... 5
      2.2.2 Construct validity ............................................................................................................... 5
      2.2.3 Predictive validity .............................................................................................................. 5
   2.3 Evaluation of usability ................................................................................................................ 5
   2.4 Evaluation of reliability ............................................................................................................. 5

3. Behaviour tests ............................................................................................................................... 5
   3.1 Anhedonia .................................................................................................................................. 6
      3.1.1 Sucrose preference .............................................................................................................. 6
      3.1.2 Intracranial self-stimulation (ICSS) .................................................................................... 6
   3.2 Decreased concentration ........................................................................................................... 7
      3.2.1 Morris water maze (MWM) ............................................................................................... 7
      3.2.2 Y-maze (spontaneous alteration) ...................................................................................... 7
   3.3 Psychomotor retardation ........................................................................................................... 7
      3.3.1 Open-field test (OFT) ....................................................................................................... 7
      3.3.2 Elevated plus-maze (EPM) ............................................................................................... 8
      3.3.3 Locomotion in home cage ............................................................................................... 8
   3.4 Conclusions ............................................................................................................................... 8

4. Physiological measurements .......................................................................................................... 9
   4.1 Changes in body weight ............................................................................................................. 9
   4.2 Altered sleep patterns ............................................................................................................... 10
   4.3 Neurological measurements .................................................................................................... 10
      4.3.1 Brain markers .................................................................................................................... 10
      4.3.2 Molecular markers .......................................................................................................... 10

5. Models for depression ................................................................................................................... 10
   5.1 Stress models ............................................................................................................................. 11
      5.1.1 Unpredictable chronic mild stress (UCMS) ....................................................................... 11
      5.1.2 Chronic social defeat stress (CSDS) .................................................................................. 11
      5.1.3 Chronic restraint stress (CRS) ......................................................................................... 12
      5.1.4 Chronic foot-shock stress (CFS) ....................................................................................... 12
      5.1.5 Learned Helplessness (LH) .............................................................................................. 13
   5.2 Olfactory Bullectomy (OB) ...................................................................................................... 13
   5.3 Cytokine induced depressive-like behaviour ............................................................................ 13

6. Discussion ....................................................................................................................................... 14

References .......................................................................................................................................... 16
1. Abstract
Major depressive disorder (MDD) is a severe mood disorder in humans. The aetiology of the pathology is still unclear. In order to study the disease, animal models are often used. Over the years, different animal models for depression have been developed, but it is unclear if every model is equally suitable. Therefore it is necessary to compare these models and the tests that can be used when measuring depression. The aim of this thesis is to make a comparison of several animal models for the study of depression and try to identify the best models.

When discussing these models, the face validity (similarity of symptoms between patients and animals), the construct validity (similarity of pathogenesis) and the predictive validity (effect of antidepressants) are evaluated. Also the usability (simplicity) and reliability (consistency over time) are discussed.

In models for depression both behaviour and physiological changes should be tested. Using the sucrose preference test, the Morris water maze (rats) or Y-maze (mice) and measurements of locomotion in the home cage, depressive-like behaviour can be tested. Changes in body weight, sleep patterns and several other factors (like BDNF, neurogenesis or corticosterone) can be measured to define the physiological changes.

These different tests are used to compare the different models. Based on the studies evaluated in this thesis, the unpredictable chronic mild stress model is the best one, if enough staff is available. If not, than the chronic social defeat stress model is the best alternative. The cytokine-induced model is the most promising model for the future, but it needs more work before it really can be used.

2. Introduction
Major depressive disorder (MDD) is a severe mood disorder in humans. The risk of developing depression is 10-25% for women and 5-12% for men. After a general medical condition, 25% of the individuals develop depression.\(^1\) Of all individuals with MDD, 15% dies by suicide.\(^1\) The impact of MDD on the life of the patients is high; most individuals stop working, causing high costs for the society. It is predicted that by 2030 depression will be one of the three leading causes of disability.\(^2\)

MDD is divined by several diagnostic criteria, both cognitive and behavioural. These symptoms are chronic, most patients show symptoms for at least 2 years.\(^1\) An overview is given in table 1.

<table>
<thead>
<tr>
<th>Diagnostic criteria</th>
<th>Major Depressive Disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Depressed mood</td>
</tr>
<tr>
<td></td>
<td>Extreme negativism</td>
</tr>
<tr>
<td></td>
<td>Excessive/inappropriate guilt</td>
</tr>
<tr>
<td></td>
<td>Loss of interest/pleasure</td>
</tr>
<tr>
<td></td>
<td>Loss of energy</td>
</tr>
<tr>
<td></td>
<td>Psychomotor retardation</td>
</tr>
<tr>
<td></td>
<td>Motoric immobility</td>
</tr>
<tr>
<td></td>
<td>Early morning awakening (&gt;2H earlier)</td>
</tr>
<tr>
<td></td>
<td>Hypersomnia</td>
</tr>
</tbody>
</table>
2.1 Modelling depression

The aetiology and progression of pathology in depression is still unclear. Research is necessary to be able to explain these processes. In order to be able to experimentally study depression a lot of effort has been put in to creating animal models that are able to mimic the symptoms of depression. The problem with an animal model for depression is that not all the symptoms can be modelled. An example is mood assessment. In humans this assessment can be obtained by using a questionnaire. In animals there is no way to test this. To bypass this problem Dzirasa & Covington (2012) proposed a work frame when using animals. They suggested using only the symptoms that can be modelled in animals and divided these symptoms in three domains. The first domain contains all reward related behaviour, such as anhedonia (decreased interest in pleasant things) or decreased concentration. The second domain contains homeostatic factors, like psychomotor retardation, changes in body weight and insomnia. The last domain are the biomarkers for depression, which can be either biochemical or neurophysiological. By using at least one test of each domain, the animal model for depression will be more valid to use.

It is important to notice that when working with animals, it is never sure if the behaviour observed is really due to depression. This is because it is not possible to assess mood and feelings of animals. It is necessary to always take this into account when working with animal models for depression. Therefore behaviour changes are called depressive-like, to differentiate between animal and human behaviour.

The field of animal models and testing has become bigger over the years. Nowadays over 25,000 publications are available about models for depression and it has become hard to decide which model can best be used in research. Therefore it is necessary to compare the different models and behaviour tests available by using several criteria on each model. Although depression can be modelled in several animals; rats and mice are most commonly used. For this reason only models using rodents are used in this paper.

In the next sections the different tests to define depressive-like behaviour are summarized and discussed. In the end also a recommendation for the best tests is given. After that the different models will be discussed by using the suggested tests given before. When discussing the different tests and models, the validity, usability and reliability will be evaluated. These criteria will be discussed first.

2.2 Evaluation of validity

To correctly evaluate an animal model, it is necessary to have criteria that can be applied in every situation. For this reason Willner (1984) described several aspects a model needs concerning its validity in relation with the human condition. For depression it is harder to validate a model, because the aetiology and biochemistry are largely unclear and still being researched. Hereafter the 3 main criteria are shortly summarized:
2.2.1 Face validity
Face validity is defined as "phenomenological similarities between the model and the disorder". This means that the used model should reproduce a core symptom of depression. For an animal model there is a limitation, because assessment of mood, guilt or suicidal thinking is not possible. Usable measurements will be described below. Another feature of depression is its chronicity. Depression is a long-term syndrome, so an animal model should show changes for at least two weeks.

2.2.2 Construct validity
For a model to have construct validity, it is important that the symptoms are mediated by the same (neurobiological) mechanisms as in humans. Some factors important in the pathogenesis of MDD are stressful life events, negative emotions (cognitive hypothesis of depression) and inflammation (cytokine hypothesis of depression). Because the mechanisms are not precisely known, it will be hard to establish a high degree of construct validity.

2.2.3 Predictive validity
Predictive validity is assessed by whether a model responds in the same way to treatments as depressed people do. A model should, for example, respond appropriately to antidepressant drugs that are clinically effective. This response should only be visible after chronic treatment, because in humans antidepressants are only effective after a couple of weeks. Chronic treatment in rodents is defined as administration over at least two weeks.

2.3 Evaluation of usability
When evaluating the usability of a model, the duration and effort needed has to be taken into account. A protocol that takes one week can be high in usability but when you need 6-8 hours work per day during this week, the effort needed is high and thus the usability lowers. In this case a model that takes four weeks with only an hour work per day has higher usability. Also, when the procedures are difficult to carry out, the usability lowers.
In conclusion, when a model can be easy carried out and needs relatively low effort, this model is high in usability.

2.4 Evaluation of reliability
The reliability of an animal model depends on its consistency over time. The phenotype of the model should not change over time; this applies to both behaviour and physiology. Also it should be reproducible. When the same experiment is carried out independently, the data obtained should be similar.

3. Behaviour tests
In the following sections the different domains will be described with their appropriate tests. In this section the different behaviour tests used in depression models will be discussed. These tests will mainly be from the first domain (reward related behaviour) and a part of the second domain (psychomotor tests). In the next section the other symptoms (body weight, sleep and biomarkers) will be discussed.
The forced swim test (FST) and tail suspension test (TST) are not mentioned in this paper, although these are common tests used in depression models. This is because their validity and reliability are very low. In both tests immobility is scored and defined as helplessness behaviour. Immobility is a vague concept and used different among persons, therefore its reliability is low. Some score immobility as total lack of movement, while others score it as small movements. The validity of these tests is also low, firstly because both tests only represent an acute situation and thus not mimic the features of depression mentioned above. Secondly, both tests are high in anxiety and it has been discussed that the behaviour is more due to psychomotor activation and not a decrease in helplessness behaviour. And finally, helplessness behaviour is not a symptom for depression in humans, as described in DSM-IV.

3.1 Anhedonia
Depressed people often show a decreased interest in pleasant things or activities. This is called anhedonia, which is a core symptom of depression. In many animal models anhedonia is measured by using the sucrose preference test or intracranial self-stimulation (ICSS).

3.1.1 Sucrose preference
It is commonly assumed that rodents derive pleasure from sucrose and thus depressed-like rodents will consume less sucrose over a given time compared to the control group. This can be easily measured by providing the animal two bottles, one with normal water and the other with sweetened water. By measuring the intake over several weeks of both bottles, the average consummation of sucrose can be determined. Using chronic treatment with antidepressant drugs can inhibit this decline in consummation of sucrose in the depressive-like groups.

An important confound in this test is the caloric value of sucrose. A decrease of sucrose consumption in the depressed-like animals could also be due to a decreased appetite and not due to anhedonia. Measuring body weight, which will be discussed in section 4, can control for this. A loss of body weight indicates a decreased appetite. So, a difference in sucrose intake without changes in body weight point towards anhedonia. Also artificial sugars can be used in this test, examples are saccharine or aspartame. These are sugar substitutes and have no caloric value.

3.1.2 Intracranial self-stimulation (ICSS)
With ICSS several approaches are possible, like electrical stimulation of basal forebrain or hypothalamus and optogenetic stimulation of ventral tegmental area (VTA). In short, in this test untreated rodents are trained to reward themselves by turning a wheel to get a stable baseline. After treatment (to become depressive-like) the animal can reward itself again. It has been shown that depressive-like animals reward themselves less compared to the control group. Normal responding can be restored by using chronic treatment with antidepressants.

Although the usability of ICSS is lower compared to other tests, due to the surgical-component, its face validity is high because this test can be used for a longer time without the animals becoming tolerant towards the self-stimulation.
3.2 Decreased concentration
A secondary symptom of depression is a decline in concentration. It is believed that this is due to a weakened coupling between thalamus and cortex. Decreased concentration results in lower cognitive functioning, and this can be tested in animal models in several ways. Here, only the Morris water maze and the Y-maze are explained, because these tests have been used a lot in models for depression.

3.2.1 Morris water maze (MWM)
In this test, the animal is put into a water tank and has to find a platform in order to escape the water. During the initial training sessions, the subject is tested for five days, four times per day. Latency to escape is measured and compared between the different groups. Depressive-like animals show longer escape latency and also a slower learning curve compared to untreated animals. On the sixth day the platform is replaced (Reversal training) and animals are tested twice. Here the depressive-like animals again showed longer escape latency. Chronic administration of anti-depressants reversed this effect. Although rats are natural swimmers, mice are not. So when using mice this test might not be the best choice, because then fear also plays a role.

3.2.2 Y-maze (spontaneous alteration)
For this test a field consisting of three arms, called a Y-maze, is used. An animal is placed in one of these arms and behaviour is recorded for 8 minutes. Each arm is numbered (1-3) and the entries are scored. An entry is divined as all paws being inside the arm. A successful alternation is any entering of the three different arms in succession. Percentage of correct alternations is calculated using a prescribed formula: total of alternations/(total arm entries – 2). Depressive-like rodents show a decrease of percentage of correct alternations. Normal responding can be restored by chronic administration of anti-depressants.

3.3 Psychomotor retardation
A decrease in locomotion, called psychomotor retardation, is also a secondary symptom of depression. This can be easily tested in an animal model by tracking locomotion in either a novel arena or in the home cage. Both options will be discussed here.

3.3.1 Open-field test (OFT)
For this test any novel arena can be used, it can be either squared or round and either lines or squares can be used as floor markers. An animal is placed inside the arena and number of lines or squares crossed is measured for 5-10 minutes. Often grooming and rearing behaviour is also measured, to get some information about stereotype behaviour or vertical exploration activity (respectively). All behaviour (crossings, grooming and rearing) is decreased in depressive-like animals and chronic administration of anti-depressants can normalize these behaviours. Because hunger, novelty and fear (anxiety) are common confounds for the OFT, lately the usage of this test for depression is questioned.
3.3.2 Elevated plus-maze (EPM)
In this test an elevated arena with four arms (like a +) is used, approximately 50 centimetres above ground. Two arms have walls around it, they are ‘closed’. The other two arms are ‘open’, they have no walls and the animals can see the flour beneath. During 10 minutes the behaviour of the test subject is recorded, number of crossings and arm entries are counted. Depressive-like animals spent less time in open arms compared to control and shows less entries in the open arm.\textsuperscript{23,26} This difference can be removed by using anti-depressants.\textsuperscript{23} Because the open arms are seen as aversive for the test animals, anxiety also plays a part in this test.

3.3.3 Locomotion in home cage
In both OFT and EPM, anxiety is a common mentioned confound. To remove this confound, recently it is suggested that locomotion should only be tested in the home cage during the whole experiment. By using a tracking device (like an actometer) this can be done mechanically. Rodents are active at night, so movement should be tested across the dark cycle. Because depressed people show altered sleep patterns, the data should be corrected for total time spend awake during the dark cycle to correct for confounds.\textsuperscript{5} By using this method, it has been showed that depressive-like animals have a decrease in locomotion.\textsuperscript{27–29} Chronic treatment with antidepressants reverses this effect.\textsuperscript{30,31}

3.4 Conclusions
Although there are a lot of possibilities when testing depressive-like behaviour in rodents, here only the common used tests are described. Above the strength of each test has been discussed using the publications available. Face validity is high when the test is comparable with the human situation. Predictive validity is high when the changes can be reversed by chronic administration of antidepressants, but not by acute treatment. Usability is high when duration and effort are low and reliability is high when different papers get the same results. Construct validity is not mentioned here, because it is not applicable with behaviour tests. Evaluation of different papers results in a scoring for each test. Sucrose preference, for example, has high face validity, because it models a core symptom of depression and can be used for a longer period. It has medium reliability, because not all papers show convincing results. In this way all tests has been evaluated, which results in a scoring for each test. An overview is given in table 2.

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
& Face validity & Predictive validity & Usability & Reliability \\
\hline
Anhedonia & * & ** & \* & ** \\
& *** & *** & *** & *** \\
Intracranial self-stimulation & * & ** & \* & *** \\
\hline
Morris water maze & ** & ** & *** & *** \\
Y-maze & *** & *** & *** & ** \\
Decreased concentration & ** & ** & *** & ** \\
Open-field test & ** & ** & *** & ** \\
\hline
\end{tabular}
\end{table}

Psyc
Plom
\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
& Face validity & Predictive validity & Usability & Reliability \\
\hline
Anhedonia & * & ** & \* & ** \\
& *** & *** & *** & *** \\
Intracranial self-stimulation & * & ** & \* & *** \\
\hline
Morris water maze & ** & ** & *** & *** \\
Y-maze & *** & *** & *** & ** \\
Decreased concentration & ** & ** & *** & ** \\
Open-field test & ** & ** & *** & ** \\
\hline
\end{tabular}
\end{table}
The aim of this comparison is to find a good behaviour test for each domain. Although ICSS has high reliability, its usability is relatively low, due to complicated surgery. Therefore the sucrose preference test is recommended, because it can be carried out very easily and over a longer time. It should be noted that when using this test, it is important to take body weight into account or to use artificial sugars.

When measuring decreased concentration, the Morris water test is recommended when using rats. Because mice are not natural swimmers, this tests results in high anxiety for them. Therefore the Y-maze should be preferred here, despite of its lower reliability.

Although depressed people often show symptoms of anxiety, this is not part of the diagnostic criteria for depression in humans. To avoid anxiety, locomotion should be tested in the home cage. This also makes long-term assessment of locomotion possible.

### 4. Physiological measurements

Next to behaviour measurements it is important to also measure physiological changes to validate an animal model. Recently there is an increased attention for physiological changes in depressed patients, mainly to improve anti-depressant therapy. Therefore it is easier to measure this in animals, because the mechanisms behind these changes are clearer. From the homeostatic domain, the psychomotor retardation is already explained above. Here body weight and sleep patterns will be explained. Also the last domain, neurological measurements, will be explained.

#### 4.1 Changes in body weight

Although a small group of depressed patients show a decrease in body weight, this is not seen as a major symptom of depression. Therefore, in animal models, body weight should be used as a control for other tests and not for a measurement for depression. There are many publications available that show no difference in body weight in their model for depression. Body weight measurements can be used to control for changes in appetite when using the sucrose preference test. When no difference in weight is found, the difference in sucrose intake can be seen as anhedonia.

When using a cytokine-induced model, the immune system is triggered. This results in sickness and thus in sickness behaviour. This behaviour has a lot of similarities with depressive-like behaviour, but only sick animals show a reduction in body weight. Measuring body weight can be used to separate sickness behaviour from depressive-like behaviour. The first few days showing weight reduction can be called sickness behaviour. Only when this reduction is gone, the model shows depressive-like behaviour.

### Table 2 Overview of validity, usability and reliability of all behaviour tests discussed. Scoring is based on all papers mentioned for each test (* = low, ** = medium, *** = high). Construct validity is not mentioned here, because this is not applicable with behaviour tests.

<table>
<thead>
<tr>
<th>Test</th>
<th>Validity</th>
<th>Usability</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevated plus-maze</td>
<td>**</td>
<td>*</td>
<td>***</td>
</tr>
<tr>
<td>Locomotion in home cage</td>
<td>***</td>
<td>**</td>
<td>***</td>
</tr>
</tbody>
</table>

The values indicate the level of validity, usability, and reliability for each test.
4.2 Altered sleep patterns
Changes in the sleep pattern are common for depression. This can be insomnia, hypersonnia or both. Insomnia is divined as sleeplessness during normal sleep hours. Patients have trouble falling asleep or wake up earlier than normal (>2H earlier). Hypersonnia is divined as excessive daytime sleepiness.\textsuperscript{35,36} In rodents this can be measured by recording sleep patterns using implanted electrodes. If there is hypersonnia, rodents show more sleep during dark phase compared to control. If there is insomnia, rodents show less sleep during light phase compared to control.\textsuperscript{27,37,38}

4.3 Neurological measurements
An animal model has high construct validity, when the symptoms are mediated by the same neurobiological mechanisms as in humans. To achieve this, several factors can be measured, either in blood plasma or in the brain.

4.3.1 Brain markers
Depressed people show neuronal atrophy and decreased neurogenesis in the hippocampus and prefrontal cortex. It is hypothesized that this is due to a decrease of brain-derived neurotropic factor (BDNF) expression in the hippocampus. BDNF is critical for growth and survival of neurons in the adult brain. Expression of BDNF is regulated by cAMP-response element binding protein (CREB), which is also lowered in depressed people.\textsuperscript{39,40} Decreased neurogenesis can be measured in animal models by injecting thymidine (bromodeoxyuridine) two weeks before sacrifice. Thymidine binds to DNA during S-phase of the cell cycle. After sacrifice brains are measured for thymidine, more thymidine means more neurogenesis. In depressed brains, less neurogenesis is found.\textsuperscript{25} BDNF and CREB levels can be measured with reverse transcription polymerase chain reaction (RT-PCR). Both are lowered in depressed brains. It is also shown that anti-depressants can reverse this.\textsuperscript{21,41}

4.3.2 Molecular markers
Depressed people also show a dysfunctional hypothalamic-pituitary-adrenal axis (HPA axis). This includes impaired inhibition of cortisol release, higher baseline corticosterone values and higher levels of ACTH.\textsuperscript{42} In animal models ACTH, corticosterone and cortisol can be measured by using ELISA (which uses blood plasma).\textsuperscript{25} The HPA axis is involved in the regulation of immune responses, and thus the production of cytokines. It is logical that when the HPA axis is dysfunctional, there will be an altered cytokine production. In humans an increase of proinflammatory cytokines in serum is found.\textsuperscript{7,43} In animal models TNF-\(\alpha\) and IL-6 are often measured in plasma. An elevation of both is found in depressive-like animals. This can be brought back to normal by using anti-depressants.\textsuperscript{25}

5. Models for depression
When using a model for depression, a lot of different designs are possible. Scientists have tried to create better models, by using everything that is known about depression in humans. The largest group of models is based on stress-induced depression, which is a common cause of depression in humans. Changes
in HPA axis activity are also important in the pathogenesis of depression, by mimicking these changes it is tried to create a different, not stress based model. Lately, the immune system is thought to play a big role in depression, although it is still not clear if it is a cause or consequence of depression. What is known, is that triggering the immune system, mainly in the brain, can cause depressive-like behaviour in animals. Of all three categories, several models will be discussed. Although there are a lot more models available, only the most used and accepted models are discussed here.

5.1 Stress models
The main cause of depression in humans is long-term stress. This stress is often a combination of increasing pressure at work and trouble in private live. Because stress is often mentioned in depression, it is logical that several scientists use this to induce depressive-like behaviour in rodents. The most used model uses unpredictable chronic stress, but social defeat, foot shocks and learned helplessness are also used often. All these models will be explained in this section. Also the validity, usability and reliability of each model will be discussed.

5.1.1 Unpredictable chronic mild stress (UCMS)
In this model rodents are exposed to 6-9 mild stressors in a random order over 2-5 weeks. By using a random order, the stressors can be used more than once without the animal getting accustomed to the stressor. The UCMS model results in an anhedonic state, showed by a reduced sucrose intake without weight differences compared to the control group. There is also a longer escape latency in the Morris water maze. Reduced locomotion in the home cage and both insomnia and hypersomnia are found. This model also results in a reduced BDNF and CREB expression and reduced neurogenesis in the hippocampus. Higher levels of plasma ACTH, IL-6 and corticosterone are also found.

The differences observed in this model compared to control can be reversed by chronic treatment with antidepressants. Altogether, this model scores high in validity. In humans chronic stress is a common cause for depression, therefore this model has high construct validity. Also, the neurological changes in this model are also shown in humans. Face validity is high, because an anhedonic state, longer escape latency (MWM) and reduced locomotion is found in this model. Finally, this model has high predictive validity because chronic treatment with antidepressants reverses the depressive-like behaviour and neurological changes of this model.

The usability of this model can be questioned. Most stressors can be carried out without being present, but altogether it takes a lot of work before you have depressive-like animals.

Next the reliability is also questionable. Although there are a lot of publications about the UCMS model, every laboratory uses a different setting. Some use different stressors and others use different time spans. Therefor comparison between these different settings is debatable.

5.1.2 Chronic social defeat stress (CSDS)
The CSDS model uses a defeat setup to induce depression. Here, rodents are exposed to a novel aggressive animal for 10 minutes per day over 10 days.
Hereafter the animals show depressive-like behaviour, like anhedonia\textsuperscript{47,48}, reduced alternation\textsuperscript{49}, reduced locomotion in the home cage and altered circadian rhythm.\textsuperscript{29} A reduced BDNF and CREB\textsuperscript{49–51} expression is found, plus higher plasma corticosterone\textsuperscript{52}, IL-6 and TNF-\textgreek{a}.\textsuperscript{51} All these changes can be removed by chronic administration of antidepressants.\textsuperscript{8,49} The predictive validity of the CSDS model is high; several studies have shown a chronic (but not acute) effect of antidepressants. The face validity of this model is moderate. Although a lot of publications demonstrate the presence of anhedonia, results for reduced cognitive function and reduced locomotion are not strong. The construct validity is also moderate. There is good evidence that in this model the same neurological pathways are affected, but it is not sure it models depression. It could also be related to psychiatric syndromes like social phobia.\textsuperscript{8} More research is necessary to exclude this. The model is high in usability; it is relatively easy to induce depressive-like behaviour. It does not take much time or effort. It is also a reliable model, the publications available show the same results and all use the same protocol to induce depressive-like behaviour.

5.1.3 Chronic restraint stress (CRS)
A different way of creating stress-induced depression is by restraining rodents in a tube for 4–6 hours per day for 3 weeks. The animals have no way to escape, and this results in depressive-like behaviour, like anhedonia.\textsuperscript{8,53} This model also creates a lower BDNF expression\textsuperscript{53,54}, increased plasma corticosterone\textsuperscript{55,56} and increased TNF-\textgreek{a}.\textsuperscript{55} These changes can be reversed by chronic administration of antidepressants. Lately this model has been questioned, because it is high in anxiety. The animals have no way to escape the stress, which results in subordinate behaviour. Also this model uses the same stressor for several weeks and the animals could get used to this. This is proven by several experiments and in the end it results in increased locomotion and increased neurogenesis.\textsuperscript{8,55,57,58}
Although there are some results corresponding to depression, there are also contradicting results. As a consequence there is a lower face validity and construct validity. The predictive validity is good, because the papers that use antidepressants all show an effect after chronic treatment. The usability of this model is medium. It does take several weeks to induce depression, but the protocol does not take much effort. The reliability of this model is high, the same protocol is used in the different papers and these show the same results.

5.1.4 Chronic foot-shock stress (CFS)
By exposing rodents to chronic foot-shock stress (CFS), depressive-like behaviour can be induced. In this model rodents are exposed to daily foot-shocks for 3 weeks. On each day, the animal is placed in a novel cage for 2 hours and in this time 5 shocks are applied, 8 seconds per shock. On each day the starting time and the interval between the shocks are randomized, so the stress is seen as unpredictable. After this procedure, rodents show depressive-like behaviour. Reduced sucrose preference, reduced successful alternations (y-maze), longer escape latency (MWM) and reduced locomotion are found in this model.\textsuperscript{59–61} Higher ATCH,
higher corticosterone and reduced neurogenesis are also seen.\textsuperscript{62–64} All these changes can be reversed by chronic treatment with antidepressants. Most publications use CFS to research neurological symptoms of depression, only a couple of studies test behaviour changes in this model. So both face validity and reliability of this model can be questioned. The construct validity of this model is low, although the same neurological changes are found in this model as with depression in humans. This is because CFS does not seem to be ethologically relevant.\textsuperscript{8} Both predictive validity and usability of this model are high. Chronic administration of antidepressants is effective in this model and it does not take much effort to induce depression.

5.1.5 Learned Helplessness (LH)

One of the earliest animal models for depression was learned helplessness (LH). This model is based on the fact that depressed people often feel helpless and do not respond to positive events. In this model rodents are exposed to uncontrollable, inescapable shocks for an hour per day over three days. After this, the animals are placed in a controllable, escapable situation but fail to escape or show effort to escape this situation. The depressive-like animals show a reduced sucrose preference, longer escape latency (MWM) and reduced alternation (Y-maze).\textsuperscript{5,9,21,65} Also a reduced expression of BDNF and CREB are found.\textsuperscript{21} After chronic treatment with antidepressants these changes disappear. Major criticism is found for this model, because only a small percentage of animals indeed develop learned helplessness next to the depressive-like behaviour. Also, this model only works with a selective number of strains.\textsuperscript{8} It looks like this model is high in face, construct and predictive validity, but this can be questioned, because not all animals subjected to this protocol become depressive-like (as mentioned). Therefore the usability and reliability are low.

5.2 Olfactory Bulbectomy (OB)

Reduced or loss of smell is common found in neurological disorders, like Alzheimer’s disease. Lately it is found also in depressed people.\textsuperscript{66} It has been proposed that changes in olfaction are caused by HPA axis dysfunction.\textsuperscript{8} In rodents, removal of the olfactory bulb has the same effect, and is therefore often used as a model for depression. A longer escape latency (MWM) and reduced alternation (Y-maze) is found in OB rodents, as is increased nocturnal activity. Also a reduction of CREB expression and a lower neurogenesis is found.\textsuperscript{22,24,67} These changes can be reversed by chronic administration of antidepressant. It could be argued that both face and construct validity of this model are good, but because this model mainly shows neurological changes it is questionable. The predictive validity is good, as is the usability and reliability of this model. Although it has to be mentioned that for this model a skilled person is necessary to preform the bulbectomy without damaging other brain structures.

5.3 Cytokine induced depressive-like behaviour

There is a growing interest in the role of the immune system in the pathogenesis of MDD. This gave rise to the ‘cytokine hypothesis of depression’, which suggests that proinflammatory cytokines are the key factor for the behavioural, neuroendocrine and neurochemical changes in depression.\textsuperscript{7,68–71} This is supported by the fact that several medical illnesses characterised by chronic inflammatory responses are often accompanied by depression. Although this
hypothesis is able to account for most of the symptoms occurring in MDD, it is still not clear if cytokines are the cause of depression or are just a side effect.\textsuperscript{7,8,68–71}

By injection of Lipopolysaccharide (LPS) into the brains of rodents, depressive-like behaviour can be induced. LPS triggers the immune system and activates the microglia, part of the brain immune system. Firstly this results in sickness behaviour, which can be accompanied by fever and results in body mass loss. After a few days the fever and the loss of weight disappear, but behaviour changes stay. From this moment the rodents show depressive-like behaviour, like lowered sucrose preference and reduced alternation (Y-maze). Higher IL-6, higher TNF-\(\alpha\) and reduced BDNF expression is also found,\textsuperscript{33,72–75} after chronic treatment with antidepressants, the depressive-like behaviour disappears. Although this model is high in validity, the usability and reliability of this model remains questionable. Although there are enough good results available, there are also a lot of contradicting results. There are not many studies that could successfully reproduce other publications. Also, the injection of LPS in the brain is a difficult procedure and often a couple of animals die as a result.

6. Discussion
In this paper it has been tried to give an overview of the common used tests and animal models for depression. It has to be noted that this is just a small overview, so not every possible test or model is discussed here. When discussing the different behavioural tests, a summary of the best ones is given (see table 2).

This summary and the psychological measurements have been used as a guideline when discussing the different models. For each model several papers have been used, each containing at least one behavioural and one psychological test. With this information the validity, usability and reliability of each model are assessed and discussed. An overview is given in table 3.

The most used model in the literature is the UCMS model. It is believed that this model works the best, because it uses the main cause of depression, namely stress, to induce it. Also, the same behaviour and physiological changes have been seen, compared to humans. These changes can all be treated with chronic antidepressants. With this knowledge, it is concluded that the validity of this model is high. But not all papers show the same results, so the reliability of this model can be questioned. This could be because not each paper uses the same protocol. There is variation in both time and number of stressors used. Therefore it is recommended that a more general protocol should be used, to increase the reliability of this model. It should be noted that this model is low in usability, because it takes a lot of effort to induce depression. When working alone, the protocol takes 4-6 hours per day over 4 weeks. But when working in a group, it should not be a problem.
The other stress models all show a lower validity. These models have lower construct validity, because the stressors used are less relevant in depression (foot-shock stress for example) or because there are not the same neurological changes present. These models have lower face validity, because not all behaviour tests show good results or have been preformed.

In conclusion the UCMS model is probably the best stress induced model, although it is labour intensive (and therefore has a lower usability). When it is impossible to use this model due to its effort, the CSDS model is the best next thing. CSDS has high reliability, usability and face validity, but face validity and construct validity are moderate (see discussion in section 5.1.2).

Concerning models that do not use stress, the OB model is the most used. This model has a lot of contradicting results concerning behaviour changes. Therefore this model is only recommended to use when interested in neurological changes in depression.

The cytokine-induced model shows good promise for a future model of depression, but more research is necessary before being able to use it. A good protocol for this model is not available, and therefore a lot of different ones are used. It is believed that with some effort, this model could be a good replacement for the UCMS, it also has the advantage that it is much less labour intensive.

In this paper, no distinction has been made between different strains of rats or mice. In humans, depression is not only dependent on the situation, but also on the individual. An event only becomes negative if the caused emotions are negative. These emotions depend on the individual’s cognitive appraisal. Therefore it is logical that in animals this also plays a role and it is to be expected that different strains show different results. In the future it is recommended to compare different strains for each model, to find the best strain to work with.

Also the usage of other animals beside rodents is not discussed here. There are a couple of reasons for this. Firstly, the research with other animals beside rodents is not extensive and making a good comparison is not cumbersome. Moreover, if other animals should be discussed, it is hard to decide which animals should and which should not be mentioned. But it could be interesting to look into the models available that do not use rodents. In potency these animal models are even better than the ones discussed here.
References


68. Maes, M. Depression is an inflammatory disease, but cell-mediated immune activation is the key component of depression. *Progress in neuro-psychopharmacology & biological psychiatry* 35, 664–75 (2011).


