Journal of Neuroscience Methods 180 (2009) 82–86

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Automated behavioral analysis of limbs' activity in the forced swim test

Roman Gersner, Michal Gordon-Kiwkowitz, Abraham Zangen*

Department of Neurobiology, The Weizmann Institute of Science, Rehovot 76100, Israel

ARTICLE INFO

Article history:
Received 19 October 2008
Received in revised form 31 December 2008
Accepted 5 March 2009

Keywords:
Swim test
Depression
Antidepressant
Gaussian distribution
Animal model
Computer vision and automatic tools

ABSTRACT

The forced swim test in rodents has been widely used to evaluate potential effectiveness of antidepressant medications since it was described in 1977 by Porsolt. In this test, a rodent is placed in a water container, and its immobility time is measured. The immobility time indicates the level of inactivity, interpreted as 'hopelessness', and has been shown to decrease when the rodent is treated with antidepressant medications. The simple measure of immobility time does not take into account intermediate behaviors during testing (ranging from total immobility to extensive 'struggling' behavior) and does not show normal Gaussian distribution in tested groups of rats.

We have previously developed a software allowing an observer to assign scores to the full range of intermediate behaviors by continuously reporting the motion of the limbs using a joystick. Based on the joystick score, we have now developed an automatic tool that uses computer vision algorithm (CVA) to analyze specifically the motion of the limbs and generate an objective, reproducible and automated score.

In the current study we have analyzed data obtained during the swim test using the traditional immobility time, the joystick analysis and the new CVA method to test the distribution of scores in a group of 20 rats. In addition, we tested the effects of various medications using these different scoring methods. The CVA method has been validated and positively correlated with the joystick score. Data obtained using the CVA method is objective, reproducible, and significantly reduces the time required for human analysis.

1. Introduction

For several decades, research communities and pharmaceutical industries have used the forced swim test to evaluate potential new antidepressant medications (Borsini et al., 1991; Dhir and Kulkarni, 2008; Porsolt et al., 1977a; Wong et al., 2000). The paradigm consists of a pretest session, in which a rat is placed into a cylinder filled with water for 15 min and a test session 24 h later, in which the rat is placed again in the same tank for 5 min. The test is based on the observation that rats, following initial escape-oriented movements, develop an immobile floating posture in the water cylinder. When they are placed again in the testing apparatus 24 h later, they resume this posture quickly. This posture was interpreted by Porsolt et al. (1977b) as reflective of the animal's state of despair, elicited by its perception of the hopelessness of the situation learned during the first session. The total amount of time in which the animal demonstrates this behavior is therefore measured and termed 'immobility time'.

Acute administration of candidate compounds between the two exposures to the swimming tank may reduce or prevent the development of such immobility (Porsolt et al., 1978). Tricyclic antidepressants, monoamine oxidase inhibitors, and atypical antidepressants such as mianserine and iprindole, are all effective in this paradigm (Borsini and Meli, 1988). However, there are several shortcomings with the traditional type of measurement and several compounds have been identified as generating false positive (Borsini and Meli, 1988; Delini-Stula et al., 1988; Panconi et al., 1993) or negative (Lucki, 1997) results. In his original paper, Porsolt described the immobility posture of the rat "only [as] those movements necessary to keep its head above the water" (Porsolt et al., 1977b). The amount of time that the animal spends immobile can indicate, according to Porsolt, a state of despair in which the rat has learned that escape is impossible. In order to remain afloat, the animal makes certain, slight swimming movements that are less relevant to its escape behavior, than active swimming using its limbs. The observer in the traditional Porsolt test does not account at all for the time the animal spends during the active behavior. In order to monitor and measure other behaviors during the swim test (including 'intermediate' behaviors that cannot be defined as clear struggling or clear immobility), a more flexible measure is needed.

Some efforts to overcome these limitations and to enhance the sensitivity of the traditional Porsolt paradigm have been previously reported. Some methods try to give a broader definition to the active part of the behavior. For example, Detke and Lucki (1996) and Cryan et al. (2005) measured different types of behaviors, divided into swimming, climbing and immobility. They define the behavior for each interval of 5 s during the swim test, as one of the

* Corresponding author. Tel.: +972 8 934 4415; fax: +972 8 934 4131.
E-mail address: a.zangen@weizmann.ac.il (A. Zangen).

0165-0270/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2009.03.003

above-mentioned three behaviors according to which was more prominent in the 5 s period. This method was shown to allow detection of antidepressant effects of medications that were not effective in the traditional scoring method.

Other methods attempted to make the measurement more physical and objective. Shimazoe et al. (1987) used tremor sensors surrounding the cylinder to record water vibrations while rats were swimming in it. De Pablo et al. (1989) measured the variation in the frequency of the natural electromagnetic field of water induced by movements of rats. Hedou et al. (2001) measured the distance that the animal moved using a special tracking system and software (Ethovision by Noldus). The advantage of such automated methods is that they are not biased by subjective or inexperienced observers. However, the first two methods can measure only a defined area of the water tank, and the third method measures only the total activity or distance that the animal moved. There are situations in which the animal makes vigorous movements with its paws within one area of the swimming tank for long periods without significant changes in the position of the entire body, or with relatively small effects on water vibrations in other parts of the tank. Such periods are not evaluated properly using the above-mentioned approaches. In addition, all of these methods require complex, expensive, and dedicated equipment.

In our recent study (Gersner et al., 2005), we proposed the joystick analysis, which allows measurements of the full range of the intermediate behaviors (from total immobility to an extensive struggling behavior), and was validated against the traditional Porsolt paradigm. The ability to generate any number between 0 and 100 creates a continuous score rather than a binary score of the activity and avoids labeling the behavior into three types as in the method suggested by Detke and Lucki (1996). Unlike the immobility time measures, which did not show a normal Gaussian distribution in tested groups of rats, the joystick score generated a normal Gaussian distribution, which is of a statistical value. Although the joystick score is valid, dependable and has been successfully used in recent works (Lewitus et al., 2008; Toth et al., 2008), it has disadvantages: it is time-consuming, and like all other measures that are not automated, it involves subjective decisions of the raters and therefore might be less reproducible and less accurate.

To overcome these problems and still measure accurately specific activity of the limbs only (and not the whole body due to the caveats mentioned above), we have established a novel automated method, which is based on the sensitive joystick score of limbs' activity. Due to its length, measurement of the tail movements would be the dominant factor of activity and limbs movements would become negligible in the final score. Moreover, tail movements is mostly required for animal's balance and the aim of our measurements is to monitor activity that is interpreted as animal's attempt and 'motivation' to struggle and escape from an unpleasant situation. We have therefore considered the tail movements largely as unspecific behavior. Considering these goals, we cooperated with *ProTrack*, a vision technology company, to create an automated measure based on a computer vision algorithm. Given a video file shot from below the water tank we developed a software based on the following specifications: (1) score high frequency movements of the front and hind limbs; (2) disregard whole-body turns and movements of the tail; (3) analyze videos taped from a position below the water tank; (4) ignore noise factors caused by water movement, differences in illumination and animal feces in the water tank, so that these would not contribute to the final score.

Methods from computer vision using proprietary algorithms of Protrack Ltd. were used to score the animal movement, thus achieving an objective and reproducible score that required only the video file to be analyzed.

2. Materials and methods

2.1. Animals

Male Sprague–Dawley (SPD) rats (300–310 g, 10 weeks old) were supplied by the Animal Breeding Center of the Weizmann Institute of Science. They were maintained under a 12 h/12 h light dark cycle (lights on at 8 a.m.). Food and water were supplied *ad libitum*. Animals were housed in groups of three in Perspex cages (18 cm × 26 cm × 40 cm). Testing was performed between 11 a.m. and 2 p.m. All animal experiments were conducted according to the Institutional Care and Use Committee, in complete accordance with NIH guidelines for care and use of laboratory animals.

2.2. Forced swim test procedure

The water tank was similar to that described by Detke and Lucki (1996). The cylinder tank (49 cm high and 19 cm in diameter, constructed at the Weizmann Institute of Science) contained water at a temperature of 27 °C (2 °C above room temperature). The water depth was about 30 cm and was individually calculated according to the animal's weight so that only the animal's tail reached the bottom of the tank.

In the first experiment for evaluation the distribution of activity or immobility during the swim test, 20 rats experienced a single, 10 min exposure to the water tank.

In the second experiment for evaluation of medication activity, an additional group of 24 rats were exposed to the water tank for 15 and 5 min on two consecutive days as described in the Porsolt paradigm. After each exposure, rats were dried and returned to their home cages. Medications were prepared and administered during the period between the two sessions.

Desipramine, promethazine and fluoxetine were administered i.p. in doses of 15, 10 and 20 mg/kg, respectively. Desipramine was chosen as a tricyclic antidepressant drug whose action is known to be detected in the Porsolt paradigm by simple measurement of immobility time. Fluoxetine was chosen as a representative of the SSRI family which was previously shown to be a false negative in the traditional Porsolt paradigm (Lucki, 1997). Promethazine was chosen as a stimulant drug which was shown to be a false positive in the Porsolt paradigm (Delini-Stula et al., 1988).

Desipramine and fluoxetine were given 23, 5 and 1 h before the test. Promethazine was given 1 h before the test (Nomura et al., 1982) and animals in this group received saline injections 23 and 5 h before the test. Three saline injections were administered to the control group. Each treatment group included six male SPD rats.

2.3. Scoring the behavior

Animals were observed and videotaped using a stable camera placed underneath the transparent swimming tank. The video conditions (mainly light, zoom and water transparency) were controlled. The room was dark and two lamps providing deem light were placed at opposite sides of the water tank, to illuminate the animal.

The camera and zoom were consistently positioned to include as much of the animal movement and as little of the cylinder tank walls (that might produce reflections) as possible. Therefore, we centered the camera below the cylinder so that the view width included the full diameter of the cylinder base.

Three types of scoring were applied in each session.

2.3.1. Traditional Porsolt scoring

The immobility time was counted using a timer and shown in seconds, as in the traditional Porsolt test. Measurements were validated by an additional (blinded) observer who measured immobility time using the video recording of each test.

Immobility was defined as "floating motionless or making only those movements necessary to keep its head above the water" (Porsolt et al., 1978).

2.3.2. Joystick scoring

The joystick scoring method has been described before (Gersner et al., 2005). Briefly, taped films were transferred to a personal computer using *Windows Movie Maker*. The activity of animals was scored continuously throughout the swim test by observers (blinded to the Porsolt scoring data), using a custom-designed code (Labview, National Instruments, Austin, TX). The joystick (made by the unit of logistics and research services of our department) could be moved in two directions (forward and backwards). This resulted in corresponding changes on a 0–100 visual-analog scale.

Graphs of motion-to-time were generated according to joystick movements. Motions of the front and hind limbs were scored separately. The guidelines for scoring were defined according to the frequency and the amplitude (the change in the position of the limbs in degrees) of movements (Gersner et al., 2005). The intermediate score was the sum of the areas under the curves (score integrated over time). The final score also took into account the number of diving attempts. Each dive added 2500 points to the score (a factor that adds approximately 5% to the average score for each diving attempt). Therefore the final score was:

$$\text{score} = \text{area(front)} + \text{area(hind)} + \text{dives no.} \times 2500.$$

2.3.3. Computer vision algorithm (CVA) scoring

The same taped films (30 frames per second) were analyzed using the novel CVA software developed with *the ProTrack* company. Thus, a numeric score was obtained automatically.

The video analysis process consists of a number of steps. First, a high-pass image filter is applied in order to reduce changes caused by illumination variability. Second, image filters were used to verify that effects of water turbidity that contributes to movements in the video, would not contribute to the final score. Then, the motion analysis is accomplished by filtering out the slow movement, which usually originates from the whole-body motion, leaving the fast movements of the limbs and the tail. The slow movement filtering also removes most of the movement produced by dirt and feces, as their movement is slow as well. The movement detection is performed in predefined time windows, measuring the motion frequency in each time window and filtering out the tail movement, as it is a fast-moving element that should not contribute to the final score, as explained in Section 1. The score for each frame is the sum of the fast movements excluding the tail area.

The output of the program is a text file for each movie with a score for each time interval as configured by the user, which sums the score for the frames in the interval. The time intervals can range from 1 s to the full movie length. We display in our results a single interval of the test length (5 or 10 min according to the two experiments).

The final score we present takes into account the total CVA score and the number of diving attempts. Each dive adds 40,000 points to the score (a constant factor based on CVA score range), as in the joystick paradigm. Our tests show that the automatic score given to a struggling animal and to a diving animal is similar to the joystick score (data not shown):

$$\text{score} = \text{automatic score} + \text{dives no.} \times 40,000.$$

2.4. Statistical analysis

In the first experiment, we used the Shapiro–Wilk test for normal distribution (Shapiro and Wilk, 1965) that rejects a normal distribution for *p*-value less than 0.05.

In the second experiment, one-way ANOVA followed by Fisher's least significant difference (LSD) post hoc test was applied. The post hoc analysis of the desipramine, fluoxetine and promethazine groups was always compared to the corresponding saline group. *p*-Value of less than 0.05 was considered statistically significant. Statistical analysis was done with a StatView 5.0 software.

3. Results

3.1. Distribution of the swimming scores using the different analysis methods

In the first experiment we measured correlations between the joystick score and the computer vision algorithm (CVA) score and the distribution of results within a group of 20 rats that were placed for 10 min in the swimming tank. All three scores: immobility time, joystick score and the CVA score were computed for each rat in the experiment. Because the CVA method was created to automate the joystick score, we first tested the correlation between these two methods. We found a significant positive correlation ($p < 0.0001$) between the two scores (Fig. 1).

In accordance with our previous findings (Gersner et al., 2005), we found that the immobility time (Fig. 2A) scoring did not show a normal Gaussian distribution as revealed by the Shapiro–Wilk test ($p < 0.0001$). In contrast, the joystick score was not significantly different from a normal Gaussian distribution (Fig. 2B). Similarly the CVA score resulted in a normal Gaussian distribution of the data (Fig. 2C).

3.2. The effect of antidepressant medications on activity during the swim test as measured by the various methods

In the second experiment we have used the traditional Porsolt paradigm to evaluate the effect of desipramine, fluoxetine and promethazine on the swimming behavior during a 5-min test (that was performed 24 h after a 15-min exposure to the swimming tank). The effect was measured by the three scoring methods: immobility time, joystick score and CVA score (Fig. 3).

One-way ANOVA performed on the immobility times (Fig. 3A) indicated a significant main effect of treatment ($F(3, 20) = 11.35$, $p < 0.0001$). Post hoc analysis revealed that the desipramine and the promethazine treatment groups had significantly less immobility time than that of the saline group ($p = 0.001$, $p = 0.0001$,

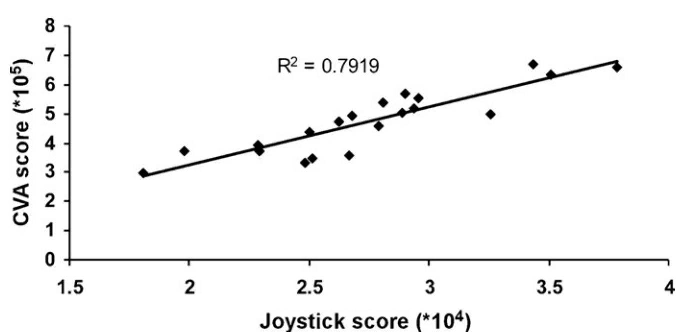

Fig. 1. Correlation between the joystick and the CVA scores. The joystick and the CVA methods were applied on video recordings of the same rats ($n = 20$). The best linear fit is presented.

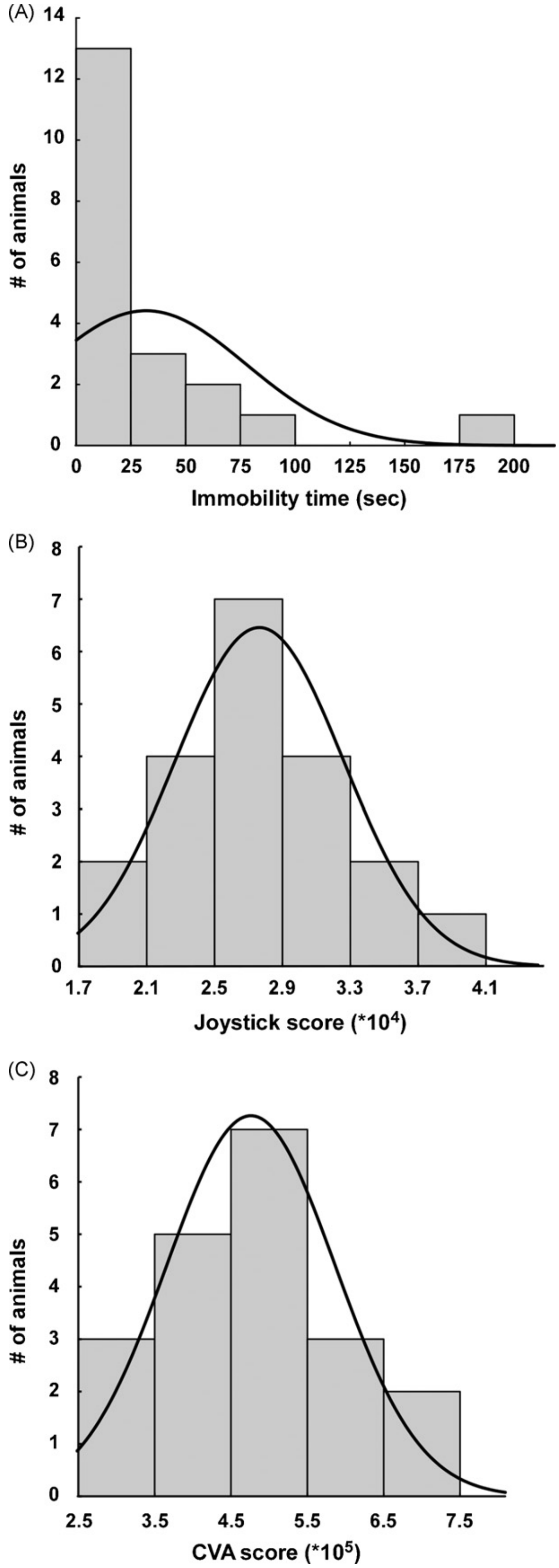

Fig. 2. The distribution of results based on measurements of the immobility time (A), the joystick score (B), and the CVA score (C) in a population of 20 rats, is presented. The line presents the best fitting normal curve, based on the average and the standard deviation of the raw data obtained from 20 rats.

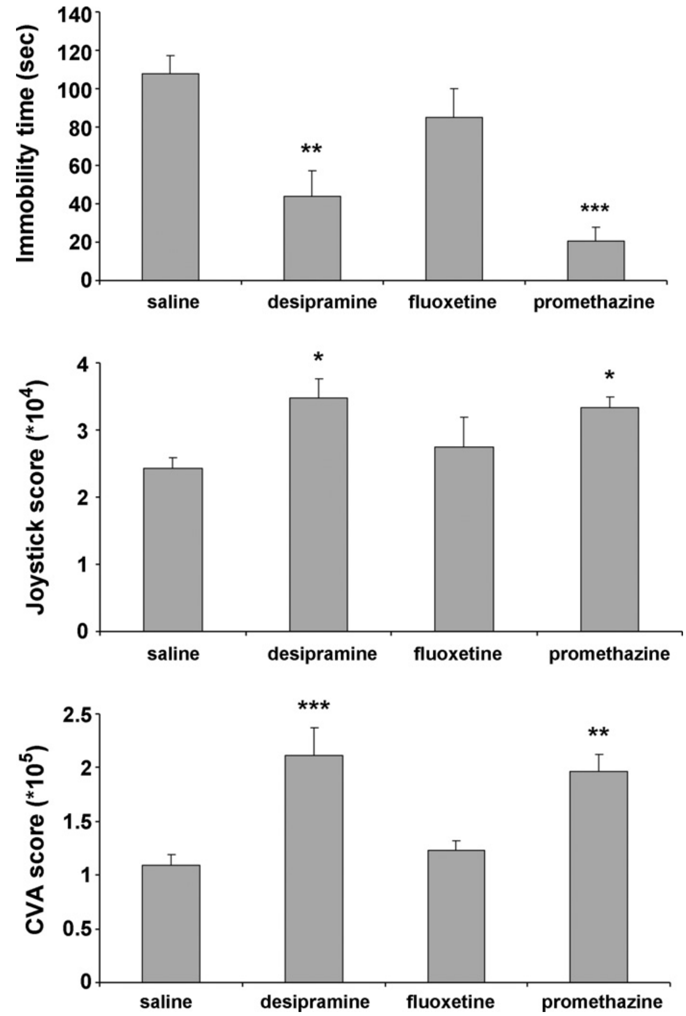

Fig. 3. Measurements of immobility time (A), joystick score (B) and CVA score (C) in the Porsolt swim test. The acute effect of treatments with desipramine, fluoxetine or promethazine are compared to saline-treated controls. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

respectively). The immobility time of the fluoxetine group was not significantly different from that of the saline group.

One-way ANOVA performed on the joystick scores (Fig. 3B), also revealed a significant treatment effect ($F(3, 20) = 2.96$, $p = 0.05$) and the post hoc analysis revealed that activity in the desipramine and the promethazine groups was significantly increased relative to the saline group ($p = 0.017$, $p = 0.038$, respectively). However, in accordance with the immobility time scores, the fluoxetine treatment did not have a significant effect (relative to saline).

One-way ANOVA performed on the CVA scores (Fig. 3C) showed a significant main effect of treatment ($F(3, 20) = 9.79$, $p = 0.0004$). Post hoc analysis revealed that that activity in the desipramine and the promethazine groups was significantly increased relative to the saline group ($p = 0.0003$, $p = 0.0012$, respectively), while similarly to the immobility and joystick scores, fluoxetine treatment did not show a significant effect.

4. Discussion

Since it was first developed in 1977, the traditional forced swim test has contributed significantly to the development of effective antidepressant drugs (e.g. Hudson et al., 2003; Palaska et al., 2001). However, the guidelines provided for determining immobility, and thus for evaluating depressive-like behavior, overlook many intermediate behaviors which also deserve attention. Porsolt described the immobility posture of the rat "only [as] those movements necessary to keep its head above the water" (Porsolt et al., 1977b). However, this definition fails to differentiate between animals that struggled vigorously during their active periods, and those engaged in small swimming movements. Moreover, the reliability of this test is hampered to a certain extent since it generates a subjective score based upon binary interpretation by a rater of animal behavior. For this reason, the new software analysis tool for the FST, an extension of the joystick FST analysis, is suggested here as an alternative that allows automatic (not depending on subjective interpretations of different raters) analysis of the full range of activity levels during the swim test. Since the FST was developed, many researchers have modified either the test itself (Detke and Lucki, 1996; Lucki, 1997), or the way it was scored and analyzed (De Pablo et al., 1989; Hedou et al., 2001; Shimazoe et al., 1987; Detke and Lucki, 1996) in order to overcome the difficulties in the test described above.

The joystick FST was devised to enable detailed and continuous scoring while accounting for the range of intermediate behaviors. Although it is useful (e.g. Lewitus et al., 2008; Toth et al., 2008) and the scores provide a normal Gaussian distribution (unlike the immobility time scoring), the joystick FST is prone to the problems of subjective scoring as well. Moreover, the analysis of activity of front and hind limbs performed by a rater in front of the computer screen is time-consuming. To overcome these problems, the CVA software was developed to automate the joystick score. This software automatically measures the various movements of an animal that has been videotaped during the FST. The difficulty in defining behavior as 'swimming' or 'struggling' as attempted in some software (e.g. Biobserve, Germany) is avoided by providing a continuous score based on motion analysis centered on the relevant parts of the animal (the limbs, that are likely more relevant to expression of the attempt to struggle and escape from the unpleasant situation).

The method distinguishes the high frequency movement from the lower frequency movement and allows separating the tail from the limbs according to the object size. Theoretically the method should be useful for any application with a similar general idea, including measuring mice behavior in the swim test. However, since mice are smaller, adjustments of sensitivity parameters might be necessary, and a setup with mice requires validation.

The new method replaces the traditional immobility score by an active mobility score which allows rating the full range of active behaviors, rather than the traditional binary (immobility or mobility) score. In the setup we use, our method did not show effectiveness of fluoxetine, as does the method of Detke and Lucki (1996). However, since the immobility score is widely used, an objective test may be of great value, even if it is not sensitive to all antidepressant drugs and despite the false positive and false negatives of the swimming test that are known for several decades. The Detke and Lucki method groups behavior into three categories: immobility, swimming and climbing which can be considered as intermediate behaviors, but can also be viewed as a way to label different types of behaviors and not necessarily the full range of 'intermediate' activities. The joystick and CVA methods, on the other hand, create a continuous score ranging from totally inactive to fully active, without labeling the behavior into different types, avoiding the necessity to determine the subjective boundaries between the behaviors and allowing an automated objective measure.

An important goal of this study was to automate the joystick-based scoring system to allow faster, more reliable and reproducible analysis. It is important to note that the swim test is not only used to screen potential antidepressant drugs, but also to measure depressive-like behavior in animal models. Therefore, an automated measure of limbs' activity can be useful for evaluating or developing animal models for depressive behavior.

In conclusion, the CVA software may allow researchers to conduct an accurate, objective, comprehensive, rapid and reproducible behavioral analysis of the swim test.

Acknowledgments

We would like to thank the ProTrack Ltd. and specifically Meir Zilkha for his programming of CVA. This study was supported by the Israel Science Foundation (ISF) and National Institute for Psychobiology in Israel (NIPI). Dr. Zangen is an incumbent of the Joseph and Celia Reskin career development chair.

References

Borsini F, Meli A. Is the forced swimming test a suitable model for revealing antidepressant activity? Psychopharmacology (Berl) 1988;94:147–60.

Borsini F, Volterra G, Lecci A, Evangelista S, Mancinelli A, Cutrufo C, et al. Potential antidepressant activity and enhancement of serotonin uptake of a new dibenzothiadiazepine derivative. Arzneimittelforschung 1991;41:573–80.

Cryan JF, Page ME, Lucki I. Differential behavioral effects of the antidepressants reboxetine, fluoxetine, and moclobemide in a modified forced swim test following chronic treatment. Psychopharmacology (Berl) 2005;182:335–44.

De Pablo JM, Parra A, Segovia S, Guillamon A. Learned immobility explains the behavior of rats in the forced swimming test. Physiol Behav 1989;46:229–37.

Delini-Stula A, Radeke E, van Riezen H. Enhanced functional responsiveness of the dopaminergic system—the mechanism of anti-immobility effects of antidepressants in the behavioural despair test in the rat. Neuropharmacology 1988;27:943–7.

Detke MJ, Lucki I. Detection of serotonergic and noradrenergic antidepressants in the rat forced swimming test: the effects of water depth. Behav Brain Res 1996;73:43–6.

Dhir A, Kulkarni SK. Antidepressant-like effect of 17beta-estradiol: involvement of dopaminergic, serotonergic, and (or) sigma-1 receptor systems. Can J Physiol Pharmacol 2008;86:726–35.

Gersner R, Dar DE, Shabat-Simon M, Zangen A. Behavioral analysis during the forced swimming test using a joystick device. J Neurosci Methods 2005;143:117–21.

Hedou G, Pryce C, Di Iorio L, Heidbreder CA, Feldon J. An automated analysis of rat behavior in the forced swim test. Pharmacol Biochem Behav 2001;70:65–76.

Hudson AL, Tyacke RJ, Lalies MD, Davies N, Finn DP, Marti O, et al. Novel ligands for the investigation of imidazoline receptors and their binding proteins. Ann N Y Acad Sci 2003;1009:302–8.

Lewitus GM, Wilf-Yarkoni A, Ziv Y, Shabat-Simon M, Gersner R, Zangen A, et al. Vaccination as a novel approach for treating depressive behavior. Biol Psychiatry 2009;65:283–8.

Lucki I. The forced swimming test as a model for core and component behavioral effects of antidepressant drugs. Behav Pharmacol 1997;8:523–32.

Nomura S, Shimizu J, Kinjo M, Kametani H, Nakazawa T. A new behavioral test for antidepressant drugs. Eur J Pharmacol 1982;83:171–5.

Palaska E, Aytemir M, Uzbay IT, Erol D. Synthesis and antidepressant activities of some 3,5-diphenyl-2-pyrazolines. Eur J Med Chem 2001;36:539–43.

Panconi E, Roux J, Altenbaumer M, Hampe S, Porsolt RD. MK-801 and enantiomers: potential antidepressants or false positives in classical screening models? Pharmacol Biochem Behav 1993;46:15–20.

Porsolt RD, Anton G, Blavet N, Jalfre M. Behavioural despair in rats: a new model sensitive to antidepressant treatments. Eur J Pharmacol 1978;47:379–91.

Porsolt RD, Bertin A, Jalfre M. Behavioral despair in mice: a primary screening test for antidepressants. Arch Int Pharmacodyn Ther 1977a;229:327–36.

Porsolt RD, Le Pichon M, Jalfre M. Depression: a new animal model sensitive to antidepressant treatments. Nature 1977b;266:730–2.

Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika 1965;52:591–611.

Shimazoe T, Shibata S, Ueki S. A new forced swimming test for the evaluation of antidepressants in rats by recording vibration of a water tank. J Pharmacobiodyn 1987;10:639–43.

Toth E, Gersner R, Wilf-Yarkoni A, Raizel H, Dar DE, Richter-Levin G, et al. Age-dependent effects of chronic stress on brain plasticity and depressive behavior. J Neurochem 2008;107:522–32.

Wong EH, Sonders MS, Amara SG, Tinholt PM, Piercey MF, Hoffmann WP, et al. Reboxetine: a pharmacologically potent, selective, and specific norepinephrine reuptake inhibitor. Biol Psychiatry 2000;47:818–29.