



Research report

The use of a test battery assessing affective behavior in rats: Order effects

Arjan Blokland*, Sanne ten Oever, Dennis van Gorp, Michael van Draanen, Theodor Schmidt, Emily Nguyen, Alexandra Krugliak, Anthony Napolitano, Sarah Keuter, Inge Klinkenberg

Department of Neuropsychology and Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands

ARTICLE INFO

Article history:

Received 1 November 2011

Received in revised form

25 November 2011

Accepted 28 November 2011

Available online 6 December 2011

Keywords:

Anxiety

Depression

Animal model

Order effect

Forced Swim Test

Open Field

Zero Maze

ABSTRACT

Many studies have used test batteries for the evaluation of affective behavior in rodents. This has the advantage that treatment effects can be examined on different aspects of the affective domain. However, the behavior in one test may affect the behavior in following test. The present study examined possible order effects in rats that were tested in three different tests: Open Field (OF), Zero Maze (ZM) and Forced Swim Test (FST). The data of the present study indicated that the behavior in ZM was the least affected by the order of testing. In contrast, the behavior in the FST (and to a less extend the OF) was dependent on the order of the test in the test battery. Repeated testing in the same test did not change the behavior in the ZM. However, the behavior in the OF and FST changed with repeated testing. The present study indicates that the performance of rats in a test can be dependent on the order in a test battery. Consequently, these data caution the interpretation of treatment effects in studies in which test batteries are used.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

There are various tests models that can be used for the evaluation of treatments on anxiety- and depressive-like behavior. These animal models measure different aspects of the affective domain. Thus, tests have been developed for assessing anxiety-related behaviors [e.g. 1–3], depressive-like behavior [e.g. 4–6], and fear/panic [e.g. 7, 8]. Although there are various tests for the assessment of anxiety, it appears that each test may assess a different aspect of the anxiety domain. This has been demonstrated by correlating the behavior across a range of tests [9,10], and the evaluation of drug effects in different tests [11]. Since treatments may affect various aspects of affective behavior, typically a test battery is used to capture treatment effects on the different affective behavioral domains [e.g. 12, 13]. The order of tests is mostly guided by the intuitive idea that animals should be tested first in the least invasive test and subsequently in more invasive tests. However, it could be argued that testing history could influence the performance in the following tasks and that therefore treatment effects should be interpreted with caution.

* Corresponding author at: Neuropsychology and Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Tel.: +31 43 388 1903.

E-mail address: a.blokland@maastrichtuniversity.nl (A. Blokland).

Previously, McIlwain et al. studied order effects by comparing mice that were used in a behavioral test battery with a group of animals that were naïve [14]. The test battery consisted of the following paradigms: neurological testing, Open Field, light–dark task, rotarod, prepulse inhibition, acoustic startle habituation, conditioned fear, Morris task and hot plate test. The comparison between the experienced and age-matched naïve group revealed that some tasks were sensitive to test experience (Open Field, rotarod, hot plate), whereas other paradigms appeared to be insensitive (prepulse inhibition, acoustic startle habituation, conditioned fear, Morris task). In addition, they examined whether the order of testing affected the performance of two mouse strains. It was found that only an order effect was found for the light–dark box.

In the present study we examined the order effects of three different tests assessing anxiety- and depressive-like behavior. We applied a test schedule comparable to the one employed in the second experiment of McIlwain et al. [14], the only difference being that we performed different tests and used rats instead of mice. Since we assumed that the Forced Swim Test would have the greatest impact on behavior in subsequent tests, we varied the position of the Forced Swim Test according to three different orders. The outcome of the present study would be very helpful in determining the most suitable testing order which can be used for this test battery in rats. In addition, this study examined the effects of repeated testing in three different tests.

2. Methods

2.1. Animals

Experimental procedures were approved by the Animal Ethics Board of Maastricht University on animal welfare according to Dutch governmental regulations. Three-month-old male Wistar rats ($n=27$) were used, housed in groups of two animals (one group of three animals) under standard laboratory conditions, temperature 21–23 °C. The rats had free access to water and food. They were kept under a reversed 12/12 h light–dark cycle (lights off between 7 am and 7 pm). Background noise was provided by a radio that was playing softly. During all testing also indirect white light was present (3–5 lux).

2.2. Behavioral procedures

The rats were randomly assigned to one of three testing orders: (1) Open Field (OF) – Forced Swim Test (FST) – Zero Maze (ZM); (2) ZM–OF–FST; (3) FST–ZM–OF. The tests were all administered during one week in which the OF and ZM were administered in four and the FST on 3 consecutive days. The animals from the different orders were tested in the same week in the different tests and each test was always given at the same time of day (see below for testing times). Thus, in the three successive weeks the rats of one group/order were tested at different times during the day in different tests.

2.3. Open Field

The OF test is a widely used test to measure anxiety-related behavior [e.g. 2]. The apparatus consisted of a square base of black polyvinylchloride (100 × 100 cm) with 30 cm-high enclosure and was virtually divided in 25 equal squares. The four corner squares were called the corner area, the remaining 12 squares were called the wall area center, and the center squares were called the center area. Testing was conducted under dim light conditions (3–5 lux). Rats were placed in the middle of the OF and allowed to explore for 5 min. The movements of the rats were recorded with a camera which was connected to a computer running the Ethovision™ program (Noldus Equipment, Wageningen, The Netherlands). Between trials the apparatus was cleaned with a 5% ethanol/water solution. The following measures were taken as indices for anxiety-related behavior [2]: the time spent in corners (TIC) and total distance traveled (DM). Furthermore, the rearing frequency was scored by two observers. Testing in the OF was always conducted between 9 and 12 am.

2.4. Zero Maze

The Zero Maze was developed to measure anxiety-related behavior [3]. The apparatus consisted of a black Perspex annular platform (105 cm diameter, 10 cm width) elevated to 20 cm above ground level. The ZM was divided in four equally sized corridors in which the two opposite corridors were enclosed by black polyvinylchloride walls (27 cm high) on the inner and outer edges. The two other parts had a rim of 1 cm high. Testing was carried out under dim light conditions (3–5 lux). Rats were placed on the open part of the apparatus and allowed to explore for 5 min. The movements of the rats were recorded with a camera connected to a computer running the Ethovision™ program (Noldus Equipment, Wageningen, The Netherlands). The time spent and frequency in the open arms (TOA and FOA respectively) and time in the entry zones (defined as the regions located at the intersect of the open/enclosed arm, both 10 cm; TEZ) were measured. These measures have been defined as behavioral parameters for anxiety-related behavior [3]. Also the rearing frequency was scored by two observers. Between test sessions the maze was cleaned with approximately 5% ethanol/water solution. Testing was always conducted between 12 am and 3 pm.

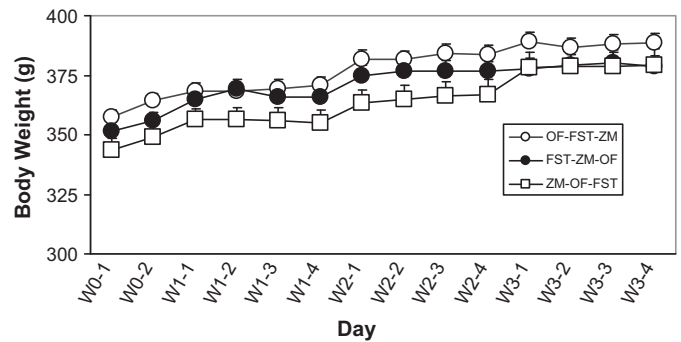


Fig. 1. Body weight of rats from the three experimental groups that were tested in a different order in three different tests.

2.5. Forced swimming test

The forced swimming task has been used for the assessment of depressive-like behavior in rodents [4,5]. Between 3 and 6 pm the rats were tested in the FST. This test consisted of a glass cylinder (40 cm × 17 cm), filled with 30 cm deep water (20–23 °C). After rats were placed in the water, time-sampling began for the behaviors immobility, swimming and climbing during 3 min sessions. Immobility was defined as floating in the water with no other movements other than those necessary to keep the head above water. Swimming was scored when rats were actively making swimming motions, more than those necessary to keep their heads above the water. Climbing was defined as movements with the forepaws in and out the water, usually against the walls. All these three behaviors have been reported to reflect depressive-like behavior [5]. The behavior were videotaped and afterwards rescored so that each behavior was scored by two observers independently. In case the inter-observer reliability was low (more than 10% deviation from largest observation), the behavior was scored again by three observers and the median of the five measures was taken.

2.6. Statistical analysis

All parameters were tested in a mixed design ANOVA with within-subjects factor Day (four levels for OF and ZM parameters and three levels in FST) and between-subject factor Order (three levels: FZO, OFZ and ZOF). A Dunnett post-hoc test was used to examine group differences in more detail. For the evaluation of the interdependency of the different variables within and between tests, Pearson correlation coefficients were calculated. For this purpose the data of the different days were aggregated according to the following formula: $\ln((\text{sum of scores per day})/\text{number of days})$. This transformation was used in previous studies and shown to reveal correlations between behavioral scores [2,9,10]. For the statistical analyses we used the Statistical Package for the Social Sciences (SPSS 15.0 software, Chicago, IL, USA).

3. Results

3.1. Body weight

The average body weight increased during the course of testing ($F(3,72)=88.75, p<0.001$, see Fig. 1). The increase in body weight

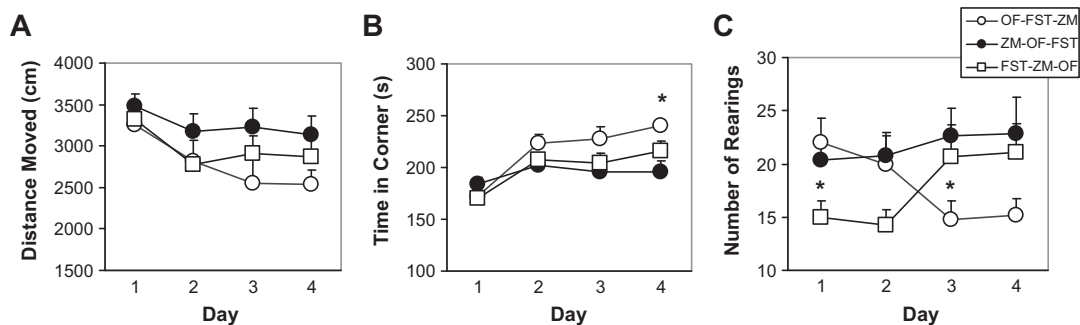


Fig. 2. Behavior of Wistar rats in an Open Field test: (A) the total distance moved, (B) total time spent in the corner areas, and (C) total number of rearing. Rats were tested in three different orders. The rats were tested on four successive days in 5 min trials. Data represent mean (+SEM). Asterisks indicate differences between groups as revealed by a post-hoc test. On day 4 OFZ spent more time in the corners as compared to rats in the order ZOF. On the first day, rats in the order OFZ showed more rearing behavior as compared to rats in the order FZO.

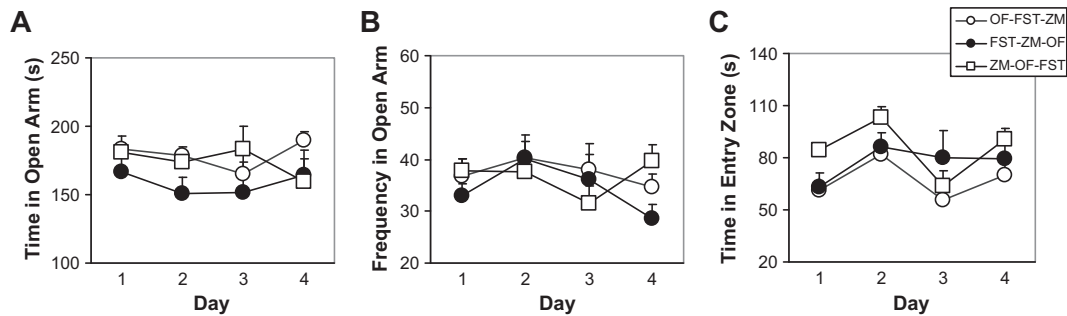


Fig. 3. Behavior of Wistar rats in a Zero Maze: (A) the total time spent on the open arm, (B) frequency of entering the open arm, and (C) total time spent in the entry zones. Rats were tested in three different orders. The rats were tested on four successive days in 5 min trials. Data represent mean (+SEM).

was not influenced by the order of testing ($F(2,24)=0.88$, n.s.) and there was no Order by Day interaction effect ($F(6,72)=5.08$, n.s.).

Day and interaction: all F s < 1.91, n.s.; data of rearing not shown, see Fig. 3C)

3.2. Open Field

The total distance moved decreased across testing days ($F(3,72)=6.94$, $p < 0.001$; see Fig. 2A). The activity of the rats was independent of the order of testing ($F(2,24)=2.32$, n.s.) and no interaction between Order and Day was found ($F(6,72)=0.65$, n.s.). The time the rats spent in the corner areas increased during the course of testing ($F(3,72)=20.99$, $p < 0.001$; see Fig. 2B). The order of testing did not influence the time the rats spent in the corner areas ($F(2,24)=2.43$, n.s.). However, a Day by Order interaction effect was found for the time spent in the corner areas ($F(6,72)=2.97$, $p < 0.05$). Post-hoc analysis revealed that on day 4, rats in the order OFZ spent more time in the corners as compared to rats in the order ZOF. The rearing behavior of the rats did not change over the days of testing ($F(3,72)=0.28$, n.s.; see Fig. 2C). Also, the order of testing did not influence the rearing frequency ($F(2,24)=1.88$, n.s.). However, there was an Order by Day interaction effect ($F(6,72)=3.59$, $p < 0.01$). Post-hoc analysis revealed that on the first day, rats in the order OFZ showed more rearing behavior as compared to rats in the order FZO.

3.4. Forced swimming test

In the FST the immobility did not change over days ($F(2,48)=1.61$, n.s.; see Fig. 4A), nor by the order of testing ($F(2,24)=2.10$, n.s.). However, a Day by Order interaction effect was found ($F(4,48)=4.70$, $p < 0.01$). Post-hoc analysis showed that on day 2, the rats that were tested in order FZO were less immobile in the FST as compared to animals which received order ZOF and OFZ. The swimming time increased over the 3 days ($F(2,48)=14.93$, $p < 0.001$; see Fig. 4B). The swimming was affected by the order in which the rats were tested in ($F(2,24)=3.56$, $p < 0.05$). Furthermore, an Order by Day interaction was found ($F(4,48)=9.71$, $p < 0.001$). Post-hoc analysis revealed that on day 2, animals in the FZO order spent more time swimming than those that were tested in the orders ZOF and OFZ. Climbing behavior increased over the testing days ($F(2,48)=36.12$, $p < 0.001$; see Fig. 4C). There was no Order ($F(2,24)=0.14$, n.s.) or Order by Day interaction effect ($F(4,48)=1.78$, n.s.).

3.5. Correlations

Table 1 shows the results of the correlation analysis for parameters within tests per experimental group. The Table does not show the correlations between tests, as only sporadic correlations between measures between tests were found. The most consistent correlation between measures in different paradigms was the rearing frequency in the OF and the ZM (all animals, 0.51 , $p < 0.01$; order OFZ, -0.45 , n.s.; order FZO, 0.67 , $p < 0.05$; ZOF, 0.82 , $p < 0.01$)

3.3. Zero Maze

The TOA ($F(3,72)=0.62$, n.s.; see Fig. 3A) and FOA ($F(3,72)=1.50$, n.s.; see Fig. 3B) did not change across days and was not affected by testing order ($F(2,24)=1.27$, n.s.; $F(2,24)=0.50$, n.s., respectively). Also, no interactions between Order and Day were found ($F(6,72)=0.97$, n.s. and $F(6,72)=1.43$, n.s., respectively). For the TEZ and the rearing frequency also no effects were found (Order and

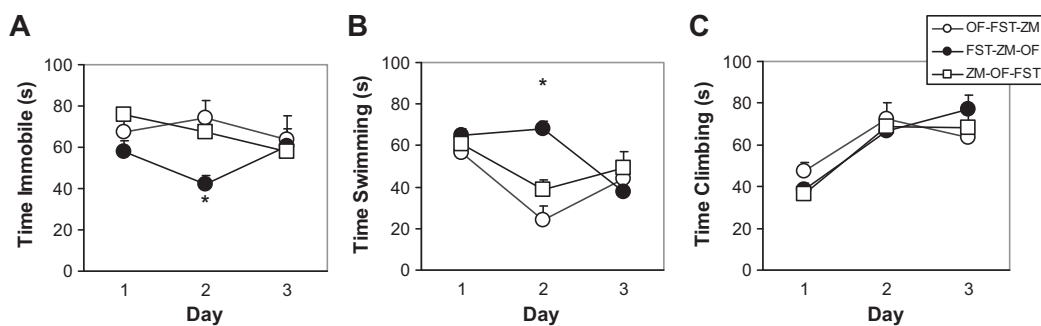


Fig. 4. Behavior of Wistar rats in a Forced Swim Test: (A) time spent immobile, (B) total swimming time and (C) total climbing time. Rats were tested in three different orders. The rats were tested on four successive days in 5 min trials. Data represent mean (+SEM). Asterisks indicate differences between groups as revealed by a post-hoc test. On day 2, the rats that were tested in order FZO were less immobile than animals which received order ZOF and OFZ. On day 2, animals in the FZO order spent more time swimming than those that were tested in the orders ZOF and OFZ.

Table 1

Correlation coefficients between the different measures of the Open Field (OF), Zero Maze (ZM) and Forced Swim Test (FST). For all animals/testing orders (A), order OF–FST–ZM (B), order ZM–OF–FST (C), and order FST–ZM–OF (D). Bold entries represent correlation coefficients with an associated probability $p < 0.05$.

(A) All groups	OF		ZM			FST	
	DM	TIC	TIOA	FOA	TIEZ	IMMO	SWIM
OF							
TIC	-0.87	-	-	-	-	-	-
Rear	0.74	-0.67	-	-	-	-	-
ZM							
FOA	-	-	0.28	-	-	-	-
TIEZ	-	-	0.48	0.52	-	-	-
Rear	-	-	0.51	0.18	0.27	-	-
FST							
SWIM	-	-	-	-	-	-0.62	-
CLIMB	-	-	-	-	-	-0.66	-0.02
(B) OF–FST–ZM	OF		ZM			FST	
	DM	TIC	TIOA	FOA	TIEZ	IMMO	SWIM
OF							
TIC	-0.97	-	-	-	-	-	-
Rear	0.61	-0.73	-	-	-	-	-
ZM							
FOA	-	-	-0.58	-	-	-	-
TIEZ	-	-	-0.32	0.50	-	-	-
Rear	-	-	0.06	-0.20	0.14	-	-
FST							
SWIM	-	-	-	-	-	-0.82	-
CLIMB	-	-	-	-	-	-0.70	-0.02
(C) ZM–OF–FST	OF		ZM			FST	
	DM	TIC	TIOA	FOA	TIEZ	IMMO	SWIM
OF							
TIC	-0.89	-	-	-	-	-	-
Rear	0.93	-0.82	-	-	-	-	-
ZM							
FOA	-	-	0.70	-	-	-	-
TIEZ	-	-	0.77	0.80	-	-	-
Rear	-	-	0.75	0.66	0.41	-	-
FST							
SWIM	-	-	-	-	-	0.43	-
CLIMB	-	-	-	-	-	-0.85	-0.77
(D) FST–ZM–OF	OF		ZM			FST	
	DM	TIC	TIOA	FOA	TIEZ	IMMO	SWIM
OF							
TIC	-0.87	-	-	-	-	-	-
Rear	0.75	-0.49	-	-	-	-	-
ZM							
FOA	-	-	-0.55	-	-	-	-
TIEZ	-	-	0.42	0.00	-	-	-
Rear	-	-	0.53	-0.15	0.27	-	-
FST							
SWIM	-	-	-	-	-	-0.55	-
CLIMB	-	-	-	-	-	-0.66	-0.13

DM = distance moved; TIC = time in corner; Rear = number of rearings; TIOA = time in open arms; FOA = frequency in open arms; TIEZ = time in entry zones; IMMO = immobility; SWIM = swimming; CLIMB = climbing.

4. Discussion

The current study tried to identify whether rats behave differently in tests measuring affective behavior depending on the testing order being used. It was assumed that when rats are first put in a relatively invasive test (FST), this would have the greatest impact on behavior in subsequent less invasive tests (OF and ZM). The present data did not corroborate this assumption as the data showed that the ZM appeared to be insensitive to testing orders used in the present study. On the other hand, the OF and FST test did show

some order effects. Rats with previous testing experience spent less time in the corner areas in the OF and displayed changes in rearing frequency. In the FST, a higher swimming score and a lower immobility score was found on the second day of testing in animals that were first tested in the FST. Of note, the pattern of the different FST measures over days may vary with other studies. This may be related with the duration of the swim times on each day.

Unexpectedly, the FST seemed to be most influenced by the order in which the rats were tested; animals that were first tested in the FST showed less immobility and more swimming on day 2

as compared to the other groups. In addition, the inter-relationship between the different parameters varied strongly between orders as shown by the within-test correlations. Taken together, these findings were in contrast with our hypothesis that tests which are considered as more invasive are less affected by previous testing history as compared to less invasive tests. Another important point to consider in respect to the FST data is related to the evaluation of drug effects in this test. Thus, if the FST is used in a test battery for the evaluation of drug treatments, effects are usually assessed at the second trial [5,15]. The current experiment shows that exactly at that time point training history of the rats has the strongest influence on behavior. Therefore, the effects of drug treatment in the FST within a test battery should be evaluated with caution.

Although behavioral test batteries are widely used, only few studies have specifically investigated the effects of testing order [14,16–18]. The majority of these studies were conducted with mice and revealed some effects of repeated testing on performance in different tests for anxiety and depression. These findings are in line with other studies showing an effect of housing conditions [19,20] and testing history of animals [21–23] on a subsequent test battery. Taken together, these data provide further evidence that behavioral performance can be influenced by a range of pre-test experience factors.

The present study is a first demonstration of a testing order effect in rats. It should be noted that in previous studies the performance of mice in each task was only assessed for one single day [14,16]. In our study we tested the rats for 3–4 days in each test. This was based on previous studies showing that repeated testing in the same test increases the validity of the test [2,9,24,25]. It has been reported earlier that successive trials in behavioral test may measure different stages of affective behavior [e.g. 26, 27]. Therefore, the use of several testing days may overcome the ambiguity in interpretation of data of only one single day. Remarkably, the behavior in the ZM was not influenced by repeated testing or by testing order. This makes the ZM a very robust test for the assessment of anxiety-related behavior.

The correlation analysis of the aggregated data in each test provides further support for the use of repeated testing of rats as the analysis revealed inter-dependencies between the measures within one task. The paradigms used in the present study were not expected to measure the same aspect of anxiety/depression and consequently no correlations between tests were expected. Apart from the correlation between rearing frequency in the ZM and OF, no consistent correlations were observed. These findings are supportive of a valid assessment of the anxiety/depressive ‘trait’ of animals using the procedure of repeated testing.

In addition, the correlation analysis showed some interesting features when evaluating the correlations within each test order group and per test. It appeared that the different measures of the OF were correlated in a similar way for all testing orders in the overall analysis. Thus, although the behavior in the OF was affected by testing order, the relation between the behavioral measures changed in a comparable manner, independent of testing order. On the other hand, the correlations between the ZM measures were dependent on testing order. The measures only showed a correlation in the testing order in which the ZM was tested first (ZOF). Assuming that the inter-dependency between the ZM measures reflects a valid assessment of anxiety-related behaviors in this test (as is the case for comparable tests, see [2,9,24,25]), this may indicate that this test does not reflect anxiety-related behavior in rats first tested in the FST and/or OF.

Examining the correlations between the FST measures demonstrated a relation between these measures in the overall analysis. Remarkably, these correlations were somewhat different from the rats that were tested in the FST for the first time (order FZO). Thus, in the order FZO a negative relation was found between immobility

and climbing and between swimming and climbing. No reliable relation was found between swimming and immobility. In the overall analysis a negative relation was found between climbing and immobility but the other correlations were different. Interestingly, there was no relation between the different FST measures in the ZOF order. This suggests that the interdependency of the measures was no longer found when the FST was administered at the end of the test battery.

One drawback of the present study was that not all possible orders were tested. On basis of our assumption that the FST would have the strongest effect on behavior in the OF and ZM test, we only controlled for the order of the FST. A complete design with all testing orders could evaluate order effects for each test separately. However, the main aim of this study was to evaluate possible order effects of different tests in a test battery. The present findings suggest that order effects can be found using different testing orders.

Another issue that should be mentioned is that in the present study test order effects may have been influenced by the repeated testing in one test. As mentioned above, this was done to enhance the validity of the behavioral trait in rats [2,9,24,25]. On the other hand, testing experience will probably affect all designs aiming to examine order effects since one study cannot control for all order effects and training effects. Further, the order effects in the present study were found in animals with a comparable test experience history. This may suggest that the order effect could be considered as a true effect.

Also, it was not possible keeping all factors constant in the present study. Because the rats were tested in the same week and room, a choice had to be made between (1) testing the experimental groups at the same time of day, or (2) keeping the time of day for each test constant. It was decided that it was more important that the animals were tested in the same test at the same time of day. This was done because of possible circadian rhythm effects on behavioral activity in the elevated plus-maze [28,29]. Further, if the rats were tested at various times of day in the different tests we could not dissociate time of day effects versus order effects. Although the design used cannot rule out possible effects of testing the animals at different times of day, it was assumed that this approach was less likely to confound the data.

Finally, it should also be noted that in the present study we used Wistar rats. It is well known that the performance in these tests is strain dependent [e.g. 10, 30, 31]. It remains to be demonstrated whether comparable order effects can be reproduced with other strains. This also holds for other anxiety and depression paradigms that are being used in test batteries.

In summary, the present study clearly showed that order effects can be found in a test battery that measures affective behavior in an OF, ZM and FST. In contrast to our expectations the FST appeared to be the most sensitive to order effects whereas the ZM was not affected by position in the test battery. Although the measures in the OF changed across days, the test still appeared to measure anxiety-related as indicated by the correlation analysis. The present study suggests that, although test batteries are very useful in measuring different aspects of affective behavior, results from test batteries could be liable to order effects.

References

- [1] File SE. Recent developments in anxiety, stress, and depression. *Pharmacol Biochem Behav* 1996;54:3–12.
- [2] van der Staay FJ, Kerbusch S, Raaijmakers WGM. Genetic correlations in validating emotionality. *Behav Genet* 1990;20:51–62.
- [3] Shepherd JK, Grewal SS, Fletcher A, Bill DJ, Dourisch CT. Behavioural and pharmacological characterisation of the elevated zero-maze as an animal model of anxiety. *Psychopharmacology* 1994;116:56–64.
- [4] Porsolt RD. Animal model of depression. *Biomedicine* 1979;30:139–40.

- [5] Cryan JF, Markou A, Lucki I. Assessing antidepressant activity in rodents: recent developments and future needs. *Trends Pharmacol Sci* 2002;23:238–45.
- [6] Holmes PV. Rodent models of depression: reexamining validity without anthropomorphic inference. *Crit Rev Neurobiol* 2003;15:143–74.
- [7] Blanchard DC, Griebel G, Blanchard RJ. The Mouse Defense Test Battery: pharmacological and behavioral assays for anxiety and panic. *Eur J Pharmacol* 2003;463:97–116.
- [8] Lim LW, Blokland A, Visser-Vandewalle V, Vlamings R, Sesia T, Steinbusch H, et al. High-frequency stimulation of the dorsolateral periaqueductal gray and ventromedial hypothalamus fails to inhibit panic-like behaviour. *Behav Brain Res* 2008;193:197–203.
- [9] Blokland A, Prickaerts J, Raaijmakers W. Reduced level of anxiety in adult Lewis rats after chronic ethanol consumption. *Physiol Behav* 1992;51:245–8.
- [10] van der Staay FJ, Blokland A. Behavioral differences between outbred Wistar, inbred Fischer 344, Brown Norway, and hybrid Fischer 344xBrown Norway rats. *Physiol Behav* 1996;60:97–109.
- [11] Griebel G. 5-Hydroxytryptamine-interacting drugs in animal models of anxiety disorders: more than 30 years of research. *Pharmacol Ther* 1995;65:319–95.
- [12] 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;99:301–16.
- [13] Kalueff AV, Wheaton M, Murphy DL. What's wrong with my mouse model? Advances and strategies in animal modeling of anxiety and depression. *Behav Brain Res* 2007;179:1–18.
- [14] McIlwain KL, Merriweather MY, Yuva-Paylor LA, Paylor R. The use of behavioral test batteries: effects of training history. *Physiol Behav* 2001;73:705–17.
- [15] Porsolt RD, Le Pichon M, Jalfre M. Depression: a new animal model sensitive to antidepressant treatments. *Nature* 1977;266:730–2.
- [16] Paylor R, Spencer CM, Yuva-Paylor LA, Piekie-Dahl S. The use of behavioral test batteries, II: effect of test interval. *Physiol Behav* 2006;87:95–102.
- [17] Bouwknecht JA, van der Gugten J, Groenink L, Olivier B, Paylor RE. Effects of repeated testing in two inbred strains on flesinoxan dose-response curves in three mouse models for anxiety. *Eur J Pharmacol* 2004;494:35–44.
- [18] Voikar V, Vasar E, Rauvala H. Behavioral alterations induced by repeated testing in C57BL/6J and 129S2/Sv mice: implications for phenotyping screens. *Genes Brain Behav* 2004;3:27–38.
- [19] Lewejohann L, Reinhard C, Schrewe A, Brandewiede J, Haemisch A, Gortz N, et al. Environmental bias? Effects of housing conditions, laboratory environment and experimenter on behavioral tests. *Genes Brain Behav* 2006;5:64–72.
- [20] Abramov U, Puussaar T, Raud S, Kurrikoff K, Vasar E. Behavioural differences between C57BL/6 and 129S6/SvEv strains are reinforced by environmental enrichment. *Neurosci Lett* 2008;443:223–7.
- [21] Salam JN, Fox JH, Detroy EM, Guignon MH, Wohl DF, Falls WA. Voluntary exercise in C57 mice is anxiolytic across several measures of anxiety. *Behav Brain Res* 2009;197:31–40.
- [22] Schmitt U, Hiemke C. Combination of open field and elevated plus-maze: a suitable test battery to assess strain as well as treatment differences in rat behavior. *Prog Neuropsychopharmacol Biol Psychiatry* 1998;22:1197–215.
- [23] Igarashi E, Takeshita S. Effects of illumination and handling upon rat open field activity. *Physiol Behav* 1995;57:699–703.
- [24] Blokland A, Lieben C, Deutz NEP. Anxiogenic and depressive-like effects, but no cognitive deficits after repeated moderate tryptophan depletion in the rat. *J Psychopharmacol* 2002;16:39–49.
- [25] Prickaerts J, Raaijmakers W, Blokland A. Effects of myocardial infarction and captopril therapy on anxiety-related behaviors in the rat. *Physiol Behav* 1996;60:43–50.
- [26] File SE, Zangrossi Jr H, Viana M, Graeff FG. Trial 2 in the elevated plus-maze: a different form of fear. *Psychopharmacology* 1993;111:491–4.
- [27] Cook MN, Crouse M, Flaherty L. Anxiety in the elevated zero-maze is augmented in mice after repeated daily exposure. *Behav Genet* 2002;32:113–8.
- [28] Jones N, King SM. Influence of circadian phase and test illumination on pre-clinical models of anxiety. *Physiol Behav* 2001;72:99–106.
- [29] Verma P, Hellemans KG, Choi FY, Yu W. Circadian phase and sex effects on depressive/anxiety-like behaviors and HPA axis responses to acute stress. *Physiol Behav* 2010;99:276–85.
- [30] Paré WP. Open field, learned helplessness, conditioned defensive burying, and forced-swim tests in WKY rats. *Physiol Behav* 1994;55:433–9.
- [31] Izidio GS, Lopes DM, Spricigo Jr L, Ramos A. Common variations in the pretest environment influence genotypic comparisons in models of anxiety. *Genes Brain Behav* 2005;4:412–9.