Neuroscience and Biobehavioral Reviews 37 (2013) 2318–2330

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

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

Review

Classical and novel approaches to the preclinical testing of anxiolytics: A critical evaluation

J. Haller*, M. Aliczki, K. Gyimesine Pelczer

Institute of Experimental Medicine, Budapest, Hungary

ARTICLE INFO

Article history:
Received 12 April 2012
Received in revised form 21 August 2012
Accepted 5 September 2012

Keywords:
Anxiety
Preclinical
Testing
Validity

ABSTRACT

Over 80% of current anxiety studies employ one of the tests that were developed earlier than, or concurrently with the elevated plus-maze, i.e. before 1985. Considering 1985 as a historical reference point, we briefly review here 115 new tests and models of anxiety, the development of which was likely prompted by the poor predictive validity of classical tests as shown here by the comparison of preclinical and clinical findings with putative novel anxiolytics. The new approaches comprise major innovations to classical tests, the pre-test application of manipulations that mimic etiological factors of anxiety disorders, and entirely new approaches including anxiety disorder-specific tests. Thus, intensive test development over the last 27 years created a large pool of novel approaches. However, these are infrequently used and as such, their impact on anxiolytic drug development remains low. We suggest here that test/model development should step over the intensive phase when several new methods are proposed each year and should start selecting and establishing the methodologies that would successfully replace or complement classical tests. We propose here a novel strategy for improving the validity of anxiety testing that includes the retrospective analysis of the predictive validity of new procedures (as opposed to classical pharmacological validation), and a call for concerted international efforts at both the conceptual and practical levels. Similar endeavors proved recently successful with other psychiatric disorders.

Contents

1. The predictive value of the classical tests of anxiety

The classical tests of anxiety had a tremendous contribution to anxiety research by making possible the quantification of an internal emotional state by behavioral means in animals. We depict here as 'classical' the approaches that were developed between 1920 and 1985 (the last being the elevated plus-maze test) and which are still in use today. Such classical tests are the active and passive avoidance, conditioned fear, defensive burying, elevated plus-maze, fear potentiated startle, four plate, holeboard, hyponeophagia, light–dark, open-field, punished drinking, punished feeding, social interaction, and ultrasonic vocalization tests (Aron et al., 1971; Brown et al., 1951; Crawley and Goodwin, 1980; File and Hyde, 1978; Gardner and Piper, 1982; Gardner, 1985; Geller

* Corresponding author at: Institute of Experimental Medicine, Hungarian Academy of Sciences, Szigony u. 43., Budapest H-1083, Hungary. Tel.: +36 1 2109406.
E-mail address: haller@koki.hu (J. Haller).

0149-7634/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.neubiorev.2012.09.001

et al., 1962; Hall, 1934; Pellow et al., 1985; Price, 1972; Shephard and Broadhurst, 1982; Treit et al., 1981; Vogel et al., 1971; Watson and Rayner, 1920). Although a large number of tests were developed after 1985, 8 out of the 14 classical tests remain the basic means of anxiety research (see below). Therefore, we consider the year 1985 the endpoint of the period when the methodological bases of current laboratory research were laid down. The variety of ways by which classical approaches allow the testing of anxiety is of great importance. Such tests quantify fears and anxieties elicited by (i) natural aversive stimuli, e.g. open spaces (Hall, 1934; Pellow et al., 1985), light (Crawley and Goodwin, 1980), unfamiliar social partners (File and Hyde, 1978), which all suppress exploratory activity in rodents, (ii) acute stressors, e.g. novel environments where feeding is suppressed (Shephard and Broadhurst, 1982) or maternal separation that elicits ultrasonic vocalization (Gardner, 1985), (iii) imminent punishment, e.g. electric shocks that suppress feeding and drinking or limit locomotion (Geller et al., 1962; Vogel et al., 1971; Aron et al., 1971), and (iv) by learning as in the case of conditioned responses (Brown et al., 1951; Price, 1972). Thus, early efforts resulted in a large number of tests that allowed a multifaceted evaluation of fear and anxiety, and were relatively simple to employ.

One can hypothesize that behaviors monitored in classical anxiety tests are suitable indirect measures of anxiety albeit the feeling per se is difficult to measure directly. However, anxiety testing is not solely used to understand a feeling; anxiety research is very often connected with anxiolytic drug development. One of the main aims of this research line is to predict the clinical efficacy of putative novel anxiolytics based on preclinical findings. In this respect, classical anxiety tests show a poor predictive validity (Table 1). About 40% of compounds that were "successful" in preclinical testing were ineffective in placebo-controlled, double-blind clinical trials. One can hypothesize that the real figures are even worse, as only a fraction of compounds belonging to particular drug classes were clinically tested, e.g. only 3 CRF_1 antagonists were tested in the clinic out of the 14 which produced promising effects in preclinical tests (Kehne and Cain, 2010). One can hypothesize that the remaining 11 would have produced similarly poor results if subjected to clinical trials. Yet, all CRF_1 antagonists were considered as one single mismatch when the percentage of poor predictions was calculated. The same applies to other drug classes that were once considered hope for the treatment of anxiety (e.g. NK_{1-3} antagonists; 5-HT_{2-3} antagonists, CCK_2 antagonists) or drug classes that provided mixed effects in the laboratory, but clinical experience clearly proved their utility as anxiolytics (e.g. serotonergic antidepressants). The poor predictive value of anxiolytic drug testing becomes especially blatant when compared with the predictions supplied by models of a related psychiatric disorder namely depression (Table 2). The idea that certain agents have antidepressant activity may not have always originated in the laboratory, and the antidepressant potential of certain compounds may not necessarily lead to their regular clinical application. Yet, the data summarized in Table 2 clearly shows that there is a tight correlation between laboratory and clinical findings on antidepressant activity. One possible explanation for this success is that tests of antidepressant activity tend in the main to be focused on cardinal symptoms elicited by manipulations akin to human etiological factors which, with the exception of conditioning models, is not true for classical tests of anxiety.

There is a basic difference between the anxiety disorders and anxieties monitored by preclinical tests: the former are chronic states that entail responses that are inadequate to the situation, while the latter monitor the acute manifestations of natural fears (Rodgers, 1997; Fuchs and Flugge, 2006; Steimer, 2011). DSM-IV (American Psychiatric Association, 1994) expressly specifies that fears are "excessive" or "unreasonable" with almost all anxiety disorders, and that this condition should last at least 1 month (post-traumatic stress disorder, panic disorder) or 6 month (social phobia, generalized anxiety, specific phobias) for a diagnosis. Classical tests of anxiety performed under classical conditions are different: these tests reveal acute responses that are entirely adequate to the testing situation. Fear of open spaces or light are natural responses in species that live in burrows and are widely subject to predation (e.g. laboratory rodents). Fear from unfamiliar individuals, and maternal separation-induced fears in pups are common fear responses in mammals, similar to fears elicited by imminent physical pain. None of these experimental situations involve "unreasonable" fears and all are transient or situation-dependent. Certain compounds reduce both natural anxiety responses and unreasonable fears associated with anxiety disorders, showing that the mechanisms underlying these two distinct phenomena are overlapping. The fact that other compounds selectively reduce one or the other demonstrates that their mechanisms are not identical.

One is inclined to assume that in such a situation the laboratory testing of anxiolytics undergoes rapid and profound changes to produce more reliable predictions. However, this does not seem to be the case. A thorough Medline search demonstrated that preferences for anxiety tests did not change over the past 22 years (Fig. 1). On one hand, classical tests remain just as frequently used as earlier. The elevated plus-maze test alone was employed in about 50% of all studies, the most frequently used 4 tests (elevated plus-maze, open-field, social interaction, and light/dark box) account for about 80%, while the 9 most "popular" classical tests together account for about 90% of anxiety testing. In addition, 4 of the 9 less frequently but still significantly used tests are also classical. On the other hand, none of the novel tests seem to break through. Although their share in anxiety testing varied over years, their usage shows no clear trend of increase.

In spite of the high stability of test choices, the need of change was felt for a long time, e.g. Green and Hodges (1991) stated that animal tests of anxiety would be better named as animal tests of benzodiazepine psychopharmacology. Similarly, Rodgers (1997) suggested that "test development strategies which, by emphasizing pharmacological (i.e. benzodiazepine) validation, have yielded models predictive of a specific type of anxiolytic activity", notably models that specifically predict the efficacy of benzodiazepines. The need of change is also shown by the pace at which novel tests are developed. In a recent review, we showed that in just 18 month (in 2010 and the first half of 2011), 36 'non-classical' anxiety tests (i.e. tests developed after the elevated plus-maze) were employed and many novel tests were proposed (Haller and Aliczki, 2012). In addition, classical tests were often performed under non-classical conditions, in the sense that basal levels of anxiety were increased by various means that range from stress exposure to genetic manipulations. Overall, more than half of anxiety studies contained at least one innovative step in the period covered by the above-mentioned review, all such steps having the ultimate aim of improving the predictive, face and construct validity of anxiety testing (Haller and Aliczki, 2012). Yet, anxiolytic drug development gained little from these developments. Firstly, only 11 of the 36 'non-classical' tests were used more than once, while the amendments to 'classical tests' were almost entirely laboratory specific. Thus, novel approaches appear to have a minimal impact on the field in general. Secondly, papers proposing novel anxiolytics appear to be especially conservative. In experiments where developmental issues, neural mechanisms, etiological factors, etc., were addressed, innovative approaches were rather frequent (~65% of all studies). We mention that any novel aspect – even minor ones – were considered in the above-mentioned review, hence the large share of innovative approaches. Such approaches were considerably less frequent in studies where the anxiety-related effects of novel compounds were studied (~35%) and even fewer in studies