Neurotoxicology and Teratology 52 (2015) 109–115

Contents lists available at ScienceDirect

Neurotoxicology and Teratology

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

# Assessment of learning, memory, and attention in developmental neurotoxicity regulatory studies: synthesis, commentary, and recommendations

Charles V. Vorhees [a,*], Susan L. Makris [b]

[a] Cincinnati Children's Research Foundation, Division of Neurology, 3333 Burnet Avenue, Cincinnati, OH 45229, United States
[b] U.S. Environmental Protection Agency, National Center for Environmental Assessment, 1200 Pennsylvania Ave., NW, (8623P), Washington, DC 20460, United States

ARTICLE INFO

*Article history:*
Received 27 August 2015
Received in revised form 14 October 2015
Accepted 16 October 2015
Available online 23 October 2015

*Keywords:*
Learning and memory
Developmental neurotoxicity
Radial arm maze
Morris water maze
Cincinnati water maze
Cognitive tests in rodents

ABSTRACT

Cognitive tests of learning and memory (L&M) have been required by U.S. Environmental Protection Agency (EPA) developmental neurotoxicity test (DNT) guidelines for more than two decades. To evaluate the utility of these guidelines, the EPA reviewed 69 pesticide DNT studies. This review found that the DNT provided or could provide the point-of-departure for risk assessment by showing the Lowest Observable Adverse Effect Level (LOAEL) in 28 of these studies in relation to other reported end points. Among the behavioral tests, locomotor activity and auditory/acoustic startle provided the most LOAELs, and tests of cognitive function and the Functional Observational Battery (FOB) the fewest. Two issues arose from the review: (1) what is the relative utility of cognitive tests versus tests of unconditioned behavior, and (2) how might cognitive tests be improved? The EPA sponsored a symposium to address this. Bushnell reviewed studies in which both screening (locomotor activity, FOB, reflex ontogeny, etc.) and complex tests (those requiring training) were used within the same study; he found relatively little evidence that complex tests provided a LOAEL lower than screening tests (with exceptions). Levin reviewed reasons for including cognitive tests in regulatory studies and methods and evidence for the radial arm maze and its place in developmental neurotoxicity assessments. Driscoll and Strupp reviewed the value of serial reaction time operant methods for assessing executive function in developmental neurotoxicity studies. Vorhees and Williams reviewed the value of allocentric (spatial) and egocentric cognitive tests and presented methods for using the Morris water maze for spatial and the Cincinnati water maze for egocentric cognitive assessment. They also reviewed the possible use of water radial mazes. The relatively lower impact of cognitive tests in previous DNT studies in the face of the frequency of human complaints of chemical-induced cognitive dysfunction indicates that animal cognitive tests need improvement. The contributors to this symposium suggest that if the guidelines are updated, they be made more specific by recommending preferred tests and providing greater detail on key characteristics of such tests. Additionally, it is recommended that guidance be developed to address important issues with cognitive tests and to provide the information needed to improve the design, conduct, and interpretation of tests of higher function within a regulatory context. These steps will maximize the value of cognitive tests for use in hazard evaluation and risk assessment.



## 1. Introduction

As noted in the Introduction to this Special Issue, developmental neurotoxicity testing (DNT) for the regulation of environmental chemicals is conducted in accordance with U.S. Environmental Protection Agency (EPA) and Organization for Economic Co-operation and Development (OECD) testing guidelines. The guidelines of particular relevance to this discussion are OPPTS 870.6300 (U.S. EPA, 1998) and OECD 426 (OECD, 2007). Because OECD 426 was published 10 years after the last revision of the EPA DNT guideline (U.S. EPA, 1998), it incorporates changes recommended on the basis of years of additional experience and scientific discourse, as compared with the EPA guideline. OECD 426 includes several changes, including the extension of the dosing period [i.e., from gestation (embryonic) day (GD or E) 6 to postnatal day (PND or P) 21 rather than GD 6 to PND 11], consideration of the need for direct administration of the test substance to the offspring postnatally, explicit recommendation for measures of behavioral

*Abbreviations:* CNS, central nervous system; CWM, Cincinnati water maze; DNT, developmental neurotoxicity; EOGRTS, extended one-generation reproductive toxicity study; EPA, Environmental Protection Agency; FOB, functional observational battery; L&M, learning and memory; LOAEL, lowest observed adverse effect level; MOG, modified one-generation; MWM, Morris water maze; OECD, Organization for Economic Co-operation and Development; PND, postnatal day; RAM, radial arm maze.

* Corresponding author.
*E-mail addresses:* Charles.vorhees@cchmc.org (C.V. Vorhees), makris.susan@epa.gov (S.L. Makris).

http://dx.doi.org/10.1016/j.ntt.2015.10.004
0892-0362/

ontogeny, and increased numbers of offspring assigned for neuropathology evaluation. Studies conducted with the procedural enhancements of OECD (2007) are accepted (and generally preferred) by the EPA.

The DNT study design is shown in Fig. 1 of the Introduction to this Special Issue (Makris and Vorhees, 2015). With regard to tests of learning and memory (L&M) in the offspring, the guideline specifies that it be assessed twice: first at an early age (ca. PND 21) and later in adults (ca. PND 60). The protocol provides guidance for most procedures (e.g., for auditory startle, motor activity, and neuropathology) in moderate detail. The exception is for cognitive testing for which the only guidance in both the EPA or OECD guidelines are:

1. "Learning must be assessed either as a change across several repeated learning trials or sessions, or, in tests involving a single trial, with reference to a condition that controls for non-associative effects of the training experience."
2. "The tests should include some measure of memory (short-term or long-term) in addition to original learning (acquisition)."

OECD (2007) additionally notes that the "measure of memory cannot be reported in the absence of a measure of acquisition obtained from the same test."

Although both the EPA and OECD DNT guidelines indicate that the selection of tests should be based on demonstrated sensitivity to the class of compound under investigation, they recognize that often such information is unavailable. In that case, a list of example tests is provided, but without guidance regarding selection of the appropriate test paradigm. That lack of guidance regarding cognitive test selection extends to the EPA Guidelines for Neurotoxicity Risk Assessment (U.S. EPA, 1998) which hardly addresses any aspect of developmental neurotoxicity testing. The same holds true for the OECD Guidance Document on Mammalian Reproductive Toxicity Testing (Number 43) (OECD, 2008) which does include a section on postnatal functional/behavioral neurotoxicity, and specifically addresses cognitive testing (pp. 43–46), providing descriptions of the various types of tests that might be utilized in a guideline DNT study, yet fails to discuss how to select the most appropriate cognitive test to use.

The concept behind the decision to allow flexibility in study design for cognitive testing in the DNT study guidelines has merit, in that no single (one-size-fits-all) assessment paradigm could be designed to comprehensively screen for any possible cognitive impairment. Nevertheless, the overly-broad criteria contained in the guidelines permit almost any form of learning to be used without regard to what aspect of cognitive function should be assessed. In addition, they do not require prior establishment of reliability, sensitivity, and validity based on the published literature (although the guidelines do require the laboratory to demonstrate proficiency with the tests by means of studies with positive control reagents).

As noted in the Introduction (Makris and Vorhees, 2015), an EPA retrospective review of 69 DNT studies that had been submitted in support of pesticide registration assessed how those studies were used in risk assessment (Raffaele et al., 2010). The apparent relative insensitivity of the cognitive tests in these studies raised concerns about the selection of cognitive methods used in the studies (typically passive avoidance and mazes), how rigorously the submitted cognitive tests were conducted and/or reported, the adequacy of the guideline language, and the lack of helpful guidance regarding study design. Given that the review found that only 4 of the 69 adequate pesticide DNT studies used a test of cognition to derive points-of-departure for risk assessment suggests that one of several explanations may account for the low 'hit' rate of these tests: (1) It is possible that the 4 findings at the LOAEL represent false positives, i.e., spurious findings caused by chance, that were not true effects on cognitive function, (2) that the 4 findings represent true positives, i.e., only 4 of the pesticides truly affected cognitive function; or (3) that these 4 findings underestimated the true rate because the tests were insensitive such that the true rate should be higher. Of greatest concern is the possible false negative rate, i.e., the case where the tests missed true effects. How can one determine if this was the case? Unfortunately, with a given data set, this is impossible to determine, but there are reasons to suspect that 4/69 may underestimate the true adverse effect rate for cognitive function based on the methods reported and insufficient positive control data used to support the submitted DNT studies.

Methodologically, for example, whether a T-maze or passive avoidance test is valid depends on some inherent limitations of these methods and of how they are executed. For T-mazes, some of the following issues may be of concern. Appetitive T-mazes require food restriction, provision of appropriate rewards, use of appropriate consequences for incorrect responses, appropriate numbers of trials appropriately spaced, training prior to assessment of learning, and a host of apparatus and procedural details. Water T-mazes also have characteristics that make them problematic in some instances. One such factor is that rats show stronger position preferences in water than in appetitive or spontaneous T-mazes. Side biases can interfere with reinforcement contingencies and complicate results. In addition, water T-mazes can be learned so rapidly that they reflect rudimentary learning and memory. There are also issues concerning reinforcement. In appetitive mazes, a wrong choice results in no reward and a consequence such as empty arm confinement, but in water T-mazes a wrong turn usually does not result in confinement of the subject to the incorrect arm but leaves the animal to continue searching until it finds the escape. Such self-correction procedures are commonly used in water T-mazes but can have the effect of mitigating consequences of errors. Because rats are able to swim to the opposite side of a T-maze quickly if they turn to the wrong way first, there is little consequence if the animal makes an error. This impedes the animal's learning of the correct initial choice. Rodents are thigmotaxic and more so in water mazes than in appetitive ones. If a rat can reach the escape by following a wall, it will usually do so and little learning will occur. All of the above, complicate the use of simple T-mazes.

For passive avoidance, there are different concerns. Rats are placed in a two-compartment apparatus connected by a door with one side dark and one side light. The concept is that rats prefer the dark such that when placed on the lighter side they will spontaneously cross to the darker side. The objective of the test is to teach rats to avoid the dark side that they prefer. On day-1 of the test, crossover times are generally short (and variable). Once they cross, the door is closed and they are given a brief foot shock. After a specified delay interval they are placed back on the lighter side to measure latency to cross a second time. The test can show increased latency in response to drugs and brain lesions, but only when deficits are large. Its ability to detect developmental neurotoxins has never proven to be very good. Some have suggested that the test can be improved by using a trials-to-criterion method (Wise et al., 1997), but no comparison studies have been done to support this assumption. However, even with this change, the potential remains for effects to occur because of differences in shock sensitivity rather than to memory per se. This can be addressed to some extent by conducting a separate experiment to test for the animals' shock threshold to ensure no group differences in sensitivity that might confound interpretation of learning differences, but in practice this is seldom done. Typically, it is assumed that if no differences on passive avoidance are found, that a shock threshold test is not necessary. However, if an animal's learning is impaired but their sensitivity to shock is increased, the two effects could cancel one another. Aside from shock sensitivity differences, the variability in crossover times makes it difficult to detect differences of any kind.

There is direct evidence that passive avoidance is insensitive. In an experiment in which methamphetamine was administered from PND 6–15 to rats and later tested on four cognitive tests, methamphetamine-treated progeny were impaired for spatial learning in the Morris water maze, egocentric learning in the Cincinnati water maze, working memory in the radial water maze, but showed no effects on 1-trial passive avoidance (Vorhees et al., 2015). Given how reliable neonatal methamphetamine is at inducing lasting cognitive deficits (Vorhees et al., 1994,