The Invalidity of the Forced Swim Test

Emily R. Trunnell, Ph.D.
People for the Ethical Treatment of Animals

February 2020

Development of the Forced Swim Test

The forced swim test (FST), also called the Porsolt Swim Test, has been used since at least the 1950s and popularized in 1977 by Roger D. Porsolt as a potential method for screening antidepressant drugs. During this test, a small animal, typically a rat or mouse, is placed in a container of water with no way to escape nor any place to rest out of the water. Naturally, the animal will spend some time swimming and trying to find a way to get out of the stressful situation in which they find themselves but will eventually become immobile and float. The experimenter records the time the animal spends swimming vs. the time they spend immobile, floating in the water. Sometimes the swimming behavior is divided into climbing behavior, where the rodent is obviously trying to climb up the sides of the tank or beaker, and swimming behavior, which typically involves the rodent swimming around and not trying to climb out of the container.

A similar test to the FST is the tail suspension test (TST), which operates on similar principles. An animal (typically a mouse) is held upside down by his or her tail, typically affixed to a stationary bar or object with a piece of tape. For a while, the mouse will struggle and try to correct this frightening and uncomfortable position but will eventually become immobile.

Porsolt (1977) and others found that when an experimenter acutely administers some commonly used antidepressant drugs to the animal prior to the FST or TST, the animal may swim (or struggle) for longer and spend less time floating (or remaining still). This was taken to mean that longer swimming times indicate a less depressed mouse (since the antidepressant is what caused the change in this behavior). Animals who spent more time immobile were thought to be in 'despair,' as if they had 'given up.' However, this interpretation is incorrect for several reasons.

Immobility is a learned behavior

There is considerable evidence to suggest that immobility in the FST is a learned or adaptive behavior, not one representing an internal state of despair. In some FST protocols, typically ones involving rats, the same animal is made to participate in the test more than once, usually before and after administration of a particular substance, so that the animal serves as their own control. In this case, immobility becomes a learned behavior. De Pablo et al demonstrated that rats generally show less mobility on the second day of testing than they do on the first day (De Pablo 1989). When a group of rats was administered anisomycin, a substance known to disrupt consolidation of memories, the anisomycin-treated group stayed more active on the second day of the test than rats who had not been given anisomycin (De Pablo 1989), meaning that disrupting the learning process impacts behavior during the FST. The untreated rats may have learned that there was no way to escape their situation, and that they would eventually be removed from the water by the experimenter, facts that the anisomycin-treated rats did not learn. The anisomycin had no effect on the rats' behavior during the first day of the test.

Proponents of FST immobility being a reflection of behavioral despair equate the behavior to those exhibited in learned helplessness paradigms (West 1990). To create a state of learned helplessness, an experimenter exposes an animal to a series of inescapable shocks. At first, the animal will actively look for ways to escape the shocks, but over time, he or she will exhibit fewer escape behaviors and sometimes the animal will not attempt to escape the shocks, even if they are provided the means to. Experimenters say that these animals have 'given up' and resigned themselves to the fate of being shocked.

When the same animal is subject to the FST more than once, it is thought by some that prior exposure to the testing situation acts as a stressor for the animal, and that increases in immobility on later testing days reflect a sort of learned helplessness caused by the inescapable FST. However, experiments by O'Neil and Valentino (1982) showed that prior exposure to the FST has no effect on behavior in other stress paradigms, such as inescapable shock, and also that allowing rats a way to escape the water container during the first FST does not affect their behavior on subsequent exposures to the test (they're still more immobile on later days – an observation not consistent with learned helplessness paradigms). This is further evidence that immobility in the FST is a learned behavior and not indicative of learned helplessness.

Immobility is an adaptive behavior

As reviews by West (1990) and Molendijk and de Kloet (2015) have explained, immobility in the FST is likely a beneficial behavior for these animals. Swimming and climbing expend unnecessary energy and animals who are quicker to realize this have a greater chance for survival in extended submerged situations. In experiments described by Nishimura, et al (1988), rats were forced to swim until sinking, for as long as two hours. Experimenters found that the amount of time spent immobile within the first 15 minutes of the test could predict sinking: the rats who struggled for longer were quicker to sink while the rats who conserved their energy were able to float for longer before sinking. The experimenters also noticed that rats who struggled and swam for longer defecated more, potentially signifying increased fear in the 'less adaptive' group.

Molendijk and de Kloet (2015) argue that the FST lacks two essential forms of validity used to assess animal models of human diseases or conditions: construct validity and face validity. Because the development of depression is a slow process, a test of 15 minutes or even tests conducted over a 24-hour period cannot be used to determine depression (Belmaker and Agam 2008); therefore the FST lacks construct validity. The FST lacks face validity because "there is no single sign or symptom of depression modeled apart from the anthropomorphic interpretation of floating behavior in terms of despair" (Molendijk and de Kloet 2015) and because there is "little similarity between the clinical symptoms of depression in humans and the behaviors measured in the test" (Commons 2017).

Another way to interpret the adaptive behavior of immobility during the FST, is to consider that the animal's actions may represent their individual response to the stressor of being immersed in water and of not knowing when or if escape will be possible. Some animals will cope with this situation actively, by struggling, and some will cope passively, by floating. In a recent review by Commons and colleagues (2017), the authors claim that

> [T]he term "depression-like" to refer to FST behavior is pointedly incorrect for three key reasons.

> (1) The FST actually measures coping strategy to an acute inescapable stress, not something like a pathological internal state of mind. Coping strategy is measured in the FST; "depression-like" is an inference that may or may not be correct…
>
> (2) "Depression-like" is jargon, used to acknowledge the limitations of the model system. However, "depression-like" is easily misunderstood by those less familiar with animal research including students, researchers in other fields, clinicians, patient advocates, and funding agencies.
>
> (3) The neurobiology underlying stress-coping strategy revealed in the FST is likely relevant to additional clinical conditions where there is poor behavioral response to acute stress…

The authors go on to note, "[w]hile it could be argued that passive coping strategies to stress are characteristic of depression, the connection between swimming and the human condition begs an abstraction at best. Behavior in the FST is a reaction to the acute stressful stimulus of being placed in a container without an escape route, and human depression reflects a chronic subjective emotional state rather than a reaction to an individual stimulus" (Commons 2017).

Initial interpretations of the forced swim test were at odds with biochemical reality

The methodology by which the FST was discovered is cause to doubt that immobility can be equated with 'despair.' Experimenters noted that acute administration of antidepressants decreased immobility, however antidepressants do not work acutely in humans to relieve depression. As noted by O'Leary and Cryan (2003), "The FST and TST [tail suspension test] have been criticised because they are sensitive to acute treatment with an antidepressant drug, whereas several weeks or months of antidepressant treatment is required before a clinical response is reported. Because the inducing factor (acute stress of swimming or suspension) is intrinsically coupled with the readout (time spent immobile), these tests also muddy the water between definitions of test versus model." The acute immobility response of mice and rats to antidepressant treatment compared with repeated exposure required for humans to note antidepressant effects indicates that these drugs act on, and these behaviors reflect, different mechanisms between species.

More recently, experimenters have shown that *chronic* treatment with the antidepressant fluoxetine also reduces time spent immobile in mice (Dulawa 2004). However, the immobility response also occurs after treatment with drugs that do not have antidepressant effects at all, such as antihistamines and other miscellaneous drugs (Arai 2000), putting the entire premise on unstable ground.

Experimental and strain effects

Experimental details such as water temperature and depth can alter an animal's behavior during the FST and potentially confound results. Jeffrys and Funder (1994) conducted an experiment designed to test whether or not water temperature influenced a rat's mobility. They found that when the water was 20°C, rats spent less time immobile, and were slower to learn immobility behavior over the course of the experiment (which included four exposures to the water tank), compared to when the water was 25 or 30°C. A different outcome has been observed for mice, with immobility decreasing in warmer water (Peeters 1991, Arai 2000).

The depth of water used by experimenters also influences results in the FST. De Pablo and colleagues (1989) found that placing rats in water with a depth of 35 cm increased swimming and decreased immobility compared to rats place in water with a depth of 15 cm. Presumably at 15 cm, the rats could detect the bottom of the container with their tails.

Importantly, mice show different behaviors in the FST depending on their strain. When comparing 11 commonly used strains of mice, Lucki and colleagues (2001) found that time spent immobile differed over 10-fold between the strain which swam the most and the one which swam the least. Strains also differed in their sensitivity to antidepressant drugs administered before the FST. Dulawa, et al. (2004) noted strain differences in the response to chronic fluoxetine treatment, where the drug regimen affected swimming and immobility times in BALB/c mice, but not in three other strains, including the ubiquitous C57BL/6 mouse.

The reality that variables such as water temperature, water depth, and strain alter FST results so dramatically and have the potential to confound interpretation further invalidate the FST as a reliable measure of 'despair' or behavior in general.

The forced swim test is used to make false conclusions

The problem with misinterpretation of the FST is that it has led to a false assumption that the FST can be used to measure depression in animals. Frighteningly, it has sometimes led to the assumption that the FST can serve as the sole measure used in a study to describe an animals' mood. In a recent commentary in *Psychoneuroendocrinology*, Molendijk and de Kloet (2015) estimate that in the 4300 papers reporting use of the FST at the time of publication, "[n]o less than [2,020] papers label the phenotype of the floating rodent as depression-like behavior—sometimes with a remark that the validity of the test is debated but often without discussion." Additionally, 7.5% of these (320 papers) had "used the FST to monitor the outcome of genetic manipulations of signaling pathways suspected to be involved in the precipitation of depression-like symptoms. Most of these studies (we estimate 70%) indeed infer a depression phenotype from the immobility response displayed by the rodent" (Molendijk and de Kloet 2015).

In a 2018 follow-up to this analysis, Molendijk and de Kloet found that, in the three years prior, "the popularity of the FST [was] still increasing" (Molendijk and de Kloet 2019). Of the papers they analyzed, 72 percent qualified the behavior of a floating mouse or rat as "depressive-like, but without evidence for face, predictive, or construct validity" (Molendijk and de Kloet 2019).

Industry and Academia Abandon the Forced Swim Test

In December 2018, PETA scientists began analyzing publicly available data on pharmaceutical companies' use of the FST to test novel compounds for their potential value as human antidepressants. PETA found that the predictive efficacy of the FST in these cases was less than 50 percent (Trunnell 2019). After being presented with their own data and asked by PETA to reconsider their position on the use of this test, 13 companies have committed to no longer conducting, commissioning, or funding the forced swim test: AbbVie, Astraea Therapeutics, AstraZeneca, Bayer, Boehringer Ingelheim, Bristol-Myers Squibb, DSM Nutritional Products, Johnson & Johnson, Novo Nordisk A/S, NutriFusion LLC, Pfizer, Roche, and Sage Therapeutics (PETA 2020).

In January 2020, major research university King's College London became the first known institution to end use of the forced swim test in an academic setting (Knapton 2020).

Conclusion

For decades, experimenters have been subjecting mice and rats to a stressful procedure where these vulnerable animals are made to swim in deep water with no way to escape. Experimenters have been using this procedure to make false and uninformed determinations about an animals' mood and to use these determinations to infer potentially false conclusions about biology related to human health.

The spread of the FST falsehood has wasted much in public funds, animal lives, and research hours. The onus to correct this bad science lies with several major players. Federal, state, and charity institutions can prevent these useless experiments before they occur by evaluating grant proposals for use of the FST or TST. Journals can prevent spurious conclusions based on the FST or TST from being reported and circulated in the literature by more closely scrutinizing manuscripts including animal behavioral protocols.

When the reality is that 90 percent of studies on animals fail to lead to therapies for humans (Garner 2014), there is something wrong with current methodologies. Animal experimentation has been cited as the primary source for attrition, or drug failure, in human neurobehavioral clinical trials (Garner 2014). It is time for experimenters to stop spending their time dunking mice and rats in water and focus their efforts on human-based experimental models, such as computational modelling using already well-defined biomarkers (Siekmeier 2015) and the use of patient-specific stem cells for personalized medicine, which "affords the ability to general neuronal cell-based models that recapitulate key aspects of human disease" (Haggarty 2016).

References

Arai, I., Tsuyuki, Y., Shiomoto, H., Satoh, M., Otomo, S. Decreased body temperature dependent appearance of behavioral despair in the forced swimming test in mice. *Pharmacological Research*. 42, 171-176 (2000).

Belmaker, R. H., Agam, G. Major depressive disorder. *New England Journal of Medicine*. 358, 55-68 (2008).

Commons, K. G., Cholanians, A. B., Babb, J. A., Ehlinger, D. G. The rodent forced swim test measures stress-coping strategy, not depression-like behavior. *ACS Chemical Neuroscience*. 8, 955-960 (2017).

De Pablo, J. M., Parra, A., Segovia, S., Guillamon, A. Learned immobility explains the behavior of rats in the forced swimming test. *Physiology and Behavior*. 46, 229-237 (1989).

Dulawa, S. C., Holick, K. A., Gundersen, B., Hen, R. Effects of chronic fluoxetine in animal models of anxiety and depression. *Neuropsychopharmacology*. 29, 1321-1330 (2004).

Garner, J. P. The Significance of Meaning: Why Do Over 90% of Behavioral Neuroscience Results Fail to Translate to Humans, and What Can We Do To Fix It? *ILAR Journal*. 55, 438-456 (2014).

Haggarty, S. J., Silva, M. C., Cross, A., Brandon, N. J., Perlis, R. H. Advancing drug discovery for neuropsychiatric disorders using patient-specific stem cell models. *Molecular and Cellular Neuroscience*. 73, 104-115 (2016).

Jeffrys, D., Funder, J. The effect of water temperature on immobility in the forced swimming test in rats. *European Journal of Pharmacology*. 253, 91-94 (1994)

Knapton, S. King's College becomes first university in the world to formally end 'cruel' mouse swimming tests. *The Telegraph*. https://www.telegraph.co.uk/science/2020/01/10/kings-college-becomes-university-world-formally-end-cruel-mouse/. Published 10 January 2020. Accessed 24 February 2020.

Lucki, I., Dalvi, A., Mayorga, A. J. Sensitivity to the effects of pharmacologically selective antidepressants in different strains of mice. *Psychopharmacology*. 155, 315-322 (2001).

Molendijk, M. L., de Kloet, E. R. Immobility in the forced swim test is adaptive and does not reflect depression. *Psychoneuroendocrinology*. 62, 389-391. (2015).

Molendijk, M. L., de Kloet, E. R. Coping with the forced swim stressor: Current state-of-the-art. *Behavioural Brain Research*. 364, 1-10. (2019).

Nishimura, H., Tsuda, A., Oguchi, M., Ida, Y., Tanaka, M. Is immobility of rats in the forced swim test "behavioral despair." *Physiology and Behavior*. 42, 93-95. (1988).

O'Leary, O. F., Cryan, J. F. Towards translational rodent models of depression. *Cell and Tissue Research*. 354, 141-153 (2013).

O'Neil, K. A., Valentino, D. Escapability and generalization: effect on 'Behavioral Despair.' *European Journal of Pharmacology*. 78, 379-380 (1982).

Peeters, B. W. M. M., Smets, R. J. M., Broekkamp, L. E. The involvement of glucocorticoids in the acquired immobility response is dependent on the water temperature. *Physiology and Behavior*. 51, 127-129 (1991).

People for the Ethical Treatment of Animals. Victories! PETA is ending near-drowning experiments on animals. https://www.peta.org/features/peta-ends-near-drowning-tests-small-animals/. Accessed 24 February 2020.

Porsolt, R. D., Le Pichon, M., Jalfre, M. Depression: a new animal model sensitive to antidepressant treatments. *Nature*. 266, 730-732 (1977).

Siekmeier, P. J. Computational modeling of psychiatric illness via well-defined neurophysiological and neurocognitive biomarkers. *Neuroscience and Biobehavioral Reviews*. 57, 365-380 (2015).

Trunnell, E. R. Are we throwing good antidepressants out with the swim test water? Labroots Laboratory Animal Sciences Conference, February 13-14, 2019. https://www.peta.org/wp-content/uploads/2019/03/LAS-2019-Poster-Dr.Emily-Trunnell.pdf

West, A. P. Neurobehavioral studies of forced swimming: The role of learning and memory in the forced swim test. *Progress in NeuroPsychopharmacology and Biological Psychiatry*. 14, 863-877 (1990).