Statement of Lori Marino, Ph.D.
Neuroscience and Behavioral Biology Program
Emory University, Atlanta, Georgia

to

The House Committee on Natural Resources Subcommittee on
Insular Affairs, Oceans and Wildlife
regarding educational aspects of public display of marine mammals
27 April 2010

Good morning. I am Lori Marino and I am a Senior Lecturer in Neuroscience and Behavioral Biology at Emory University and a faculty member in the Emory Center for Ethics. I am also an Adjunct Faculty member in the Department of Psychology at Emory University and former Research Associate at The Smithsonian Institution National Museum of Natural History. I have taught a variety of courses and been involved in course development and evaluation for the past fifteen years.

I wish to thank you, Chairwoman Bordallo, and members of the Subcommittee, for inviting me to testify on this panel addressing the educational aspects of public display of marine mammals. I very much appreciate the opportunity to share my professional experience and knowledge on this issue. Over the past seventeen years, I have published over eighty papers in the field of animal behavior, neuroscience, and human-animal interactions including forty peer-reviewed scientific papers on dolphin and whale brains, biology, intelligence and cognition and have studied dolphins in captivity and in the wild. Along with my colleague Diana Reiss I published the first definitive study demonstrating mirror self-recognition in bottlenose dolphins in 2001[1], and have published several in depth studies of brain structure, growth and complexity in bottlenose dolphins, orcas and several other marine mammal species. A list of my peer-reviewed publications in these areas is included as Attachment 1. I have also published several peer-reviewed papers on dolphin assisted therapy and human-dolphin interaction programs, as well as analyses of the educational claims of the zoo and aquarium community[2].

Introduction

The Marine Mammal Protection Act (MMPA) Section 104 (c) (2) (A) (i) requires that public display facilities provide a program of education or conservation for visitors that meets professionally recognized standards. Permission to display marine mammals rests upon meeting this criterion. In this testimony I will evaluate the evidence for adequate current recognized professional standards for education or conservation programs at public display facilities and the need for further regulation of these standards.

In order for a program to meet even minimum standards for education or conservation two very reasonable criteria must be met.

> First, the information provided about the animals on display and their natural history, biology, behavior and conservation status must be *accurate* (Criterion 1).
>
> Second, there must be evidence, *based on valid outcome measures*, that visits to these facilities serve an educational or conservation purpose (Criterion 2).

In this testimony I will evaluate the evidence that public display facilities are meeting their own current professional standards for education or conservation programs and the need for improved agency oversight of these programs. To do this I will use the public information provided by three major representative organizations – the Alliance of Marine Mammal Parks and Aquariums ('the Alliance'), the Association of Zoos and Aquariums ('the AZA'), and SeaWorld Parks and Entertainment ('SeaWorld'). These three organizations collectively represent more than 60% of the zoos and aquariums in the U.S. holding marine mammals on public display. A list of the web

materials critiqued here, and the websites where they can be found, is appended to the end of this testimony as Attachment 2.

Criterion 1: The information provided about the animals on display must be factual.

In order to assess whether this criterion is met I will evaluate the information provided by the Alliance in the section of their website entitled Frequently Asked Questions as well as online information provided by SeaWorld.

The Alliance is an international association representing theme parks, aquariums, zoos and other marine mammal facilities. Their online FAQs include the following question.

FAQ: How do the lifespans of dolphins in the wild and those in public display facilities compare?

Much of the information offered on the Alliance website about mortality and longevity rates in captivity and the wild is incorrect. The Alliance states that: "Beluga and killer whales in our facilities live as long as or longer than those in the wild… and… live long, happy lives". In fact, the best available scientific information indicates that these two species live much shorter lives in captivity than in their natural habitat. Furthermore, the emotional statement that the animals lead "happy lives" is pure speculation.

FAQ: Do marine mammals get stressed?

The Alliance appears to downplay the possibility that captive marine mammals can become stressed. They state that: "The results of behavioral and medical evaluations of animals in public display facilities indicate the animals breed very successfully, form social groupings, eat well and exhibit the same behaviors they do in the wild."

Furthermore, in support of this claim they state that: " a recent scientific study of steroid hormones produced by the adrenal cortex, a common measure of stress in animals, demonstrates that stress is not an issue in marine mammal in-water interactive programs. This Dolphin Quest/SeaWorld study was submitted to the U.S. government in September of 2000 and provides clear evidence that the animals are in a healthy environment." [3]

Before evaluating the validity of these statements it should be noted that these, like many of the claims made by the theme park community, are not based on peer-reviewed scientific papers. For instance, the Dolphin Quest/SeaWorld study referred to above was published as a short paper in conference proceedings but not, to the best of my knowledge, with the usual full details found in the peer-reviewed literature. Therefore these claims do not meet even the most minimal professional standards of peer-evaluation and would be unacceptable in any other legitimate research area.

Furthermore, the U.S. Marine Mammal Inventory Report (2010)[4] lists numerous stress-related disorders, such as ulcerative gastritis, perforating ulcer, cardiogenic shock and psychogenic shock as 'cause of death', strongly indicating that stress is an important component of captive display in marine mammals.

With that said, we can, at the very least, evaluate the specific findings of the Dolphin Quest/Sea World study from the limited information provided in the short proceedings paper. Stress in animals and humans can be measured by assessing elevations in stress hormone levels. In this study the authors compared stress hormone levels of captive bottlenose dolphins in swim-with-the-dolphin (SWTD) programs with those of dolphins in shows and concluded that: "…there continues to be no evidence…that animals involved in interactive SWTD programs experience any measurable levels of stress greater than *any other measured population of Tursiops* (my italics)." But these findings are only relevant as support for the 'no stress in captivity' claim if the authors had included a non-captive control group as a comparison. The fact that they did not means that *at most* the findings reveal that there are no significant differences in stress hormone levels between captive dolphins in shows and captive dolphins in swim programs. But these findings are irrelevant to the claim that they are attempting to support. Stress levels in captive dolphins could still be quite high compared to wild dolphins and this study would not be able to determine that.

Finally, the Alliance website states that: "In addition, symptoms commonly referred to as stress indicators, such as ulcers, are more common in wild animals that have been found stranded than in animals in responsible public display facilities."

But again, we are asked to take this information on faith as there is no way to evaluate its validity as one would normally be able to do in any other peer-reviewed research domain. Moreover, even if we accept their proposition the comparison to stranded animals is not the appropriate one. Stranded animals would be *expected* to have higher rates of pathologies, including ulcers, as a matter of course. The proper control group would be a random healthy sample from a wild population. Since the authors did not examine such a group their conclusion is uninterpretable at best.

As to the general assertion that captive dolphins and whales experience little if any stress, the scientific literature is absolutely clear on this issue: stress and its associated health problems are a recognized concern for captive dolphins. There is an abundant literature showing that stress in captive wildlife is a source of aberrant behavior, hyperaggressiveness, illness and mortality. [5] Recent work shows that handling and transportation of captive dolphins is so stressful that it can affect their immune system function.[6]

Stress derives from many aspects of captivity, not the least of which is stress associated with the many changes in social groupings and isolation that occurs in captivity. Social relationships play a critical role in the lives and well-being of dolphins and whales. In the wild individuals can have very strong and long-lasting relationships.[7] Conflict in the wild is resolved through various effective means that often require shifting alliances within large groups of animals[8], an opportunity not afforded by captivity. And social group composition is dynamic and fluid with individuals exerting choice about their associations. In the confines of captivity where social groups are often artificially constructed and transferred in and out of different pools and facilities without choice, and there is not enough room or social support to resolve conflict, dolphins and whales suffer extreme stress that has led to deaths and reduced life expectancy. [9]

Moreover, stress can be a result of the physical conditions and risks associated with the conditions in these facilities. For instance, ingestion of foreign objects is listed several times as a cause of

death in the U.S. Marine Mammal Inventory Report[4], a situation that arises due to the public's lack of supervision at poolside in many facilities.

In a policy paper on dolphin-human interaction programs my co-author and I reviewed the scientific evidence for stress in captive cetaceans, and were led to conclude that: "Many captive dolphins display physiological and behavioral indicators of stress, including elevated adrenocortical hormones, stereotypies, self-destruction, self-mutilation and excessive aggressiveness. [10]

Despite the implication in the Alliance FAQ that marine mammals in captivity do not experience stress, there is ample scientific evidence to the contrary. Captive dolphins and whales not only experience stress, they are often very seriously affected by it.

FAQ: Do dolphins and whales have unique intelligence?

The Alliance and SeaWorld

On the one hand, the Alliance seems to suggest that the intelligence of dolphins is high enough to make them suitable "subjects" in various human-driven activities, while, on the other hand, it downplays that same intelligence so as to undermine concerns about keeping these intelligent animals in captivity. Like the last bowl of porridge in the Goldilocks fairytale dolphin intelligence is *just right*.

SeaWorld publishes online information booklets (Animal Info Books) on bottlenose dolphins, orcas, beluga whales and other animals in their parks. These pamphlets and other information resources, such as their teacher's guides, are littered with inaccuracies – all aimed at biasing perceptions of dolphins and other cetaceans as interesting but rather ordinary animals in an attempt to allay any notions that they are animals with such a complex intelligence that they may not be suitable for captivity.

In the FAQ section of their website the Alliance makes a number of misleading and erroneous statements about dolphin intelligence. For example, they claim that: "…dolphins are large animals with proportionately sized brains." This statement is patently false. In the scientific community, brain size is evaluated by taking body size into account. Large animals have large brains and small animals have small brains. Many animals have brains that are proportionate to their body size. However, some animals have brains that are much larger than would be expected for their body size. Humans, for instance, have brains that are seven times larger than they should be for our body size – they are out of proportion. The same is true of dolphins. Many dolphin species have brains three, four or five times larger than expected for their body size.[11] Therefore, like humans, *dolphin brains are out of proportion for their body size*. They do not have proportionately sized brains, as the Alliance website claims. This fact is relevant because those species that have larger brains than expected tend to show exceptional intelligence in many ways. Just as human intelligence is, at least partly, due to our larger-than-expected brain, so is dolphin intelligence. In fact, dolphin relative brain size is second only to that of modern humans. The Alliance apparently wishes to hide this similarity along with any concerns that dolphin sensitivities may be too similar to that of humans for them to be in captivity.

SeaWorld also makes several statements about dolphin intelligence in their online Animal Info Books that are misleading. In their book on bottlenose dolphins, SeaWorld admits that dolphins have larger brains than many other animals of their body size but follow with: "One likely theory is that a larger brain size in dolphins may be at least partially due to an increased size of the auditory region to facilitate sound processing." And in their book on beluga whales: "The auditory cortex of the brain is highly developed". These statements are, by themselves, not false. However, they clearly are meant to imply that dolphin and whale brains are large simply in order to process sound and not because they are processing more complex information at more abstract levels. This notion (which is sometimes referred to as "the dolphin brain as a large radar screen") is an outmoded theory that is not based on current scientific knowledge of dolphin and whale brains. Researchers have identified the parts of the dolphin and whale brain that process sound information alone (the auditory cortex). These structures do not account for most of the large mass of dolphin and whale brains. In fact, most of the dolphin and whale cortex is not associated with sensory processing and is apparently involved in higher-level information processing and thinking, just as our cortex is.[12] The fact that the Alliance and SeaWorld neglect to mention this fact suggests that either they do not know the current facts on dolphin and whale brains or are attempting to bias readers' ideas about intelligence in dolphins and whales. In either case they are not meeting best current scientific knowledge standards.

It is interesting that, despite their claims about the ordinary nature of dolphin and whale brains, they hedge their bets with a litany of misleading and plainly incorrect statements. The Alliance claims that: "… brain size does not indicate intelligence" and "… it is impossible and inappropriate to compare the intelligence of different species." And in their online book on bottlenose dolphins SeaWorld states: "Hypotheses that large brain size in dolphins indicates high intelligence are untested and disputed." And "Rating the intelligence of different animals is misleading and extremely subjective. In fact, a reliable and consistent intelligence test for humans has yet to be developed." All of these statements range from false to misleading. It is neither impossible nor inappropriate to compare different aspects of intelligence (learning, memory, problem solving, behavioral flexibility, etc.) across species. The established scientific fields of comparative psychology, cognitive ethology and behavioral neuroscience are based on the comparison of brains and behavior across species. What we currently know is that, while brain size is not a perfect predictor of intelligence and there are other aspects of the brain that relate to intelligence as well, brain size is correlated with a host of behaviors and cognitive abilities that are considered components of intelligence. These include feeding complexity, social complexity, frequency of innovation and tool use, behavioral flexibility and variability, and self-awareness (see below). [13] Few people have trouble with the fact that our own prodigious level of intelligence is related to our large brains. It would be inconsistent to think otherwise for dolphins and whales or any other species. And, although a side point, the claim that we do not currently have a reliable intelligence test for humans is also false. There are currently several well-constructed, valid and reliable cross-cultural intelligence tests for humans. [14]

The claim that "Hypotheses that large brain size in dolphins indicates high intelligence are untested and disputed" is, again, misleading. Although there have been some theories put forth that dolphin brains and intelligence are limited, these ideas have not stood up to scientific scrutiny. The claim that the brain-intelligence hypothesis has never been tested in dolphins and whales is ludicrous and ignores several decades of scientific work.

The Alliance website also states that: "… people continue to infer that dolphins and whales are uniquely intelligent" in their continued effort to refute the views of many observers regarding dolphin and whale intelligence. In fact, we do not need to *infer* that dolphins and whales are uniquely intelligent at all because we have decades of scientific research that *demonstrates* the complex intelligence of dolphins and whales. The scientific evidence from decades of research clearly shows that dolphin intelligence cannot be characterized as average or ordinary in any way. It is, in fact, exceptional in a number of ways and very similar, in many respects, to that of our own. Dolphins and whales possess sophisticated learning, problem solving, communicative and even cultural abilities, including the possession of some capacities that are extremely rare. [15]
One example is mirror self-recognition, the ability to recognize oneself in a mirror. In 2001 my colleague Diana Reiss and I showed that bottlenose dolphins are capable of using a mirror to investigate their own bodies – an ability that even human children do not reliably possess until they are two years old. More importantly, this and other studies show that dolphins have self-awareness, a sense of themselves not unlike our own. [1]

In summary, the claim that dolphins and whales are ordinary in brain size and intelligence is far from correct. All of the scientific evidence points clearly to the conclusion that dolphins and whales have brains that are larger and more complex than expected and intellectual capacities that few other animals possess and are strikingly similar to our own.

Criterion 2: There must be evidence, based on valid outcome measures, that visits to these facilities serve an educational or conservation purpose

If we accept, for argument's sake, that the information being given by the theme park community is valid, then we must ask whether it is, in fact, effective education. As someone with over 15 years experience as an educator I am well aware that what students say they know and what they actually know are often very different. When an educator wishes to determine if education (or learning) has taken place the standard practice is to test the student's knowledge. It is not proper to simply ask students whether they have learned or what they *think* they have learned or how much they enjoyed the class, in order to determine if learning has taken place. Only by directly testing knowledge can learning and education be assessed with any validity. *Tests of knowledge are the most direct and authentic outcome measures in education.* We then may ask whether any valid outcome measures exist for learning through public displays of marine mammals.

The literature that the captivity community relies upon to support their claims of education are characterized by the very weakness described above; the studies typically involve asking zoo and aquarium visitors whether they *think* they have been educated. But they do not actually test knowledge.

The Roper Poll

As an example, the Alliance poses the following question on its FAQ site: *Are people learning about marine mammals from zoos and aquariums?* They answer that:

"A 1998 [it was actually 1995] Roper Starch poll[16]… provides clear evidence that programs at Alliance member marine life parks, aquariums, and zoos are educational and provide the public with a heightened appreciation of the importance of conserving marine mammals. Ninety-four percent (94%) of the park visitors interviewed for the poll said, "I learned a great deal about marine mammals today. Responses to the poll indicate that seeing living marine mammals enhances the educational experience for the visitors to these zoological parks and aquariums. Almost everyone (97%) interviewed said their experience with living marine mammals had an impact on their appreciation and knowledge of the animals. The impact was greater for those visiting facilities where they actually had an opportunity to interact with marine mammals."

They conclude that: "The Roper poll shows that Alliance member marine life parks, aquariums, and zoos successfully teach visitors about marine mammals and, additionally, serve to inform visitors about environmental issues that may have an impact on the animals."

In fact, no such conclusions can be drawn from the data provided by the Roper poll. The intent of the study was to identify public attitudes and opinions toward animal facilities. And that is what it did. The Alliance seems to suggest that the poll shows that visits to zoos and aquariums create "heightened appreciation" of marine mammals but, once again, this is not what the Roper poll asked so there is no possibility that this poll could provide support for the claim that Alliance theme parks *teach* visitors about marine mammals and *inform* about environmental issues. (The Alliance website also fails to inform the reader that the Roper poll was commissioned by SeaWorld.)

The Harris Poll

In the same section the Alliance states that a 2005 online survey conducted by Harris Interactive Poll®[17] found that, "…97 percent of respondents agree that marine life parks, aquariums and zoos play an important role in educating the public about marine mammals they might not otherwise have the chance to see. In addition, 96 percent agree that marine life parks, aquariums and zoos provide people with valuable information about the importance of oceans, waters and the animals that live there. The poll also shows that if looking for educational information about marine mammals, 75 percent of the survey participants would either visit a marine life park, aquarium or zoo or go to their Web sites." In its online information books, SeaWorld also makes the case that captive animal facilities are educational by citing this 2005 Harris poll.

The Alliance concludes that: "Results of the Harris Interactive® and Roper polls indicate that visitors are coming away from their marine mammal experiences with a heightened overall environmental concern and additional interest in taking environmental action."

Like the Roper poll, the Harris poll apparently only assesses visitors' beliefs and perceptions, not whether they actually learned. The poll does not ask respondents about the specific knowledge they have gained, or what specific conservation actions they will undertake after visiting a public display facility.

The NAIB Study

Another similar study[18] was conducted at the National Aquarium in Baltimore (NAIB), a non-profit dolphin display facility. The authors used entry and exit polls to assess four key aspects of the visitor experience: (1) incoming conservation knowledge, attitudes, and behavior of NAIB visitors; (2) patterns of use and interaction with exhibition components throughout the NAIB; (3) exiting conservation knowledge, attitudes, and behaviors of visitors; and (4) over time, how the NAIB experience altered or affected individuals' conservation knowledge, attitudes, and behaviors. They concluded that there were changes in visitors' conservation knowledge, understanding, and interests. However, this study was riddled with numerous flaws and potentially confounding variables that undermine its validity.

First, the entry and exit interviews were conducted face-to-face and no information is offered on the details of these interviews and, most importantly, how, or if, they minimized the confounding demand characteristics inherent in this method. In other words, we do not know whether the questions asked were leading or biased and we do not know whether there were inadvertent cues from the interviewers that influenced the visitor's responses.

Second, all of the questions asked were about the conservation message of the NAIB – not about actual conservation knowledge per se. They found that visitors were able to pick up on the intended conservation message of the aquarium. But this does not show that the visits impacted conservation knowledge. It simply shows that the NAIB was successful in making clear to visitors that their message is one of conservation. It would be very surprising if they did not find this given all of the signage and efforts put forth at the facility to impart this message. But the message is more about NAIB and the way it wants to be perceived than about real conservation.

Third, the authors report that the participants were a self-selected population and were generally more knowledgeable about, more concerned about, and more involved in conservation-related issues than the general public. Therefore, it is unclear that these findings are relevant to the general public, which is the main population of visitors to aquariums around the country. The authors admit that the visitors' general working knowledge and associations with conservation did not tend to be impacted by the aquarium visit.

Fourth, importantly, there was no evidence that a visit to the NAIB changed the visiting public's conservation actions. In fact, as the authors report, after a few weeks, their enthusiasm and emotional commitment to conservation generally fell back to original levels. The visit had no lasting impact on behavior.

All of this is not to say that sound conservation education cannot take place in aquaria. However, the link between aquaria and meaningful and lasting education, effect on attitudes, and impact on behavior is unclear. Certainly, the link between the above and captive marine mammal exhibits is even less clear.

Why Zoos Matter

The largest and most recent visitor research study conducted is a multi-institution research program entitled "Why Zoos Matter: Assessing the impact of a visit to a zoo or aquarium" published online by the AZA in 2007[19]. In this paper the authors referred to a comprehensive review article by

Dierking et al. [20] which revealed that there are few studies demonstrating actual changes in behavior as a result of a zoo or aquarium visit. The AZA concluded that visitor research up to that point gave only "an incomplete picture about the impact zoos and aquariums have on conservation-related knowledge, attitudes and behavior." (p.5). The AZA study was conducted to address this deficit of information and provide more conclusive evidence that visits to zoos and aquaria impact knowledge and conservation attitudes. Over three years, more than 5,500 visitors and 12 zoos and aquariums participated in this government-funded study. On the basis of their findings the authors concluded that visits to zoos and aquariums have a measurable positive impact on the conservation attitudes and understanding of adult visitors. Moreover, this study, which was not published in a peer-reviewed journal, was heralded as the first to validate the idea that zoos and aquariums are having a strong positive impact on visitor attitudes. Cynthia Vernon, vice president of conservation programs for the Monterey Bay Aquarium and one of the investigators in the study stated that: "The Visitor Impact Study shows that zoos and aquariums are enhancing public understanding of wildlife and the conservation of the places animals live." And AZA President and CEO Jim Maddy boldly asserted that: "For the first time we have reliable data validating the positive impact zoos and aquariums have in changing visitors' feelings and attitudes about conservation". These conclusions are cited on the AZA website and several other theme park websites as the "holy grail" that the captivity community has been searching for to validate their educational and conservation claims.

However, in a methodological critique[2] of the AZA study recently published in the peer-reviewed journal, *Society & Animals*, my co-authors and I found that these conclusions are entirely unfounded. I am appending a copy of our paper as Attachment 4, so I will not go into detail here about all of the methodological weaknesses of the AZA study. However, I would like to point out some of the more significant flaws in the study and their impact on the validity of the authors' conclusions.

Several of the weaknesses in the AZA study had to do with a lack of control over confounding factors that could potentially bias the results. These include but are not limited to non-random sampling, lack of control for general effects of being in a novel environment (being on vacation in a new place, for example), and demand characteristics (again, the interviewers' inadvertent cuing of the responders, which would lead to response bias). All of these flaws have an impact on both the validity and the ability to generalize from this study. Furthermore, setting all of the multiple methodological weaknesses aside for the moment, the study suffers from the same conceptual weakness previous visitor research studies possess: The AZA study did not assess attitudes or knowledge. They only evaluated what visitors *believed* they felt or learned. Finally, when all was said and done the actual reported gains in stated visitor knowledge were disappointing. The authors found that: "there was no overall statistically significant change in understanding…" (p. 10). That is, the authors of the AZA study found no significant gains in general knowledge resulting from zoo or aquarium visits.

Taken together, it is abundantly clear that the conclusions of the AZA study are unfounded. It is, frankly, surprising that the authors based such strong claims on these flawed findings.

Given that the AZA study was the most comprehensive and recent visitor research study to have the potential to show that visits to zoos and aquariums are educational and given that this was not

accomplished, there is no compelling or even strongly suggestive current evidence that visits to zoos and aquariums promote positive attitude change, learning or conservation actions.

Representatives of the theme park community would like the public to think that they are doing cutting-edge science and that *academic* standards for science and research methodology are outdated. While methods for collecting and analyzing data may evolve, the criteria for good science *have always and will always be the same*. They are based on logic and logic does not change. Two plus two will always equal four regardless of *when* you do the math!

Conclusion

At the outset of my testimony, I offered that two minimal criteria must be met in order for an education program to be considered valid. First, the information provided about the animals on display must be *accurate* (Criterion 1). Second, there must be evidence, *based on valid outcome measures*, that visits to these facilities serve an educational or conservation purpose (Criterion 2). A review of a subset of the online materials published by the Alliance and SeaWorld shows that much of the information provided to the public is either misleading or incorrect. And the analyses of the visitor research studies above demonstrates that, to date, there is no compelling evidence that visiting zoos and aquariums is an authentic educational experience. Therefore, neither of the two criteria are met. It is difficult to understand how claims about effective education can be made when there is so little evidence to support them.

Given that the captivity community has clearly not met minimal educational standards it is urgent that the NMFS work to ensure compliance with the education standards set by the MMPA for display permit holders. NMFS is urged to exert greater control over this important issue and its very serious consequences.

Attachment 1

List of relevant peer-reviewed publications, from most recent on, dolphin and whale brains and intelligence authored (or co-authored) by Lori Marino, Ph.D., Emory University

Sarko D, Domning D, Marino L, Reep R. (in press) Estimating body size of fossil Sirenians. *Marine Mammal Science*.

Marino L, Butti, C, Connor RC, Fordyce, RE, Herman LM, Hof PR, Lefebvre L, Lusseau D, McCowan B, Nimchinsky EA, Pack AA, Reidenberg JS, Reiss D, Rendell L, Uhen MD ,Van der Gucht E, Whitehead H. (2008) A claim in search of evidence: Reply to Manger's thermogenesis hypothesis of cetacean brain structure. *Biological Reviews of the Cambridge Philosophical Society,* 83: 417-440.

Montie, EW, Ketten DR, Schneider G, Marino L, Touhey KE, Hahn ME. (2008) Volumetric neuroimaging of the Atlantic white-sided dolphin (*Lagenorhynchus acutus*) brain from *in situ* magnetic resonance images. *The Anatomical Record*, 291: 263-282.

Montie EW, Ketten DR, Schneider G, Marino L, Touhey KE, Hahn ME. (2007). Neuroanatomy of the subadult and fetal brain of the Atlantic white-sided dolphin (*Lagenorhynchus acutus*) from in situ magnetic resonance images. *The Anatomical Record*, 290: 1459-1479.

Marino L, Connor RC, Fordyce, RE, Herman LM, Hof PR, Lefebvre L, Lusseau, McCowan B, Nimchinsky EA, Pack AA, Rendell L, Reidenberg JS, Reiss D, Uhen MD ,Van der Gucht E, Whitehead H. (2007) Cetaceans have complex brains for complex cognition. *Public Library of Science (PLOS) Biology*, 5(5): e139.

Marino L (2007) Cetacean brains: How aquatic are they? *The Anatomical Record*, 290:694-700.

Marino L (2006) Absolute brain size: Have we thrown the baby out with the bathwater? Invited commentary in *Proceedings of the National Academy of Sciences USA*, 103(37): 13563-13564.

Marino L, Sol D, Toren K, Lefebvre L, (2006). Does diving limit brain size in cetaceans? *Marine Mammal Science*. 22(2): 413-425.

Marino L, Hof P (2005) Nature's experiments in brain diversity. *The Anatomical Record.* 287A: 997-1000.

Hof P, Chanis R, Marino, L. (2005). Cortical complexity in cetacean brains. *The Anatomical Record.* 287A: 1142-1152.

Marino L (2005) Big brains matter in novel environments. Invited commentary in *Proceedings of the National Academy of Sciences USA.* 102(15): 5306 – 5307.

Marino L, McShea D, Uhen MD (2004) The origin and evolution of large brains in toothed whales. *The Anatomical Record*. 281A: 1247-1255.

Marino L, Sherwood CC, Tang CY, Delman BN, Naidich TP, Johnson JI, Hof PR (2004) Neuroanatomy of the killer whale (*Orcinus orca*) from Magnetic Resonance Imaging. *The Anatomical Record*. 281A: 1256-1263.

Marino L (2004) Dolphin cognition. *Current Biology*. 14, R910-R911.

Marino L, Sudheimer K, Pabst D A, McLellan WA, Arshad S, Naini G, Johnson JI (2004) Anatomical description of an infant bottlenose dolphin, *Tursiops truncatus,* brain from Magnetic Resonance Images. *Aquatic Mammals* 30: 315-326.

Marino L, Pabst DA, McLellan WA, Sudheimer K, Johnson JI (2004) Neuroanatomical structure of the spinner dolphin (*Stenella longirostris orientalis*) brain from magnetic resonance images. *The Anatomical Record* 279A: 601-610.

Marino L (2004) Cetacean brain evolution – multiplication generates complexity. *International Journal of Comparative Psychology* 17: 1-16.

Marino L, Pabst DA, McLellan WA, Sudheimer K, Johnson JI (2003) Magnetic resonance images of the brain of a dwarf sperm whale (*Kogia simus*). *Journal of Anatomy* 204: 57-76.

Marino L, Sudheimer K, Sarko D, Sirpenski G, Johnson JI (2003) Neuroanatomy of the harbor porpoise (*Phocoena phocoena*) from magnetic resonance images. *Journal of Morphology* 257: 308-347.

Marino L, Uhen MD, Pyenson ND, Frohlich BF (2003) Reconstructing cetacean brain evolution using computed tomography. *The New Anatomist* 272B: 107-117.

Sarko D, Marino L, Reiss D (2003) A bottlenose dolphin's (*Tursiops truncatus)* responses to its mirror image: further analysis. *International Journal of Comparative Psychology* 15: 69-76.

Marino L, Sudheimer K, Pabst DA, McLellan W A, Filsoof D, Johnson J I (2002) Neuroanatomy of the common dolphin *(Delphinus delphis*) as revealed by magnetic resonance images (MRI). *The Anatomical Record* 268: 411-429.

Ridgway SH, Marino L, Lipscomb T (2002) Description of a poorly differentiated carcinoma within the brainstem of a White whale (*Delphinapterus leucas*) from magnetic resonance images and histological analysis. *The Anatomical Record* 268: 441-449.

Marino L (2002) Convergence in complex cognitive abilities in cetaceans and primates. *Brain, Behavior and Evolution* 59: 21-32.

Marino L, Uhen MD, Frohlich B, Aldag JM, Blane C, Bohaska D, Whitmore Jr FC (2000) Endocranial volume of mid-late Eocene archaeocetes (Order: Cetacea) revealed by computed tomography: Implications for cetacean brain evolution. *Journal of Mammalian Evolution* 7: 81-94.

McCowan B, Marino L, Vance E, Walke L, Reiss D (2000) Bubble ring play of bottlenose dolphins: Implications for cognition. *Journal of Comparative Psychology* 114: 98-106.

Marino L (1998) Brain growth patterns in the La Plata River dolphin (*Pontoporia blainvillei*). *Aquatic Mammals* 24: 111-116.

Marino L (1998) Quantifying brain-behavior relations in cetaceans and primates. Correspondence in *Trends in Ecology and Evolution* 13 [148]: 408.

Marino L, Sudheimer K, Murphy TL, Davis KK, Pabst DA, McLellan W, Rilling JK, Johnson JI (2001) Anatomy and three-dimensional reconstructions of the bottlenose dolphin (*Tursiops truncatus*) brain from magnetic resonance images. *The Anatomical Record* 264: 397-414.

Reiss D, Marino L (2001) Self-recognition in the bottlenose dolphin: A case of cognitive convergence. *Proceedings of the National Academy of Sciences USA* 98 (10): 5937-5942.

Marino L, Murphy TL, Gozal L, Johnson JI (2001) Magnetic resonance imaging and three-dimensional reconstructions of the brain of the fetal common dolphin, *Delphinus delphis*. *Anatomy and Embryology* 203: 393-402.

Marino L, Murphy TL, DeWeerd AL, Morris JA, Ridgway SH, Fobbs AJ, Humblot N, Johnson J I (2001) Anatomy and three-dimensional reconstructions of the brain of a white whale (*Delphinapterus leucas*) from magnetic resonance images (MRI). *The Anatomical Record* 262: 429-439.

Marino L, Rilling JK, Lin SK, Ridgway SH (2000). Relative volume of the cerebellum in the bottlenose dolphin and comparison with anthropoid primates. *Brain, Behavior, and Evolution* 56: 204-211.

Marino L (1998) A comparison of encephalization between odontocete cetaceans and anthropoid primates. *Brain, Behavior, and Evolution* 51: 230-238.

Marino L, Stowe J (1997) Lateralized behavior in a captive beluga whale *(Delphinapterus leucas*). *Aquatic Mammals* 23: 101-103.

Reiss D, McCowan B, Marino L (1997) Communicative and other cognitive characteristics of bottlenose dolphins. *Trends in Cognitive Sciences* 1: 140-145.

Marino L, Stowe J (1997) Lateralized behavior in two captive Bottlenose dolphins (*Tursiops truncatus*). *Zoo Biology* 16: 173-177.

Marino L (1997) The relation between encephalization level, gestation length, and body weight in Odontocetes (toothed whales). *Marine Mammal Science* 13: 133-138.

Marino L (1996) What can dolphins tell us about primate evolution? *Evolutionary Anthropology* 5: 81-85.

Attachment 2

List of website materials analyzed for testimony of Lori Marino, Ph.D., Emory University

The Alliance for Marine Mammal Parks and Aquariums FAQs
All materials were accessed as recently as April 2010
http://www.ammpa.org/faqs.html

SeaWorld Educational Materials
All materials were accessed as recently as April 2010

Information Books:

- Beluga Whales: http://seaworld.org/animal-info/info-books/beluga/index.htm
- Bottlenose Dolphins: http://seaworld.org/animal-info/info-books/bottlenose/index.htm
- Killer Whales: http://seaworld.org/animal-info/info-books/killer-whale/index.htm

Association of Zoos and Aquariums
All materials were accessed as recently as April 2010
http://www.aza.org/uploadedFiles/Education/why_zoos_matter.pdf

Attachment 3

Endnotes for testimony of Lori Marino, Ph.D., Emory University

[1] Reiss D, Marino L (2001) Self-recognition in the bottlenose dolphin: A case of cognitive convergence. *Proceedings of the National Academy of Sciences USA* 98 (10): 5937-5942.

[2] Marino L, Lilienfeld S, Malamud R, Nobis N, Broglio R (2010). Do zoos and aquariums promote attitude change in visitors? A critical evaluation of the American Zoo and Aquarium study. *Society and Animals,* 18: 126-138.

[3]Sweeney J C, Dold C, Riederson T, Reddy M (2001) Circulating levels of cortisol and aldosterone in *Tursiops truncatus*: a comparative look at display animals and animals in SWTD programs. Presented and published in the proceedings at the 2001 Conference of the International Association for Aquatic Animal Medicine.

[4] U.S. Marine Mammal Inventory Report (2010) National Marine Fisheries Service, Office of Protected Resources.

[5] Fowler ME (1978) in *Zoo and Wild Animal Medicine*, ME Fowler (ed). WB Saunders: Philadelphia, pp 33-34.

[6]Noda K, Akiyoshi H, Aoki M, Shimada T, Ohashi F. (2007) Relationship between transportation stress and polymorphonuclear cell functions of bottlenose dolphins, *Tursiops truncatus*. *Journal of Veterinary Medical Science, 69(4)*: 379-383.

7See, for example, Shane SH, Wells RS, Wursig B. (1986) Ecology, behavior and social organizations of the bottlenose dolphin: A review. *Marine Mammal Science, 2(1)*: 34-63.

[8] See, for example, Connor RC, Heithaus MR, Barre LM. (1999) Superalliance of bottlenose dolphins. *Nature, 397*: 571-572.

[9]Waples KA, Gales NJ. (2002). Evaluating and minimising social stress in the care of captive bottlenose dolphins (*Tursiops aduncus*). *Zoo Biology, 21*: 5-26.

[10] Stewart K L, Marino L. (2009). Dolphin-human interaction programs: policies, problems, and practical alternatives. Policy paper for *Animals and Society Institute* (42 pgs).

[11] Marino L, Sol D, Toren K, Lefebvre L, (2006). Does diving limit brain size in cetaceans? *Marine Mammal Science*. 22(2): 413-425.

[12]Marino L. (2006) Cetacean brain evolution. In *Evolution of nervous systems in mammals,* V 3 Ed by L Krubitzer, JH Kaas. Elsevier. Oxford, UK, pp. 261-266.

[13] Marino L (2005) Big brains matter in novel environments. Invited commentary in *Proceedings of the National Academy of Sciences USA. 102(15)*: 5306 – 5307.

[14] Phelps RP. (2009) *Correcting Fallacies About Educational and Psychological Testing,* American Psychological Association, Wash DC.

[15] Marino L, Butti, C, Connor RC, Fordyce, RE, Herman LM, Hof PR, Lefebvre L, Lusseau D, McCowan B, Nimchinsky EA, Pack AA, Reidenberg JS, Reiss D, Rendell L, Uhen MD ,Van der Gucht E, Whitehead H. (2008) A claim in search of evidence: Reply to Manger's thermogenesis hypothesis of cetacean brain structure. *Biological Reviews of the Cambridge Philosophical Society,* 83: 417-440.

[16]Roper Starch Worldwide (1995) Public attitudes towards zoos, aquariums and animal theme parks.

[17]Harris Interactive Poll® (2005) (www.harrisinteractive.com)

[18] Adelman LM, Falk JH, James S. (2000) Impact of National Aquarium in Baltimore on visitor's conservation attitudes, behavior and knowledge. *Curator, 43*(1): 33-61.

[19] Falk JH, Reinhard EM, Vernon CL, Bronnenkant K, Deans NL. Heimlich JE. (2007). Why Zoos & Aquariums Matter: Assessing the Impact of a Visit. *Association of Zoos & Aquariums.* Silver Spring, MD.

[20]Dierking LD, Burtnyk K, Buchner KS, Falk JH. (2002) Visitor Learning in Zoos and Aquariums: A literature review. Silver Spring, MD: American Zoo and Aquarium Association.

Attachment 4

BRILL

Society and Animals 18 (2010) 126-138

Society & Animals Journal of Human-Animal Studies

brill.nl/soan

Do Zoos and Aquariums Promote Attitude Change in Visitors? A Critical Evaluation of the American Zoo and Aquarium Study

Lori Marino,[a] Scott O. Lilienfeld,[a] Randy Malamud,[b] Nathan Nobis,[c] Ron Broglio[d]

[a] Emory University
lmarino@emory.edu
[b] Georgia State University
[c] Morehouse College
[d] Arizona State University

Abstract

Modern-day zoos and aquariums market themselves as places of education and conservation. A recent study conducted by the American Zoo and Aquarium Association (AZA) (Falk et al., 2007) is being widely heralded as the first direct evidence that visits to zoos and aquariums produce long-term positive effects on people's attitudes toward other animals. In this paper, we address whether this conclusion is warranted by analyzing the study's methodological soundness. We conclude that Falk et al. (2007) contains at least six major threats to methodological validity that undermine the authors' conclusions. There remains no compelling evidence for the claim that zoos and aquariums promote attitude change, education, or interest in conservation in visitors, although further investigation of this possibility using methodologically sophisticated designs is warranted.

Keywords

aquarium, attitude, conservation, education, marine park, methodology, validity, zoo

Background

Displays of captive animals have existed since ancient times. Zoos and aquariums (which include marine parks) were until recently generally accepted forms of entertainment, with little thought given to their purpose or the trade-offs associated with the capture and confinement of animals. Since the 1970s, however, public awareness of nature and environmental and conservation issues has come to the fore. Many zoos and aquariums have responded to this shift in political winds by rebranding themselves as agents for species preservation and public education.

DOI: 10.1163/156853010X491980

Over the years a number of studies have yielded an incomplete understanding of the impact of zoos and aquariums on educational and conservation-oriented objectives (see Dierking, Burtnyk, Buchner, & Falk, 2002, for a review). A recent study conducted by the American Zoo and Aquarium Association (AZA) (Falk et al., 2007) titled "Why Zoos and Aquariums Matter: Assessing the Impact of a Visit to a Zoo or Aquarium," however, is being widely heralded as the first direct evidence that visits to zoos and aquariums produce long-term positive effects on people's attitudes toward other animals.

The AZA accredits, represents, and promotes 216 of America's most prominent zoos and aquariums. Of these members, a substantial number currently tout the findings of this study on their Web sites, including the Monterey Bay Aquarium, the Naples Zoo (Naples, Florida), the Fresno Chaffee Zoo, the El Paso Zoo, and many others. The press release quoted by most of these sites refers to the investigation as a "groundbreaking study" and claims that "visiting accredited zoos and aquariums in North America has a measurable impact on the conservation attitudes and understanding of adult visitors" (American Zoo and Aquarium Association, 2006). It goes on to quote Cynthia Vernon, vice president of conservation programs for the Monterey Bay Aquarium and an investigator on the study: "The Visitor Impact Study shows that zoos and aquariums are enhancing public understanding of wildlife and the conservation of the places animals live. It validates the idea that we are having a strong impact on our visitors" (American Zoo and Aquarium Association, 2006). It further quotes AZA President and CEO Jim Maddy as asserting that "For the first time we have reliable data validating the positive impact zoos and aquariums have in changing visitors' feelings and attitudes about conservation." (American Zoo and Aquarium Association, 2006). As of May, 2009, the AZA report had been cited 10 times by various zoos and aquariums (Google Scholar search, May 15, 2009) and yielded approximately 120 Web hits (Google Web search, May 15, 2009), virtually all of them providing laudatory coverage of the Falk et al. study.

For these reasons, the AZA report warrants particularly careful scrutiny. If the claims made on behalf of many zoos and aquariums regarding the AZA report go beyond its findings, consumers of zoo and aquarium Web sites and other promotional materials may come to misleading conclusions. Moreover, as Mason (2000) notes, there is a marked dearth of information on the effects of zoos and aquariums on visitors, making this report particularly noteworthy. Indeed, the questions addressed by Falk et al. are undeniably important. Although our analysis will identify significant methodological weaknesses in the AZA report, our intent is ultimately constructive. Specifically, we wish to use the AZA report as an object lesson that may aid future researchers in this

area in avoiding methodological pitfalls, some of which are shared by other visitor research on zoos and aquariums.

The Falk et al. study comprised two phases. The first focused on the motivations that lead people to visit zoos and aquariums, and the second attempted to measure changes in attitudes toward conservation as a result of visiting the institutions. The study's primary goal was to "assess the impact of a zoo and aquarium visit on adults, as well as develop a set of tools that every institution could use for assessing their conservation impact on visitors" (Falk et al., 2007, p. 6). Over three years, more than 5,500 visitors and 12 zoos and aquariums participated in the study. On the basis of their findings, the authors concluded that visits to zoos and aquariums have a measurable positive impact on the conservation attitudes and understanding of adult visitors. In this article, we address whether this conclusion is warranted by analyzing the study's methodological soundness.

Analysis and Findings

The major hypothesis of Falk et al. is that zoos and aquariums have a positive impact on visitors' feelings, attitudes, and knowledge about conservation. They tested this hypothesis in Phase Two of the study, which focused on measuring changes in visitors' short- and long-term conservation-related knowledge and attitudes. Falk et al. chose two zoos and two aquariums to "capture the most generalizable picture possible of the conservation knowledge of zoo and aquarium visitors as they enter and as they exit, as well as the responses, purposes, and general outcomes of their visit" (p. 8).

In this section, we examine whether this study was designed appropriately to address its central hypothesis. Falk et al. draws strong conclusions based unequivocally on causality: they claim that visiting zoos and aquariums has a measurable impact on visitor knowledge and attitudes. For this hypothesis-based conclusion to be supported, Falk et al. would have needed to conduct a study that provided the opportunity to adjudicate between empirical evidence for two hypotheses. In other words, a valid study must be designed to provide evidence that *disconfirms* the hypothesis if it is false. Falsifiability is a sine qua non of sound scientific research (Popper, 1959).

With these epistemic strictures in mind, we assessed the validity of Falk et al. according to standard methodological criteria put forth by four well-established sources: Cook and Campbell (1979); Shadish, Cook, and Campbell (2002); Kendall and Norton-Ford (1982); and Shaughnessy and

Zechmeister (1994). These sources describe a set of threats to validity that should be avoided in research. The presence of even one major threat to validity can render a study's findings difficult, or in some cases impossible, to interpret.

Before describing each of the threats to validity that we identified in Falk et al., we should highlight a major conceptual weakness of the study from the outset. The authors' stated goal was to assess whether zoo and aquarium experiences affect visitors' beliefs and knowledge. With regard to knowledge, however, Falk et al. assessed only what responders *said* they believed or understood; they administered no direct measures of knowledge. There is a copious literature on the inaccuracies associated with self-report measures. For instance, Ross (1989) cautioned that self-report instruments can be particularly susceptible to the effects of implicit theories (personal narratives). In particular, he noted that if people believe that their attitudes will change as the result of an experience or intervention, they may incorrectly recall their initial (pre-experience or preintervention) attitudes as more different than they actually were. Thus, without direct measures of knowledge changes, Falk et al. may at best have assessed what responders *believe* they know or understand and not, as they claimed, what they actually know. Falk et al. were presumably interested in the effects of zoo and aquarium visits on beliefs and knowledge because these attitudinal changes may promote positive behavior in visitors. But they administered no measures of behavior per se.

Setting this major weakness aside, we turn now to a detailed methodological examination of the validity of Falk et al. Specifically, we pose the question: did Falk et al. adequately assess respondents' reported *beliefs* about their attitudes and knowledge? The main methodological threats to the validity of Falk et al. concern poor experimental control. We identified seven independent threats to validity, which we outline below (see Table 1). Most of these threats relate to either *construct validity*, that is, the soundness of the measures as indicators of the constructs examined by the investigators, or *internal validity*, that is, the soundness of the relationship between the variables under study. In the interests of brevity, we limit ourselves here to the most serious threats to validity.

Nonrandom sample. Falk et al. relied on a nonrandom sampling of participants. They reported that they used a "continual ask" method to minimize sample bias by approaching the first available visitor group entering the facility, followed by the next, and so on. They maintained a refusal log to track visitors who declined to participate. Nevertheless, they did not report any results from this log, making it impossible to evaluate the characteristics of

Table 1. Major Threats to Validity of Falk et al. (2007)

Validity Threat	Definition
Construct Validity	**The soundness of the measures as indicators of the constructs purported to be examined by the investigators**
Nonspecific effects	Improvements or changes from effects not specific to the factor or treatment under study
Novelty	General energizing and uplifting effects of a new, exciting experience
Construct confounding	Failure to take into account the fact that the experience under study may include more than one component that affects outcome
Demand characteristics	The tendency of participants to alter their responses in accord with what they believe to be the researchers' hypothesis
Experimenter expectancy effects	The tendency of investigators to unintentionally bias the results in accordance with their hypotheses
Internal Validity	**The soundness of the relationship between the variables under study**
Nonrandom sampling	Unintentional sampling of subjects that introduces systematic error or bias into the results
Response bias	A bias in subject responding due to the test instrument rather than the subjects' actual beliefs

refusers and thereby evaluate the degree to which the sample was representative. Because participants in this study were self-selected, they were quite probably nonrandom. Although the researchers instructed the interviewers to be impartial in their interception of visitors, they provided little detail regarding how objectivity was achieved or measured.

Nonspecific effects. Nonspecific effects are improvements arising from generic influences that are not specific to the intended condition or primary variable under study and that can be caused by a wide variety of other experiences. Zoo/aquarium experiences are designed to be stimulating and positive. They include immersion in a sensory and physically engaging environment that

includes many novel components. Therefore, assessment of the experience is vulnerable to a host of nonspecific effects, including novelty effects. Novelty effects are the general energizing and uplifting effects of a new, exciting experience (Shadish et al., 2002). Falk et al. did not assess or control for novelty effects by comparing their results with responses to other largely novel stimuli, such as new and exciting entertainment park experiences that do not include animals. Therefore, novelty effects remain a viable explanation for their results.

Construct confounding. Construct confounding occurs when there is a failure to take into account the fact that the experience under study includes more than one component that affects outcome. The zoo/aquarium experience consists of a complex assortment of components that include interaction with other people, walking in an attractive indoor or outdoor environment, going to gift shops and food stands, and often rides, tours, and other attractions. Even the experience of visiting individual animal displays is a complex one that can be deconstructed into various components, such as interacting with a docent or trainer, and walking through a physical display contrived to contain many components of "nature" such as trees, boulders, and water. In the case of interactive animal displays and, particularly, swim programs in aquariums, a multitude of salient components can contribute to participants' overall responses (Marino & Lilienfeld, 2007, 1998).

In the psychology literature, construct confounding is typically minimized or eliminated by dismantling studies (Kazdin, 1994), which separate the potential effects of different treatment ingredients by creating different experimental conditions containing these effects. Although there is no single, ideal control for the zoo/aquarium experience, Falk et al. did not incorporate even minimally effective dismantling procedures to address this issue.

Demand characteristics. One of the most common threats to validity is the presence of demand characteristics, i.e., the tendency of participants to alter their responses in accord with what they believe to be the researchers' hypothesis. Zoo/aquarium experiences are contextualized as educational experiences in obvious ways. Modern zoos/aquariums have recently shifted the strategy of marketing their facilities as places of entertainment and amusement to marketing them as centers of education, research, and conservation (Malamud, 1998).This new message saturates many elements of the zoo/aquarium experience, including the appearance of animal displays, the kinds of items sold in gift shops, the language used in display text and by docents and trainers, and the description of visitor activities. For example, the Bronx Zoo, which was one of the participating facilities in Falk et al., refers on its Web site to some of its displays as "living classrooms." Another participating facility, the

Binder Park Zoo, introduced to their attractions a Conservation Carousel that featured a menagerie of crafted animals that can be "sponsored," claiming that "riders of the carousel will go a long way toward supporting the Zoo's conservation programs worldwide." The Florida Aquarium in Tampa announced an event in which the aquarium went "Green." The obvious conservation and education messages associated with these attractions make the intentions of the zoo, as well as those of the investigators, apparent to all who visit, thereby imbuing Falk et al.'s study with the potential for demand characteristics that may undermine its validity. Falk et al. neglected to guard against this problem; to the contrary, they informed visitors "fully and accurately of the purpose of the study" (Visitor Evaluation Toolbox, p. 10) and, on p. 13 of the Toolbox, instructed interviewers to "assure them [the visitors] that their participation will provide positive and tangible benefits to future zoo or aquarium visitors." These instructions render virtually all Falk et al.'s findings potentially suspect. Furthermore, Falk et al. instructed interviewers to hand out tokens of appreciation, in the form of small gifts, to participants. Evidence suggests that such tokens can produce mild mood-elevating effects that, in turn, may bias ratings (Westerman et al., 1996).

Experimenter expectancy effects. Experimenter expectancy effects refer to the tendency of investigators to bias the results unintentionally in accordance with their hypotheses. A large body of research shows that experimenter expectancies can influence not only how subjects' responses are coded and interpreted, but even the responses themselves (Rosenthal, 1994). Because the surveyors who administered the assessments to responders were aware of the desired outcome, the objectivity of the scoring procedure in Falk et al. is suspect. In particular, the possibility of subtle and unintentional cueing of subjects by surveyors is difficult to exclude. Falk et al. made no mention of efforts to mitigate this potential problem. Furthermore, they offered little information about how they conducted the assessments, such as where surveyors were standing and looking when the responders completed their surveys. At the very least, potential experimenter expectancy effects could have been minimized by the inclusion of raters blind or neutral to the hypothesis.

Response bias. Response bias can arise in several ways; for example, survey respondents may answer questions in the way they think the questioner wants them to answer rather than according to their true beliefs (see section on *Demand characteristics*). Such bias is especially likely if survey items are worded to make one type of response inherently more likely than another, independent of their content. Falk et al. determined the affective response of visitors with a 13-item, 7-point Likert-type exit survey (their Figure 1). Two types of response bias to which this survey is susceptible are acquiescence bias

and social desirability bias (see Paulus, 1991). In acquiescence bias or "yea-saying," respondents tend to agree with survey statements, irrespective of their content. A review of the content of the Likert-type items in Falk et al. reveals that only 2 of the 13 items were keyed negatively. The two items read, "I am part of the problem with nature" and, "There is not much I can do to help nature." The remainder of the items were keyed positively, e.g., "I am part of the solution to nature's problems"; "Animals are amazing"; and "Being at the zoo/aquarium is fun." Because most of the items were keyed in the same direction, the scale is susceptible to a potential acquiescence response bias.

Social desirability bias is the inclination to present oneself in a manner that will be viewed favorably by others. When social desirability cannot be eliminated, researchers often resort to administering an independent scale that measures socially desirable responding, with the assumption that if a participant answers in a socially desirable manner on that scale, they are in all likelihood answering similarly throughout the study. In some cases, investigators then use scores on this scale as a moderator variable or covariate in analyses (Piedmont, McCrae, Riemann, & Angleitner, 2000). There is no evidence that Falk et al. employed safeguards against social desirability or that they prescreened items for high levels of saturation with a social desirability dimension.

Weaknesses of the post-only, retrospective-pre design. Instead of an actual pre-post (i.e., enter-exit) survey, Falk et al. conducted their survey entirely on exit and asked visitors to reflect on how they *would* have answered the same items on entrance (retrospective-pre). Their stated justification for this post-only, retrospective-pre measure is that it provides a way to eliminate response-shift bias. Response-shift bias is a change in the participant's metric or context for answering questions from the pretest to the posttest that confounds the apparent effects of the program or manipulation under study (Howard, 1980). The retrospective-pre method is designed to mitigate response-shift bias by limiting participants' responses to the same time frame and context. But the retrospective-pre method is most useful in guarding against response-shift bias when assessing changes in knowledge from training programs over a relatively long period of time, not the effects of shorter-term general experiences on beliefs or affect, as was the case in Falk et al.

Falk et al. contended that a post-only, retrospective-pre measure, which has been used by some other researchers in this area, is more reliable than traditional pre/post measures for assessing attitudes. They cited two studies to support this conclusion (Stevens & Lodl, 1999; Rockwell & Kohn, 1989). Neither Rockwell and Kohn (1989) nor Stevens and Lodl (1999), however, reported a quantitative measure of reliability in their evaluation of this method, so it is unclear on what basis Falk et al. advances this claim. (Falk et al. reports a

reliability (stability) coefficient of 0.842.) Moreover, although the traditional pre-post method tends to underestimate program effect, the retrospective-pre measure tends to *over*estimate program effect (Colosi & Dunifon, 2006).

In general, although the retrospective-pre method eliminates certain sources of error, it introduces others, which are not dealt with by Falk et al. These include recall bias (the inability to accurately recall attitudes held in the past), social desirability bias (described earlier), effort justification (the reporting of change to justify time and energy invested in the experience), and cognitive dissonance (reporting improvement or change, even if it did not occur, to ease internal conflict stemming from the expectation that changes *should* have occurred).

None of these potential biasing effects were controlled or even evaluated by Falk et al.

Weaknesses of the long-term impact study. Falk et al. conducted a "long-term impact study" to assess long-term changes in visitor attitudes, beliefs, and perceptions after the initial study. They collected e-mail and phone information from participants. Due to low response rates, however, they were unable to obtain a random sample of respondents. Out of 592 participants, only 84 completed the long-term interviews, despite more than one attempt to contact some of the participants. The authors admitted that a valid response rate could not be generated. Nevertheless, they did not report how responders differed from nonresponders on potentially relevant variables. They designed "parallel assessment instruments" comprising open-ended questions designed to probe visitors' recall of the initial experience from 7-11 months earlier. Falk et al. reported that nearly all contactees recalled their experience. Most relevant to the study's aims, Falk et al. found that 61% conversed about what they had learned from their initial zoo or aquarium visit. When asked what the zoo or aquarium hoped visitors would take away from their visit, 40% mentioned conservation and 66–76% mentioned that they believed zoos and aquariums played important roles in conservation and education.

Falk et al. interpreted these findings as offering support for the education and conservation role of zoos and aquariums over the long run. But the evaluation of this conclusion is weakened by several methodological limitations. First, the relatively low rate of return on the long-term survey raises the possibility that those who responded were unrepresentative of the entire sample. Because the authors did not determine whether responders differed from nonresponders on potentially important variables, such as initial attitudes toward zoos and aquariums, this possibility (known as subject mortality) cannot be evaluated. This weakness only adds to the problems of nonrandom sampling in this study.

Second, it is well documented that memory is far more reconstructive than has traditionally been thought (Loftus, 1993) and that retrospective reports are often of suspect validity. Ross (1989) reviewed the literature on the effects of implicit theories on retrospective measures and concluded: "The biased retrospections obtained in survey research may lead, among other things, to inaccurate conceptions of human behavior" (p. 354). In an elegant series of studies, Ross (1989) showed that individuals in treatment studies often unintentionally distort their memories of improvement on the basis of their expectations concerning change. For example, if individuals expect to improve as a result of treatment but experience no objective improvement, they will often recall their pre-treatment status to be worse than it actually was (Conway & Ross, 1984). The same phenomenon could account for the reported results of Falk et al., because responders might remember their previsit attitudes as less positive than their postvisit attitudes.

Furthermore, Falk et al. never assessed or analyzed attitudes that might have *worsened* as a result of the zoo and aquarium experience. Therefore, the proportion of participants who provided negative responses, i.e., responses indicating that their zoo or aquarium visit was accompanied by a worsening of attitudes about education and conservation, is not known. For example, Falk et al. did not include items assessing the extent to which visitors view animals as objects of entertainment rather than conservation, a change that many might view as negative. Instead, the authors appear to have assumed that all effects of zoo and aquarium visits are necessarily positive, an assumption that does not appear warranted, given the dearth of systematic data on these effects.

Interpretative issues. In addition to the major threats to validity already detailed, Falk et al.'s study was compromised by a number of interpretative problems. The central weakness in Falk et al. is that the authors repeatedly draw causal conclusions from data that are noncausal in nature. Their general conclusion is that a visit to an accredited zoo or aquarium has a measurable *impact* on conservation attitudes and understanding in adult visitors. For instance, Falk et al. states that "Our three-year visitor impact study found that a visit to an accredited zoo or aquarium in North America has a measurable *impact* on the conservation attitudes and understanding of adult visitors" (p. 9; emphasis added). This statement implies that zoos and aquariums *cause* a change in visitors' attitudes and understanding, even though this statement is unwarranted, given the quasi-experimental (rather than experimental) nature of their design. In fact, the authors make similar causal claims no fewer than nine separate times in their report. As we noted earlier, because Falk et al. draws strong causal conclusions, their study can be validly criticized on the basis of whether those conclusions are supported by methodologically sound

research. Had Falk et al. not drawn causal conclusions, there would have been little reason to discuss the methodological weaknesses associated with threats to validity.

Finally, even putting aside all of these methodological threats to validity, it is sobering to note the actual reported gains in stated visitor knowledge. Falk et al. finds that "there was no overall statistically significant change in understanding seen" (p. 10). Therefore, the authors do not obtain strong supportive evidence for their hypothesis because they found no significant gains in general knowledge from zoo or aquarium visits. In response, Falk et al. speculate that their subjects might have gained more *specific* knowledge of animals or conservation, a form of knowledge they neglected to measure. Yet, curiously, they argue that "[i]f we had sought to measure this kind of knowledge, we very likely would have found significant visitor gains" (p. 10). This kind of reasoning, referred to by Dawes (1994) as "an argument from a vacuum," is problematic, because it hinges on an unverifiable—and ultimately nonscientific—assumption that changes would have been observed on dependent variables that were not measured. In summary, our methodological analysis of Falk et al. shows that their primary findings and conclusions are uninterpretable and unfounded.

Discussion and Conclusion

Falk et al. are to be applauded for examining an important issue that has heretofore received precious little attention (Mason, 2000), namely the effects of zoos and aquariums on visitor knowledge and attitudes. Nevertheless, despite the widespread acceptance of Falk et al.'s study by the zoo and aquarium community, we have shown that numerous methodological weaknesses render their findings difficult or even impossible to interpret. More important, their claims—extensively disseminated on zoo and aquarium Web sites—greatly outstrip their methodologically limited findings. We therefore urge zoos and aquariums to cease citing this study in their promotional materials as evidence that visitors' attitudes are changed for the better, as this conclusion is unwarranted and potentially misleading to consumers.

We also encourage further research that addresses the methodological threats to validity that we have identified. In particular, we urge researchers to use designs that (a) incorporate appropriate comparison groups of participants exposed to other forms of stimulating entertainment but not to zoos and aquariums, (b) administer full pre-post assessments in both groups, (c) attempt to minimize experimenter expectancy effects, ideally by using observers who are blind to hypotheses, (d) administer questionnaires that assess actual knowl-

edge in addition to beliefs and attitudes, attempt to control for response biases, and assess potential worsening effects, and (e) conduct subsidiary analyses to examine the potential impact of nonrepresentative sampling on the results. In this respect, our critique, although directed at one influential study, may serve in part as a set of guidelines for future zoo and aquarium researchers for conducting more internally valid research.

In summary, to date there is no compelling or even particularly suggestive evidence for the claim that zoos and aquariums promote attitude change, education, and interest in conservation in visitors. Some might contend that the methodologies used by Falk et al. are standard in a good deal of zoo and aquarium visitor research. There may well be some truth to this assertion, but it does not gainsay our methodological criticisms or imply that the flaws of their study need not be remedied in future zoo and aquarium visitor research. Only well-controlled research, not enthusiastic assertions that outstrip the quality of scientific evidence, can address the question of whether claims concerning the positive effects of zoo and aquariums on visitors are justified. We encourage such research with a particular eye toward remedying the methodological threats to validity we have identified.

References

American Zoo and Aquarium Association. (2006). Groundbreaking study identifies impact of zoo and aquarium visits. [Press release.] Silver Spring, MD: AZA.

Colosi, L., & Dunifon, R. (2006). *What's the difference*: *"Post then Pre" & "Pre then Post."* Prepared for Cornell Cooperative Extension. Available at: http://www.citra.org/Assets/documents/evaluation%20design.pdf.

Conway M., & Ross, M. (1984). Getting what you want by revising what you had. *Journal of Personality and Social Psychology, 47*, 738-748.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Boston, MA: Houghton Mifflin.

Dawes, R. M. (1994). *House of cards*: *Psychology and psychotherapy built on myth*. New York: Free Press.

Dierking, L. D., Burtnyk, K., Buchner, K. S., & Falk, J. H. (2002). Learning in zoos and aquariums: A literature review. Silver Spring, MD: American Zoo and Aquarium Association.

Falk, J. H., Reinhard, E. M., Vernon, C. L., Bronnenkant, K., Deans, N. L., Heimlich, J. E. (2007). Why zoos & aquariums matter: Assessing the impact of a visit to a zoo or aquarium. Silver Spring, MD: Association of Zoos & Aquariums.

Howard, G. S. (1980). Response-shift bias: A problem in evaluating interventions with pre/post self reports. *Evaluation Review, 4*, 93–106.

Kazdin, A. E. (1994). Methodology, design, and evaluation in psychotherapy research. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed.), (pp. 19-71). New York: Wiley.

Kendall, P. C. & Norton-Ford, J. D. (1982). Therapy outcome research methods. In P. C. Kendall & J. N. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 429-460). New York: John Wiley and Sons.

Loftus, E. F. (1993). The reality of repressed memories. *American Psychologist, 48*, 518-537.

Malamud, R. (1998). *Reading zoos: Representations of animals and captivity*. New York: NYU Press.

Marino, L., & Lilienfeld, S. (1998). Dolphin-assisted therapy: Flawed data, flawed conclusions. *Anthrozoös, 11,* 194–199.

Marino, L., & Lilienfeld, S. (2007). Dolphin-assisted therapy: More flawed data, more flawed conclusions. *Anthrozoös*, *20*, 239-249.

Mason, P. (2000). Zoo tourism: The need for more research. *Journal of Sustainable Tourism, 8*, 333–339.

Paulus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson & P. R. Shaver (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic Press.

Piedmont, R. L., McCrae, R. R., Riemann, R., & Angleitner, A. (2000). On the invalidity of validity scales: Evidence from self-reports and observer ratings in volunteer samples. *Journal of Personality and Social Psychology, 78*, 582–593.

Popper, K. R. (1959). *The logic of scientific inquiry*. London: Hutchinson.

Rockwell, S. K., & Kohn, H. (1989). Post-then-pre evaluation. [Electronic version]. *Journal of Extension, 27*(2). Available at: http://www.joe.org/joe/1989summer/a5.html

Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science, 3*(6), 176-179.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review, 96*(2), 341–357.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Shaughnessy, J. J., & Zechmeister, E. B. (1994). *Research methods in psychology*. New York: McGraw-Hill.

Stevens, G. L., & Lodl, K. A. (1999). Community coalitions: Identifying changes in coalition members as a result of training. [Electronic version]. *Journal of Extension, 37*(2). Available at: http://www.joe.org/joe/1999april/rb2.html

Westermann, R., Spies, K., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology, 26*, 557–580.