

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.

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 *overestimate* 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.