

The Benefits of Long-Term Evaluations of an Interactive Exhibit Area at Brookfield Zoo

Barbara A. Birney
Brookfield Zoo
Brookfield, IL

Bird Discovery Point is located in the Aquatic Bird House, which is a lovely example of neo-classical architecture at Brookfield Zoo. In 1986 the interior walls of the House were lined by a series of live exhibits and there was a large, open, tiled area in the middle of the domed building. It was not uncharacteristic for visitors to halt and turn away at the door – their first sensation of the building being its fishy odor and “emptiness.” Older couples appeared to enjoy sitting on the benches, resting and observing the birds. Younger couples with preschoolers often crossed immediately to the penguin exhibit at the end of the hall, turned next to look at the flamingoes, and then left the building, missing at least three other major exhibits. For a large exhibit space, it was used in a very narrow fashion by the public.

Evaluation studies are often associated with individual projects or exhibits. The studies conducted over a four-year period on *Bird Discovery Point*, (hereafter referred to as *BDP*) addressed four types of goals: goals related to the institution’s mission, its philosophy of interpretation, the overall exhibit, and individual components.

Bird Discovery Point is a group of interactive exhibit components. Docents convey information about bird adaptations in the Demonstration Area. The Flight Strength component humbles those humans who cannot get a cardinal off the ground. The Bird Wing Movement component guides humans through the little known motion of flight. And the Skeletal Contrasts exhibit highlights and compares the internal structure of humans to birds.

BDP was intended to accomplish several goals:

1. Increase use of an exhibit area that was under-utilized;
2. Provide visitors with an opportunity to learn about bird biology, behavior, and ecology;
3. Increase visitors’ appreciation of birds and further the Zoo’s conservation mission;
4. Promote family interaction by providing multi-generational learning opportunities; and
5. Be an entertaining introduction to ornithology.

Given that eight studies were produced on an exhibit containing six interactive stations, I could not, in this paper, do justice to the specific findings associated with those studies. What I wish to do here is describe the issues that were raised in each evaluation effort, the lessons learned from both the evaluation process and findings, and how the *BDP* evaluations impacted the decision-making processes of the institution, staff perceptions, and exhibit development in other areas.

The indirect benefits of the *BDP* studies were many. Perhaps the least talked about benefit involved the mistakes that were made. For instance, the *BDP* series significantly impacted the design of studies done elsewhere in the zoo. It became clear that the earlier studies were too ambitious for the time frame given. Progress on the larger *BDP* studies was slowed as 20 other studies were requested and conducted parallel to this series. As a result, subsequent evaluations were smaller, more tightly designed, with a more rapid turnaround. One negative consequence was that qualitative research was done less often because it simply was too labor-intensive.

A second indirect benefit of the studies was that staff began to demand evaluation for a wide variety of projects – far more than the resources of the department could handle. With this came an increased role for evaluation in many areas besides exhibit development.

The direct benefits of evaluation were both immediate and long-term. At the time that evaluation was first introduced into the team planning process, the curator had become discouraged with the exhibit development process and dropped out six months previously. Also, animal collections staff stated that my conversation with them to make arrangements for the first baseline study was the first contact anyone had made with them about the pending exhibits. Collections staff were brought back into the process and have played an integral role ever since. The needs inherent in the evaluation process have helped reinforce the Society's firm commitment to team planning.

Funding had also been requested for two interactive computers. National Science Foundation reviewers, split in their opinions, asked for more information. Reviewers maintained that the preliminary findings from the earliest prototypes that were presented to them brought the discussion to resolution and the Society received an additional NSF grant to complete the exhibit.

The studies impacted decision-making. Since many findings are *NOT* specific to a single exhibit, referring to them increases the confidence and speed with which decisions are made. Decision-making relies less on individual influence and more on group consensus. Perhaps the most important direct benefit is how decisions, based on the findings, began to reflect a melding of the institution's mission and goals with a respect for the visitors' agenda and prior knowledge.

The methods chosen for each study were as varied as the evaluation questions themselves. Baseline data were gathered on visitors' traffic

patterns within the building. Open-ended interviews were taped and transcribed to gather baseline data on visitors' knowledge of birds. Surveys containing close-ended items were administered to assess visitors' understanding of conservation issues. Formative evaluation assessed prototypes of a problematic component and graphics. Finally, summative evaluations were conducted to examine visitors' use of the overall exhibit, knowledge of the main messages, and attitudes towards bird conservation issues.

Each area of study yielded both results and a set of lessons. The traffic pattern studies conducted will be considered later in this discussion. Taped interviews probing for visitors' baseline knowledge of ornithology took 25 minutes to complete, dispelling the common assumption that visitors will not interrupt their visits for long periods. The key to doing longer interviews is to have the researcher working with a docent who attends to the rest of the group while the respondent participates in the study.

The first major prototype was developed in response to a classic dilemma – after planners presented their ideas to other individuals outside the planning process, it was suggested that the flowing design of the Bird Wing Movement component might be boring. In response to the administration's reluctance to spend \$15,000 for an exhibit that some felt lacked any appeal, a series of prototypes were tested. In brief, the public loved it and 52% of the users over six years of age understood the main message – up from 4% of non-users.

Several other issues arose. For instance, it is still the case that staff tend to react negatively to new prototypes. It seems to be a classic primate response – fear of novelty. Hence, it's important for evaluators to be considerate of and simultaneously immune to a plethora of remarks suggesting that the prototype is a failure – before any data have been gathered.

This first prototype resulted in design “overkill.” I was asked whether it was really necessary to spend \$5000 to build it – a third of its eventual cost! It became clear that these estimates were based on the use of high quality materials. Once designers became accustomed to the idea of using plywood, foam core, and other simple materials, we were able to cost-effectively prototype many exhibits.

Another consequence of prototyping has been a relaxation in curatorial standards for addressing complex biological topics. The original prototype was accurate in counteracting common beliefs of how birds fly – it was also uncomfortable for humans. Changing the curvature of the exhibit rails increased the ease with which the exhibit could be used but did not decrease visitors' ability to extract the main message. Planning teams still agonize over content – and continue to discover that simple, elegant exhibits will meet their goals. Prototyping offers the reassurance to staff that they are striking a balance between their standards and desired visitor outcomes.

Formative evaluation of the prototype's graphics taught us at least two valuable lessons which have since been verified by studies elsewhere. The first lesson involved methodology. The impact of instructional signs on visitors' behavior was so subtle, it is essential to have especially large sample sizes – over 500 individuals of observational data. Fortunately, such sample sizes are not difficult to obtain with a targeted behavioral study. In contrast, we have been able to prototype headers and advance organizers with 40-60 subjects.

The second lesson was related to content. Our visitors are consistently confused by language that is poetic or involves the use of imagery – especially if they are being asked to take the perspective of the animal which is one of the ways in which planners may try to build empathy.

The dominant stimuli of the multimedia exhibit pulled visitors into the space long enough for them to habituate to the fishy smell. Subsequent observational data would reveal that after *BDP* was installed, visitors behaved like ping-pong balls. Exhibit planners had accurately targeted appropriate age groups for each component. Group size and the presence of children significantly and positively influenced the length of time visitors spent in *BDP*. Also, it was interesting to note that the rate of use by visitors of the Bird Wing Movement component remained almost the same when installed with three large competing exhibits (plus six large live animal exhibits). Currently, if a free-standing, unattended interactive exhibit is used by 40% of a sample, we consider it to be successful.

The time spent at live exhibits, a major concern of zoo staff, was not decreased but increased. The findings on whether interactive exhibits would complement or interfere with live animal exhibits have been used by other institutions as well. Still, there were lessons to be learned.

As a consequence of these studies, I am now reluctant to accept average times for *any* behavioral study due to the enormous variation between groups. The first traffic pattern studies indicated that the average time spent in the entire building was 2.5 minutes. In our study, some groups stayed 20 minutes in the exhibit area and skewed the average times. This prompted me to examine other researchers' studies. I've seen averages used in some visitor studies to report a "doubling" of the time spent with exhibits. When median times are calculated, however, the differences are insignificant. I now believe that museum evaluators should no longer present anything but median times when conducting research on interactive areas.

Our observational studies distinguished between users and observers of other users at exhibits – direct and vicarious learning, if you will. The rates for observers of other users were extremely high (+90%). Actual use of the four components was substantially lower (approximately 35%) and very dependent on how long it engaged visitors who used it properly as well as how accessible it was.

In a summative evaluation of visitors' knowledge, we administered one of the stiffest kinds of tests in a leisure setting. A series of open-ended questions analyzed for specific scientific concepts represented in the exhibits. It was administered after they had left the building. We had no way of knowing what myriad of exhibits they had touched during their visit. Visitors were sorted by whether they reported using different exhibit components or watching others use it. For instance, in the Demonstration Area, children can work with the docent while a parent looks on. *A series of regression analyses yielded that when use, observation of other group members, education, age, and sex were taken into account, only actual use of an individual component was related to visitors' responses on that topic.*

Education was a factor for one of the exhibit components, the Flight Strength, but not for the most obvious reason. Subsequent analyses showed that all users were high-school graduates while college-educated visitors had chosen to watch others rather than use the exhibit themselves. Hence, there was some confoundment in that area.

Observational data had already indicated that the Skeletal Contrasts exhibit was used properly and by the appropriate target audiences. Yet, the main messages of this component were completely absent from visitor responses. This may have been because the visual stimuli and auditory information presented to visitors were not internally consistent.

These findings also reinforced my view that it is essential to gather more than one form of data on interactive exhibits. Observational data and interview data introduce specific biases and while one method may suggest that the exhibit is ineffective, another method may indicate that visitors are indeed acquiring the appropriate messages. Using multiple methods allows one to construct an exhibit "profile." On a wide variety of criteria, *BDP* was a demonstrated success and its weaknesses were documented so that they could be addressed.

Still, *BDP* failed in one critical area. A second set of surveys on visitors' perceptions of conservation issues were conducted. The findings were nearly identical to those obtained three years earlier. Increasingly, we are building a body of evidence that suggests that increasing visitors' awareness of wildlife biology and behavior alone are insufficient *if the primary mission of the institution is to move visitors towards a more environmentally responsible lifestyle.* As these data gather strength, new planning teams are calling into question exhibit messages that they have chosen. This issue is a critical one for our institution and related directly to a shift in our mission's emphasis.

Some of the other key ways in which the evaluation studies have had a long-term impact on institutional processes include the following. Animal managers are more likely to accept the value of multi-media exhibits and recognize how they can complement the collection. The role of evaluation in preventing costly mistakes is especially valued by operations staff responsible for guiding the institution through difficult budgetary times.

With a commitment to various target audiences, communications departments feel confident about how they can use interactive exhibits to reach multi-generational groups. Development staff find that these and other studies give them an edge in competing for funds.

For exhibit teams, an effort is made not to "reinvent the wheel." For example, the heights developed for one component intended to serve children were so successful that they are being used in the new *Habitat Africa!* exhibit under construction.

Finally, as the evaluation process is applied to other aspects of exhibit development, it is giving management assistance in determining accountability. The issues raised in these studies continue to be a standard part of staff discourse related to exhibit interpretation.