



# **Profile of an Exhibit**

## **Evaluation Summary**

of

**"Darkened Waters: Profile of an Oil Spill"**  
at Oakland Museum

for the

**Pratt Museum**

**Beverly Serrell**

**November 1991**

## Table of Contents

Introduction	1
What We Wanted to Find Out	1
Description of Exhibit Components	4
Evaluation Plan and Methods	5
Results from Comment Cards	7
Summary of General Demographics	7
Results of Tracking and Timing	8
Results of Uncued Exit Interviews	17
Results of Cued Testing	20
Results of Cued Questionnaire	22
Conclusions	27
Notes	29
Appendix	

## **Acknowledgements**

This evaluation study was initiated by the Pratt Museum in Homer, Alaska, by Betsy Pitzman and Martha Madsen, and Kathleen McLean of Independent Exhibitions in Berkeley, CA. It was organized and directed by Beverly Serrell of Serrell & Associates, ably assisted by: JFK Museum Studies students Lisa Hubbell Mackinney, Christie Davis and Dana Powell; Deborah Perry, educational testing and measurement specialist at Museum of Science and Industry, Chicago; Oakland Museum staff; and the exhibit designers and developers. Funds for the evaluation came from a grant from the National Science Foundation, Informal Science Education Division, the Oakland Museum and Serrell & Associates.

## **Profile of an Exhibit Evaluation Summary**

### **Introduction**

A multifaceted evaluation study was conducted of the 2,000-square-foot exhibit "Darkened Waters: Profile of an Oil Spill" between September 14 and October 12, 1991, at the Oakland Museum, Natural Sciences Department. "Darkened Waters" is the second generation of an exhibit originally conceived and built by the Pratt Museum in Homer, Alaska, to provide information and a forum for the environmental impact of the 1989 Exxon Valdez oil spill. The exhibit will travel to the Smithsonian Institution in December 1991, and will be available to other institutions through the Association of Science-Technology Centers.

Before "Darkened Waters" goes on the road, it will be modified and improved. Decisions about what to change have been made by the exhibit developers based on what they perceive to be necessary and desirable to enhance the esthetics, technical aspects and durability of the exhibit. In addition, changes are being made to improve the organization and presentation of information, based to a large extent on feedback from visitors during this evaluation.

How well the exhibit meets its communication goals was the main focus of the evaluation. Data about how visitors used the exhibit and what impressions they took away from it helped us make that determination.

### **What We Wanted to Find Out**

To plan any evaluation study, it is helpful to know both what the instigators of the evaluation wish to find out and what the exhibit was intended to accomplish in the first place. Exhibit goals and evaluation questions are closely linked, but making those links clear, logical and thorough is often a long and thoughtful process. The links can be established at the outset of the exhibit's development. More often, however, they have to be forged much later, after the exhibit is done and open to the public, and plans for evaluation have been added. How did that process look in this case?

When the Pratt Museum staff wrote their grant application to NSF, they included the following objectives for the exhibit project:

To communicate information about issues of development, transport, and use of petroleum

To increase peoples' awareness of marine life

To tell about the relationship of the Exxon oil spill and national and global issues

To present information about prevention and scientific inquiry

To find solutions to long-term energy needs

To a large extent these goals focus on abstract concepts and are stated in a way that emphasizes the exhibit's role in the delivery of intellectual, scientific information. Only one is stated in terms of what might happen to a visitor (their awareness of marine life will be increased), but even it is not very specific (awareness of *what* about marine life?).

In order to evaluate an exhibit, it is necessary to translate the exhibit objectives and the exhibit's communication potential (derived from the actual content of the exhibit) into more concrete, specific, measurable objectives. Taken to the extreme, these statements can become highly focused on overt, cognitive objectives, such as:

Visitors can describe accurately what happened after the Exxon Valdez disaster.

Visitors can list accurately some of the causes.

Visitors can generalize about this incident to more national and global environmental problems by relating it to and naming other events and situations.

Visitors can list common products that contain oil.

Visitors can name some ways to reduce oil use.

These are very specific learning goals, and they are useful for the planning and design of the exhibit. But they are highly inappropriate for evaluation questions. Behavioral objectives like these are often used in teaching school. Visitors, however, can not be expected to perform like students whose mastery of the content is contingent on subsequent rewards (e.g., good grades). Visitors rarely perform well on a cognitive test in a free-access, limited-exposure, short-duration and intrinsically motivated museum experience.

Better suited to the needs of both evaluation and planning were the communication objectives specified for the second-generation traveling version of "Darkened Waters" as it was being developed for the Oakland site.

Because the new, succinct communication goals were stated casually, almost conversationally, they could be related easily to what visitors might say about the exhibit after seeing it. They were very useful in designing the evaluation instruments, tabulating data and measuring success--defined as the degree to which the communication goals were achieved. These objectives were stated as six "big ideas" that visitors might see, feel, know and experience in the exhibit and take home with them:

1. Alaska is a beautiful wilderness and should be protected.
2. The oil spill was a huge disaster.
3. They couldn't clean it up.
4. It's not over.
5. There are lessons to be learned.
6. I'm personally going to do something.

These ideas were closely linked to the layout and content of the exhibit and were stated as "headers" in the panel groups (e.g., How Big Was the Spill? We Couldn't Clean It Up; Not the First, Not the Last).

From all of the above, a list of questions was developed and various strategies for collecting information to answer them were devised, giving us the scope and focus of the evaluation plan. The questions were:

- Who are the visitors to this exhibit?
- Do they have a special interest in or knowledge of the subject?
- How long do they spend in the exhibit?
- What do they do while in the exhibit?
- What parts attract and hold their attention longest?
- Do they use the interactive components?
- What do they think the exhibit is about?
- Do they agree with the exhibit's viewpoint?
- What do they get out of it?

Do they understand some of the specific topics we are trying to communicate?

How do they react to the exhibit?

Some of these questions are easier to answer than others. Some questions typically require research techniques that are not readily adaptable to the museum setting. We had a menu of strategies available to help find the answers, and we also had to develop some new approaches to suit our needs. Exhibit evaluation--especially when it involves questions of impact on people's knowledge, feelings and attitudes and objective measures of success--is, in the words of one exhibit designer "like trying to nail jello to the wall."

### **Description of Exhibit Components Considered in the Evaluation**

There were roughly 5,000 words of typeset copy (labels, text, captions, headers, etc.) and more words in the form of documents, news clippings and articles, political cartoons, lab reports and visitors' comments.

About 80 color photographs, ranging from five inches square to three feet by five feet, depicted spill-related events. Real objects on display included barrels, oil samples, cleanup equipment (oil-soaking booms and pompons), otter pelt, products made from oil, dead ducks, clothes worn by spill workers, bumper stickers and letters.

In addition to reading and looking, visitors to "Darkened Waters" (hereafter referred to as DW) visitors could:

- listen to tapes (one of the captain's voice reporting the accident, and four others, approximately three to five minutes each, of native Alaskans talking about the impact of the spill on them and their communities);

- touch the otter fur and pompon;

- look through a microscope at plankton;

- write on a comment card;

- compare the size of the spill (using a movable template) to areas on a map of North America, and lift up labels to see which products were made from oil at "the petroleum game."

## Evaluation Plan and Methods

We felt that a combination of strategies would provide a more complete description of the visitors' experience in DW than any single method alone could do. With limited time and resources, we wanted to look at what was happening from a variety of angles, rather than collect a larger sample using only one tactic.

Thus, there were five parts to the evaluation plan:

1. **Comment Cards.** Visitors could fill out a card saying what they thought about the exhibit and drop it in a barrel. Some completed cards were posted nearby on a bulletin board to give others an idea of the range and types of responses gathered. The purpose of the cards was to initiate visitors' feedback about their thoughts and reactions to the exhibit and to give visitors a voice in the exhibit content.

2. **Tracking and Timing of Visitors in the Gallery.** Unobtrusive observations were made of the paths visitors took through the gallery and the amount of time spent and stops made. By watching where visitors went and how long they spent we could gauge the relative ability of the different exhibit elements to attract and hold visitors' attention.

3. **Uncued Exit Interviews.** After they had visited DW, visitors were asked to participate in a short interview about their reactions to the exhibit, regardless of where they had gone or how much time they spent, i.e., they had not been cued to look at the exhibit in any special way.

4. **Cued Testing.** Visitors were asked to participate in looking more closely and completely at one of three parts of the exhibit and then asked several questions about it. These three areas were under special scrutiny for their communication effectiveness.

5. **Cued Exit Questionnaire.** Before they entered DW, visitors were recruited to commit more time and look more closely at the whole exhibit and then fill out a short questionnaire. We wanted to get more in-depth feedback from visitors who had been prompted to pay attention.

The number of samples in each section was limited by the realities of the amount of time and money we had to spend. We wanted a quota of 60 tracking and 75 interview samples within a three-week period and could pay data collectors to meet those quotas. Tracking takes twice as much time, on the average, as interviewing, and during September visitation was often sparse, so the evaluators had to take



samples opportunistically to make the best use of their time. We strove for a representative sample that would give us a minimum of objective data from which to draw some straightforward conclusions. No complex statistical analyses were planned or carried out.

Each part of the study (other than the comment cards) had a different data sheet, and information was collected systematically from randomly selected visitors. Data sheets were developed jointly by the evaluation team, Oakland Museum staff and exhibit developers, and they were tested in the gallery by the evaluation assistants before actual data collection began. Each of the data sheets for parts 2, 3, 4 and 5 included information about the subject's age and gender, the size and composition of his or her social group, the date and time of day and the recorder's initials. Samples of all data sheets are in the Appendix.

On the tracking and timing data sheets--part 2 of the evaluation plan--we drew a floor plan of the exhibit and indicated the numbers of the different panel groupings or other logical sections of the exhibit to calculate how many of the exhibit's sections were attended. (These areas were given equal values in our summaries, even though each section was not equal in size, amount of content, or time required to "see/look/hear it all." These discrepancies will be discussed more below.) A line indicating the visitor's path through the exhibit was recorded on a data sheet for each subject, noting direction, stopping points, behaviors and time.

Two general questions were asked for all the samples collected in parts 3, 4 and 5 regarding the number of times they'd seen the exhibit and if they had any special interest in the topic. The evaluator filled out the sheets while talking to visitors and then added comments or observations before beginning the next interview. The questionnaire for part 5 had similar spaces at the top for demographics and two questions with plenty of room for visitors to fill in responses themselves.

Data were entered on a Macintosh computer using Excel spreadsheet software. Numerical responses were tabulated, while open-ended answers to questions and comments were entered verbatim, then categorized, sorted and tabulated. Data from the cued questionnaire were reviewed and processed by hand.

## **Results of Comment Cards**

In response to the prompting label, "Tell us what you think," visitors filled out comment cards and dropped them in a barrel at the end of the exhibit. Responses ranged broadly, from emotional reactions (sad, angry, upset), to praise for the exhibit, to judgments on various issues (Exxon, big business, politics), to new awareness and personal commitment, to totally irrelevant scribbles (swear words, drawings).

Because there was no way to know anything about who filled out the cards, when, why or how many were filled out by one person or group, it was not possible to get much quantifiable data from this method. However, the comments did provide a rich source of qualitative data about visitors' responses to the exhibit. They gave us ideas about the kinds of behaviors we could watch for during tracking and the kinds of responses we could quantify in the interviews. About a third of the 1,000+ cards collected were summarized. The results are in the Appendix.

## **Summary of General Demographics**

Collection of data for evaluation parts 2, 3 and 4 took place between the hours of 11 a.m. and 5 p.m., Wednesday through Sunday. (The Oakland Museum is closed on Mondays and Tuesdays.) A total of 190 visitors participated in these three parts. Roughly the same number of males and females were observed. All adult age groups were represented; the largest group was in their 30s. A large percentage of visitors toured the exhibit alone (39%). The majority of visitors in groups were with other adults (74%); 26% of the groups were adults accompanied by children.

About 7% of the visitors sampled (during interviews) reported that it was not their first visit to the DW exhibit. All others were seeing it for the first time. Several mentioned that they were planning to come back, or were going to tell someone else to come.

During the 72 uncued exit interviews, 76% (55) said that they had no special interest, knowledge or training about oil spills. Of the 17 who said that they did, 5 mentioned that they had read about the spill in the newspaper or had a general concern for environmental issues. "I follow it in the papers." "I'm interested, but just as a citizen." Others (15 people) mentioned interests related to their jobs, home town or hobbies (e.g., cartographer, draftsman, environmental writer, deep-sea

diver, seaman, resident of Homer here on vacation, resident of Seattle). One person had worked for Shell Oil and had participated in spill cleanups.

### **Results of Tracking and Timing**

Sixty-four visitors were sampled in the tracking and timing study. All of the percentages in this section are figured on 61 visitors; three samples were dropped because those people either didn't stop or stopped at only one part of the exhibit and stayed less than one minute, which met our criteria of being a "nonvisitor."

#### **•Total Time Spent**

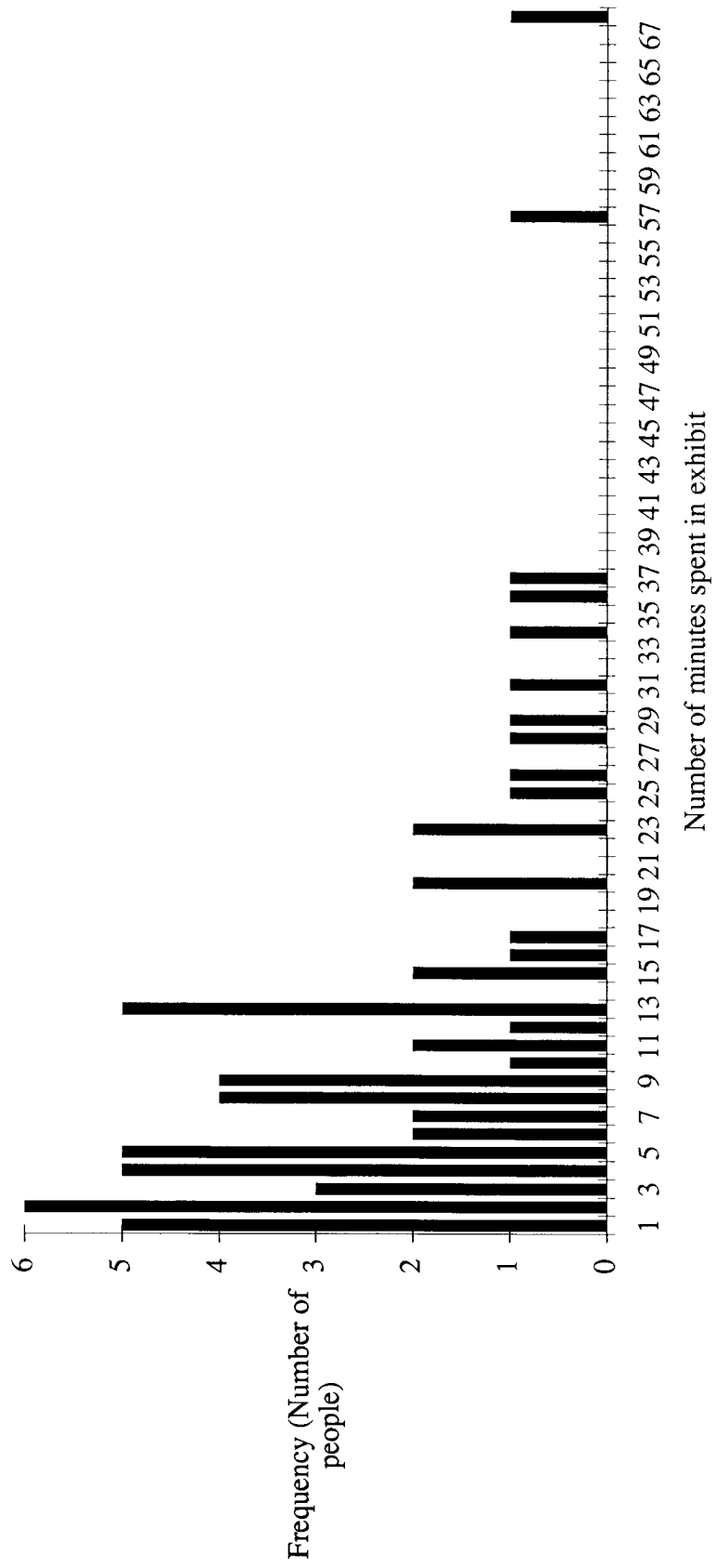
Overall, the median amount of time spent was nine minutes in the exhibit, and the median number of panel groups seen was nine (out of a possible 17) (53%).<sup>1</sup> The shortest time was one minute; the longest 68 minutes. For half the sample to spend more than nine minutes in an exhibit of this size compared favorably to other much larger exhibits which had average times of less than 10 minutes.<sup>2</sup>

Graph A of the frequency distribution of the number of minutes spent shows that the range of times for people who spent less than nine minutes was much narrower than the range of times for people spending more than nine minutes. It also shows the anomalies of the two highest times and the comparatively large number of people who spent 13 minutes. The patterns seen here are not atypical for museum times, although there are not many published data for comparison. The average time per panel group was 60 seconds, far more time than many other studies have reported.

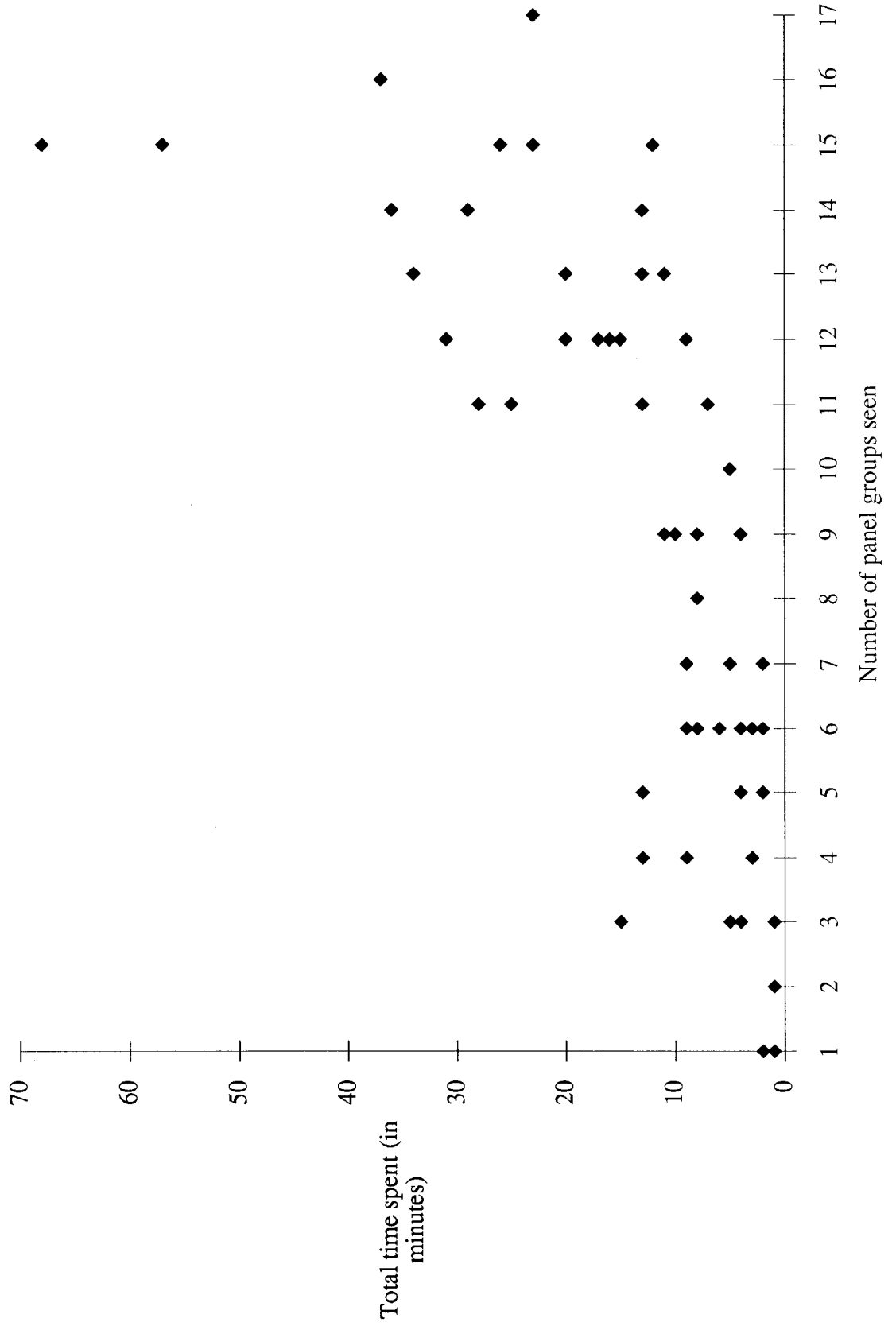
More interesting than time alone is what people did with their time--where and at how many panels they stopped.

Graph B, plotting visitors' total time and number of panel groups seen, shows a very wide scattering of points. In the upper right quadrant, the people who spent the most time and looked at the most cases do not form any discernible pattern, i.e., it's not a clear linear relationship between time and number of panels seen. Of the five people who each spent 13 minutes, they saw 4, 5, 11, 13 and 14 panels respectively.

A. Number of Minutes Spent



B. Time vs Panel Groups



Graph C shows a frequency distribution of the number of people who saw the different panel groups. It is interesting to note that although the median number of panel groups seen was nine, the distribution is not a normal curve. In fact, it looks more bimodal, with six panels typical for a short-term visit, and 12 panels typical for a longer stay in the exhibit.

Graph D of mean time per panel and number of panels seen shows an even wider scattering of points, accurately reflecting the many different styles of viewing the exhibit. Could any of these be called "typical"? Yes. The short visit was typified by #14, a male in 30s with two children (approx. 8 and 13) who glanced at four panels, read seven, touched the fur, looked in the microscope and talked with his kids about the exhibit, spending a total of six minutes. A more intense visiting style was exemplified by #62, male in 20s, who stopped and read 19 times and zigzagged around, spending 20 minutes total. Copies of their data sheets are in the Appendix.

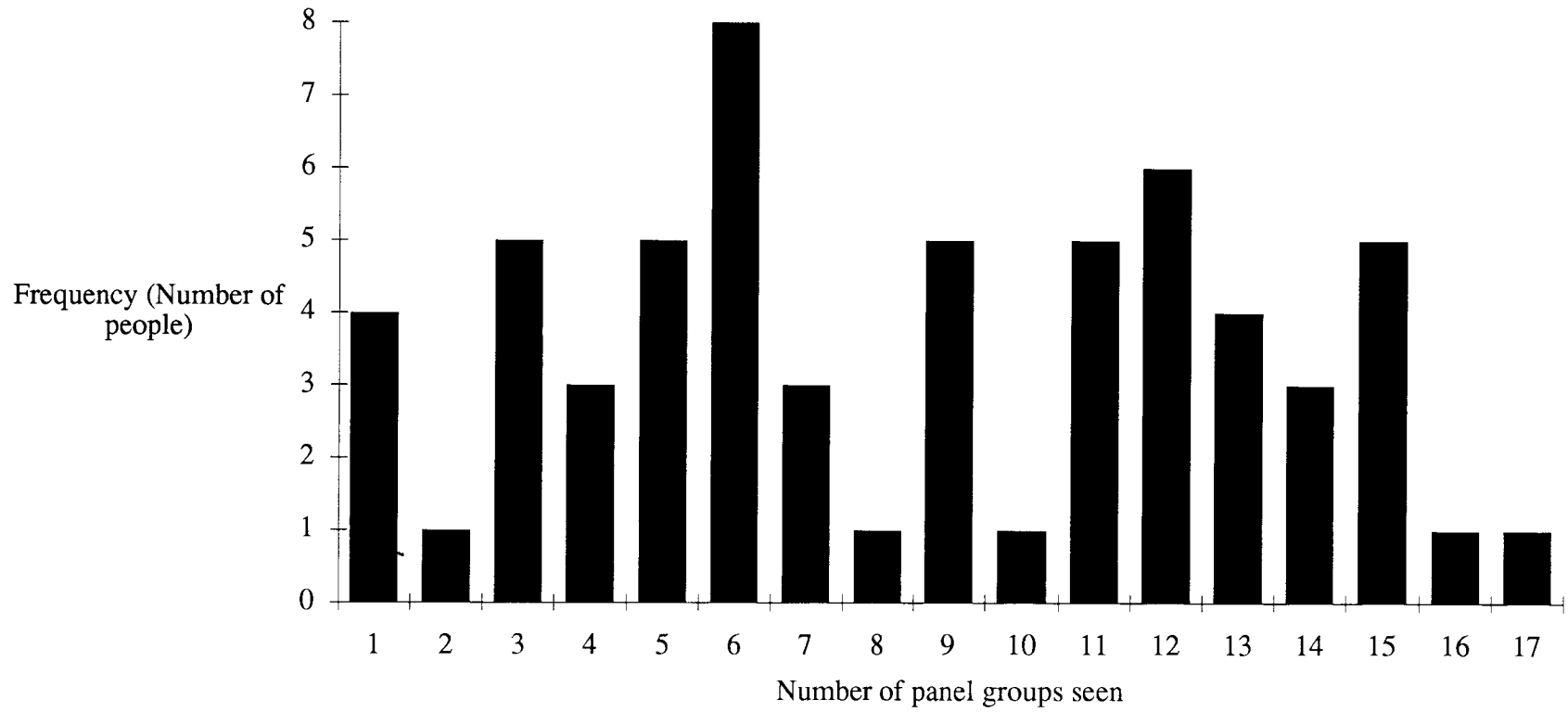
#### •Characteristics of "Serious Visitors"

Before the data were tabulated, we tried to predict how many visitors would stay more than 10 minutes, read, interact and see more than 70% of the exhibit, that is, fall into our definition of "serious visitors." (Guesses by staff and friends ranged from a pessimistic 2% to an optimistic 65%, with the mode being about 10%. )

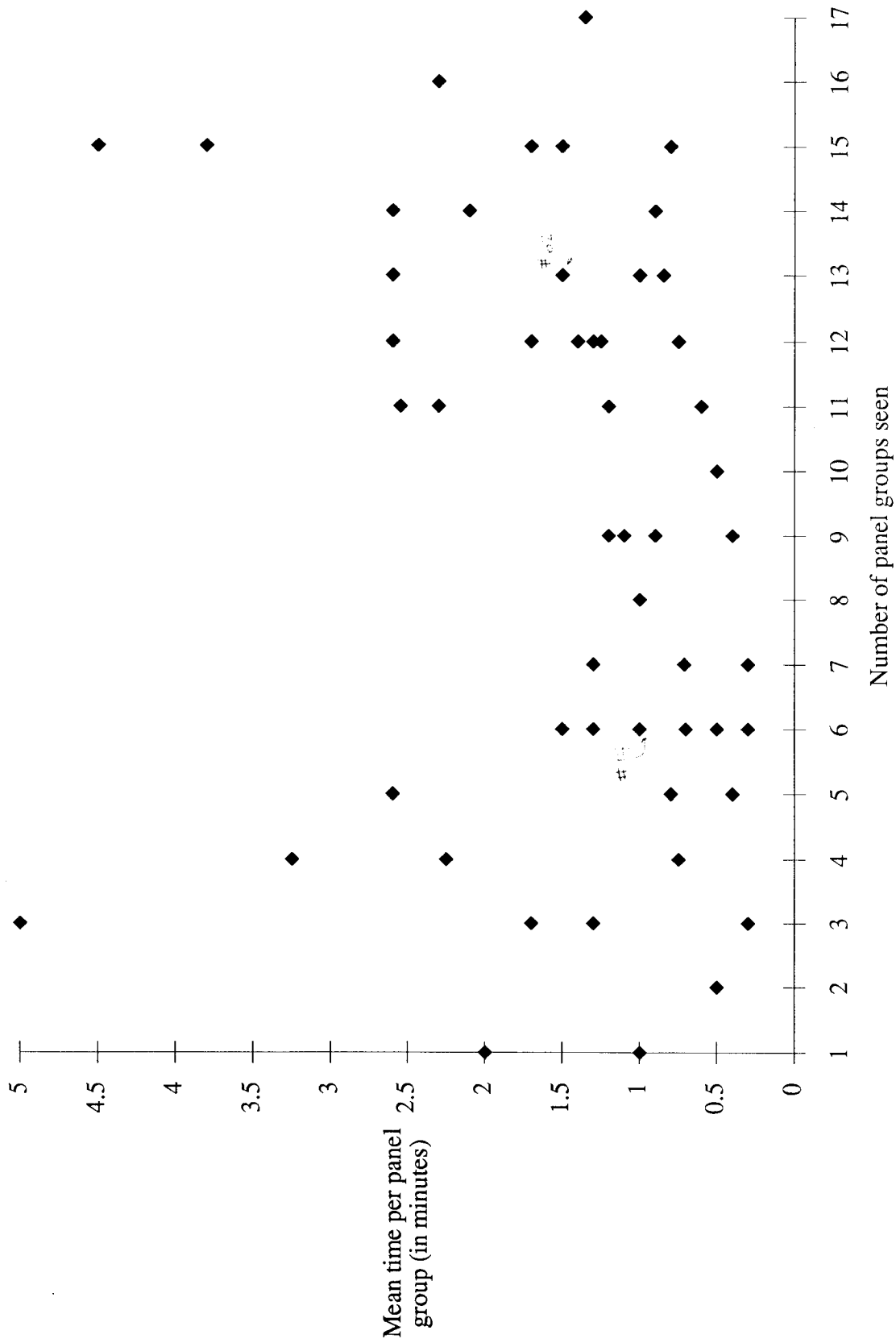
Since the data showed that the median amount of time was nine minutes, and the median number of cases seen was nine, we revised the definition to include visitors who ranked above the medians in both cases, i.e., spent more than 10 minutes and saw at least 65 % of the exhibit.

In all, 36% (22 people) met the criteria for "serious visitor." This seems like a sizable proportion, compared to other exhibits.<sup>3</sup> As a group they did not seem to be significantly different in age and sex ratio from the total group tracked and timed, but there were more adults alone and adult-only groups among the "serious visitors." Does this indicate that other people, especially children, inhibit one's ability to concentrate on reading, looking and spending more time? The data suggest that this might be true.

C. Number of Panel Groups Seen



D. Mean Time vs # of Panel Grps





- Traffic Flow

Thirty percent (18 of 61) went through the exhibit “backwards,” that is, they went to the left and proceeded clockwise through the room, whereas the exhibit was laid out to be read in a counterclockwise story--going somewhat chronologically in time. The exhibit developer had worried that “everybody” was turning left, which wasn’t born out by the data; yet 30% is a substantial number. Several design alternatives to help change that tendency can be implemented in laying out the exhibit in other installations if museum staff feel it’s important to have visitors go the “right way.” In terms of how long they stayed, it didn’t make any difference which way they went: the median time for both groups was nine minutes. But, in fact, a few panels toward the intended “end” of the exhibit seemed to get more attention by the people who went backwards. This follows the generally understood trend in visitor viewing patterns to spend more time at the beginning of the exhibit than at the end.

- Stops at Specific Areas

Graph E shows the relative “popularity” of different areas of the exhibit. The introduction and the credit panel were the least popular, but they were not expected to be. In addition, the intro panel could be read without stopping at it. Another area that didn’t rank high, but didn’t need much time to “take in” was the display of cleanup clothes.

The most popular areas dealt with the animals and how the oil affected them. Specimens, a touchable pelt and a microscope helped make this area interesting, as did the many color photos of high emotional impact.

Eleven of the 13 panels with text, photos and artifacts (not counting the intro and the credits) were attended by more than 50% of the subjects. This indicates that most of the panels are attracting visitors attention.

The political cartoons and the newspaper update sections each attracted 38% of the total sample. These two are the weakest compared to the other panels and seem to have two very different possible explanations:

--Political cartoons are a very sophisticated, often subtle, form of humor that requires some "work" to understand. Perhaps they do not appeal to the broad range of museum visitors.

--The newspaper bulletin board lacked a title to call attention to the fact that this was the "update" section, something many uncued visitors said they would like to see added to the exhibit. Corrections in the location and display of this information are being made.

Thirty-three people (54%) stopped at the "California component," and 13 of them picked up handouts in that area. This degree of attention to the local story will be of interest to other museums that are thinking about adding a locally relevant component.

#### •Behaviors

We were interested in some of the specifics of what people do in the exhibit and had decided ahead of time to record instances of the following behaviors: Read, glance, read out loud, point, talk (about exhibit), use an interactive, laugh and cry. "Cry" was based on comment cards which said, "The exhibit made me cry," and we were curious if we would witness it. During the tracking and timing, no one was seen crying.<sup>4</sup>

Four people were observed laughing--two who were next to the political cartoons, one by the spill map interactive, and one in an instance not related to an exhibit.

Reading out loud was observed in 16 of the groups (26%). In general, visitors tended to be quiet in the gallery, and in the museum as a whole, we noticed.

Reading (silently) was seen in 95% of the total sample (58 of 61), an appropriate behavior, we think, for a highly text-oriented exhibit.

#### •Interactives

Overall about 81% of the groups used at least one interactive. The opportunities

were listed above and included touching, listening, manipulating, etc. All groups that included children participated in at least one interactive. Groups of adults interacted 72% of the time.

Five people (7%) were observed filling out a comment card. The decision to eliminate this area could be justified by this low number.

Trying to guess how many visitors have seen the exhibit based on the number of cards filled out (e.g., five cards for every 61 people) probably is not valid. School groups, for example, contribute a much higher number of cards per visitors than families.

- **Very Short Visits and "Nonvisitors"**

Three people sampled during the tracking visited the exhibit but did not read anything. One was a man who walked through the whole room in one minute without stopping, glancing at only two of the panels directly. (In an uncued exit interview, he reported that he'd "breezed through," yet he thought that the exhibit covered the topic well. He noted that he'd like to have "more graphic pictures of animals that died," and he thought "...it's good that you guys have done it"--not an undesirable impression in one minute.) One woman with two young children came in, went to the otter fur and patted it, and left. The third subject was a teenage girl with four young children who did not seem to be engaged by the exhibit at all.

A few visitors approached the exhibit or entered briefly, then walked away. When asked, "What made you decide not to visit this exhibit?" two of them said there looked like too much reading and text, and one said, "When I see little dead animals...I don't like to see them."

### **Results of Uncued Exit Interviews**

Visitors were encountered as they left DW with an invitation to participate in an interview that would take about three minutes. Most visitors were cooperative, and we obtained 72 samples. A few people refused, due to lack of time, impatient children, or a lack of confidence speaking English. A summary of their answers about first visit and prior interest has been reviewed above. Here we will

summarize the answers to the other seven questions. (See data form in Appendix.)

We are assuming that the patterns of viewing by visitors who participated in the interviews were similar to those we saw in the tracking and timing sample, i.e., an average time of nine minutes, including typical patterns of short and more intense visits.

Everyone was able to answer the question about “the main idea,” and they collectively identified 134 ideas which broke down into several categories that each had similar ideas:

- **Most often mentioned (53 times)** was the intent of the exhibit to make people aware of or **to show what had happened.**

“To show the various issues that went into the spill and the interrelated causes.” “To increase our awareness.” “To make people aware of the damage it caused.”

- **Damage to the environment** was mentioned frequently (30 times).

“How much it affects the environment.” “Destruction [sic] of the birds and of our sea life.” “The social and environmental results of the oil spill.”

- **Personal responsibility and prevention** (23 times)

“We have to be more ecologically aware.” “We can do more to prevent this.” “How the average person can do more to prevent spills and use less petroleum.”

- **Unpreparedness and the scale of the disaster** (17 times)

“To show the scale of the problem and the difficulty of solving problems like this.” “That we’re not prepared for oil spills.” “It brought home to me what happened. Especially seeing the size of spill on the map.”

- **Complexity and tragedy** (11 times)

“It has so many aspects! If oil has to spill, be prepared and do a good job of cleaning up, because the dangers can go on for years and years.”

Visitors were asked to tell us “one new idea” gleaned from DW. About 97 new ideas were mentioned, many which fell into categories similar to those above:

- **Personal responsibility, use of oil and action** (27)

“It made me think about how much I use my car.” “We should contact legislators who have the power to enact preventive measures that can be enforced.”

- **Animals and people hurt and damage done** (21)

"I didn't realize how devastating it was to the wildlife until I saw all those pictures." "I didn't realize the pervasiveness of the oil...the amount of suffering." "I feel sad that it can't be like it was before, with the oil destroying the animals, the landscape and the people who are making a living."

- **Size of the spill** (11)

"The size of the spill was so dramatic." "I didn't realize how large the spill was in comparison with the state of California."

- **Remarkable that the museum has this exhibit** (4)

"It hasn't all sunk in yet. The exhibit is overwhelmingly depressing. It's not the kind of thing you come to a museum for." "I like to see that the museum is bringing the information to the public."

- **Other spills** (4)

"I didn't realize there were so many others." "How often they happen."

- **Cleanup costs** (2)

- **Other** (9)

Twenty-five percent (18) of the people interviewed answered the new idea question with "nothing." Of these, five said they knew it already. "Really hard to say. I followed the spill closely in the news. Nothing new really." One person said that he had gone through fast (our "breezer" mentioned earlier.)

To determine if visitors agreed with the point of view of the exhibit, we asked them to select one statement that they "agreed with most" from a pair of statements. One of each pair better represented the exhibit's point of view. Ninety-two percent agreed with the "right one" of the first pair, and seven people commented about their choice: "I'd like to think that the second one is true, but I'm afraid it's the first one. "That's a hard choice isn't it?" "It's going to take a long time to repair." Eighty-eight percent (63) agreed with the "right one" of the second pair, and seven people couldn't choose, or said "both."

Responses to the interview questions about "explain more clearly" and "want more information" were similar in several ways. The majority of cases said "nothing" (74% and 56%). "No. That's the thing that impressed me so much. They covered it all so thoroughly."

In response to both questions, visitors mentioned: what people can do personally to help; more about what happened to the animals; and what the long-term effects are. Other responses to the "more clearly" included how it happened and cleaning up. To "more information," they wanted more about people, prevention, and consequences for Exxon. In both cases, several people remarked that it was hard to say because they had not looked at everything, or, instead, they volunteered a positive comment about the exhibit.

It is not possible to tell if visitors had seen the areas in the exhibit that related to these sections and actually *did* want more, if they'd missed them, or if they were just thinking of something to please the interviewer. It is clear from the data on tracking visitors that the majority of people did not stop at or read more than half of the panels available. The similarity of responses to both questions, the fact that most of those topics are covered somewhere in the exhibit and the lower response of "nothing" to the second question suggest that visitors were trying to come up with something to say so the interviewer would go on to another kind of question.

The last question was open-ended for any other comments visitors wanted to share. Sixty-eight percent were positive statements about the exhibit, its importance, its impact on the visitor and the need for other people to see it, too. "It is good that the exhibit is so concise. It offers a good overview of the subject." "It should travel all over the country." "It shouldn't be stuck down here--it should be out on the sidewalk!" "...or in shopping centers (in rich suburbs) where there are the kind of people who could make a difference."

Suggestions for changes or additions (that were not already part of DW) included a video to explain the evolution of the spill and more interactives. Both of these are either planned or already under consideration for the traveling version.

### **Results of Cued Testing**

Exhibit developers were concerned about visitors' understanding and comprehension of a specific message in three areas of the exhibit: at the microscope; at a group of labels and a photo about bioremediation; and at the panel on oil in the

environment. Separate cued testing evaluation strategies were used to look more closely at visitors' reactions to and suggestions for those areas.

Cued testing usually involves a small sample of visitors who give very detailed, specific feedback. We were trying for a quota of 25 participants in each of the tests, but did not achieve that many. Nevertheless, the responses we did gather proved to be very useful.

Visitors for the cued testing were recruited with the statement, "We're asking some visitors to give us feedback about a particular part of this exhibit because we are planning to make some changes. Would you help?" The demographics of those recruited were similar to the visitors sampled for the other parts of the evaluation study.

After looking at the component, visitors to each were asked the same basic three questions: What is this about? Do you think visitors will understand it? What changes would improve it?

At the microscope, visitors (n=16) were asked to read the label and look in the microscope (a Wentzscope, equipped with a large ocular and single-control focus knob, showing a slide with preserved plankton, mostly copepods) and then answer the questions. We wanted to know if they were able to make the connection from the specimens viewed to the notion of food chains, with plankton at the base of the food chain. From their comments, we concluded that more than half were "getting it," with answers such as, "The oil spill really destroyed the food chain. If the plankton died, the organisms that feed on it die, and so on all the way up," and "I guess it's trying to show that an organism as small as plankton was affected by this, as were all the other animals." The recommendation here was to simply add another sentence that would help visitors feel more certain that they had understood it. This was a case where visitors correctly thought they understood the message, but they needed to be given the chance to feel more confident about it.

At the section that talked about bioremediation (a photo, a label with two paragraphs and four quotes), visitors (n=9) were asked to read the label out loud to the evaluator (an exercise guaranteed to strike terror into the hearts of most visitors)

and then discuss the questions. Six of the people had some difficulty reading the word "bioremediation." Two people didn't get the point. Six were hesitant that other visitors would understand it, and they suggested simplifying it. We knew from these data that a major rewrite was called for. One visitor's suggestion was very helpful: "Start with 'an experimental process called...'" which allows readers to skip the difficult word and still get the main idea.

With the panels on oil in the environment, we tried out a new quarter-scale mock-up of an alternate idea to the panels that were there because the exhibit developers felt that this story was weak in its original form. Visitors (n=29) took from 30 to 120 seconds to look at it. None seemed to have many problems or questions with the mock-up. Twenty-four thought that visitors would understand it and 10 had no suggestions for improvements. Of the suggestions made, several related to the difference between a mock-up and the final form (e.g., make it bigger, add real photos) changes that were already planned. No other major alterations were made based on this feedback.

In all three cases, due to time limitations, the changes cannot be tested again to assure that what we assume will be better will actually prove to be.

### **Results of Cued Questionnaire**

Part 5 of the evaluation was added after we were well into tracking and interviewing, and it looked like a good opportunity to do a more in-depth evaluation of visitors' reactions to the exhibit. We decided to work with a sample of visitors who would be asked to look at the exhibit more thoroughly than the average visitor.<sup>5</sup> The questionnaire was carried out on Saturday afternoon, October 12, 1991.

Forty-one visitors were recruited as they approached the exhibit. We used this invitation: "We are doing an experimental evaluation of this exhibit today, and we are looking to recruit visitors who would be willing to spend a little more time than ordinary in this exhibit and then fill out a questionnaire. Would you be willing to help?" If they said yes, we told them to go into the exhibit and spend as much time as they wanted and to be sure and look at everything. We offered a free t-shirt as an



incentive for spending the extra time and effort, mentioning that fact when visitors seemed somewhat reluctant (and it usually worked). People who declined (about 15%) were thanked anyway.

By cuing visitors this way, we *created* a sample of "serious visitors" that we could then get feedback from--i.e., visitors who were planning to visit DW in the first place, but were now committed to seeing all of it and spending more than nine minutes. In fact, this sample averaged 26 minutes. The longest time was 58 minutes, while two people spent just eight minutes.

It is interesting to note that the percentage of visitors who reported a special interest is higher in this part of the evaluation than in the uncued part. When asked if they had "any special interest, knowledge, or training about oil spills," 15 visitors answered affirmatively-- 38% here vs. 24% in the uncued interview sample. This may be due to the fact that these visitors were selected to perform a specific task, and by "studying" the exhibit, they adopted a higher level of interest. Also, willingness to take part in this study may in itself be an indicator of higher interest in the subject matter.

Visitors filled out a questionnaire (see sample in Appendix) that asked them to complete the statements "After seeing this exhibit, I now realize that..." and "After seeing this exhibit, I was reminded that..." in their own words. We felt that these two "questions" would help us to understand the visitors' experience of the exhibit and the extent to which they could interpret or extrapolate from that experience, thus giving us a measure of the educational potential and a way to rate the success of the exhibit. Rather than right or wrong answers to factual questions, the results were evaluated according to the framework of the six "big ideas" (see page 3).

From the 41 visitors surveyed, each answer written on the questionnaire sheet was transcribed verbatim and identified as one or more complete thoughts that were considered to be discrete, measurable units. This produced a total of 192 units (the average was four per person, with a range of one to eight). Each unit was then sorted into six categories corresponding to the six "big ideas."

During the sorting process, it became apparent that the six communication goal categories needed to be expanded in order to accommodate the majority of visitor responses. An example of this is the first idea: "Alaska is a beautiful wilderness and should be protected." In their comments, very few visitors spoke directly of Alaska. However, many spoke about wilderness and nature in general being beautiful, and that we should protect our environment. In all six cases, visitors seemed to be taking away ideas that were broader generalizations than the original ideas. Each category was thus restated to reflect its broadest interpretation.

The comments made were definitely not equally distributed over the six ideas. Out of the 192 responses, the highest was 77 and the lowest was 12.

•Categories of Comments from Cued Questionnaires

**The devastation caused by the oil spills was mentioned by 40% of visitors (77 comments).** By far, most responses dealt with the fact that oil spills are bad. Many visitors remarked on the quantity of oil that was spilled, but only one reported this as a fact about how many gallons were spilled; the rest of the comments demonstrated a much deeper appreciation of how the amount of oil spilled had an impact on wildlife, people and the environment in general. It appears from these data that visitors overwhelmingly developed a sense of appreciation about the devastating effects of oil spills, and were able to generalize and apply what they had learned at this exhibit to the harmful effects of oil spills.

"The impact of small spills is much more devastating than I first realized."

"Not only were the animals, nature, and the environment affected by the spill, but the entire world as a whole in one way or another is affected. "

"Many animals both on land and in the sea were greatly harmed. The ones who did survive experienced stress and the possibility of early death."

**Visitors expressed a desire "to do better" in resolving the problems caused by oil spills.** In this category 25 comments (13%) included things that Exxon could have done differently as well as things in general that could be improved. Few people spoke directly of *lessons* that we should have learned from the tragedy, but many made specific recommendations about what we could have done and should be doing. Some spoke of business and industry roles, some mentioned preventive measures and methods for improved cleanup--among which were suggestions for

double-hulled tankers, safer shipping standards, security and supervision of ships and routes and more money devoted to developing better oil recovery techniques.

"I hope Exxon will become more concerned about the environment as well as their profit margin."

"The Exxon Valdez did very little to be prepared for such an incident."

"[I am] sad that we are not doing anything about it."

Two visitors were pessimistic about learning anything from the spill.

**Eleven percent (22 comments) indicated that visitors believed the effects of oil spills are far-reaching and long-term, and that it could easily happen again.** Three themes emerged in these responses: oil as one part of a larger interconnected picture; a sensitivity to our vulnerability to another spill; a quiet despair and disappointment in the human race.

"Oil is a very important natural resource, but at the same time it can be used to destroy lives."

"The Bay area is as vulnerable as ever. Very little has been done to prepare us even though a lot was learned about the inadequate methods for cleanup."

"There has got to be something very wrong when we feel that money fully compensates for the loss of life of any creature and a way of life for any man."

**Nature is beautiful and we have a responsibility to protect it.** Eight percent, or 16 comments addressed the issue of human beings as stewards of this planet in strongly worded, philosophical and almost religious statements. Almost half of these comments came from the prompt, "I was reminded that..." on the questionnaire. Although only one person mentioned Alaska specifically, the exhibit seemed to stimulate global thinking that all of nature is beautiful and that it is our responsibility as people to protect the whole of the planet Earth.

"This is my planet even though I don't live in Alaska."

"I'm sure the Lord didn't die for us so that we could do this to these beautiful creatures. I wish America would wake up and smell the oil."

**Visitors felt futility and despair, although some expressed the desire to make changes in order to help the environment.** Thirteen comments (7%) fell into the What can I do? category. Few visitors stated specifically that they would go out and change a particular behavior on their part. Instead, they appeared to acknowledge

that oil is an issue that needs to be addressed at various levels. The exhibit appears to develop in visitors a deep understanding of the harmful consequences of our use of oil products, but does not give many specifics about what can be done by individuals. More than half (7 comments) mentioned sadness and a feeling of futility, questioning what, in fact, they *could* do. A few of these blamed the U.S. government.

“More conservation on my part would mean less possibility of an oil spill.”

“I am very sad and will leave this museum very unhappy, wondering what I can do.”

“Our country . . . is in deep trouble and needs decent leadership.”

**The theme of inability to cleanup oil spills did not come through very clearly in visitor comments, appearing in only 6%, or 12 instances.** This may not be due to a lack of understanding of the difficulty of cleanup, but rather that this understanding is overshadowed by comments which fell into the category of “We can do better,” which takes the causes and consequences of oil spills a step further. Two comments referred to the expenditure of time and energy needed to cleanup the spill; three mentioned the difficulty of cleanup and two of these referred to inadequate tools. One visitor specifically mentioned the clothes people wore during the cleanup (examples of which were displayed in one of the exhibit sections), and two commented on the money made during the cleanup. One person cited high pressure hoses forcing oil into the ground, but did not elaborate on what that implied about cleanup effectiveness.

•**Mentions of specific parts of the exhibit.**

The “petroleum game” proved to be an attractive interactive, and five visitors mentioned they had learned about the many things that are made from oil, specifically naming several of the products displayed there.

Seven other areas were referred to, with seven of 16 comments about the “How Big is Big?” panel that had a map with a movable overlay of the spill which allowed size comparisons. This simple interactive brought home the relative size of “big” very effectively by showing visitors how much area the Exxon Valdez spill could cover in an area more familiar to them than Alaska.

Listening to the tapes of the Native Americans and of the voice of the ship's captain to the Coast Guard just after the accident made a big impression--mentioned by several people.

It is important to remember that an assessment tool such as the one used in this section documents visitors' immediate recall, prompting it in a particular way. This is not to say that they did not learn or carry other ideas away with them, and it does not tell us what they will recall in their long-term memory.

As can be seen from the data presented in the cued questionnaire section, all of the six communication goals are represented to some extent. What was most surprising about visitors' responses in this part of the evaluation study was the depth and richness of their language in expressing their realizations and recollections. The content of their comments was not all that different from the kind we had gathered during the uncued exit interviews, but it was much more detailed and explicit.

## **Conclusions**

The data collected in the five different parts of this evaluation gave us broad as well as detailed descriptions of visitors' behaviors, interactions, demographics, thoughts, opinions, interests and reactions to "Darkened Waters." Although our sample sizes were modest by marketing standards, the quality of information and the patterns that emerged gave us sufficient evidence to conclude that the exhibit is successfully meeting its communication goals and compares favorably with other exhibits in terms of its ability to attract and hold visitors' attention.

The findings from the cued questionnaire are particularly interesting. What seems to emerge there is the visitors' tendency to report information (in the form of facts, impressions, opinions) that reflect a broader understanding of the concepts than are actually presented in the exhibit. Perhaps this is because the exhibit presents the details of one (albeit major) case study/documentary in a clear and concise way which then allowed visitors to make their own generalizations, abstractions and value judgments by integrating the exhibit's information with their own prior

knowledge. The decision on the part of the exhibit developers to limit the content appears to have a very positive effect on the evidence of visitors' learning.

Efforts are in progress to make a good thing even better. In response to the feedback from this study, changes are being made to DW before it travels. Other changes, planned before the evaluation, are also being made to enhance visitors' enjoyment and their ability to grasp the exhibit's messages quickly, answer their questions clearly, allow them to experience more interactive exhibit devices and move them through the space comfortably.

## Notes

1. Medians alone do not give an accurate or descriptive picture of what visitors did in "Darkened Waters." Medians are useful for comparing data from one exhibit to another, but not for describing visitors within one space, because there is no "average visitor." Looking at frequency distributions and graphs is more helpful to see what's going on.

2. Average time in exhibits has been reported infrequently in museum literature. A few examples for comparison include the following:

Field Museum's old animal halls--eight minutes in a hall containing 65 cases of stuffed birds. Eleven seconds per case.

California Academy of Sciences' new "Life Through Time"--12 minutes in a large, varied hall with dioramas, live animals, video, computers, specimens and graphics. Most exhibit stops were less than 20 seconds.

Shedd Aquarium's "Spirits in the Wilderness" exhibit about art and culture of the Northwest Coast--median of five minutes in a 2,000-sq.-ft. area with photos, artifacts and video.

3. Where might you find a higher proportion of visitors who meet these "serious viewer" criteria (spend longer than 10 minutes, see more than 65% of the exhibit, read labels)? The Mexican Fine Arts Museum in Chicago might be one place with its annual exhibit, "Dia de los Muertos," or "Day of the Dead." This exhibit has high intrinsic interest, a high frequency of repeat visitors each year, small overall attendance (less than 300 a day average), a large percentage of people who have a special interest in the topic, which is also easily understood by someone without any special prior knowledge.

4. Crying was actually observed once during the cued exit interview study. The visitor said, with tears in her eyes, "I'm too overwhelmed to talk."

5. The whole notion of looking for "learning" and the question, What did visitors learn from the exhibit? is an elusive and inappropriate search using casual visitors. See Lucy Nedzel's valiant attempt and her conclusions in *The Motivation and Education of the General Public Through Museum Experiences*, a University of Chicago dissertation in 1952. We felt that it would be unfair to embarrass visitors by

asking them to tell us what they learned after they looked at the exhibit for nine minutes. The point is not what *did* they learn, but what *could* they learn if we motivated them to participate in a situation where the question made more sense. We carefully avoided using the word "learn." Rather, we picked up on the words they'd used in the uncued interviews: "It made me think...," "I never realized that...," "I was reminded that...," and "I didn't know that...," for developing the questionnaire.