

# National Air and Space Museum

## All Exhibitions Study

### Summary Report

Key Findings	1
Introduction	2
Frequencies	3
Audience Differences	6
Exhibition Differences	7
Bottom Line	11

Office of Policy and Analysis  
Smithsonian Institution  
June 2015

## Key Findings

- Overall Experience Ratings across all 20 exhibitions (0% Poor, 2% Fair, 24% Good, 56% Excellent, 18% Superior) were very close to the ratings for the museum as a whole. They are close to the Smithsonian average, although less-than-Excellent ratings are lower than the SI average.
- *America by Air* and *Pioneers of Flight* had the highest ratings; *Looking at Earth* the lowest. *Time and Navigation* was closest to the average.
- Overall Experience Ratings for the 20 exhibitions were very similar to one another.
- Audiences differed in overall experience ratings. Females were more likely to give lower ratings, US non-local residents were more likely to rate their overall experience Superior, compared to those from other countries.
- Audience composition differed for some exhibitions, e.g., *Space Race* and *Golden Age of Flight* drew higher percentages of first-time visitors.
- For three of the 20 exhibitions (*America by Air*, *Pioneers of Flight*, and *Wright Brothers*), Superior ratings exceeded less-than-Excellent ratings.
- The differences in both Superior and Less-than-Excellent overall experience ratings between *America by Air* and *Looking at Earth* are statistically significant, but not large enough to identify them as true outliers.
- The exhibitions likely to yield valuable information in an in-depth study are *America by Air*, *Pioneers of Flight*, and *Looking at Earth*, together with one exhibition that received an average rating, as a control.

# Introduction

## The Study

Between 29 May and 15 June, 2015, the Office of Policy and Analysis (OP&A) conducted surveys of visitors exiting every one of the 20 exhibitions in the museum. Overall 4,089 adult visitors completed simple, 5-question surveys. The cooperation rate was 80%. Approximately 200 visitors were surveyed at each exhibition. The purpose of the study was to identify exhibitions that were particularly successful or unsuccessful with visitors. It was further proposed that these outlier exhibitions could be studied in greater depth to gain insight into what works best and what should be avoided in the future as the museum is re-done.

## The Questions

The central measure was the Overall Experience Rating:

**Please rate your overall experience at this exhibition [exhibition name], today.**

**Poor    Fair    Good    Excellent    Superior**

This measure has been used by OP&A for over a decade at all Smithsonian museums and has been proven to be a reliable and accurate measure of the quality of visitor experience. Overall across the Institution, about half of visitors rate Smithsonian exhibitions and museums as “Excellent.” Those who are particularly excited about their experience mark the highest level on the scale, “Superior.” Those who have some concerns that prevent them from rating their experience as Excellent choose one of the three lower ratings, primarily “Good.” The Excellent rating appears to function as a kind of default setting – the variability across programs and museums lies primarily with ratings of Superior (on the positive side) and ratings of less than Excellent (i.e., Poor/Fair/Good) on the negative side.

Additional variables in the survey questionnaire were: First visit; visiting alone, with other adults, or with youth under 18; age; male/female; live in local DC area, live elsewhere in US, live in another country.

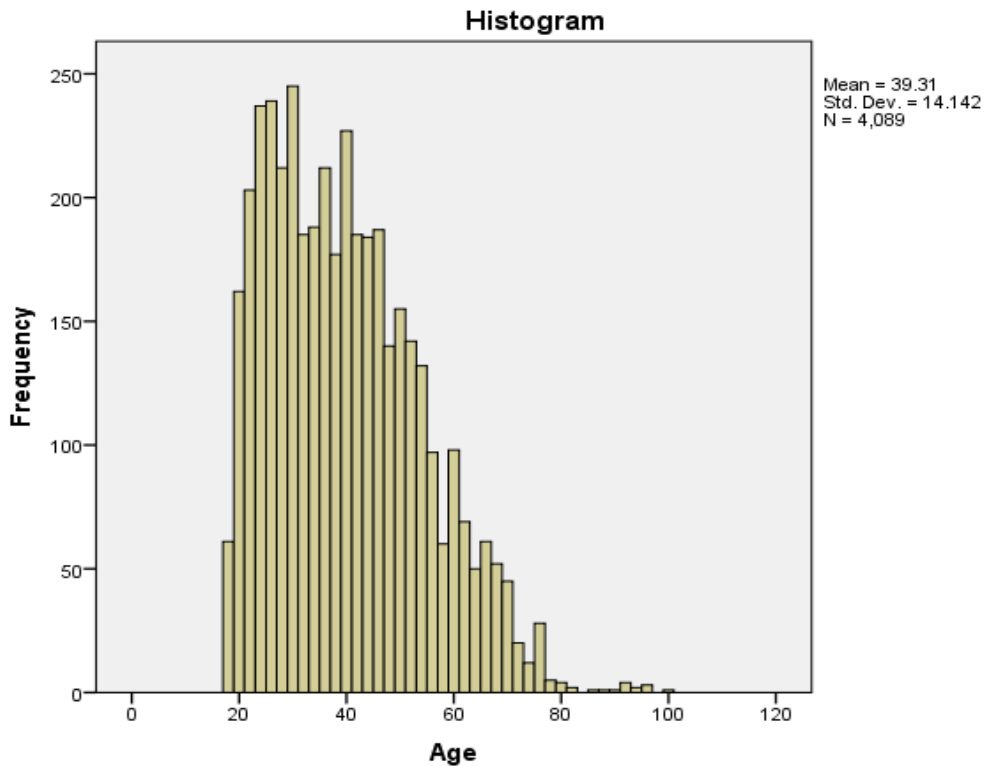
# Frequencies

## Visitor Characteristics

Across the entire dataset:

- 65% were making a first visit to the museum.
- 8% were visiting alone.
- 69% were visiting with at least one other adult.
- 29% were visiting with at least one youth under age 18.
- 58% of visitors were male.
- 8% live in the local DC area.
- 65% live elsewhere in the U.S.
- 26% live in another country.
- The median age of visitors was 38 (mean age: 39.3). The most common age was 29. See Figure 1.

**Figure 1**  
**Ages of Visitors Ages 18 and Over**

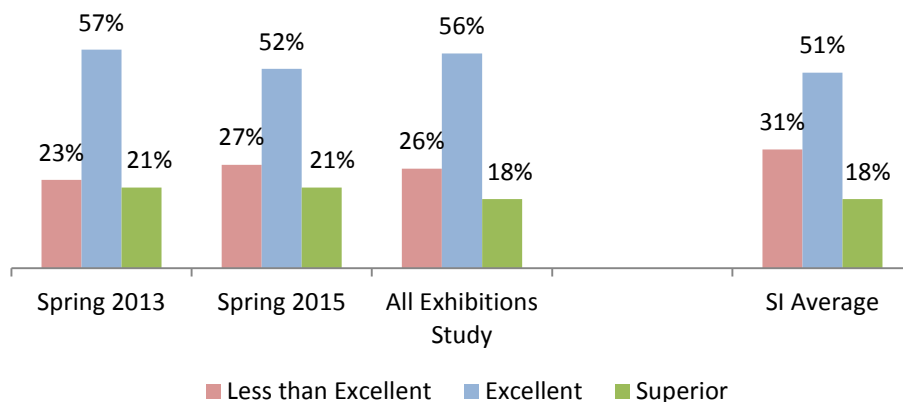


## Overall Experience Ratings

The average rating across all the exhibitions was 0% Poor, 2% Fair, 24% Good, 56% Excellent, 18% Superior. This is very similar to the rating for visitors exiting the museum as a whole in Spring of 2013 (0% Poor, 1% Fair, 22% Good, 57% Excellent, 21% Superior), and also in the survey completed one month before the all-exhibitions study began. See Figure 2.

**RATINGS ACROSS ALL THE EXHIBITIONS WERE VERY CLOSE TO THE RATINGS FOR THE MUSEUM AS A WHOLE. THEY ARE CLOSE TO THE SMITHSONIAN AVERAGE, ALTHOUGH LESS-THAN-EXCELLENT RATINGS ARE LOWER THAN THE SI AVERAGE.**

**Figure 2**  
**Overall Experience Ratings**

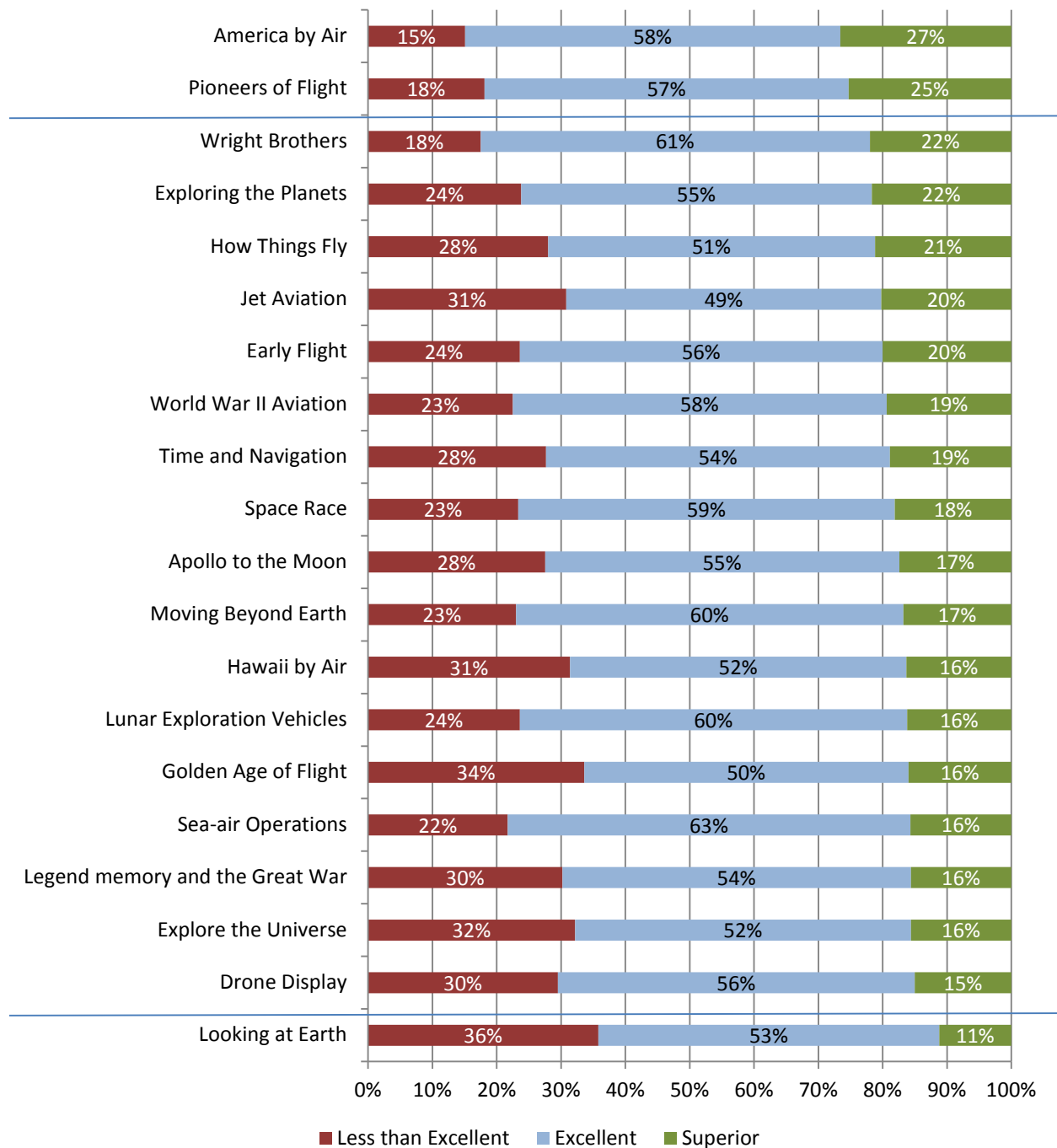


**AMERICA BY AIR AND PIONEERS OF FLIGHT HAD THE HIGHEST RATINGS, LOOKING AT EARTH THE LOWEST. TIME AND NAVIGATION WAS CLOSEST TO THE AVERAGE.**

*America by Air* was rated Superior by 27% of visitors and rated less than Excellent by 15%. *Pioneers of Flight* was rated Superior by 25% and less than Excellent by 18%. At the other end of the rating list, overall experience in *Looking at Earth* was rated as Superior by 11% and less than Excellent by 36%.

Except for three exhibitions – *America by Air*, *Pioneers of Flight*, and *Looking at Earth* – the percentages of respondents rating their overall experience Superior were within one standard deviation of the mean rating. See Figure 3.

**Figure 3**  
**Comparison of Overall Experience Ratings for All NASM Exhibitions**



## OVERALL EXPERIENCE RATINGS FOR THE TWENTY EXHIBITIONS WERE VERY SIMILAR TO ONE ANOTHER.

As suggested by Figure 2, the majority of the 20 exhibitions received ratings that were fairly close to one another. In other words, the overall experience ratings are all quite close to the average.

### Audience Differences

#### AUDIENCES DIFFERED IN OVERALL EXPERIENCE RATINGS

Across the entire dataset there are some significant associations<sup>1</sup> between characteristics and overall experience ratings:

- Females were more likely to rate their overall experience as less than Excellent (30% vs. 24% for males).
- Visitors who live elsewhere in the U.S. were more likely to rate their overall experience Superior (20% vs. 13% for those who live in another country).

#### AUDIENCE COMPOSITION DIFFERED FOR SOME EXHIBITIONS

Some exhibitions drew more visitors with certain characteristics:

- *Space Race* and *Golden Age of Flight* drew higher percentages of first-time visitors (73% and 74% vs. an average of 65%).
- *Explore the Universe* and *Moving Beyond Earth* drew lower percentages of visitors with youth (19% and 22% vs. an average of 29%).
- *Drone Display* drew a higher percentage of males (69% vs. an average of 58%).
- *Wright Brothers* and *How Things Fly* drew higher percentages of visitors from elsewhere in the U.S. (75% and 72% vs. an average of 65%).
- *Looking at Earth* drew higher percentages of foreign visitors (32% vs. an average of 26%).

---

<sup>1</sup> This report identifies significant differences among sub-groups with a chi-square test where significance is set as  $p < .01$ . In such cases the likelihood that a difference is an accident of the sample is less than one in one thousand.

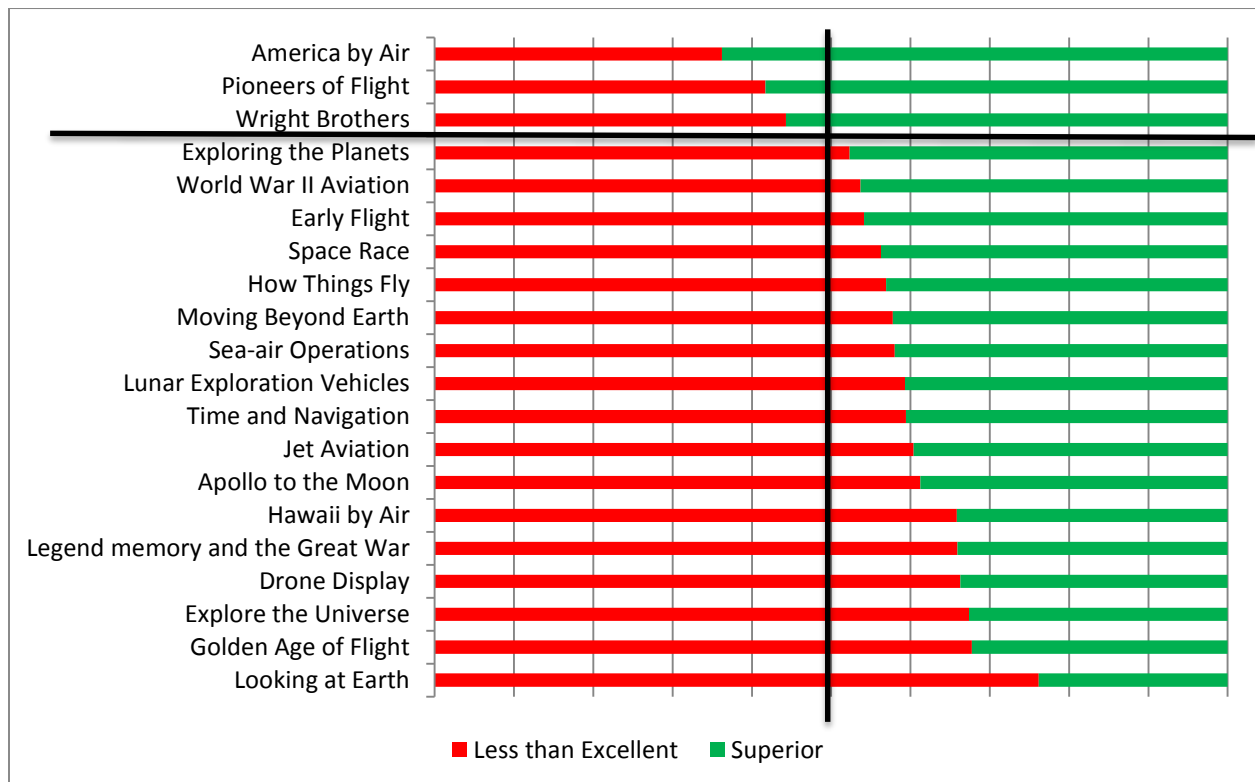
## Exhibition Differences

**FOR THREE OF THE 20 EXHIBITIONS, SUPERIOR RATINGS EXCEEDED LESS-THAN-EXCELLENT RATINGS.**

### Superior vs. Less than Excellent

One simple way to compare overall experience ratings across exhibitions is to examine the relationship between Superior ratings (strongly positive visitors) and ratings that are less than Excellent (critical or negative visitors). The relationship among the 20 exhibition is presented graphically in Figure 4, which shows the ratio between Superior percentages and less than Excellent percentages. Note that for three exhibitions, *America by Air*, *Pioneers of Flight*, and *Wright Brothers*, the ratio favors Superior. In *Exploring the Planets*, *World War II Aviation*, and *Early Flight*, the percentage that rated their experience Superior is approximately equal to the percentage that rated their experience less than Excellent. In the remaining exhibitions the more negative ratings dominated, reaching a high point with *Looking at Earth*.

**Figure 4**  
**Ratio of Superior to less than Excellent Overall Experience Ratings**





## THE DIFFERENCE IN EXPERIENCE RATINGS BETWEEN *AMERICA BY AIR* AND *LOOKING AT EARTH* IS STATISTICALLY SIGNIFICANT, BUT NOT LARGE ENOUGH TO IDENTIFY THEM AS TRUE OUTLIERS.

### Outliers

The central aim of the study was to search for outliers among the exhibitions with respect to visitors' quality of experience. Should we think of the three exhibitions with the highest ratings, *America by Air*, *Pioneers of Flight*, and *Wright Brothers*, and the one with the lowest rating, *Looking at Earth*, as outliers? In order to address this question we need to go a bit deeper into the statistics.

### Confidence intervals

All survey statistics have confidence intervals. When we choose a sample of visitors from a population and measure some characteristic among them, there is always the question of how confident we can be that the measure we obtained from our sample (the sample measure) is the same as the population measure, i.e., the measure that we would have obtained if we had surveyed every visitor who was in the exhibition during the entire period of the survey, not just during the hours when we were surveying. The size of the confidence interval is based on the size of the sample (in our case 200 for each exhibition), the percentage in question, and the level of probability desired (usually either 95% or 99%).

The simplest way to grasp the relationship among confidence intervals is to examine them visually, as shown in Figures 5 and 6. Figure 5 shows 99% confidence intervals for Superior ratings and Figure 6 shows 99% confidence intervals for less than Excellent ratings.<sup>2</sup> In each exhibition we can feel 99% certain that the measure in the population (not just in our sample) is somewhere in that bar. When two bars overlap, we know that there is some possibility that the population measure is in fact the same for those exhibitions.

For both Superior and less than Excellent ratings nearly all the bars overlap. Only two bars, *America by Air*, and *Looking at Earth* do not have any overlap. We can say with 99% confidence that the difference in ratings between these two exhibitions is statistically significant.

---

<sup>2</sup> In this case 99% confidence intervals have been chosen rather than the more usual 95%, because there are 20 exhibitions. In the case of a 95% confidence interval there is a probability that in one out of 20 exhibitions the "true" value would be outside our confidence interval. With a 99% confidence interval, however, we can be more secure that the bars in our graph include the measure in the population as a whole.

**Figure 4**  
**99% Confidence Intervals for Superior Ratings**

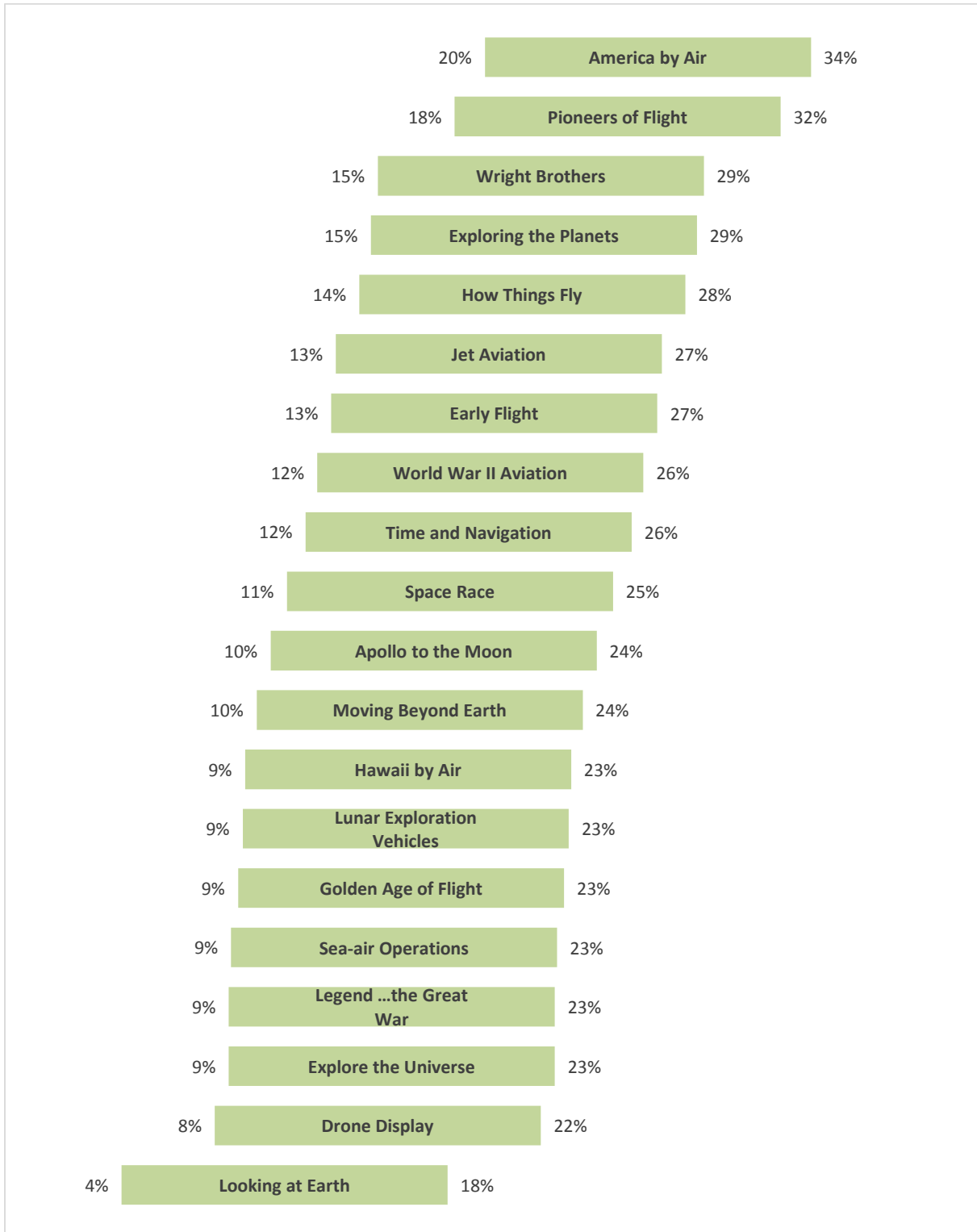
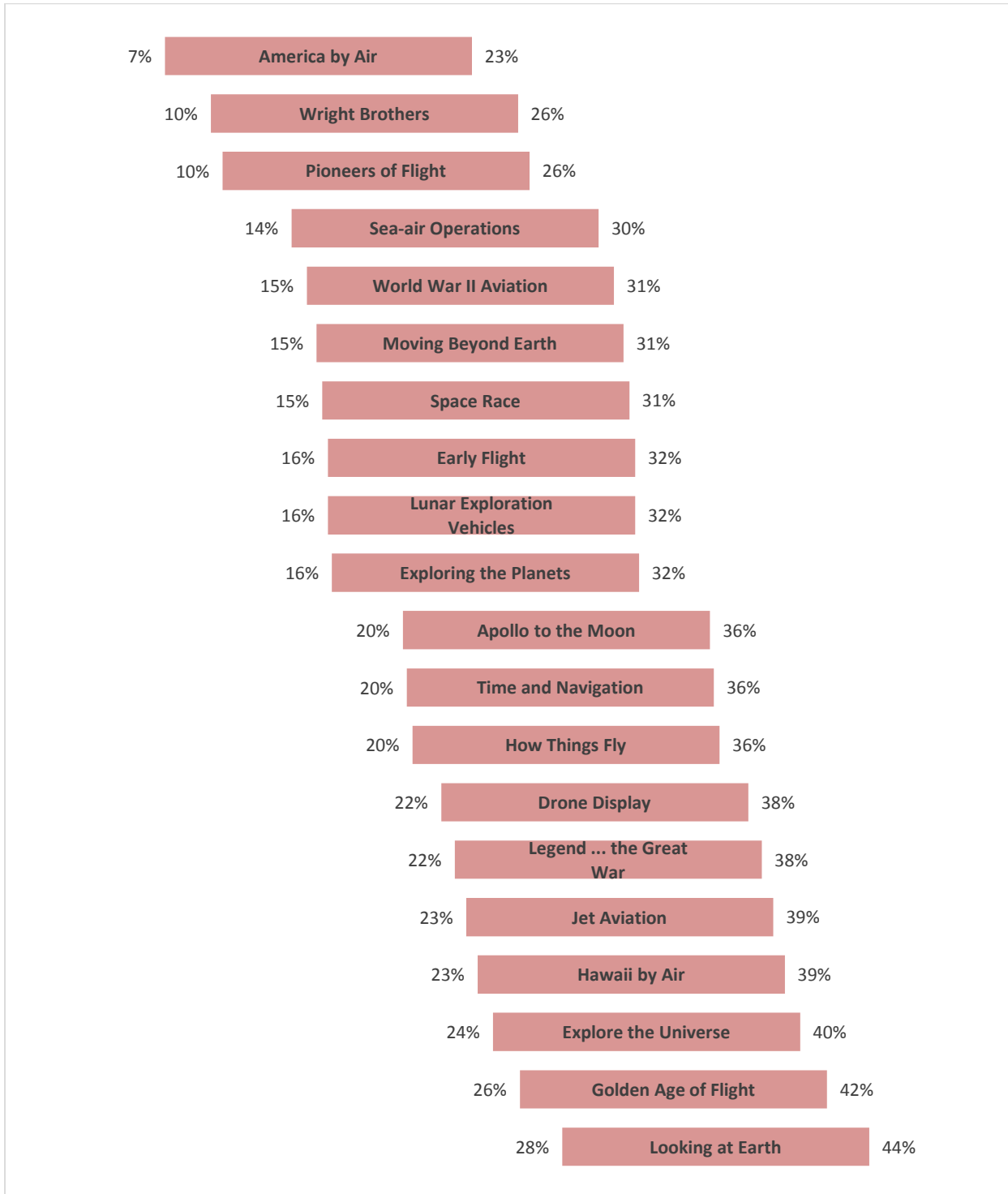


Figure 5

99% Confidence Intervals for Ratings Less Than Excellent



As demonstrated above, *America by Air* and *Looking at Earth* are clearly different from one another, but neither of them is that different from other exhibitions. Thus they are different from a scientific viewpoint, but whether or not one calls them “outliers” is a more subjective judgment.

There could be real value in studying the exhibitions at the extremes (*America by Air*, *Pioneers of Flight*, and *Looking at Earth*) together with one in the center, such as *Time and Navigation*.

## **Bottom line**

**THE EXHIBITIONS LIKELY TO YIELD THE MOST VALUABLE INFORMATION IN AN IN-DEPTH STUDY ARE *AMERICA BY AIR*, *PIONEERS OF FLIGHT*, AND *LOOKING AT EARTH*, TOGETHER WITH ONE EXHIBITION THAT RECEIVED AN AVERAGE RATING, AS A CONTROL.**

In view of the museum’s desire to make exhibitions that excite the public and the need to consider which new exhibitions and features to create, insights into what visitors find so exciting about *America by Air* and *Pioneers of Flight* or so disappointing about *Looking at Earth* might be valuable. The data from the current study cannot provide those kinds of answers, but it does help us know that these exhibitions are the places where investigation is most likely to be rewarded with new understanding.