SAE TECHNICAL PAPER SERIES

2008-01-2027

Testing the Celentano Curve: An Empirical Survey of Predictions for Human Spacecraft Pressurized Volume

Marc M. Cohen
Northrop Grumman Corporation

38th International Conference on Environmental Systems
San Francisco, California
June 29-July 2, 2008

SAE International™

400 Commonwealth Drive, Warrendale, PA 15096-0001 U.S.A. Tel: (724) 776-4841 Fax: (724) 776-0790 Web: www.sae.org

By mandate of the Engineering Meetings Board, this paper has been approved for SAE publication upon completion of a peer review process by a minimum of three (3) industry experts under the supervision of the session organizer.

ISSN 0148-7191

Printed in USA

2008-01-2027

Testing the Celentano Curve: An Empirical Survey of Predictions for Human Spacecraft Pressurized Volume

Marc M. Cohen
Northrop Grumman Corporation

ABSTRACT

In 1963, Celentano, Amorelli, and Freeman of North American Aviation described a set of curves as an *Index of Habitability* that can predict the amount of pressurized volume necessary per crewmember to conduct a mission at "tolerable, performance, or optimal" levels. This paper presents an analysis of the "Celentano Curve" that depicts a relationship between spacecraft pressurized volume and the duration of a space mission.

Since Yuri Gagarin flew in Vostok 1 in 1961, the US, Russia, and China have launched more than 250 human spaceflights. This survey collects the empirical data and tests the Celentano curves against it. The statistical approach treats the Celentano curve as the hypothesis stating a causal relationship between mission duration and volume. Many authors have published variations of the Celentano curve, and this author considers nine interpretations, plus three versions of the crew size hypothesis and one functional operations hypothesis.

This analysis shows that pressurized volume increases as a function of mission duration, both as a power curve and a logarithmic curve. This volume trend does not level off but continues to rise throughout the historic envelope of human spaceflight.

INTRODUCTION

Since Celentano, Amorelli, and Freeman's seminal paper, habitability researchers published dozens of citations, interpretations, and variations of the Celentano curves. This paper presents an analysis of the Celentano hypothesis and its acolytes, testing them empirically against the historical data of human spaceflight. In each case, the authors frame their predictions in terms of "meeting crew needs." Most notable among these predictions of volumetric "requirements," NASA adopted an embodiment of the Celentano curve in the 1987 Man-System Integration Standard (MSIS).

The history of human spaceflight -- from the earliest flights in tiny capsules to the capacious International Space Station -- provides the baseline from which to evaluate and test these predictions. What the historical record affords is a metric to analyze how pressurized volume varies with mission duration, crew size, and other causes.

FIGURE 1 shows the original Celentano plot that features the volume prediction rising steeply over the shorter missions, but leveling out after six months at about 700ft^3 (19 to 20 m^3). Celentano et al posited three levels of accommodation "tolerable, performance, and optimal," but did not define them clearly.

OBJECTIVES

The objectives of this research are:

- To determine what the facts are and what is true about volume and mission duration.
- To provide an empirical baseline of historical spacecraft and missions against which to compare and perhaps validate volumetric designs in the future, and
- To identify the spacecraft volume and mission envelope issues to which architectural design research may make the greatest contribution.

CAVEAT: VOLUME ESTIMATION FROM FIRST PRINCIPLES

This research does not address design guidelines and methodologies to size pressurized spacecraft and space habitats. Certainly, those topics are essential future steps, but they exceed the scope of the present effort.

- The scope of this study extends only to identifying, clarifying, and assessing the historical and empirical record of human spaceflight. It is not a substitute for calculating volume and mass requirements from first principles in the design of crewed spacecraft. What this caveat means is that the human spaceflight community must develop and validate the tools to predict and plan these spacecraft parameters and requirements to meet crew needs across a broad range of functions, missions, and operations.

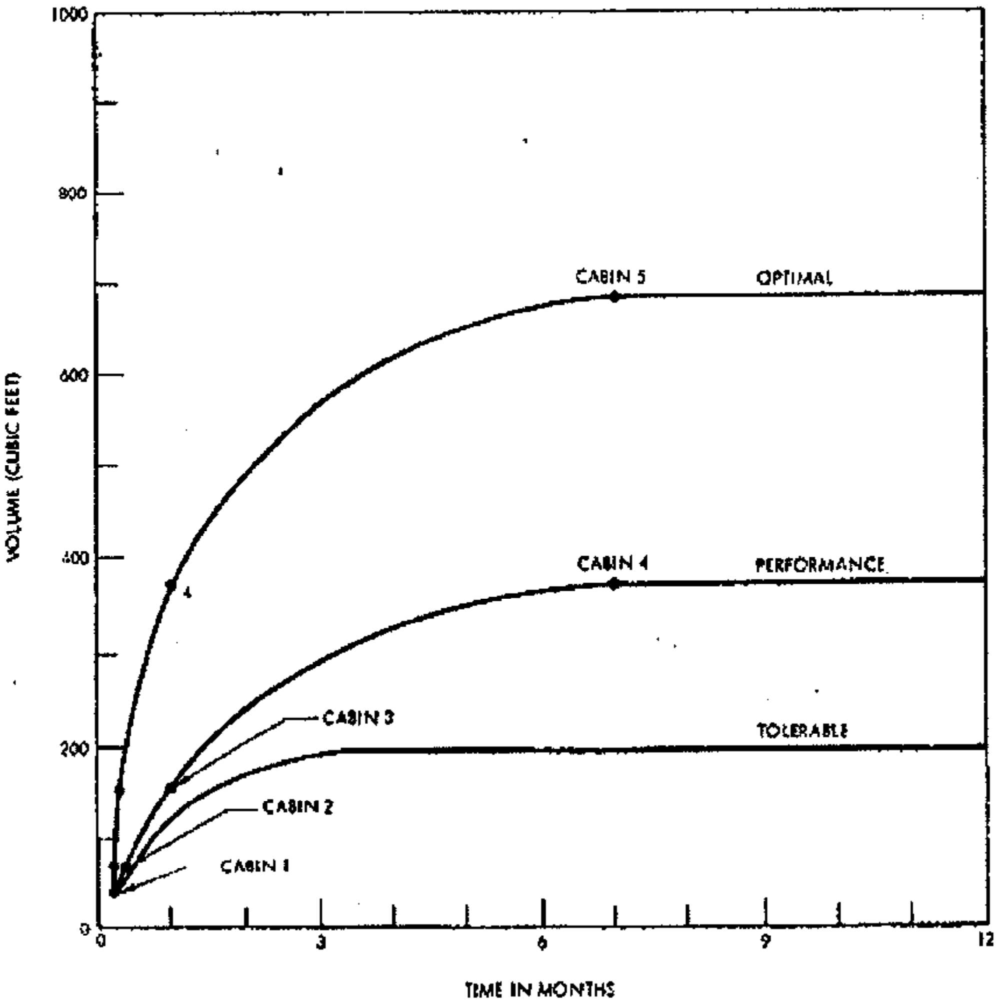

Figure 11. Living Space Per Man (Volume)

FIGURE 1. The original "Celentano Curve" 1963 shows Volume in ft^3 on the Y-axis and Mission Duration in months on the X-Axis. The three curves appear from top to bottom as "Optimal, Performance, and Tolerable."

A further caveat is that this study addresses only gross pressurized volume. It does not make distinctions among subtractive properties within the pressurized cabin, known variously *as habitable volume, free volume, or living space.* There are two reasons for this choice:

1. There is good documentation available only on actual pressurized volume. Very few spacecraft have any measurements available for the subtractive volumes. At the same time, there are sometimes inaccurate and conflicting values published for certain spacecraft volumes, so it may take real detective work to find the true data.
2. There is no agreement on how to measure habitable volume or free volume or even how to define it, but there is universal agreement that pressurized volume is the entire space within the pressure vessel that contains the crew cabin.

HISTORICAL CONTEXT

At the time Celentano, Amorelli, and Freeman posed their hypothesis, human spaceflight was in its earliest period of the Vostok, Mercury, Voskhod, and Gemini spacecraft. The experience of zero gravity was limited to very small cabin volumes. Even the Apollo and Soyuz programs seemed far in the future, so that they had few data points for their theory. Never the less, by predicting quantitative relationship between mission duration and spacecraft size their paper became highly influential. This relationship is very important because the ability to predict cabin size will minimize huge potential variations in mass and volume for long duration missions such as a Mars Transfer Vehicle (MTV) or at a moon or Mars surface base.

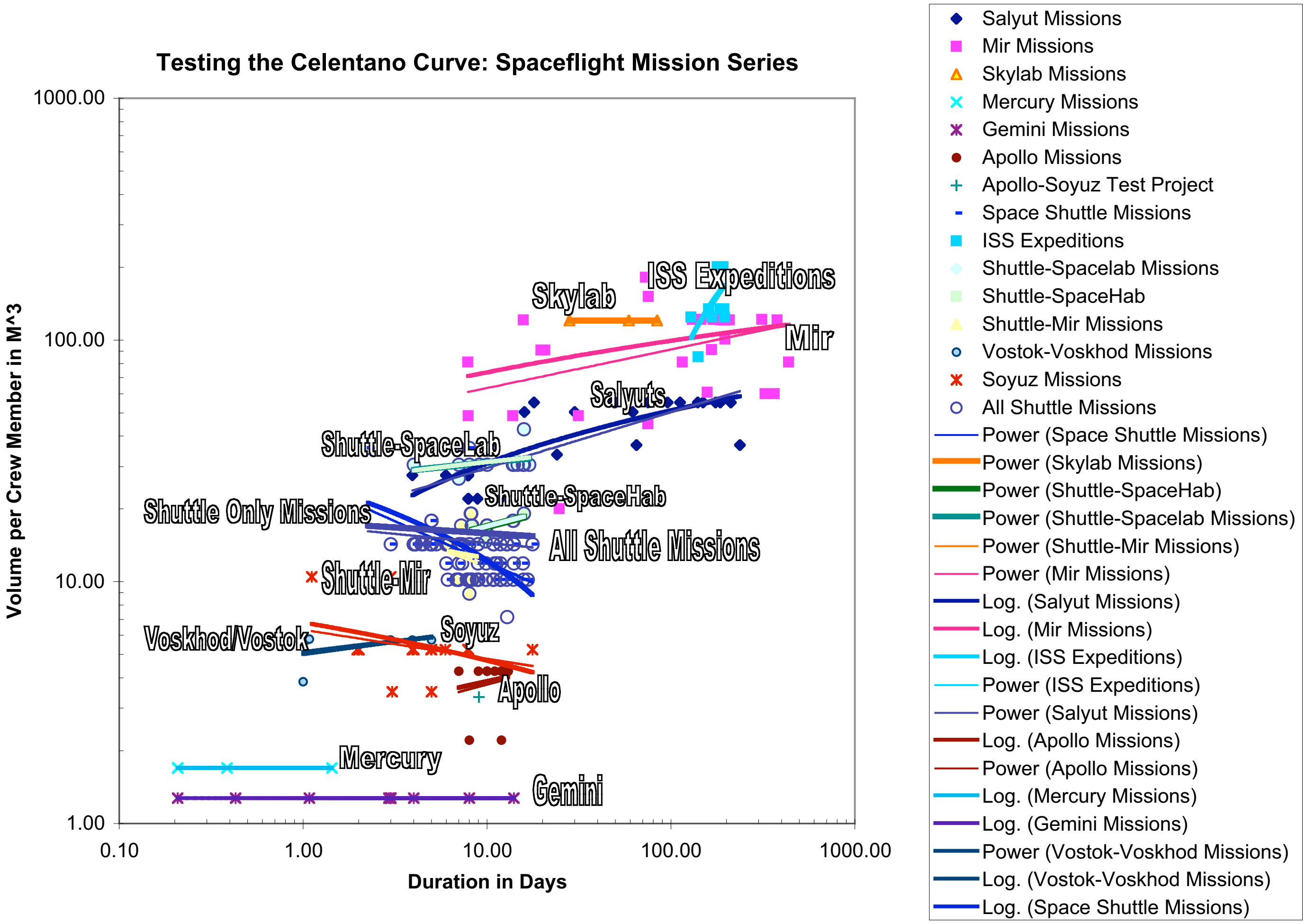

FIGURE 2 shows the complete historic spaceflight data set for all spaceflights through the period of this study in 2006, omitting the outliers of Mercury and Vostok test flights for t<0.1 days.

FIGURE 2 shows the record of human spaceflight by spacecraft. How far has the space habitability community progressed since Celentano? Christopher S. Allen, et al. (2003, p. 49) offered this stark assessment:

> There is currently no method available to determine with absolute certainty, the amount of habitable space needed per crewmember for missions beyond LEO. *Until better data is available, designers should plan on allocating a minimum of 16.99m^3 (600ft^3) of usable space per crewmember* (original emphasis. Allen, et al, 2000, p.47).

Researchers increasingly are calling the Celentano habitability index into question. Most recently, Marianne Rudisill, a psychologist in the Space Mission Analysis Branch at LaRC offers this critique:

On reviewing the original Celentano, et al. paper, here is a summary of their experimental method:

"…They did a good job in describing the multiple factors that impact living space and 'habitability'; however, my primary concern with their work is that their 'habitability index' was based on studies done with very few subjects under controlled laboratory conditions for very short durations from which they extrapolated to multiple months. In particular, they based their habitability index on studies including the following conditions:

Cabin A**:* living volume = 200 cu ft, living space = 39 cu feet, 13 sq ft/man, 3 subjects, 7 days' duration (at, essentially, bed rest) = ***Tolerable.

***Cabin B**: living volume = 1500 cu ft, living space = 150 sq ft, 37 sq ft/man, 4 subjects, 7 days' duration (at sedentary activity level) = **Performance**.*

Cabin C**: living volume = 1600 cu ft, living space = 400 sq ft, 200 sq ft/man, 2 subjects, 4 days' duration (at average office worker activity level) **= Optimal.

So, yes, their work was done in gravity, but under conditions that were different to the extreme from the long durations on the lunar surface and it leads me to question its generalizability…"

This is what I was referring to in my talk [at the 2008 ASCE Earth and Space Conference, March 4, 2008]. I have found that many people quote Celentano, et al. (because the curve they generated is in MSIS and it's the ONLY reference to volume in MSIS) without having read the original paper and, therefore, without being aware of the conditions in their study; as an experimental psychologist, I find their conditions very short in duration (the maximum was 7 days and they extrapolated to several months), very controlled, and with very few subjects (a total of 9 subjects across 3 conditions where there should have been at least 10 subjects in each of the 3 conditions), ***which leads me to question the generalizability of their findings*** [emphasis added; e-mail from Marianne Rudisill, April 7, 2008 published by her permission].

This meta-analysis is essential to appreciate what Celentano et al actually did to create their "Habitability Index." Assuming the design of a spacecraft is never frivolous, the volumes represent what the designers consider at least minimal to accomplish the mission. For example, in the design studies for the Orion Crew Exploration Vehicle, one can see how it is possible to suffer crowding from too little volume. FIGURE 3 gives a glimpse of an early concept for the Orion with a crowded cabin that expresses the criticality of simple volume.

RESEARCH DESIGN

The research design treats the Celentano curve and its variants as hypotheses about the relationship between mission duration and volume. The elements to distinguish such a hypothesis consist of:

- A theory of causation for spacecraft pressurized volume; e.g., mission duration drives volume per crewmember.
- Mission duration as the independent variable.
- Pressurized volume as the dependent variable.
- A quantifiable relationship between independent and dependent variables.
- Specific properties of the plot of the dependent variable that relate back to the hypothesis.

FIGURE 3. Andrews Aerospace 10 Crew CEV Mockup for NASA-Johnson Space Center, Courtesy of Andrews Aerospace Corp.

COLLECTING THE DATA

The starting point for this study is simply to collect the data for three key parameters:

1. Defining pressurized volume for each spacecraft,
2. Stipulating the mission duration in days, and
3. Identifying the number of crewmembers in the spacecraft for each mission.

TABLE 1 presents the summary of human spaceflight data collected and considered for this study.

TABLE 1. Summary of the Human Spaceflight Data Set as of July 18, 2006.

Spacecraft Type	Category	Number of missions	Max. Mission Duration Days	Min. Mission Duration Days	Max. Volume Per Crew m^3	Min. Volume Per Crew m^3	Max. Crew	Min. Crew
Mercury	Capsule	6	1.43	0.02	1.70	1.70	1	1
Gemini	Capsule	10	14.00	0.21	1.28	1.28	2	1
Apollo CM with and w/o LM	Capsule	11	12.75	6.00	4.27	2.22	3	3
Apollo LM	Lander	7	3.21	1.00	3.33	3.33	2	2
Apollo-Soyuz	Capsule	1	9.04	9.04	3.33	3.33	5	5
Vostok	Capsule	6	5.00	0.07	5.73	5.73	1	1
Voskhod	Capsule	2	1.08	1.00	2.87	1.91	3	2
Soyuz	Capsule	42	14.00	0.43	1.28	1.28	2	2
Shenzhou	Capsule	2	5.00	1.00	17.00	8.50	2	1
Space Shuttle	Shuttle	89	17.67	2.25	35.75	8.94	8	2
Shuttle-Spacelab/SpaceHab	Shuttle	25	16.90	4.00	42.70	14.66	8	5
Skylab	Station	3	84.00	28.00	120.33	120.33	3	3
Salyut	Station	17	237.00	16.00	55.25	33.50	3	2
Mir	Station	25	437.75	72.82	181.35	45.00	3	2
ISS	Station	12	195.82	128.86	201.13	85.17	3	2

Defining the Spacecraft and its Volume -- The first decision point was how to define spacecraft and their missions. Although this task may seem self-evident, it is far from obvious. For example, the Apollo Command Module (CM) flew in five mission/configurations: Apollo orbital missions, Apollo lunar missions without and with the Lunar Module (LM), Apollo-Skylab, and Apollo-Soyuz. In the first four cases, there were three astronauts; Apollo-Soyuz had also the two cosmonauts. In each case, the volume varied, with the last three adding the volume of the LM, Skylab, and Soyuz. The Skylab volume was so much greater than the CM that the reasonable approach is to treat Skylab as its own volume, with the docked CM as a secondary volume.

Stipulating the Mission Duration -- The second decision was the mission duration. For missions where the crew launches and lands in the same spacecraft, this stipulation is simple because the space agency records the mission elapsed time (MET) to the second. However, when a crew rotates through a space station, it becomes a matter for analysis and judgment.

- Does the mission extend for the entire time that the space station is continuously inhabited?
- Does the mission duration include the time from launch to docking or from undocking to landing?
- How does the mission change when one crewmember arrives to replace another who then returns to earth?

For clarity, this survey uses the responsible space agency's stipulation. This way implies that, for example, ISS mission duration means the time that an *Expedition* occupies the station, and does not include the transportation time in a Shuttle or Soyuz.

Identifying the Crew Size –- Crew size can vary for post-Apollo and post-Vostok spacecraft. For example, Soyuz flies with two or three crewmembers. The Shuttle flies with two to eight. The CEV will fly with two to six crewmembers. There are instances of partial crew rotation on space stations that necessitate an educated judgment about what was the crew size. Given a spacecraft volume, ***changing the number of crew changes the volume per crewmember.***

47 MAXIMA-UNIQUE DATA POINTS

The pivotal research design decision was to identify the unique maxima values for each combination of crew and volume in each spacecraft. Rather than use the entire 254 data points, many of which (e.g. shuttle flights) are nearly identical and produce a dense "square plot", this analysis selects one data point for the maxima of each combination of crew size and a spacecraft configuration for a total of 47. This maxima-unique approach means that there is one data point for a shuttle mission with a single SpaceHab module for a given crew size and another data point for a shuttle/single SpaceHab mission with a different number of crew. Similarly, there would be two separate maxima for a Shuttle flights with the same number of crew, but one with a single SpaceHab module and the other with a double SpaceHab module.

QUESTIONS FOR THE HYPOTHESES

In approaching these hypotheses, the research design poses a set of questions as a basis for testing them.

1. Has the evolution of spacecraft from Vostok and Mercury to the International Space Station followed the path predicted by the Celentano Curve?

2. Specifically, does the volume prediction follow a curve that rises then levels out at six months?

3. Which curve pattern best fits the data under each hypothesis?

4. Can we evaluate this "best fit" by the R^2 value, or do we need to test for correlation significance among the curves?

5. Does this curve pass through the origin or otherwise show no minimum value?

6. How does the aggregation or disaggregation of the data affect the results?

7. Is there any substantiation in our years of human spaceflight for the guidelines for pressurized volume in terms of tolerable, performance, and optimal limits or levels?

HYPOTHESIS TESTING

The research design postulates a null hypothesis against which to compare all the Celentano curve variations as alternate hypotheses, including the original curve.

<u>The Null Hypothesis, H_0</u> -- H_0 always states that there is no relationship between the hypothesized independent variable and dependent variable; there is no relationship – no effect – between mission duration and volume.

<u>The Alternate Hypothesis, H_n</u> -- H_n states that there is an effect between the independent the dependent variables; mission duration affects pressurized volume.

THE MISSION DURATION HYPOTHESES

This section presents the dozen hypotheses, explicating in detail the first one, Celentano (1963), and the second, Fraser (1966), to demonstrate the methodology. TABLE 2 shows the hypotheses in chronological order.

<u>H_1 Celentano, Amorelli, Freeman (1963)</u> – FIGURE 1 shows this famous graph. The minimum occurs at about 45 ft^3 (1.28m^3), for the Gemini spacecraft. They use the term "living space" as "volume per-man requirements." This graph follows a discussion of breathable atmospheres, so it is clear that they mean total volume. Their principal assertion states that this volume requirement increases over mission duration to an upper limit after six months.

This accommodation occurs within three "levels:"

a) A "tolerable" level of about 5.6 m^3 (~200 ft^3),
b) A "performance" level of about 10.6 m^3 (~375 ft^3), and
c) An "optimal" level of about 19 or 20 m^3 (which the authors state at 700 ft^3).

To summarize the parts of the Celentano Hypothesis H_1:

- H_{1a}: There is a minimum value of approximately 1.25m^3 pressurized volumes per crewmember.
- H_{1b}: The curve levels off beyond a duration of six months; no more volume is required.
- H_{1c}: There are three levels of volumetric habitability, defined as tolerable, performance, and optimal.
- H_{1d}: The optimal curve for volume requirements levels off at maxima of about 20m^3.

<u>H_2 Fraser (1966)</u> – T. H. Fraser portrays three zones of impairment to the crew: *no impairment, detectable impairment,* and *marked impairment* in FIGURE 4. Fraser built his data from habitat analog and simulator studies as curves on a two-axis logarithmic graph. Two years later, he published another version of the graph, but logarithmic only on the X-axis for time as shown in FIGURE 5. In *LIVING ALOFT* (1985, p 61), Connors, Harrison, and Akin discuss Fraser's contributions:

> Fraser (1968a) evaluated the results of 60 confinement studies to determine at what point physiological or psychological impairment occurred which was related to spatial restriction. He found that impairment (which he defined as the demarcation between "no impairment" and "marked impairment") occurred at between 50 ft^3 [1.42 m^3] for very brief confinement, and 150 ft^3 [4.25 m^3] for 60-day confinement.

TABLE 2. Summary of Null and Alternate Hypothesis Features:

	Volume Description/ Crewmember	Log Scale	Type of Curve	Minimum Value	Number of Curves	Base on Empirical Data?	Based on Study Predictions?
MISSION DURATION DRIVES VOLUME HYPOTHESES							
H_0 No Effect	Pressurized Volume	No	None	No	None	No	No – Null Hypothesis
H_1 Celentano, Amorelli, Freeman 1963	"Living" Press. Volume, 3 Limits: *Tolerable, Performance, Optimal.*	No	Quasi-log, flattens to upper limit	~1.28m^3	3	Partially	Limited data
H_2 Fraser 1966, 1968	Pressurized Volume	Yes	Zones of Points	~25-30 ft^3 = ~.71-.86 m^3	3, not H_1	Yes	Partially
H_3 Manned Space Center, 1966	Habitable Living Volume/Man	No	Quasi-log	Yes	2 with a zone between	Yes	Partially
H_4 Marton 1971; MSIS, 1987, 1995; Woolford, Bond, 1999;	Habitable Volume (Undefined)	No	Same as H_1	No. Passes through 0.	3, same as H_1	Same as H_1	Same as H_1
H_5 Gore, Martin, Trust 1978	Puts Mission Duration Limits on Celentano	No	Quasi-log that does not flatten	No, varies with criteria curve.	3 plus 8 limit slopes	Crew stowage Req'ts.	Space Shuttle Studies.
H_6 Sherwood, Capps, 1990.	Pressurized Volume	Yes	Curve rises.	~1m^3	2	Mostly	Partially
H_7 Petro; Perino; Kennedy; Rudisill 1999-2008,	Pressurized Volume	Yes	Straight line rises	~1m^3	1	Mostly	Partially
H_8 Sforza 2004	Free & Habitable Volume(undefined)	Yes	Power S-curve rises	40 ft^3 = 1.13m^3	2	Partially	Partially
H_9Hofstetter, de Weck, Crawley 2005.	Habitable & Pressurized Volume	No	Same as H_1, but part fit with poly.	No. Passes through 0.	Show 3 but test only 1.	Modeling MSIS	No, relies on MSIS.
CREW SIZE DRIVES VOLUME HYPOTHESES							
H_{10} Davenport, Congdon, Pierce, 1966.	Minimum Volume / Man	No	4 Rising Straight Lines	1.43 m^3, increases with crew	4	Not Clear	Yes
H_{11} Reynerson, 2004	Facility Volume	No	Straight line rises	No	1	No	Yes
H_{12} Kennedy, Toups, Smitherman, 2008	Three limits: *tolerable, performance, preferred.*	No	3 Straight lines rising	No	3	Unclear	Unclear
MISSION FUNCTION REQUIREMENTS DRIVE VOLUME HYPOTHESIS							
H_{13} Schwartz, 2005	Pressurized and Habitable Volume	No	N/A	~1.25 m3,	2	Yes	Yes

He concludes that a volume of 250-700 ft^3/person [7.08-19.82m^3] [depending upon] length of confinement, is adequate [annotations based upon personal conversation with Mary Connors, NASA-Ames Research Center, July 14, 2006].

Fraser offers his variations on the Celentano Hypothesis:

- H_{2a}: The levels of tolerance of confinement and acceptable crew cabins occur as zones (rather than as precise curves).
- H_{2b}: These zones are:
  - H_{2b1} No Impairment,
  - H_{2b2} Detectable Impairment, and
  - H_{2b3} Marked Impairment.
- H_{2c}: There is a minimum required volume of about 0.7 m^3 (25 ft^3) for a space cabin, associated with the bottom limit for "Marked Impairment."

Fraser's minimal volume of 0.7m^3 is impossibly small for a spacecraft, and *marked impairment* would be an unsurprising consequence.

H_3. Manned Space Center (1966) -- The staff at the newly constructed Manned Space Center in Houston prepared a version of the Celentano curve in FIGURE 6. The zones of habitability between upper and lower bounds are consistent with Fraser's (1966) concept, which they cited. Unfortunately, the anonymous authors did not define these upper and lower bounds. One inconsistency is that MSC shows the Gemini capsule as having more volume per crewmember than Mercury capsule, when the opposite was true. Their choice of analog environments is equally curious. They chose the Triton, a conventionally powered submarine, unspecified nuclear submarines, and a 1957 Antarctic Sleeping Area. The only spacecraft they predict beyond Apollo is a Space Station that falls below the lower bound.

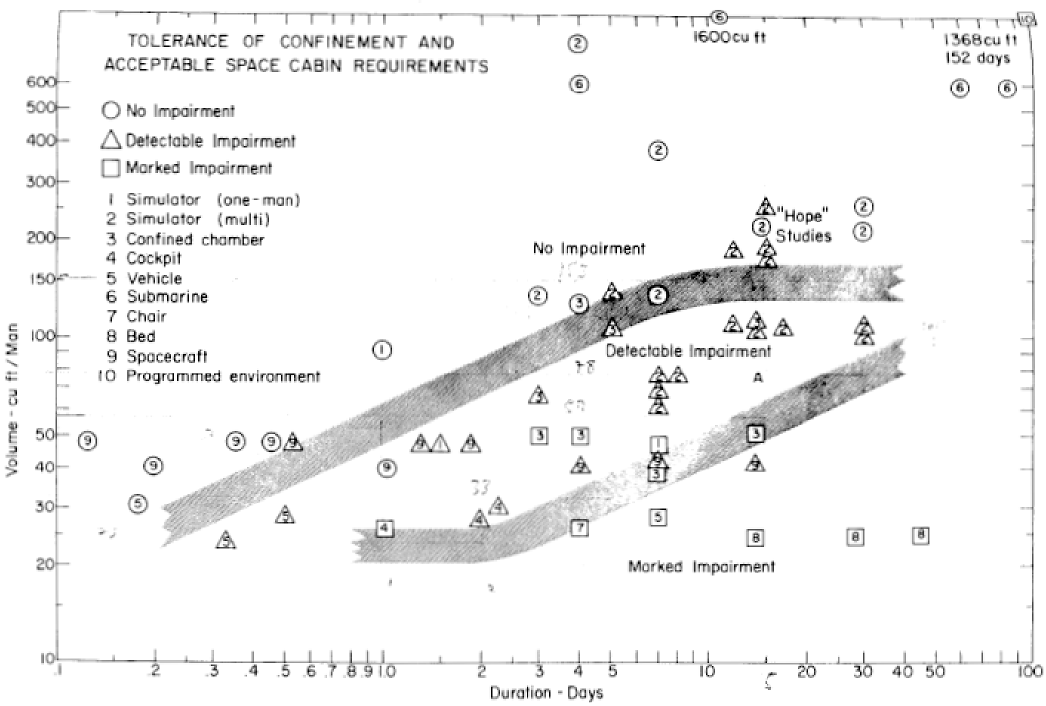

Figure 8. Tolerance of Confinement and Acceptable Space Cabin Requirements.

FIGURE 4. Fraser's 1966 plot of No *impairment, Detectable Impairment,* and *Marked Impairment* from confined volumes from NASA CR-511.

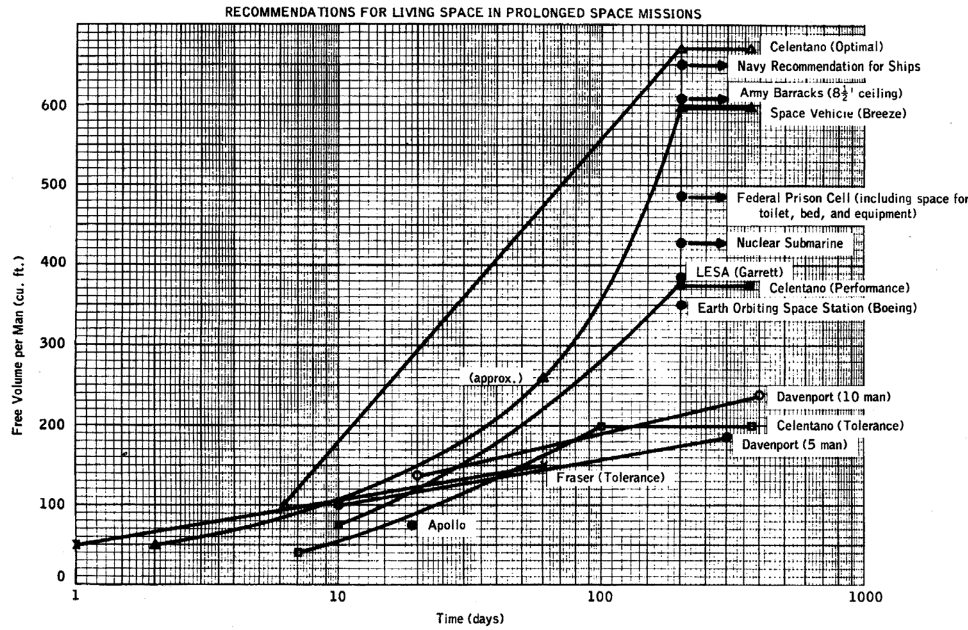

Figure 1

FIGURE 5. T. M. Fraser's 1968 "half-logarithmic" plot of *Free Volume* versus Time (NASA CR-1084, p. 2)

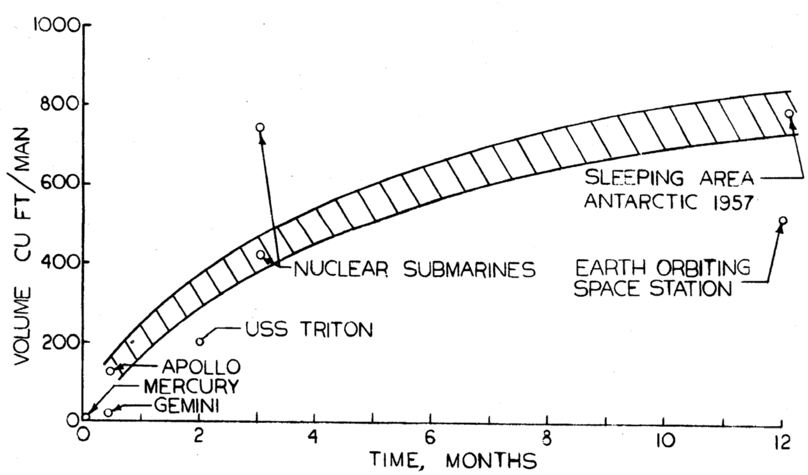

FIGURE 2.3 TOTAL HABITABLE LIVING VOLUME

FIGURE 6. Manned Space Center (1966) curve suggesting a habitable zone between two Celentano-like curves.

H_4 Marton et al., (1971) etc. – In 1971, Marton, Rudek, Miller, and Norman published a version of the Celentano Curve but placed the minimum value at the origin -- zero -- making a *kind* of arithmetic logic:

Zero mission duration = zero volume = zero crew.

NASA reproduced the Marton et al plot in the first two editions of the *Manned Systems Integration Standard*

(MSIS), NASA STD-3000 (1987, 1995), in FIGURE 8. Later, Barbara Woolford and Robert Bond in the Human Factors organization at JSC, who were major contributors to MSIS, published this chart in Wiley, Pranke, Eds (1999). The Y-axis bears the label "habitable volume." However, the lack of a definition of "habitable volume" has led to widely varying interpretations. As Marianne Rudisill notes, this MSIS version is the most widely cited in the literature.

What this analysis can test is whether this curve passes through zero, its slope, and whether it levels out horizontally at any of the three value limits for tolerable ($5m^3$), performance ($10m^3$), and optimal ($20m^3$). However, it cannot test the hazy distinction between pressurized and habitable volume.

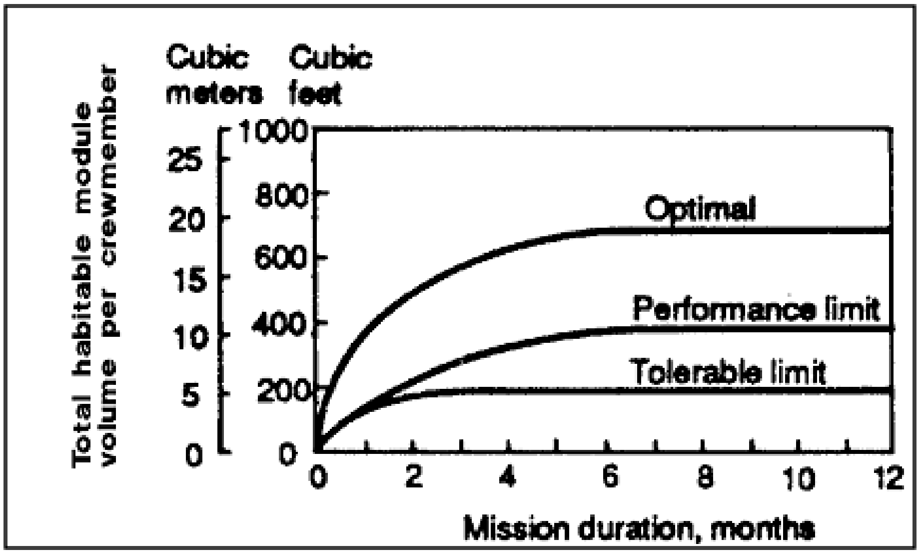

FIGURE 7. Marton, Rudek, Miller, Norman (1971), as reproduced in NASA Std. 3000, Man-System Integration Standard, Figure 8.6.2.1-1. Guideline for determination of total habitable volume per person in the space module.

H_5 Gore, Martin, and Trust: Celentano with Duration Limits (1978) -- During the period of developing the Space Shuttle, Gore, Martin, and Trust of Rockwell International applied the Celentano curve to planning Shuttle missions. In FIGURE 8, they recognize that available volume could impose mission duration limits, anticipating Spacelab and SpaceHab. They defined eight duration limits as lines of negative slope, based upon their estimates of delta down weight, subsystem capability and extension requirements, and the stowage volume a crew of four would need above what Celentano indicates. The curves resemble MSC's. Gore et al retain the terms *Optimum* and *Performance* for the upper curves, but label the lower curve *Minimum*. In the legend, they equate the *Optimum* curve to a *Maximum Requirement* and the *Performance* curve to a *Median Requirement*, (they do not imply a statistical meaning). They apply Celentano with reservations:

> For a crew of four, habitability [sic] or "free" volume does not become a limiting factor until durations of about 90 days, assuming the Celentano "performance" criteria provide such a measure. Furthermore, orbiter free volume improves well before the Celentano "performance" threshold by the needed addition to stowage volume (Gore et al, 1978, p. 9).

Gore, Martin, and Trust's goal was to prepare for the Extended Duration Orbiter (EDO) project. NASA planned EDO in at least two phases, first to extend the original seven to ten day orbiter missions to the range of 16 to 20 days, and second to 30 to 32 days (also called the 30-day orbiter). These two extensions appear in FIGURE 8 as the AL-20 and the AH-30 lines that intersect the "Optimum" curve. The EDO project involved much more than augmenting the Shuttle cabin volume; "The heart of the EDO is the cryogenic pallet," (Saucier, 1992, p. 1). The cryogenic pallet added LH2 and LOX to extend the life support, power, and propulsion systems capabilities.

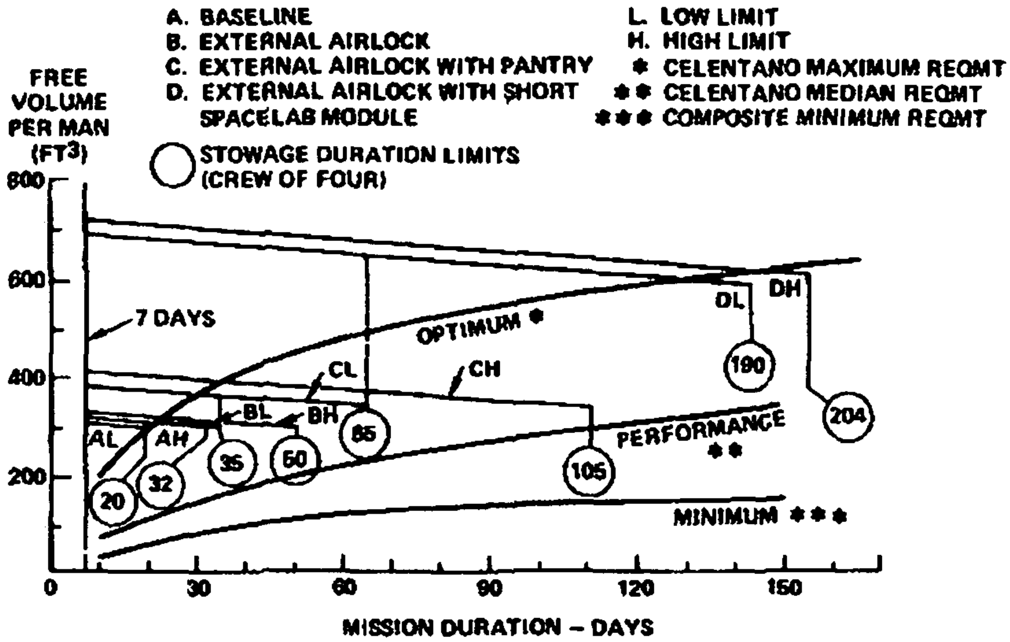

Figure 25. Duration Limits as Function of Habitability Requirements (Crew of Four)

FIGURE 8. Gore, Martin, and Trust's (1978) overlay of mission duration limits upon the Celentano curves.

<u>H_6 Sherwood, Capps (1990)</u> – Brent Sherwood and Stephen Capps of Boeing Space and Electronic Systems in Huntsville, AL performed the first analysis of spacecraft volume that distinguished among different spacecraft types. Their most significant finding was to identify two distinct populations for pressurized volume in the human spaceflight data: launch and landing capsules defined by aerothermal geometry and "other habitable spacecraft" including the Apollo Lunar Module (LM), the Space Shuttle, and all space stations. In parsing the data, they identified the parameters for the types of spacecraft, individually and in various combinations.

FIGURE 9 shows Sherwood and Capps' graph, with one curve for Capsules and the other for all "other habitable vehicles." Only the curve for Space Shuttles and Stations evokes the shape of the Celentano curve. Brent Sherwood, a Space Architect, who is now head of the JPL Strategic Planning & Project Formulation Office recalls:

> Let me tell you how it really happened when I did the hab analysis. I assembled data on prior designs from easily available sources, which inevitably meant total pressurized volume, as you note elsewhere. When I plotted the data, it struck me (just visually) that one way to make sense out of the messy *(i.e., not highly correlated) data at the low end* was to divide it into two classes. I convinced myself that this made sense because atmospheric entry vehicles (not the LM, which is why I didn't include it in that set) have an overriding geometry constraint due to shape that orbital systems don't have (e.g., the LM). Cones are horrible for packaging efficiency.
>
> So I reasoned that the overwhelming nature of the aeronautical shape constraint would contaminate any thoughts about how much space people needed...besides, who cares about how much space they need when they're only in it for a few hours? Even I can stand flying United overseas, because although I

hate it, I am only trapped for a few hours. The real issue is orbital space, so I wanted to eliminate the distraction of the atmospheric entry systems. There was no deeper statistical reasoning than that. I still believe my logic is valid...CEV might care about the lower curve, but nobody else should (emphasis added, e-mail from Brent Sherwood, March 19, 2008). The challenge of testing the Sherwood and Capps' hypothesis will be to determine if the data supports the above observation of "not highly-correlated data at the low end."

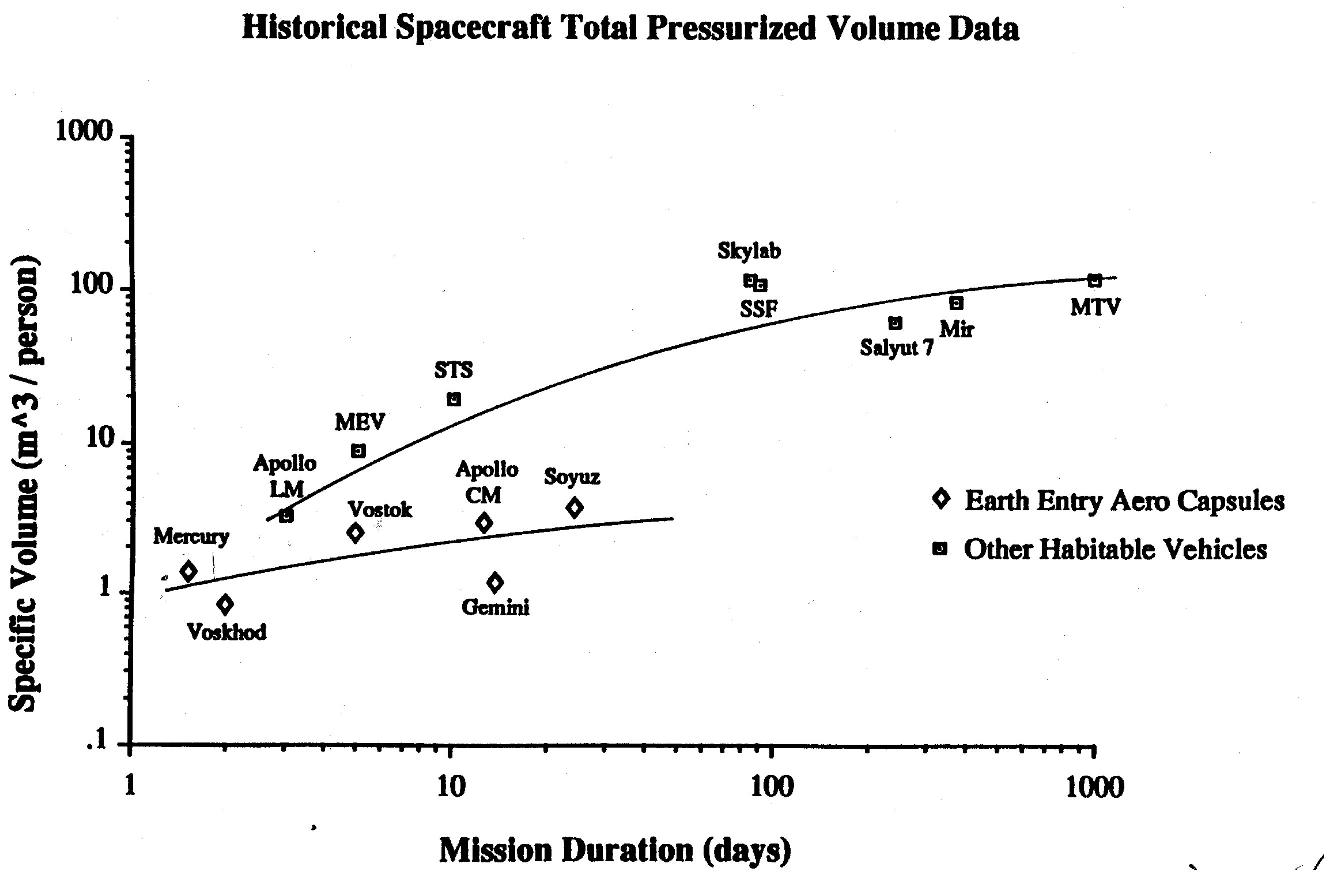

FIGURE 9. Sherwood and Capps, 1990, separation of two curves that distinguish aero-entry capsules and all other "habitable vehicles" (courtesy of Brent Sherwood).

H_7 Petro, (1999), Perino (2005), Kennedy (2006), Rudisill et al (2008). – In the same anthology in which Woolford and Bond reproduced the MSIS chart, Petro (1999) published a different plot of the historical data. He shows the curve as a single straight power curve of positive slope on a logarithmic scale (it is reasonable for data that forms a curve in a linear scale to appear as a straight line on a logarithmic scale). FIGURE 10 shows Rudisill et al's (2008, p. 4) most recent version of this plot. Rudisill et al explain their approach:

> Dimension and volumetric data from past and present spacecraft (both US and Russian vehicles) were gathered and evaluated. Spacecraft vary across a number of parameters relevant to volume estimation, such as era of development, crew size, and mission duration. In addition, all spacecraft operate in a microgravity environment (other than the Apollo Lunar Module, the only vehicle for which we have crew operations data from the lunar surface, albeit for very short durations, on the order of three to four days), while we were deriving estimates for a 1/6th g environment. However, gathering and comparing spacecraft provided a broad view of volumes of built and operated vehicles.

We identified another relevant factor: mission "type." That is, all spacecraft evaluated were found to group into either of two categories, *"transportation-like"* or *"station-like"*; understandably, vehicles used primarily to "ferry" crews to a destination serve a rather different function from those designed primarily for long-duration crew operations. This grouping of vehicle "type" can be seen in [FIGURE 10], showing total pressurized volume as a function of mission duration (note that the predicted maximum lunar outpost mission duration of 180 days is indicated on the X-axis as a reference).

> Given that we were estimating volume required for a lunar surface habitat serving as an "outpost," we focused our assessment on "station-like" spacecraft, given the long duration nature of these missions (original emphasis, (original emphasis, Rudisill et al, pp. 3-4).

In making this distinction between *transportation-like* and *station-like* vehicles, Rudisill et al align with Sherwood and Capps' 1990 findings, although they include the Space Shuttle and LM with the capsules.

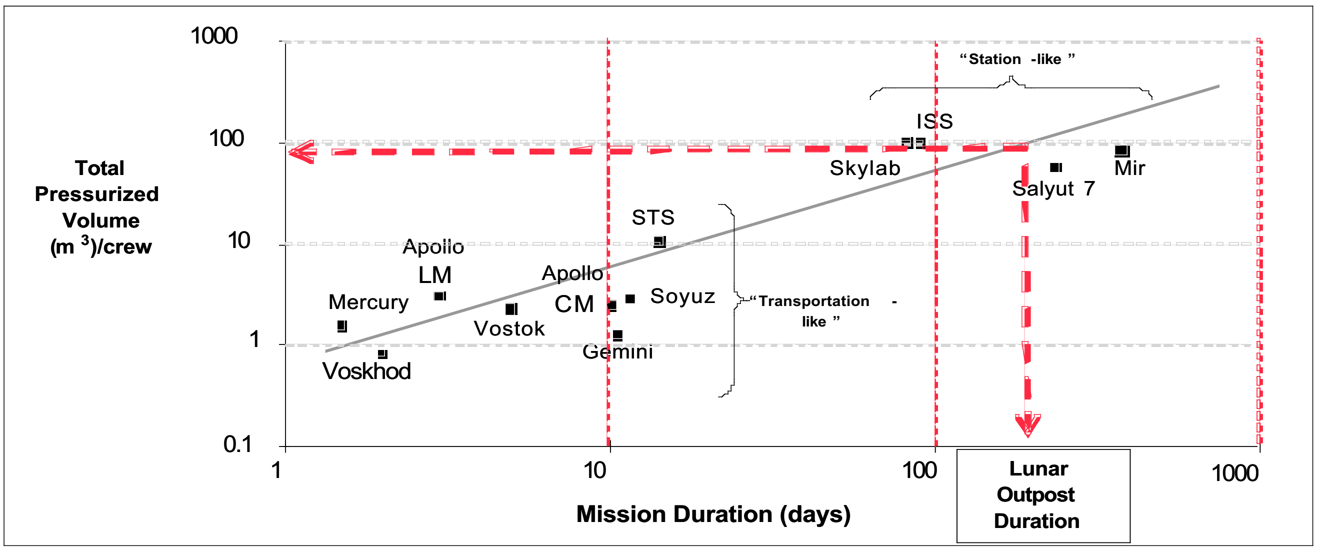

FIGURE 10. Rudisill et al's (2008) version of the straight-line power curve on a logarithmic scale, after Petro (1999); Kennedy; and Perino. (Courtesy of Marianne Rudisill, NASA-Langley Research Center).

H_8 Sforza (2004) – Prof. Pasquale Sforza at the University of Florida published on his website a version of the Celentano curve with several new features in FIGURE 11. He introduces the term "free volume" that he seems to use interchangeably with "habitable volume." He sets the minimum volume limit by putting a reverse curve at the bottom creating an overall "S-curve" effect.

In Sforza's "S-curve," the flat curves at the top and the bottom represent the minima and maxima of pressurized volume per person. The middle section is a power curve from four to 180 days. Sforza explains his approach:

> The number of crewmembers and mission duration are basic specifications that strongly influence the vehicle configuration. The habitable volume required for each crewmember may be estimated by applying the Celentano volume criterion (ref.), which is an s-shaped, "learning curve" shown in [FIGURE 11].
>
> [FIGURE 11] suggests that a doubling of the Celentano values is more representative of current experience. . . . The International Space Station will have 15,000 cubic feet ($425m^3$) of habitable

volume but a crew complement of only six, yielding 2500 ft^3/person (70.8m^3/person), much more than even double the Celentano criterion. Note that 100 ft^3 corresponds to a box roughly 4ft by 4ft square and 6 ft high.

Sforza states, "the Celentano criterion curve may be approximated in a piece-wise fashion according to the following equations," in EQUATION 1.

EQUATION 1

$$0 < t < 4days : v_{free} = 40 ft^3 / person$$

$$[1.13m^3 / person]$$

$$4 < t < 180days : v_{free} = 20t^{0.58} ft^3 / person$$

$$t > 180days : v_{free} = 400 ft^3 / person$$

$$[11.3^3 m / person]$$

Sforza states his definition of "Free Volume" and offers two curves or limits – one roughly replicating the "Celentano volume criterion" (although he does not specify which one), and another that tracks the Kliper, Shenzhou, and Space Shuttle at their maximum crew size.

The analysis can test whether the curve starts level at the lower end as Sforza suggests, runs straight, and levels out near the top. Sforza's minimum for all spacecraft is the unprecedentedly small 40 ft^3 (1.13 m^3). What he does is create the upper curve by moving it to the left along the timeline so that where 1.13 was sufficient for less than 4 days in a capsule on the lower curve; it is acceptable for only one day in a spacecraft in the upper curve.

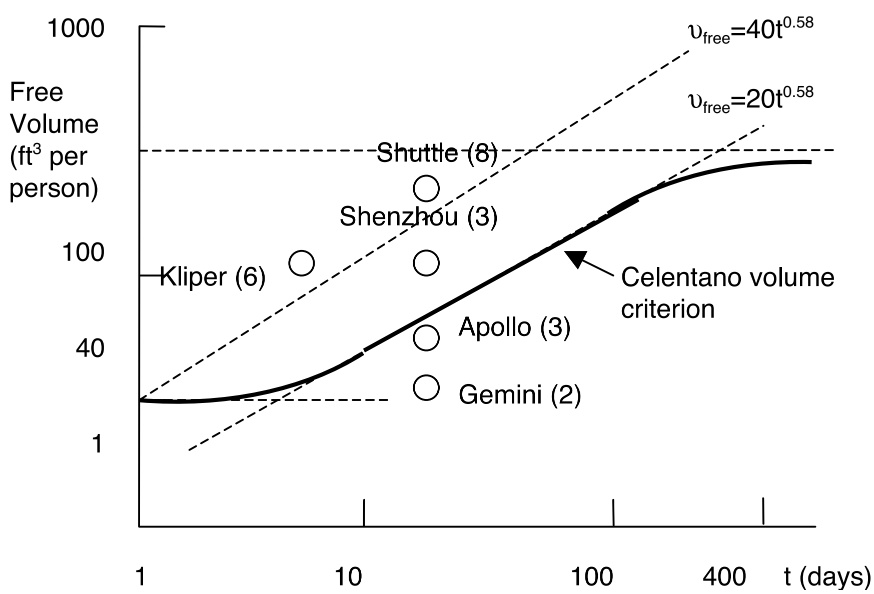

FIGURE 11. Sforza's interpretation of Celentano and his proposal for a doubled requirements curve.

H_9 Hofstetter, de Weck, Crawley (2005) – These MIT professors took the MSIS chart and constructed a curve from the origin to a maximum plateau at 270 days that produces an "analytical interpolation" that falls below the Celentano optimal criterion in FIGURE12. They drew this curve from EQUATION 2a (Hofstetter et al, 2005, p. 4).

EQUATION 2a

$$V_{Habitable}(\Delta t_{mission} \leq 270d) =$$
$$= 19m^3 * N_{Crew} * \left(1 - \left(\frac{\Delta t_{Mission} - 270d}{270d}\right)^4\right)$$

Where, N_{crew} is the number of the crew, $t_{mission}$ is the mission duration in days, and d is days. They propose a multiplier to size pressurized volume in EQUATION 2b.

EQUATION 2b

$$V_{\text{Pressurized}} = 3 * V_{\text{Habitable}}$$

While this method of estimating may fall in the ballpark, the design and engineering challenges of calculating spacecraft size are far more complex. Yet, Hoffstetter De Weck and Crawley do not shy away from complexity:

> There are several approaches to compute the necessary pressurized volume; here, a polynomial of fourth order shall be used to estimate the habitable volume required as a function of mission duration. . . . For mission durations longer than 270 the habitable volume is assumed to stay constant at about 19 m^3 per crewmember (Hofstetter et al, 2005, p. 4).

CREW SIZE HYPOTHESES

This survey found three versions of an alternate hypothesis that ***crew size is the primary driver of volume per crewmember.*** Connors, Harrison, and Akin (1985, p. 162) discuss this issue with regard to *whether having more people allows the vehicle to have smaller volume per person* (e.g., the transition from Mercury to Gemini). They cite T.M. Fraser and conclude that the number of crewmembers is unproven as a driver of volume per person. However, three sets of authors argue the opposite in the Crew Size Hypothesis.

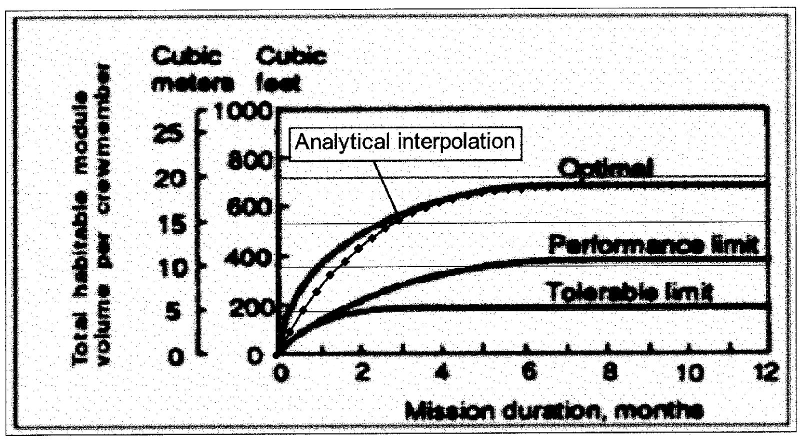

FIGURE 12. Hofstetter, de Weck, and Crawley's "analytical interpolation" of a partial logarithmic trend line function below the Celentano curve from MSIS, NASA STD 3000.

H_{10} Davenport, Congden, Pierce (1963) Crew Size – Davenport et al proposed "preliminary requirements for crew volume versus space mission duration," in which they posited crews of four sizes: 1, 3, 5, and 10 crewmembers. They present their recommendations as a series of lines – one for each crew size in FIGURE 13. Each line of increasing crew size appears progressively higher and steeper on the graph.

H_{11} Reynerson (2005): Volume function of crew size – Charles Reynerson of Boeing presented criteria for mission duration of 180 days, in FIGURE 14, that are about par for ISS and is moderate compared to historical Salyut and Mir missions. However, the proposed number of crew and the facility volume is much greater – by up to 1.5 orders of magnitude.

- Crew Number: 0 - 100
- Endurance: 0 - 180 days
- Facility Weight is Most Sensitive to Endurance

Reynerson argues that mass scales primarily with mission duration while volume scales more closely with crew number.

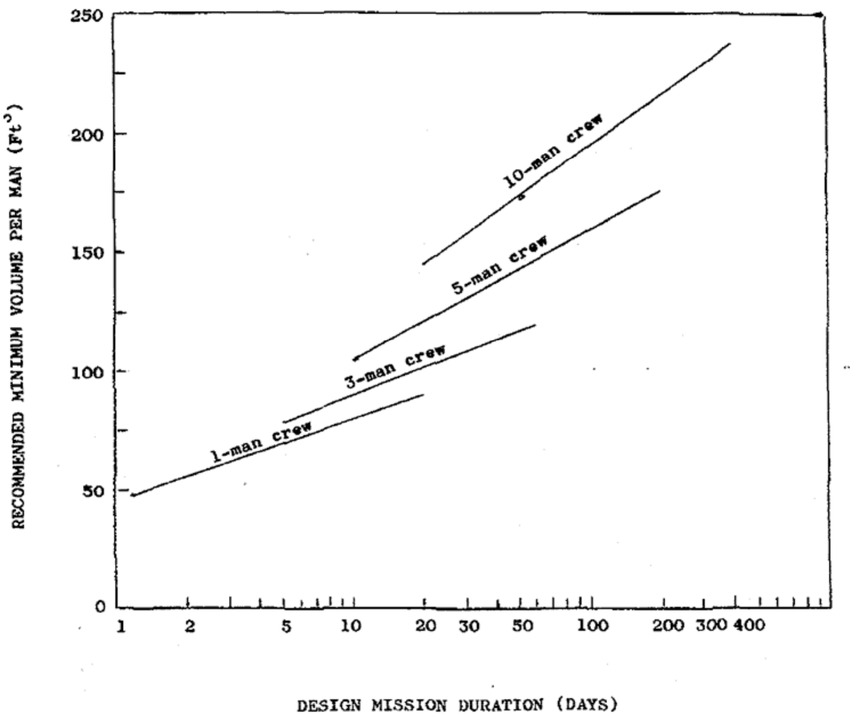

FIGURE 13. Davenport, Congden, Pierce (1963). “Preliminary requirements for crew volume versus mission duration.”

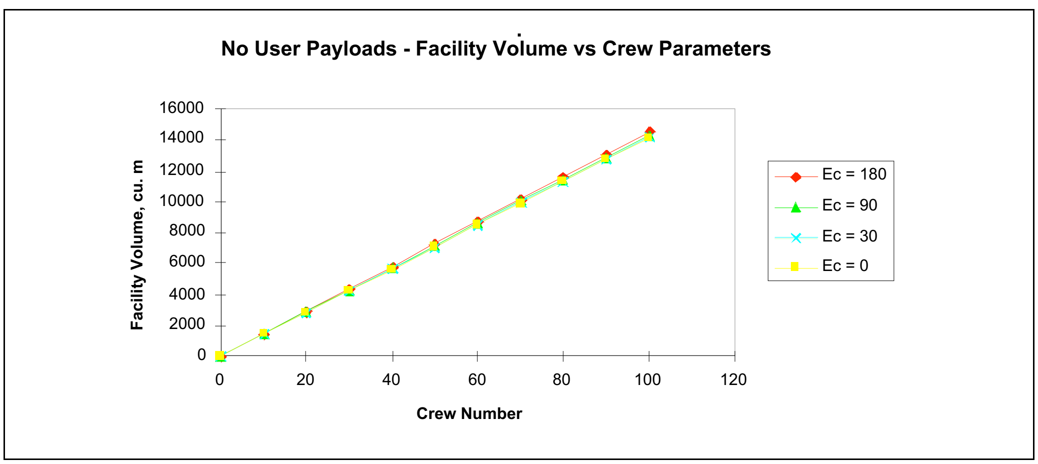

FIGURE14. Reynerson’s plot of a linear relationship between number of crew and pressurized volume of a spacecraft.

H_{12} Kennedy, Toups, Smitherman (2008) Three Limits – The Space Architects Kriss Kennedy and Larry Toups at JSC and David Smitherman at MSFC framed the crew size hypothesis in Celentano’s terminology (Kennedy et al, 2008, p.2). In FIGURE 15, they plot three curves of tolerable, performance, and *preferred* criteria as straight lines of positive slope, in a continuum from one to six crewmembers. All Kennedy, Toups, and Smitherman’s curves give larger volumes than the corresponding criteria in Celentano or MSIS, peaking at 30m^3. They explain their approach:

> The gross pressurized volume required for space habitats can be estimated based on historical data about human space exploration and remote environments on earth . . . . A first order parametric volume estimation based on crew size and mission duration gives the designer a starting point for the space habitation system. Historical data combined with ISS data show the habitation volumes divided into three categories; minimum tolerable limits, minimum performance limits, and preferred limits . . . . These rules-of-thumb are applicable for medium duration missions. Short

duration missions will be roughly analogous to the Shuttle, and for long duration missions there is little data available to make a determination. When determining the initial volume required, one should consider a parametric range of volumes based on the mission objectives and requirements (Kennedy, Toups, Smitherman, 2008, p. 2),

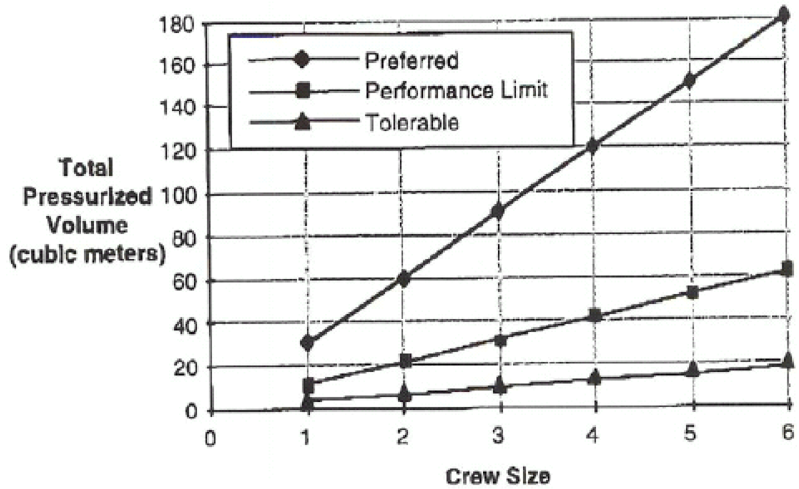

FIGURE 15. Kennedy, Toups, & Smitherman's adaptation of Celentano, Amorelli, and Freeman's three limits to the crew size as the explanatory variable for volume.

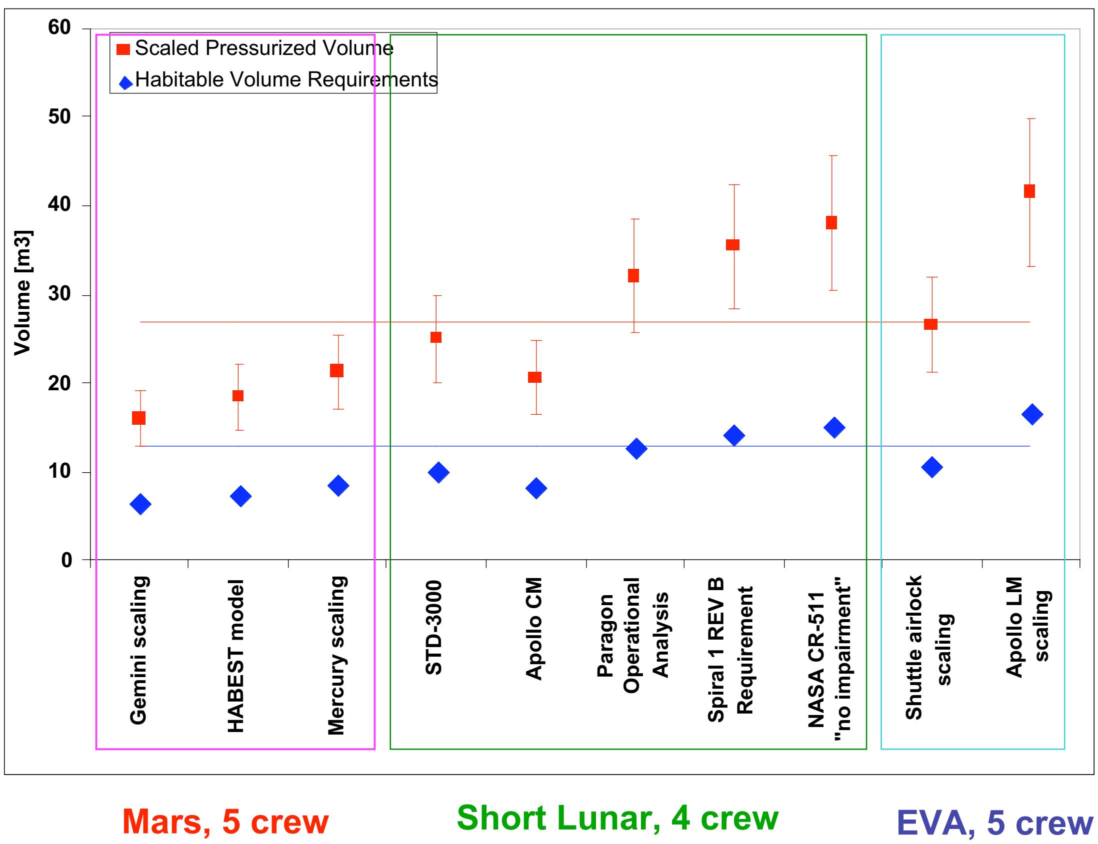

FIGURE 16. Jana Schwartz (Raytheon, 2005) Multi-ConOps Operational Analysis.

Kennedy et al describe their three mission durations:

Short duration: For a few days to a couple of weeks, crews can share personal quarters by rotating shifts, as is done when the Space Shuttle carries Spacelab . . .

Medium duration: For up to six months, crews require their own private personal quarters for sleeping as well as private recreation (reading and communication with relatives), and will require more room for grooming and personal hygiene . . .

Long Duration: For mission durations of six months or more, crews require all the necessary "comforts of home." Each crewmember will need a private sleeping area with personal storage, a dressing area, and a sitting area. More generous recreational and exercise facilities will be required as well as a complete health maintenance facility (Kennedy et al, 2008, p. 2). Functional Operations Hypothesis

H_{13} Schwartz (2005) – In the Draper Lab's Concept Exploration and Refinement Study for the CEV, Jana Schwartz tries to bridge between the minimum "habitable volume" and the total pressurized volume. She argues that function drives volume more than mission duration her graph in FIGURE 16 shows the pressurized volumes multiplied from habitable volume within a confidence interval for specific missions. Most fascinating is her attempt to correlate prediction methods for volume sizing to specific missions and crew sizes. Although Schwartz's argument is incomplete, it has a compelling aspect: to tie pressurized volume to the purposes for which the crew uses it in conducting a space mission.

TABLE 3. Descriptive Statistics for 254 Human Spaceflight Missions

Days		*Volume Per Crew Member*	
Mean	41.64	Mean	31.48
Standard Error	4.73	Standard Error	2.76
Median	8.90	Median	11.92
Mode	5.00	Mode	14.30
Standard Deviation	75.35	Standard Deviation	43.96
Sample Variance	5,677.50	Sample Variance	1,932.15
Kurtosis	6.37	Kurtosis	3.45
Skewness	2.47	Skewness	2.03
Range	437.68	Range	199.85
Minimum	0.07	Minimum	1.28
Maximum	437.75	Maximum	201.13
Sum	10,575.65	Sum	7,996.55
Count	254.00	Count	254.00
Confidence Level (95.0%)	9.31	Confidence Level (95.0%)	5.43

HYPOTHESIS TESTING

The research design takes the parts of each hypothesis, and plots their salient characteristics within the spaceflight data set to test each hypothesis. FIGURE 17 represents the maxima-unique data points. The research design then evaluates the threats to validity.

DESCRIPTIVE STATISTICS

TABLE 3 shows the data for 254 human spaceflights until mid-2006. The mean duration was 42 days and the mean volume was 31.4 m^3. These means fall very far above the medians of 8.9 days and 11.9m^3. The standard deviations (SD) are huge, especially for mission duration at 75 days. Thus, the SD for duration is about an order of magnitude larger than the fraction of the range from the median to the minimum of 0.07 days. This observation agrees with Sherwood and Capps, and Rudisill et al for the distinction between the capsules that fall below the median and space stations above it.

COEFFICIENT OF DETERMINATION

Given non-random data such as mission duration and volume, one of the most effective tools is the coefficient of determination, known as the R-squared value (R^2). The National Institute of Standards (NIST) defines R^2 as:

> A statistic for a predictive model's lack of fit using the data from which the model was derived (NIST, accessed 8 April 2008).

EQUATION 3 gives the mathematical definition of R^2, where SSreg is the Sum of the Squares of the *Regression to the Mean;* SSerr is the Sum of the Squares of the *Errors of the mean*; and SStot is the *Total* Sum of the Squares.

EQUATION 3: Coefficient of Determination

$$R^2 = 1 - \frac{SSerr}{SStot} = \frac{SSreg}{SStot}$$

$$= 1 - \frac{\text{UNEXPLAINED Variance}}{\text{Total Variance}}$$

$$= \frac{\text{EXPLAINED Variance}}{\text{Total Variance}}$$

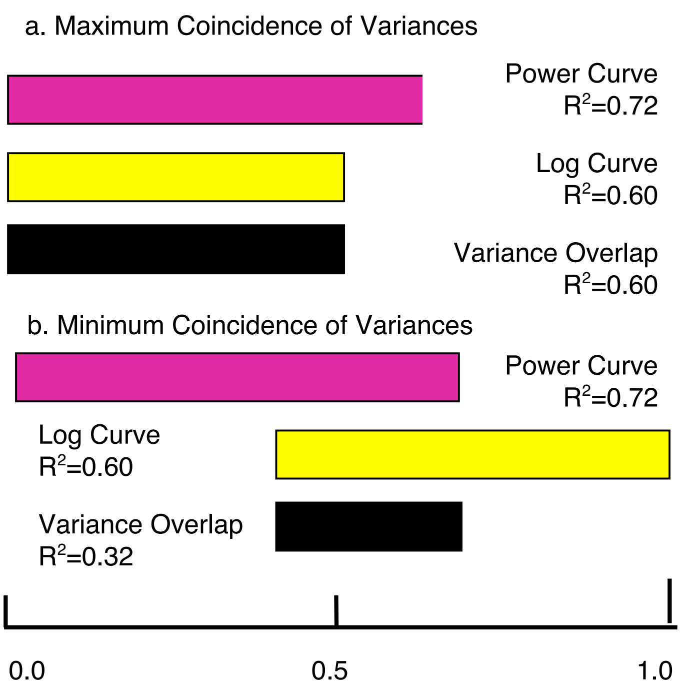

DIAGRAM 1. R^2 as a representation of differing portions of the variance in the Power Curve (Purple) and the Log Curve (Yellow). Overlapping of variances is black.

R^2 gives a measure of the variance in the dependent variable Y as a function of variance in the independent variable X, stated as a percentage. X is the *explanatory* variable and Y is the *explained* variable. When plotting the variance of Y as a function of the variance in X, the R^2 value gives the probability that the curve Y represents the variance the data. X may be the explanatory variable but is not necessarily the causative variable.

There is a rule of thumb that an R^2 value of more than 50 percent indicates an effect or relationship between the variance in X and the variance in Y. R^2 does not provide a test for significance in the way inferential statistics can using random data. Instead, estimating significance depends upon the magnitude – relative or absolute -- of the R^2 s. Another rule of thumb is if Y_1 and Y_2 are within 10 percent of one another, they represent the same variance. EQUATION 4 explains:

EQUATION 4: Differences in Y.

$$D = R^2(Y_1) - R^2(Y_2)$$

The approach to evaluating Difference in Y (D) is:

$D \leq 0.1$ means that Y_1 and Y_2 are identical.

$D > 0.1$ means that Y_1 and Y_2 diverge as D increases.

When D becomes much larger than 10% and one of the Ys is greater than 50%, it suggests that, compared to the lower Y, the effect of the higher Y may be significant.

TESTING THE MISSION DURATION HYPOTHESES

The tests of the mission duration hypotheses derive from the FIGURE 17 representation of the human spaceflights. FIGURE 17 portrays both the natural log curve in yellow and the power curve in purple for the same data. Although the simplest approach is just to declare the power curve the “winner” because of its larger $R^2 = 0.72$ compared to the log curve’s $R^2 = 0.60$, it does not explain adequately what the data reveal.

When dealing with R^2 values there is a tendency to regard the percentages of variance on an integer scale where a larger percentage of variance would contain ***all*** of a smaller percentage of variance. While this result is possible, it is not strictly correct because the two variances must overlap only by the amount shown in EQUATION 5.

EQUATION 5

$$Minimum \quad Overlap = R^2(Y_1) + R^2(Y_2) - 1$$

The Log and Power curves in FIGURE 17 return R^2 values that, at 12 percent difference, are close enough to consider them almost ***identical in magnitude but not necessarily in variance.*** DIAGRAM 1 illustrates the potential overlaps of variance in the R^2 values in FIGURE 17. The colors of the bars correspond to the respective power and natural log curves. The two R^2 values for Y may occur as a complete overlap of variance as in bar a, so that Y_2 includes all of the variance in Y_1.

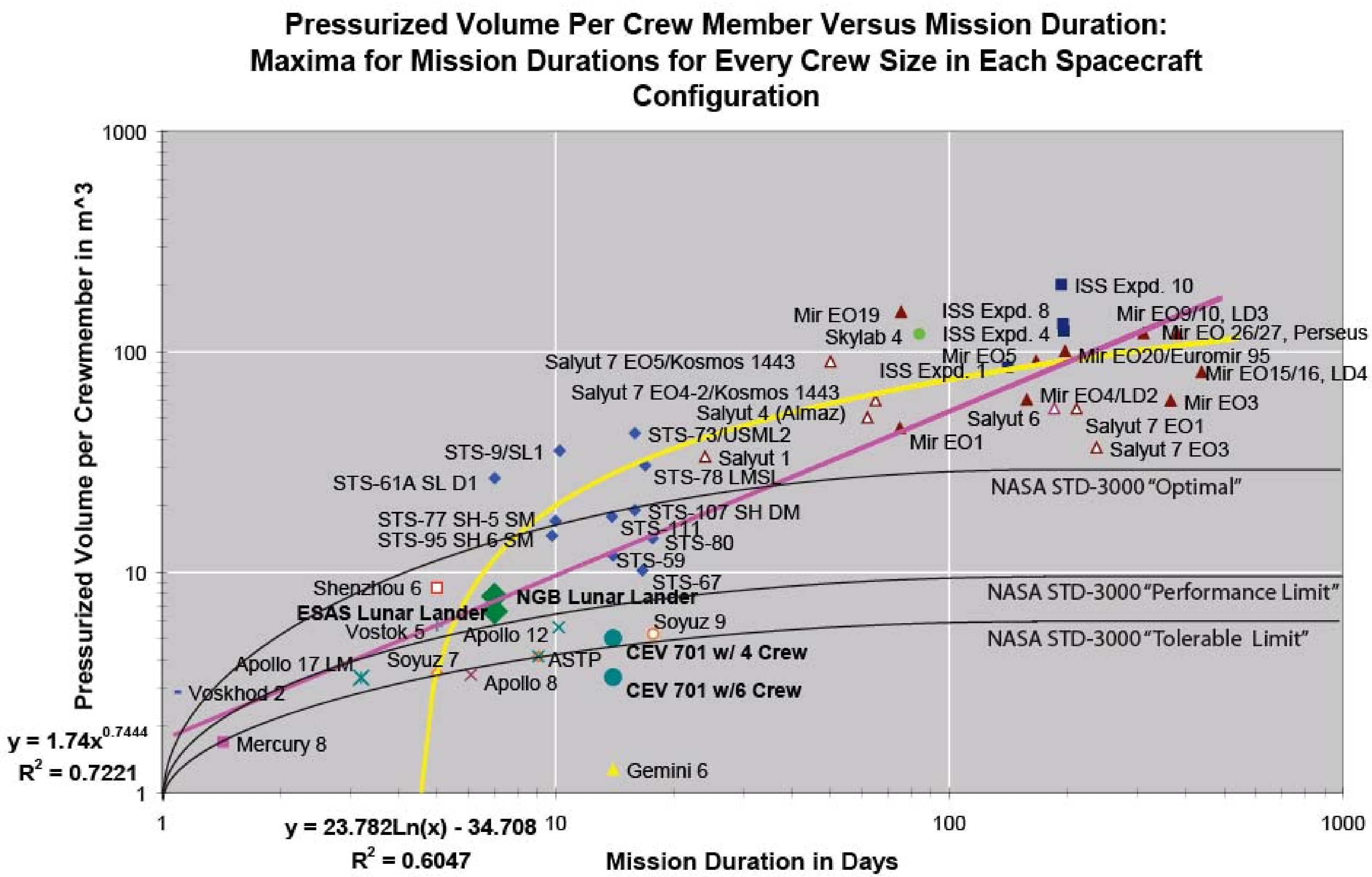

FIGURE 17. Maxima-Unique Data Points for the Historical Human Spaceflight Dataset overlay upon the Marton et al, NASA MSIS, and Woolford & Bond curves

TABLE 4a. Summary of Hypothesis Testing Results: H_0 means no effect. Rejecting H_0 means there is an effect.

Authors	Year	Alternate Hypothesis H_n	Reject H_o	Remarks
		MISSION DURATION HYPOTHESES		
H_1 Celentano, Amorelli, Freeman	1963	Pressurized volume requirement increases as a function of mission duration.	Reject H_o	R^2=0.72 power, =0.60 log.
H_{1a}		Minimum volume of about 1.25 m^3/ crewmember.	Reject H_o	(Approx. Gemini).
H_{1b}		Curve levels off flat after certain duration.		No basis.
H_{1c}		Maximum volume needed of about $20m^3$		No basis.
H_{1d}		Three levels of *Optimal, Performance, and Tolerable* volumes for crew requirements		No Evidence that it exceeds 7 days.
H_2 Fraser	1966	Zones of impairment on a log graph.		
H_{2a1}		"No Impairment Zone trendline".	Reject H_o	R^2=0.68 power
H_{2a2}		"Detectable Impairment Zone trendline."	Reject H_o	R^2=0.59 power
H_{2a3}		Marked Impairment Zone trendline.		R^2=0.05 power
H_{2b}		Minimum Volume defined as below the minima of the Marked Impairment Zone.		
H_3 MSC (JSC)	1966	Press. volume increases with mission duration.	Same H_1.	
H_{3a}		Minimum Volume/crew of about 2.8 m^3	Reject H_o	(~ Apollo CM).
H_{3b}		Slope of the curve gradually lessens.	Reject H_o	Like a log.
H_{3c}		Volume requirement occurs in a band between upper and lower bounds.		No evidence presented.
H_4 Marton; MSIS; Woolford, Bond,	1971	Press. volume increases with mission duration.	Same H_1.	
H_{4a}		No minimum volume		
H_{4b}		Trendline passes through the origin		
H_{4c}		Curve levels off flat after 6 months.		Same as H_{1b}
H_{4d}		Maximum optimal volume of about $20m^3$		Same as H_{1c}
H_{4e}		Three levels: *Optimal, Performance, and Tolerable*		Same as $H_{1d.}$
H_5 Gore.	1978	Press. volume increases with mission duration.	Same H_1.	
H_{5a}		Volume curves keep rising and do not level off.	Reject H_o	
H_{5b}		Shuttle-specific volume limit on mission duration.	Partially Reject H_0	See text
H_6 Sherwood	1990	Press. volume increases with mission duration.	Same H_1.	
H_{6a}		Earth Entry capsules have a different distribution than "other habitable vehicles."	Reject H_o	R^2=0.06 versus R^2=0.74*
H_{6b}		Low data correlation for capsules	Reject H_o	R^2=0.06
H_7 Petro; Perino; Kennedy, Rudisill.	1999-2008	Pressurized volume increases with mission duration.	Same as H_1.	
H_{7a}		Minimum volume of about 1.25 m^3/ crewmember	Same H_{1a}	(Approx. Gemini).
H_{7b}		Straight, positive power curve on a log scale.	Reject H_0	R^2=0.72 power

TABLE 4b. Summary of Hypothesis Testing Results: Rejecting H_0 means that there is an effect.

Authors	Year	Alternate Hypothesis H_n	Rej H_o	Remarks
MISSION DURATION HYPOTHESES, Cont.				
H_8 Sforza	2004	Minimum free volume of a1.13 m^3/ crewmember	**Reject H_0**	OK for capsules.
H_{8a}		Free volume increases as with mission duration.	**Same as H_1**	
H_{8b}		Straight positive power curve.	**Same H_{7b}**	
H_{8c}		"S-Curve" levels off top and bottom	**Partially Reject H_0**	Actuals not level, Equation low-end
H_{8d}		The true free volume requirement is about double the Celentano volume curve.		Insufficient data
H_{8e}		The Celentano volume curve is $v_{free}=20t^{0.58}$, while the true volume curve is $v_{free}=40t^{0.58}$		Insufficient data
H_{9a} Hofstetter	2005	The curve passes below MSIS' *Optimal* curve.	**Reject H_0.**	
H_{9b}		The curve between the minima (0) and the maxima of the MSIS ~19 to 20 m^3 is a 4^{th} order polynomial trendline.		Questionable Interpolation.
H_{9c}		The equation curve levels off at an upper bound.		
H_{9d}		The equation levels off at an upper bound of $19m^3$.		$57m^3$
CREW SIZE HYPOTHESES				
H_{10} Davenport, Congdon, Pierce	1966	Recommended minimum volume per man increases as a first order effect of crew size.		No basis. Data show no relationship.
H_{10a}		Minimum volume of ~1.43 m^3 per crew for 1 crewmember		Too small, even for Mercury ($1.70m^3$).
H_{10b}		Minimum volume of ~2.12 m^3 per crew for 3		Too small for Apollo or Soyuz
H_{10c}		Minimum volume of ~2.83 m^3 per crew for 5		Too small for Shuttle
H_{10d}		Minimum volume of ~4.25 m^3 per crew for 10		Surely too small.
H_{10e}		Volume is a second order effect of mission duration.		No evidence.
H_{11} Reynerson	2005	Volume scales as a linear function of crew size.		
H_{11a}		No minimum volume with zero crew.		Math correct but. . .
H_{12} Kennedy, Toups, Smitherman	2008	Volume scales as a linear function of crew size, on three limit curves		No Relationship, No Effect
H_{12a}		Three curves: *tolerable, performance, & preferred.*		No evidence.
H_{12b}		Three distinct and different mission durations: *short, medium, and long.*		No evidence.
FUNCTIONAL CON-OPS HYPOTHESIS				
H_{13} Schwartz	2005	Volume requirements correspond to mission type, which implies specific tasks.		Insufficient data
H_{13a}		Habitable volume requirement is a scaled fraction of the required pressurized volume.		Insufficient data

Alternatively, these same curves may entail a minimum overlap of variance as shown in bar b. Bar b implies that most of the variance in Y_1 is not included in Y_2. The two Ys would paint different portions of the total variance in the dependent variable.

Looking at the curves, the flatter part of the yellow Log curve comes close to the slope of the purple power curve. However, the steep portion of the left side of the yellow Log curve clearly is not contained in the variance of the Power curve. For this reason, the plot of the two curves for pressurized volume may resemble bar b more than it does bar a. EQUATION 6 gives the unexplained variance as the probability of 28 percent that the power curve does not represent the variance in Y:

EQUATION 6

$$\text{UNEXPLAINED} \quad \text{Variance}_{\text{Power}} =$$

$$1 - 0.72 = 0.28$$

Simultaneous with EQUATION 6, EQUATION 7 gives the unexplained variance as the probability of 40 percent that the log curve does not represent the variance in Y.

EQUATION 7

$$\text{UNEXPLAINED} \quad \text{Variance}_{\text{Log}} =$$

$$1 - 0.60 = 0.40$$

The ***unexplained*** *variance* in the Power curve of 28 percent is roughly half the ***explained*** *variance* of 60 percent in the Log curve. Thus, DIAGRAM 1b shows the black area of overlapping variance between the power and log curves, with a minimum overlap of 32 percent. Therefore, the difference of R^2 probability in representation of the same variance by the two curves ranges from a minimum of 32 percent to a maximum of 60 percent.

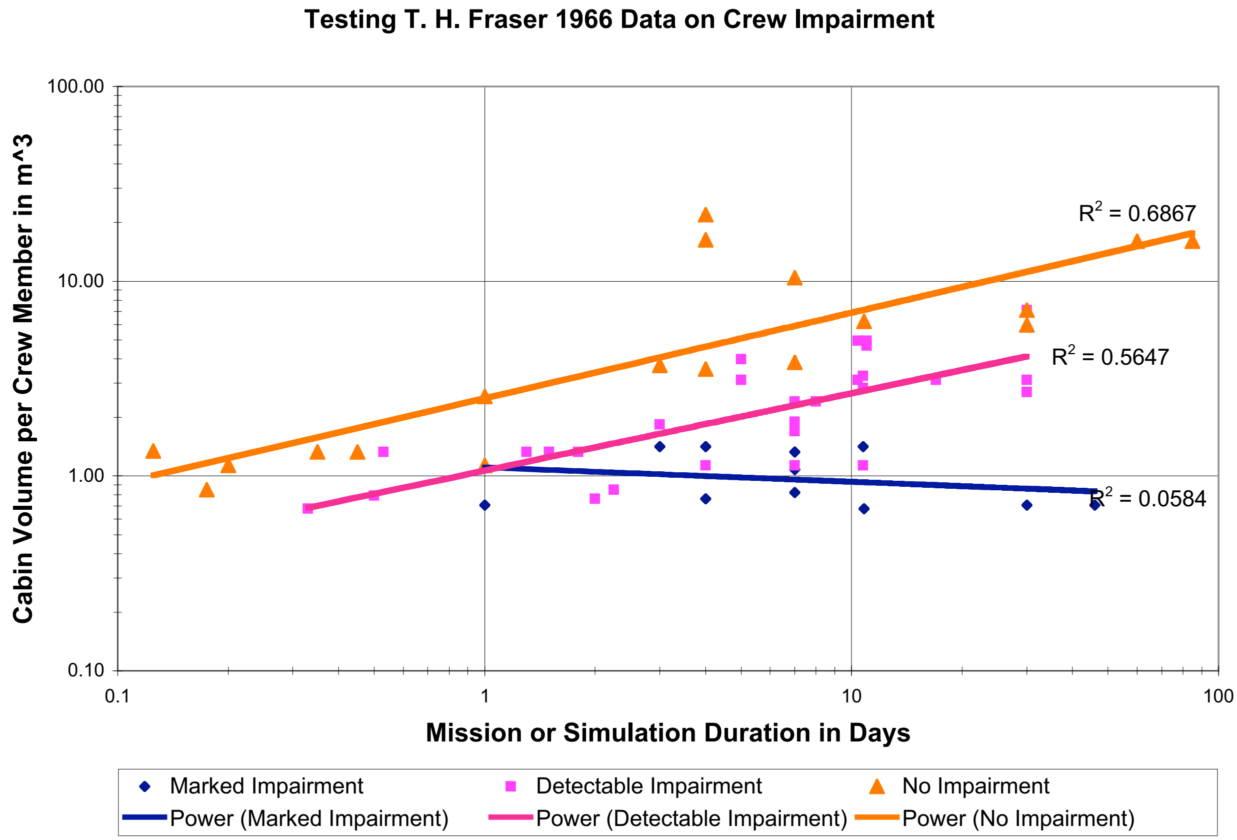

FIGURE 18. Fraser's (1966) Curves of No Impairment, Detectable Impairment, and Marked Impairment.

Thus, the two curves represent differing, but complementary views of reality. Envision the log curve that rises steeply from its minima as the 28 percent of the variance that the power curve leaves unexplained. Then imagine the start of the power curve above the origin and continuing in a straight diagonal as the 40 percent variance that the log curve leaves unexplained. The two curves overlap across a range of 32 to 60 percent probability that they represent the same part of the variance.

Testing H_1 Celentano, Amorelli, Freeman (1963) – The discussion above of Testing the Mission Duration Hypothesis answers all the questions for H_1. We reject the null hypothesis (H_0) for two parts of H_1 on the basis that there is a relationship between mission duration and pressurized volume and that there is a minimum volume of about 1.25m^3/crewmember. However, we fail to reject H_0 for the other three parts of this hypothesis.

Testing H_2 Fraser (1966, 1968) – This test plot of Fraser's data in FIGURE 18 came from a scan of his 1966 data points It was not possible to use the present survey of spacecraft data because very few of his data came from space craft. FIGURE 20 shows R^2 = 0.68 for *No Impairment* and R^2=0.56 for *Detectable Impairment*, suggesting that there may be a relationship between mission duration and volume for these two curves. However, *Marked Impairment*, R^2=0.05 shows no effect. We reject the H_0 for *No Impairment* and *Detectable Impairment* but not for *Marked Impairment.*

Testing H_3 Manned Space Center (1966) – For MSC, we reject H_0 for three of the four assertions. In addition to H_1, we find effects for the minimum volume/crew of about 2.8m^3, which approximates the Apollo Command Module, and for the gradual lessening of the slope of the curves as in the log curve in FIGURE 17. MSC does not provide evidence that the volume requirement would fall in a band between upper and lower bounds, and so we fail to reject H_0 for that part.

Testing H_4, Marton et al (1971); NASA MSIS (1987, 1995); Woolford & Bond (1999) – Except for the basic H_1 effect, we fail to reject H_0 for the other parts of this hypothesis. Neither curve in FIGURE 17 passes through the origin. At the 180 days, H_{4c} and H_{4d} claim the optimal curve levels out at maxima of 19 or 20m^3. Instead, FIGURE 17 shows the log and power curves crossing at 100m^3 – 5 times more than predicted. Finally, the results give no evidence for the optimal, performance, and tolerable criteria.

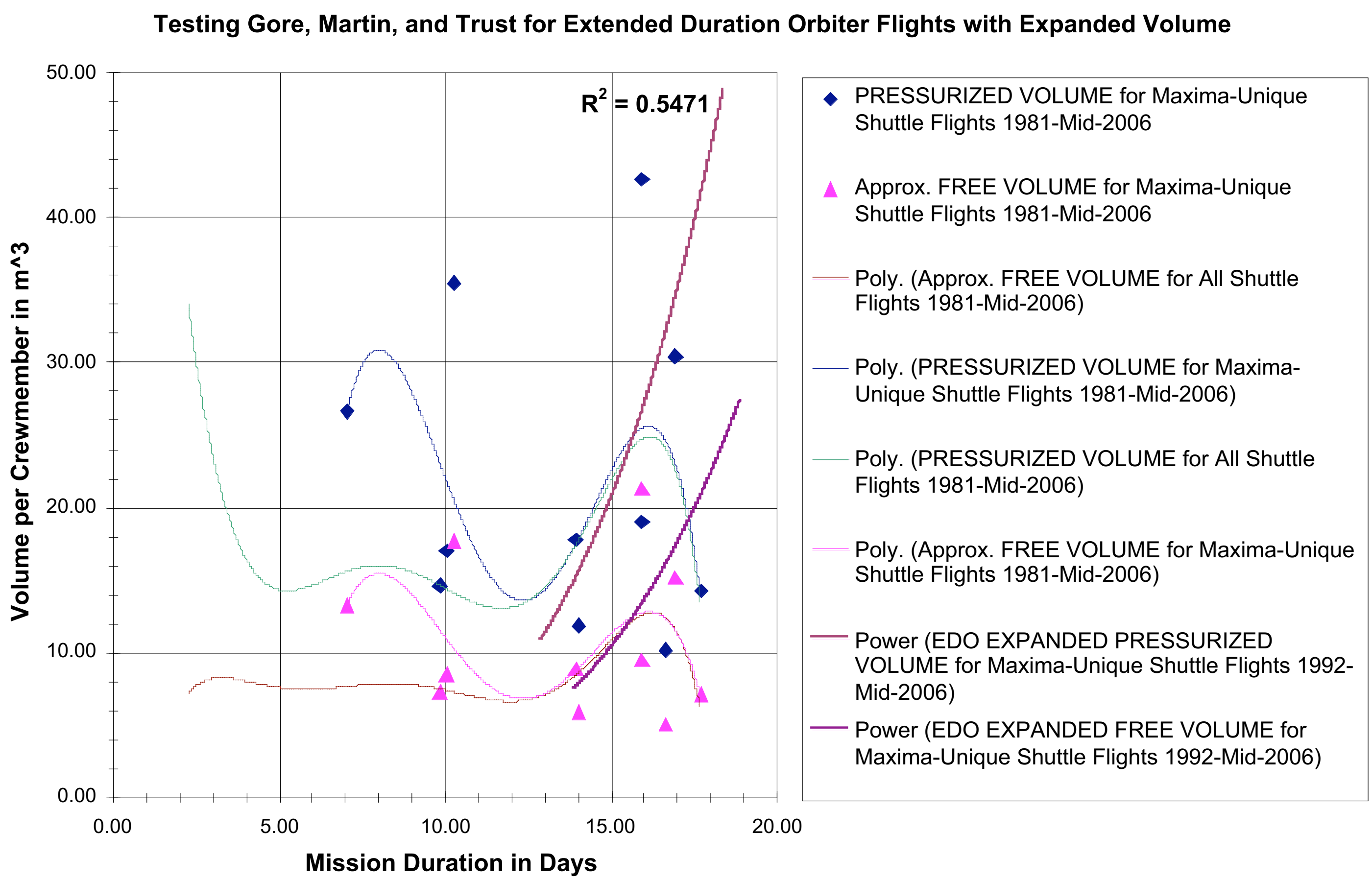

FIGURE 19. Test of Gore, Martin, and Trust for Space Shuttle Extended Duration Orbiter

Testing H_5 Gore, Martin, and Trust (1978) – We reject H_0 for the first two parts H_{5a} and H_{5b}, because they agree with the results of FIGURE 17. Gore et al addressed crews of four in terms of "free volume per man," but did not define free volume clearly. Their non-consideration of larger crews creates a difficulty because nearly all the

EDO flights involved crews of five to eight. FIGURE 19 shows a plot of the curves for shuttle flights, showing both the maxima-unique data points and all shuttle flights for actual pressurized volume and free volume estimated very roughly at 50 percent of pressurized volume. This graph shows two interesting effects. First, the maxima-unique sample and the total population begin to converge in the longer durations because that is where most of the shuttle maxima occur. Second, the part of the curves, corresponding to most of the EDO flights from 12 to 16 days follows an upward slope that appears to reflect what Gore et al intended in their chart on the "Optimum" curve. However, they did not anticipate EDO missions without expanded volume as the Columbia did for the longest flight, 17.6 days on STS-80, with five crew in the 71.5 m^3 orbiter cabin for an approximate "free" volume of 7.15m^3 per person.

Despite the fact that Gore, Martin, and Trust provide only three usable benchmarks, the seven day baseline, the 20 day EDO, and the "30-day Orbiter" that Rockwell International planned in detail but never completed, their negatively sloping duration limit lines reflect reasonably well the actual longer shuttle missions. These EDO missions included Shuttle-Spacelab and Shuttle-SpaceHab. FIGURE 19 shows the trendlines for the maxima-unique EDO flights with the actual *expanded pressurized volume* and the *estimated expanded "free" volume* that Gore, Martin, and Trust predicted. The power curve shows an effect at R^2 = 55 percent, partially supporting Gore, Martin, and Trust's assertions for H_{5c} therefore we reject $H_{0.}$ partially concerning the volume limitations on orbiter duration.

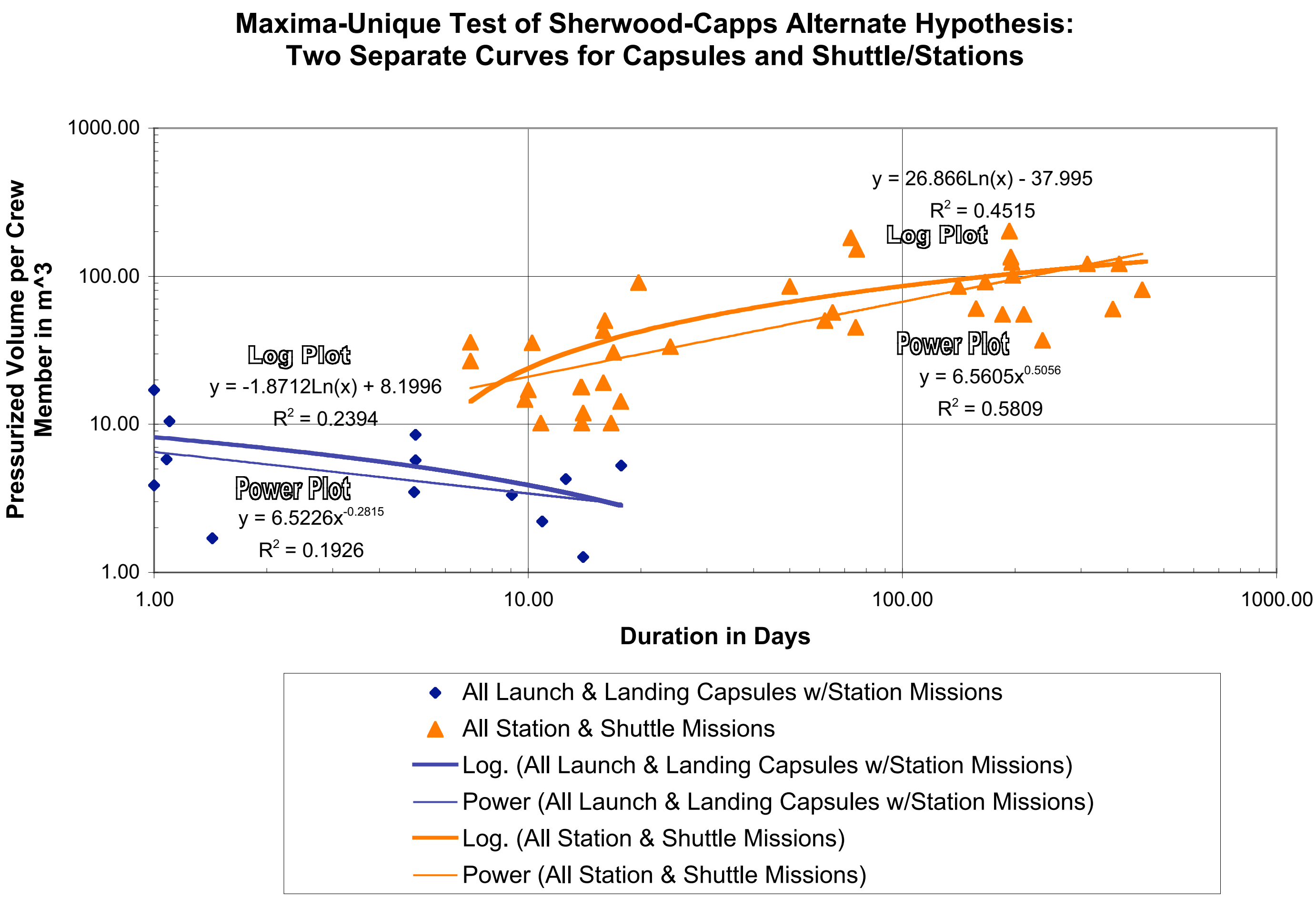

FIGURE 20. Test plot of Sherwood and Capps' dual curve hypothesis, using the empirical and historical data

<u>Testing H_6 Sherwood and Capps (1990)</u> – We reject H_0 for all the Sherwood and Capps hypotheses. TABLE 5 shows the results for an analysis based on their precept of separating spacecraft into two classes. TABLE 5 shows also three permutations of two spacecraft combined. Combining two or more widely different samples -- or types of spacecraft -- increased the R^2 by up to 50 percent, even for capsules and shuttles. Therefore, the data show that Sherwood and Capps relegated the low-scoring capsules to a second curve separate from the combined Space Stations and Shuttles. In the Effect Column, a YES means that the

results show an effect. In the Significance column (Sig), YES means that the effect may be significant.

TABLE 5. Test of Sherwood and Capps Demarcation of Spacecraft Data Sets for the Maxima-Unique Data Set

Set of Spacecraft	L(n) R^2	e^x R^2	Effect	Sig
Capsules	0.24	0.19	YES	
Space Shuttles	0.06	0.05		
Space Stations	0.06	0.03		
Capsules & Shuttles	0.08	0.09		
Shuttles & Stations	0.47	0.61	YES	YES
Capsules & Stations	0.50	0.40	YES	
Aggregated, All Sets	0.60	0.72	YES	YES

FIGURE 20 confirms Sherwood and Capps' hypothesis of two separate distributions for Capsules versus Shuttles and Stations. The relatively low R^2 value of 0.19 to 0.24 supports their finding of "not highly correlated data at the low end" when compared to the combined Shuttles and Stations with an R^2 of 0.47 to 0.61 percent.

Based on the criterion that the LM was a launch and landing vehicle of very short flight duration, it would seem to be an outlier among Stations, and might fit better on the capsule curve (where Rudisill et al put it).

Testing H_7 Petro (1999), Perino (2005), Kennedy (2006), Rudisill et al (2008) -- This hypothesis advocates for the power curve, shown in FIGURE 17, so we reject H_0 to this extent. It appears that these authors found that the power curve returns the highest R^2 value, and so adopted it as the best plot of the data. TABLE 5 shows R^2=0.08 percent for Shuttles and Capsules combined and 0.06 percent for Space Stations. These R^2 values are lower than Sherwood and Capps' segregation of the same data. However, those differences are a function of how the authors aggregate the samples rather than one being more correct than another is.

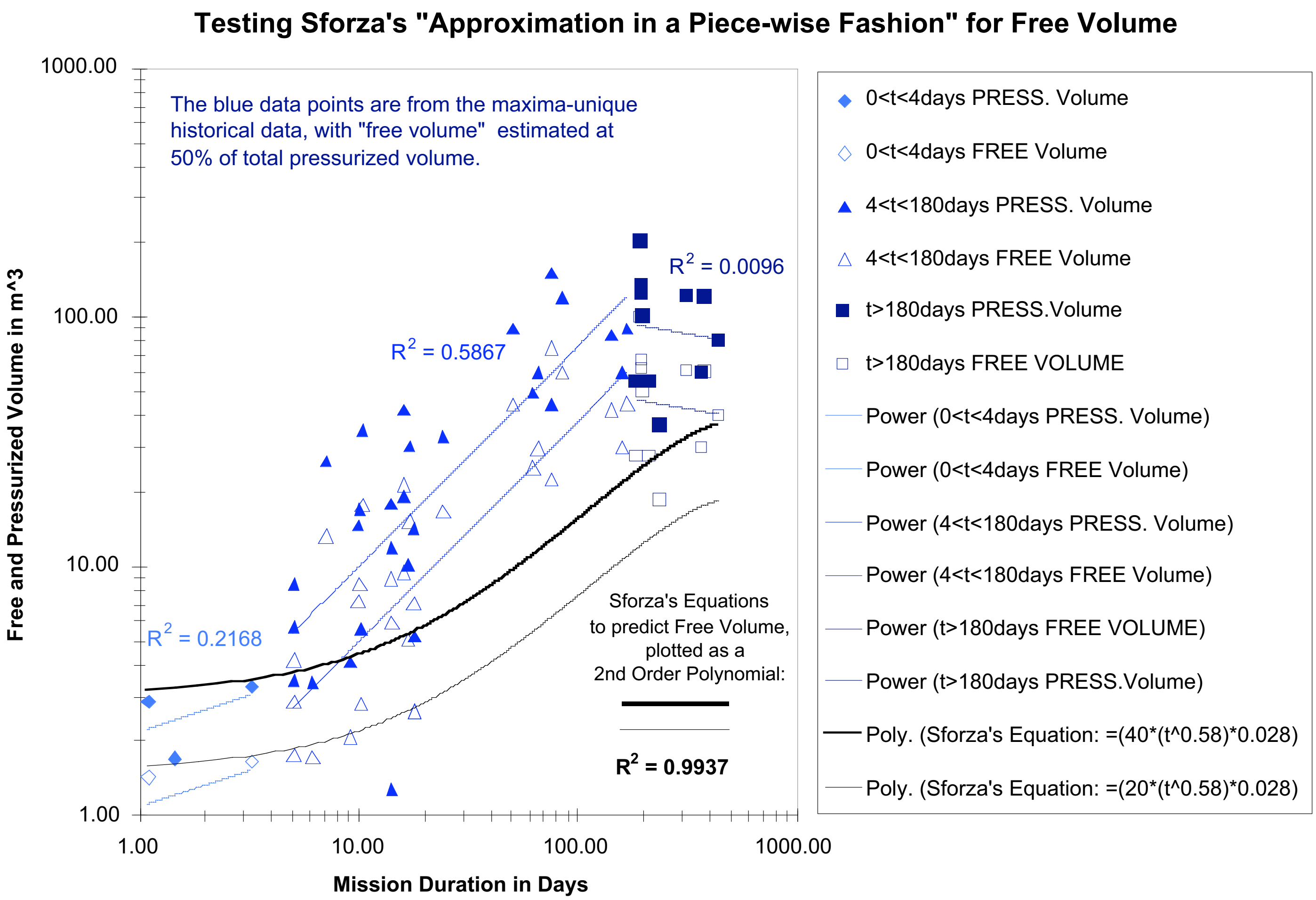

FIGURE 21. Test of Sforza's Equations to Predict "Free Volume."

Testing H_8 Sforza (2004) – FIGURE 21 plots the historical spaceflights according to the three segments in Sforza's EQUATION 1 using the maxima-unique data set. FIGURE 21 reveals that the three segments do form a sort of "S-curve" although they do not level our precisely horizontal. However, Sforza's power curve equation (converting 1 ft^3 = 0.028 m^3) for the complete maxima-unique data – when plotted as a polynomial -- does level out at the bottom according to his prediction. We reject H_0 for the first three of six parts of Sforza's hypothesis: the 1.13m^3 minimum *free volume* for capsules, *free volume* increasing as a function of mission duration and its plot as a power curve. We partially reject H Sforza's "S-curve" because it only partially does what he claims. Sforza's volumetric results are too low to reject H_0 for the other parts of the hypothesis. Sforza's moving the decimal point one place to the right from the missions of less than four days to missions of more than 180 days implies that the long duration spacecraft could be an order of magnitude larger per crewmember than the short duration transportation vehicle. Yet, even doubling his maximum volume as Sforza suggests, does not begin to explain the actual Skylab, Salyut, Mir, and ISS volumes per crewmember that are respectively ten and five times larger than his 11.3m^3 or 22.6m^3.

Testing H_9 Hofstetter, de Weck, Crawley (2005) –This survey tested Hofstetter et al on three parameters:

1. Hofstetter et al's main point was that the true value of the optimal curve falls below MSIS optimal. That is what happens in FIGURE 17 with the log curve (and also with lower order polynomials, so we reject H_0.for H_9.

2. Given the huge extrapolations already in Celentano, this interpolation means "putting a micrometer at the end of a furlong." Therefore, we fail to reject H_0 for the overly precise curve fitting of H_{9a}.

3. This test translated EQUATIONS 2a and 2b into Excel as EQUATION 8a and 8b, where DaysMax is the constant with the value of 270 days, and tmission and Ncrew are the same as EQUATION 2.

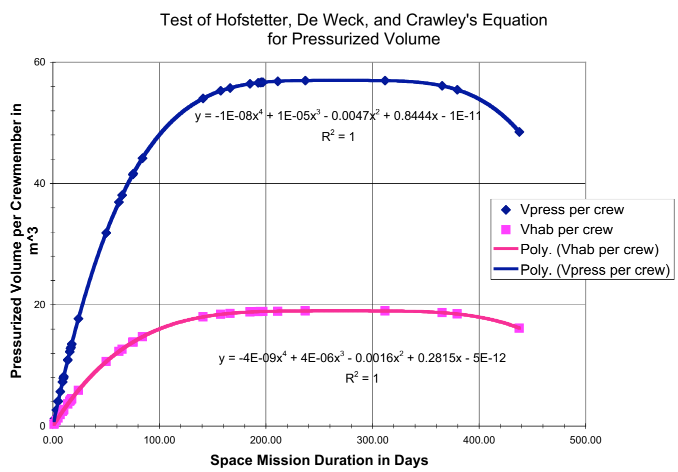

FIGURE 22a. Plot of the Hofstetter et al curves.

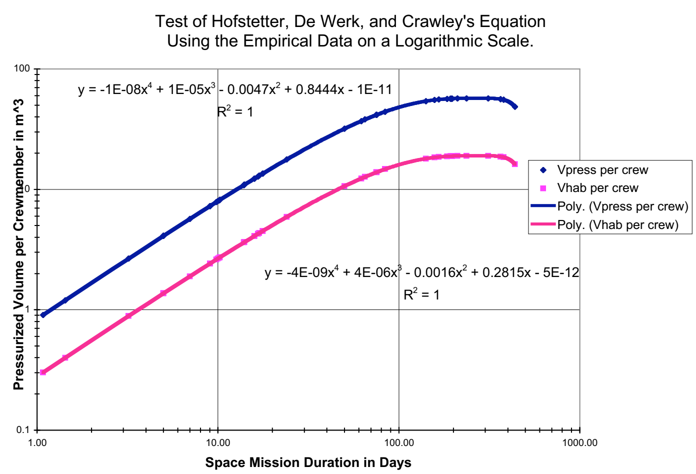

FIGURE 22b. Hofstetter et al curves on a log scale.

EQUATION 8a.

*Vhab=19*Ncrew*(1-((tmission-DaysMax)/DaysMax)^4)*

EQUATION 8b

*Vpress=Vhab*3*

FIGURE 22a plots EQUATION 8a and 8b as fourth order polynomial trendlines. This equation does what the authors claim in replicating the vaguely defined *habitable volume* in MSIS that levels off at about 19m^3. The *pressurized volume* curve rises to a maximum of about 57 m^3 per crewmember -- three times greater than MSIS. However, when using historical values for DaysMax instead of a constant, the whole proposition falls apart because the equation produces only a horizontal line at 57 m^3 for all missions. Although, Hofstetter et al may have a promising idea, it needs to be flexible enough to accommodate actual mission durations and volumes. Therefore, even though the algebra works as promised within the framework the authors defined, we fail to reject H_0.

FIGURE 22b plots EQUATION 8 on a log scale, where it resembles Sforza's graph, at least at the upper end.

The enlarged detail at the low end of this graph shows that the curves do not in fact pass through the origin, but the pressurized volume curve passes about 1m^3 above zero.

TESTING THE CREW SIZE HYPOTHESES.

The three crew size hypotheses proved testable, but the results were very different from the predictions. ***Try this thought experiment:*** Your house has a kitchen, living-dining room, kitchen, a bathroom, and three bedrooms. You want to add a bedroom, increasing the "crew capacity" by 33 percent. Will you also add 33 percent or more area to the kitchen and living-dining room -– as Davenport, Reynerson, and Kennedy imply?

Test of Davenport, Congdon, and Pierce's Crew Size-specific Volume/Duration Curves

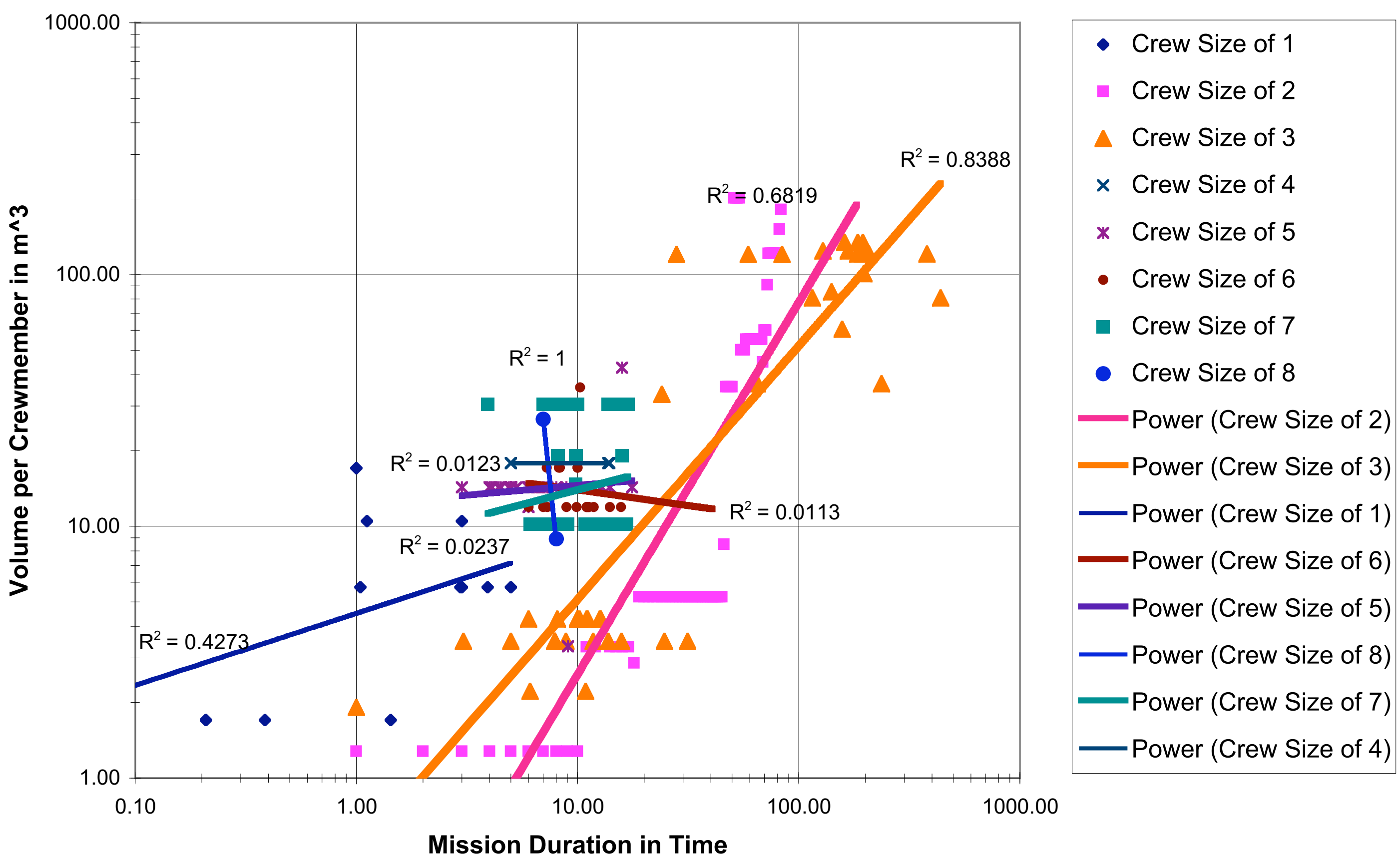

FIGURE 23. Testing H_9, Davenport, Congden, Pierce, Crew Size Hypothesis

Testing H_9, Davenport, Congden, and Pierce (1963) – It was simple to plot the trendlines for historical space missions of one, three, and five crewmembers, but it was not possible to plot a mission for 10 crewmembers, as there have been none to date. The test analysis plotted eight power curves to represent historical crews of one to eight members. FIGURE 23 shows these results for Davenport et al, for which we fail to reject H_0. Only the curves for one, two, and three crewmembers sloped positively in the general direction of Davenport's graph, with the R^2 showing an effect only for two and three crew. This result is not surprising given that 1963, the authors found data available only for those crew sizes.. However, none of these test curves follows Davenport's neat ordering, (where each larger crew appears higher and steeper than the smaller one). As the crews grow, the curves stray even farther from the prediction with negative slopes for C=6 and C=8.

Testing H_{10}, Reynerson (2005) -- We fail to reject H_0 because in FIGURE 24 neither curve shows an effect. The polynomial trendline returns an R^2 = 33 percent and resembles a slightly skewed normal curve with a tail. The linear trendline returns a vanishingly small R^2 = 0.07 percent. Thus, the Reynerson hypothesis does not connect to historical or empirical reality for volume. However, in his secondary assertion, he makes the inescapable point that mass scales as a function of mission duration.

Testing H_{11} Kennedy, Toups, and Smitherman (2008) – We fail to reject H_0. Kennedy et al assert the same idea as Reynerson, except they graph known crew sizes from one to six for three criteria lines *(tolerable, performance, preferred*). They do not provide data for the three curves so it is not possible to test this vestige of Celentano. The polynomial curve in FIGURE 24 returns an R^2 three orders of magnitude larger than the linear curve that most resembles FIGURE 15. The polynomial shows no resemblance to the predicted rising straight lines in FIGURE 15. On the contrary, ***the polynomial curve suggests the opposite:*** that for a large portion of the range from about two to seven crew members, volume per crew varies inversely to the number of crewmembers.

Kennedy, Toups, and Smitherman argue that the intrinsic character of space operations changes with mission duration. They do not provide data to test these short, medium, and long durations, but their operational definition of living and working arrangements changing with increasing duration is a valuable contribution.

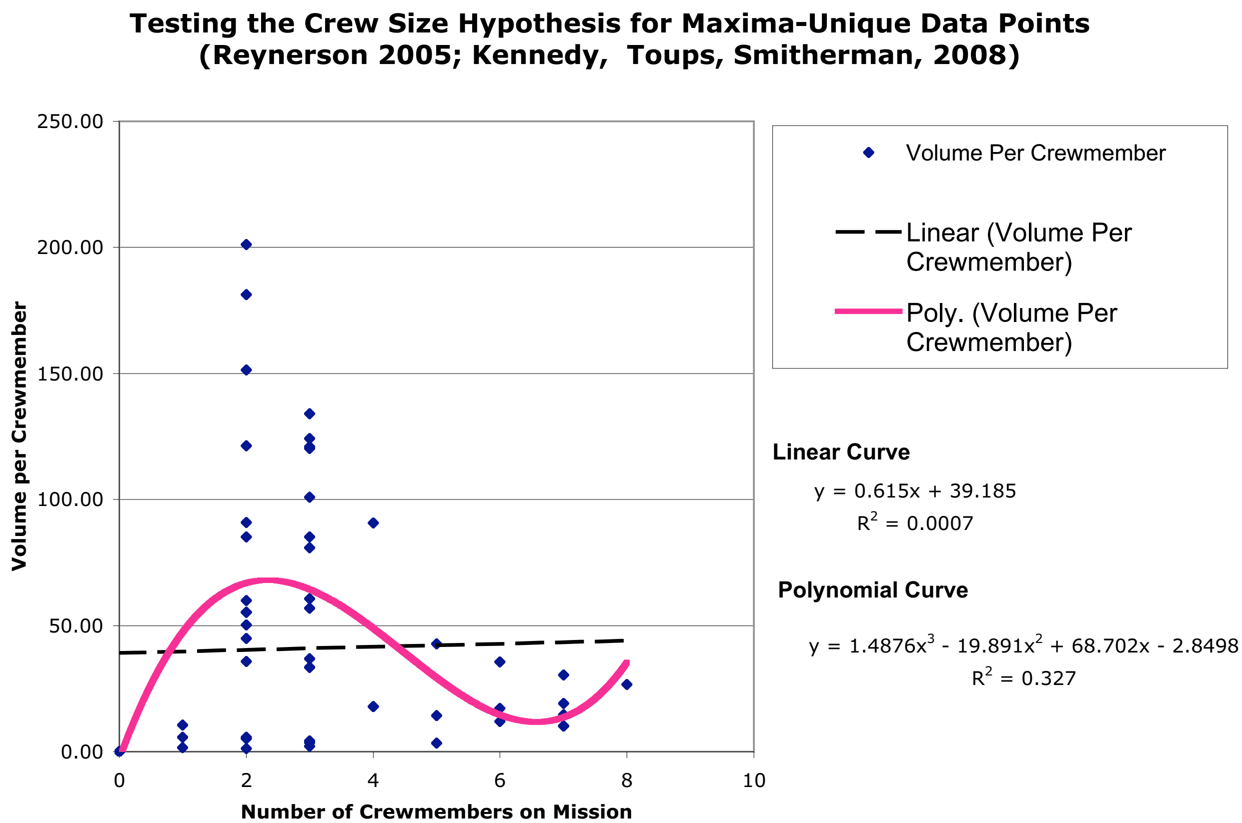

FIGURE 24. Test of H_{10} (Reynerson) and H_{11} (Kennedy et al): the Crew Size Hypothesis

TESTING THE FUNCTIONAL SIZING HYPOTHESIS

Testing H_{12} Schwartz (2005) – Schwartz was not testable, given the information available, so we fail to reject H_0. This hypothesis suffers from naiveté insofar as it assumes that operational functions make up a major portion of the spacecraft's volume. In fact, structure, mechanisms, utility chases, and life support systems make up a larger fraction of spacecraft mass and volume than operational functions do as a sizing driver.

Schwartz's concept suggests the importance of moving beyond raw pressurized volume to a more nuanced comprehension of the living and working environment in a spacecraft or space habitat. Rudisill makes this point:

> The space habitation community uses a series of terms to define types of spacecraft pressurized volumes. A primary concept is "net habitable volume," the generally accepted "usable spacecraft volume" after subsystems, stowage, outfitting, etc. have been accommodated and design inefficiencies are considered (traditionally, "net habitable volume" has equaled ~60% of total pressurized volume) (Rudisill et al, p. 2, 2008).

DISCUSSION: THREATS TO VALIDITY

A concomitant of this research design is to understand its limitations from potential threats to validity. The threats that apply are Conclusion Validity, Internal

Validity, Construct Validity of Cause, Construct Validity of Effect, and External Validity or Generalizability, (Campbell, Stanley, 1966; Cook, Campbell, 1979).

CONCLUSION VALIDITY

For conclusion validity, there are two types of error:

1. Conclude that there is no effect or relationship when there is.

2. Conclude that there is an effect or relationship when there is not.

Type 1 Error – The hypotheses evaluated in this survey did not exhibit any Type 1 errors of finding no effects when they existed.

In general, the challenge to this analysis for Type 1 error concerns the low statistical power, because the main tool available for use with non-random data is the coefficient of determination (R^2). This survey has a sufficiently large ***n*** of 254 human spaceflights, and the maxima-unique sample of 47 flights is also respectable. The effect size in the range of volumes is quite large –from 1.25m^3 per crewmember in Gemini to more than 200 times that much in the current ISS. However, these parameters do not translate into statistical power in the standard sense. It is not a random sample, which means that inferential statistics and their tests for significance are not available.

Low reliability of measures – Low reliability is a potential Type 1 threat. However, the measurement data for mission duration is precise, with timelines available to the minute for all flights and to the second for many. Measurement of pressurized volume is only slightly less precise, but still reflects engineering exactness. The issue of possible errors of *reporting* in the spacecraft literature comes under the *Instrumentation* section for *Internal Validity.*

Type 1 errors do not pose a threat to this survey.

Type 2 Error – Many of the hypotheses evaluated in this survey suffer from Type 2 errors; they find effects and relationships where there are none including many of the Mission Duration and all of the Crew Size hypotheses.

In general, the main threat from a Type 2 error would concern "fishing" for data points or "cherry picking" from the data. Conversely, this threat may incur the selective exclusion of data-points. This study handled this Type 2 error by standardizing the data on the unique maxima for each spacecraft in terms of crew size and volume.

Type 2 errors do not pose a threat to this survey.

INTERNAL VALIDITY

Internal validity asks, "Is there an effect, a relationship?" For the *mission duration hypothesis,* the results show a relationship. For the *crew size hypothesis,* the results are even clearer that there is no relationship. The key threats to internal validity include history, maturation, selection bias, experimenter bias, non-random sample, and instrumentation.

History – The history threat states spacecraft and space habitats have grown over time and that we would build bigger habitats over time anyway, regardless of mission duration. The answer of this argument is economic: that these spacecraft and stations are extremely expensive and conducting longer missions reduces the launch costs per unit of time in orbit. As Gore et al show, larger volume enables the crew to stay in space longer. The longer the crew can stay on a space station, the fewer launches needed to keep the station staffed.

Maturation – The maturation threat states that space agencies build larger spacecraft because over time they learned how to build bigger spacecraft. Although this threat may seem similar to history, it differs in the respect that it may take into account the beneficial effects of learning how to conduct long duration missions. What the maturation argument misconstrues is that the purposes of flying crews in space have matured too: the scientific experiments, construction and assembly tasks, and engineering tests have all become more sophisticated and complex. The mission maturity drives the maturation of the spacecraft.

Selection Bias and Distribution – The parsing of the data sets into various samples can skew the results. This question concerns whether to distinguish among different spacecraft types, as Sherwood and Capps did. There is a fascination in separating each class of spacecraft from the larger distribution into smaller sets and examining them and their diverse R^2s. This wide range of results for R^2 raises a question about this practice. However, Herbert Simon, Nobel Prize winner in economics, explains:

> To explain the word *distribution*, we make some assumptions that might be thought outrageous if applied in detail, but that might be plausible if only applied in the aggregate. (Emphasis added, Simon, 1989, p. 145).

Applying the analysis in the aggregate is vital to address the complete distribution of mission duration and spacecraft sizes, and it helps to avoid selection bias.

Experimenter Bias – The opportunity for experimenter bias arises in choosing which authors, hypotheses, and curves to analyze. Many more authors have cited and published the Celentano Curve and its variations than the dozen that appear in this survey. In many cases, the curve appears with little or no discussion, documentation, or explanation. The method of selecting curves for this survey was to seek the authors who advanced testable hypotheses with documented assertions and quantitative data.

Non-Random Sample – Nearly all historical data is non-random, but that is not an obstacle to quantitative historical research. The coefficient of determination (R^2) does not require random data, and this survey attempts to use it impartially and objectively.

Instrumentation – Instrumentation means the ability to measure the data accurately and consistently. The instrumentation threat arises primarily in terms of how authors describe spacecraft volume. This survey found that authors use varied definitions and protocols for living volume, living space, habitable volume, etc. It was also necessary to be careful of errors in the literature. For example, the original Apollo literature states that the CM has a pressurized volume of 366 ft^3 (10.34 m^3) and that the "unoccupied" volume is 210 ft^3 (5.94 m^3) (Spacecraft Systems Operations Branch, 1969, p. 1-10). However, a retrospective NASA publication gives the CM pressurized volume as 6.17 m^3; Wikipedia picked up this value and now it is all over the web. By using ***only documented engineering data*** for pressurized volume, this survey avoids the instrumentation threat.

Time Measurement – Although there appear to be some variations in how the US and Russian Space agencies measure mission duration, these differences are so small as not to constitute an instrumentation threat.

CONSTRUCT VALIDITY OF CAUSE

The construct of cause asks if the x that explains the variance in y is really the "x" that the alternate hypothesis says it is. The other alternative hypotheses pose the threat that crew size is the determinant of volume/person or that functional requirements drive the volume. Testing with the coefficient of determination is the primary defense against these threats.

Generalizing Across Time – This threat implies that it might not be possible to draw equally valid results for the same parameters from missions of differing duration (perhaps in the way that Kennedy, Toups, and Smitherman suggest for different classes of mission duration). This threat arises in plotting the data for the early test flights in first programs – notably Mercury, Vostok, Gemini, and Shenzhou. These outliers had very short durations; to plot the complete mission set of 254 spaceflights forces an additional two orders of magnitude ($0.01<t<0.1$ days and $0.1<t<1.0$ days) into the logarithmic scale. For example, for FIGURE 2 omits the Mercury and Vostok test flights so that the minimum of the time scale was 0.1 days rather than 0.01 days. This threat implies that for almost any spacecraft, it may not be valid to generalize from the early test flights to the full operational capability for the later, longer duration flights, evoking Sherwood and Capps. The defense against this threat is to apply the maxima-unique data set, which uses the longest mission for each spacecraft for each crew size.

CONSTRUCT VALIDITY OF EFFECT

The construct of effect asks if the y for which x explains the variance is really the "y" that the research observes.

Effect: Meeting the Crew's Needs? -- Is the effect really the volume that meets the crews' requirements? Cynthia Null, former Chief of the Human Factors Research Division at NASA-Ames and now the Human Factors Authority at the NASA Engineering Safety Center (LaRC) articulated this threat:

> The only reason we say that the volume met crew requirements is that nobody died. The crew takes what we give them and suck it up, so we say it meets their requirements. (Personal conversations, June 2006).

This construct of effect is the most serious of all the threats to validity. The space community does not yet possess data to show how well a spacecraft design met the crew's needs over the mission duration. For now, it is possible to state the effect only in the negative that the volume did not fail the crew requirements insofar as there were no catastrophic consequences.

Effect: Pressurized Versus Habitable Volume? -- The MSIS (1987) introduction of the term *habitable volume*, led to much speculation about how it differs from the pressurized volume. The lack of documentation on the constructs of free, habitable, or living volume hindered the authors from any definite assessments. Sforza and Hofstetter et al, and Rudisill et al offer algorithms as multipliers to predict pressurized volume from a vague notion of habitable volume. This construct of effect asks in essence, which is a more important measure: pressurized or habitable volume? Certainly the two are related, and the next big question to ask is how they are related in evidence-based architectural design terms. However, the fact that this relationship is yet ill-defined does not invalidate the understanding of pressurized volume.

Other Effects -- Another construct of effect (but not necessarily a threat) asks if volume is the only effect of mission duration. Reynerson asserts that mass is the effect of mission duration, and certainly launch mass is a major concern for spacecraft designers. Mass occurs predominantly the "solid" portions of the spacecraft – not the open volume where the crew live. Certainly, there is a mass penalty for free volume in primary (pressure vessel) and secondary (decks, stand-offs, partitions) structure, but it does not correspond directly with the living space.

EXTERNAL VALIDITY

External validity asks how far it is possible to generalize from a result. Sherwood and Capps challenge generalizing from capsules to space stations. Rudisill et al challenge generalizing from transportation vehicles to space stations or lunar/planetary habitats. Do the authors extrapolate too far beyond the actual data as Celentano et al did, do they or interpolate too precisely within the data, as Hofstetter et al did? External validity can also mean generalizing to other settings, known as generalizing across effect constructs. It resonates with the effect threat of ***whether the spacecraft volume meets crew requirements.*** It poses the question of generalizing from spacecraft to other habitats, including lunar and planetary bases. This survey's results may generalize only to future spacecraft within the historical envelope and to a cautious extrapolation beyond it -- perhaps as far as Mars and back in about 1000 days.

SUMMARY OF THREATS TO VALIDITY

Although they may pose formidable questions, these threats do not invalidate the main results of this survey that pressurized volume per crewmember increases with mission timeline, up to any currently contemplated duration. The main weakness arises in claims that the volume meets crew needs or serves vague levels of crew comfort or performance. The fact that the human spaceflight community cannot yet make those correlations compels the need for further research using direct observation and measurement of space habitats and analogs with the crews who live and work in them.

FINDINGS: THE QUESTIONS FOR THE HYPOTHESES

1. *Has the evolution of spacecraft from Vostok and Mercury to the International Space Station followed the path predicted by the Celentano Curve?*

Yes, to a remarkable extent it has, although the upper volume bound rises an order of magnitude higher than Celentano's $20m^3$/crewmember.

2. *Does the volume prediction follow a curve that levels out at a time limit?*

The curve does not level out, but in the logarithmic curve, the slope lessens gradually.

3. *Which curve pattern best fits the data under each hypothesis?*

The close R^2 values for log and power curves mean they represent the same variance in different but overlapping ways. The two R^2 are so close, neither curve is better than the other is.

4. *Can we evaluate this "best fit" by the R^2 value, or do we need to test for correlation significance among the curves?*

The R^2 plus the graphical form of the curve are the best indicators of what variance the curves represent and how well they represent it. "Correlation significance" did not emerge as a weakness.

5. *Does this curve pass through the origin or otherwise show no minimum value?*

If the volume curve passes through zero, it is no longer that of a crewed spacecraft, so the question is moot. In fact, the curves do not pass through the origin.

6. *How does the aggregation or disaggregation of the data affect the results?*

As shown in TABLE 5, the tricks of aggregating or disaggregating data can alter the results (R^2) by orders of magnitude. The challenge is how to handle the data in an impartial and objective way to avoid this kind of skewing.

7. *Is there an empirical basis in 47 years of human spaceflight for pressurized volume in terms of tolerable, performance, and optimal levels?*

No.

CONCLUSION

This survey demonstrates that curve fitting is a poor method for designing human spacecraft and space habitats because of its limited validity and usefulness. The graphs of historical spaceflight may provide a frame of reference for spacecraft design, they cannot substitute for designing from first principles on a both a quantitative and qualitative basis. The conclusions are:

FOR THE MISSION DURATION HYPOTHESIS,

- Spacecraft pressurized volume per crewmember increases as a direct function of mission duration.

- Unlike Celentano et al and MSIS, the pressurized volume does not level off at atop limit, but keeps rising to about 1000 days, the nominal length of a human Mars mission.

- The volume vs. mission duration results shows a statistical effect as both a logarithmic and a power curve, representing the nearly the same magnitude, with differing but overlapping portions of the variance.

- These results support Sherwood and Capps' and Rudisill et al's argument that because of their strict aerothermal shape, small capsules differ fundamentally from larger space habitats or vehicles and belong in separate data sets.

FOR MORE GENERAL RESULTS,

- Kennedy, Toups, and Smitherman argue that missions of substantially different durations are intrinsically different in character. This argument complements Sherwood and Capps' and, Rudisill at al's reasoning about capsules versus stations because of the difference in mission duration.

- The crew size does not affect volume/crewmember.

- The impact of pressurized volume upon spacecraft launch mass and system mass requires further study.

FOR FUTURE RESEARCH , AND

- Because of the prevalence of Type 2 errors in previous results, future studies should use a Bayesian analysis to avoid Type 2 Errors and other "false positives."

- Future research must handle the construct of effect threat from the difficulty of verifying that the spacecraft "met the crew's needs."

FOR FUTURE SPACECRAFT DESIGN.

- For future spacecraft sizing, it will be vital to start from first principles with functional, mission, and operational requirements, translated into volumetric units of analysis and design.

- For future spacecraft design, it is important to understand "the area between the curves" as the domain where design can be most effective.

- Sforza, Hofstetter et al, Rudisill et al, and Schwartz all suggest diverse multipliers to translate the habitable volume requirement to total pressurized volume. Future spacecraft design efforts must develop and validate evidence-based sizing methods to achieve this goal.

As the space habitability and architecture community prepares for the second half-century of human spaceflight, it must progress beyond well-intentioned but speculative predictions. Instead, the community will need to develop sound quantitative and reproducible models. The next step should be a systematic measured survey of human spacecraft interior architecture, quantifying all areas and volumes on a rigorous basis to which it is feasible to apply inferential statistics that afford tests for significance.

ACKNOWLEDGMENTS

I would like to thank Greg Zwernemann, a director of Advanced Capabilities and Integrated Solutions at Northrop Grumman, for starting me down the path to produce this study. Carl Meade, director of Space Systems at Northrop Grumman, provided encouragement and support. I would further like to thank Mary Connors, Cynthia Null and James Stevenson at NASA-Ames Research Center for their advice, encouragement, and mentorship in this study. Finally, I thank my readers Jim Berry, Mary Davis, Marianne Rudisill, Brent Sherwood, and Ed Shoenstein.

REFERENCES

Allen, Christopher S., et al (2003, Jan). Guidelines and Capabilities for Designing Human Missions, NASA TM-2003-210785.

Campbell, D. T., Stanley, J. C. (1966). Experimental and Quasi-experimental Designs for Research. Chicago: Rand McNally.

Celentano, J. T.; Amorelli, D.; Freeman, G. G. (1963, May 2). Establishing a Habitability Index for Space Stations and Planetary Bases, AIAA 1963-139. AIAA/ASMA Manned Space Laboratory Conference, Los Angeles, CA. New York: AIAA.

Connors, Mary M.; Harrison, Albert A.; Akins, Faren R. (1985) LIVING ALOFT: Human Requirements for Extended Spaceflight, NASA SP-483. Washington DC: NASA.

Cook, T. D., Campbell, D. T. (1979). Quasi-Experimentation: Design and Analysis for Field Settings. Rand McNally, Chicago, Illinois.

Davenport, E. W.; Congdon, S. P.; Pierce, B. F. (1963, June 17). The Minimum Volumetric Requirements of Man in Space, AIAA 1963-250. New York: AIAA.

Fraser, T. M. (1966). The Effects of Confinement as a Factor in Manned Spaceflight, NASA CR-511. Washington, DC: NASA.

Fraser, T. M. (1968, June). The Intangibles of Habitability During Long Duration Space Missions, NASA CR-1084.

Gore, D. C., Jr.; Martin, D. A.; Trust, R. D. (1978). Space Shuttle Orbiter Habitability and its Extensibility, AIAA 1978-1669. AIAA Conference on Large Space Platforms: Future Needs and Capabilities, Los Angeles, CA, Sept. 27-29, 1978. New York: AIAA.

Hofstetter, Wilfried; de Weck, Olivier; Crawley, Edward (2005). Modular Building Blocks for Manned Spacecraft: A Case Study for Moon and Mars Landing Systems, Seattle, WA: International Council on Systems Engineering.

Kennedy, Kriss J. (2006, Sept. 19-21). Advanced Habitation Efforts at JSC-CTSD, AIAA-2006-7333, AIAA Space 2006, San Jose, CA.

Kennedy, Kriss J.; Toups, Larry; Smitherman, David (2008, March 4). Lunar Habitation Strategies, ASCE Earth and Space Conference, Long Beach, CA.

Larson, Wiley J.; Pranke, Linda K., Eds. (1999). Human Spaceflight: Mission Analysis and Design, New York: McGraw-Hill.

Manned Space Center (1966, Nov. 7). Preliminary Technical Data for Earth Orbiting Space Station: Standards and Criteria, Vol. 2, MSC-EA-R-66-1, NASA TMX-59700. Houston, TX: NASA Johnson Space Center.

Marton, T.; Rudek, F. P.; Miller, R. A.; Norman, D. G. (1971, October). Handbook of Human Engineering Design Data for Reduced Gravity Conditions, NASA CR-1726. Washington, DC: NASA.

NASA (2005, December). Exploration Systems Architecture Study (ESAS) Report. http://www.nasa.gov/mission_pages/exploration/news/ESAS_report.html, accessed July 12, 2006.

NIST (2006, July 18). NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/index.htm accessed 8 April 2008.

Perino, M. A. (2005, May 26). Moon Base Habitation and Life Support Systems, Moon Base: A Challenge for Humanity, Venice, ITALY: May 26-27, 2005. http://www.moonbase-italia.org/PAPERS/D1S2-MB%20Assessment/D2S2-03HabitationSystem/D2S2-03HabitationSystem.Perino.pdf

Petro, Andrew (in Larson, Pranke, Eds. 1999). Transfer, Entry, Landing, and Ascent Vehicles, pp. 391-420.

Reynerson, Chuck (2003) Designing for Space Exploration: Space Hotels to Interplanetary Space Vehicles. Accessed August 1, 2006. http://www.mines.edu/research/srr/2003_Meeting/ModelingHumanSpaceflight2.ppt,

Rudisill, Marianne; Howard, Robert; Griffin, Brand; Green, Jennifer; Toups, Larry; Kennedy, Kriss (2008, March 4). Lunar Architecture Team – Phase 2 Habitat Volume Estimation: "Caution When Using Analogs," ASCE Earth and Space Conference, Long Beach, CA.

Schwartz, Jana (2005, Feb. 28). NASA Concept Exploration and Refinement Study: Crew Exploration Vehicle. http://exploration.nasa.gov/documents/reports/cer_final/Draper_MIT.pdf, accessed July 12, 2006.

Sforza, Pasquale (2004). Crew Volume Allowance, EAS4710 Aerospace Design 2 – Space Access Vehicle Design. http://aemes.mae.ufl.edu/~sforza/EAS4710/index.htm, accessed July 11, 2006.

Sherwood, Brent; Capps, Stephen (1990, March 23). Long Duration Habitat Trade Study: Space Transfer Concepts and Analyses for Exploration Missions. NASA Study Contract NAS8-37857 for Marshall Space Flight Center. Huntsville, AL: Boeing Aerospace and Electronics.

Simon, Herbert A. (1989) The Sizes of Things, in Tanur, Judith M., et al, Eds, Statistics: A Guide to the Unknown, 3rd Edition. Pacific Grove, CA: Wadsworth and Brooks. pp. 142-150.

Spacecraft Systems Operations Branch (1969, October 15). Apollo Operations Handbook Block II Spacecraft: Volume 1 Spacecraft Description, SM2A-03-Block II. Houston, TX: NASA Apollo Project Office.

Saucier, D. R. (1992, May 24-27). Extended Duration Orbiter - Meeting the challenge, AIAA-1992-1271, Space Programs and Technologies Conference, Huntsville, AL,

Trochim, William M. (2006. Oct. 20). The Research Methods Knowledge Base, 2nd Edition. http://www.socialresearchmethods.net/kb/ Accessed March 18, 2008).

Woolford, Barbara; Bond, Robert L. (in Larson, Pranke, Eds. 1999). Human Factors of Crewed Spaceflight. pp. 133-153.

CONTACT

Marc M. Cohen, Arch.D
Space Systems Business Segment
Northrop Grumman Integrated Systems
One Hornet Way, MS 9QS4/W5
El Segundo, CA 94025-2804 USA
+1 310 332-0819
marc.cohen@ngc.com

ADDITIONAL SOURCES

Barrera, Thomas (1988, July 11-13). Extended Duration Orbiter, AIAA-1988-2864, 24th AIAA Joint Propulsion Conference, Boston, MA.

Celentano, J.T.; Adams B.B. (1960). Habitability and Maintenance of Human Performance in Long Duration Space Missions, AAS 1960-83. New York: American Astronautical Society.

Celentano, J.T.; Amorelli, D. (1963). Crew Status in Various Space Cabin Configurations and Volumes, NAA Publication 543-Q. Downey, CA: North American Aviation.

Glynn, Philip; Barrera, Thomas (1993, Sept. 21-23). Space Shuttle Orbiter Extended On-Orbit Duration Implementation and Benefits, AIAA-1993-4057, AIAA Space Programs and Technologies Conference and Exhibit, Huntsville, AL.

NASA Constellation Project Office (2007). Human Systems Integration Requirements, NASA CxP-70024. Houston, TX: Johnson Space Center.

Nelson, William R.; Bagian, Tandy M. (2000, Nov). Critical Function Models for Space Operations of the International Space Station, International Topical Meeting on Nuclear Plant Instrumentation, Controls, and Human-Machine Interface Technologies, Washington, DC.

Rathert, George A., Jr.; McFadden, Norman M.; Weick, Richard F.; Patton, R. Mark; Stinnett, Glen W.; Rogers, Terence A. (1964, Feb). Minimum Crew Space Habitability for the Lunar Mission, NASA TN D-2065. Washington, DC: NASA..

Stuster, Jack (1986). Space Station Habitability Recommendations Based on a Systematic Comparative Analysis of Analogous Conditions, NASA CR-3943. Washington DC: NASA..

Wise, James (1988, August). The Quantitative Modeling of Human Spatial Habitability, NASA CR 177501. Moffett Field, CA: NASA.

DEFINITIONS, ACRONYMS, ABBREVIATIONS

AIAA: American Institute of Aeronautics and Astronautics

ASCE: American Society of Civil Engineers

Coefficient of Determination (R^2): A measure of the variance in the dependent variable Y that the variance in the independent variable X can explain.

CEV: The Orion Crew Exploration Vehicle

CM: Apollo Command Module

ESAS: Exploration Systems Architecture Study, NASA, 2005

x^e: Exponential of x or power curve.

Gemini: US 2 crew spacecraft

ISS: International Space Station.

LM: The Apollo Lunar Module

LSAM: Lunar Surface Access Module from the ESAS Report

Mir: Russian Space Station, 1986-2000

H_0: The null hypothesis

$H_{1...n}$: The alternate hypothesis or hypotheses.

Hypothesis, Null: The null hypothesis always states that there is no effect of the independent variable upon the dependent variable.

Hypothesis, Alternate: The alternate hypothesis states that the independent variable has an effect on the dependent variable.

JPL: NASA Jet Propulsion Lab

JSC: NASA Johnson Space Center

LaRC: NASA Langley Research Center

LEO: Low Earth Orbit

LH2: Liquid Hydrogen

Ln(x): Natural Log of x curve.

LOX: Liquid Oxygen

Mercury: US single crew spacecraft, 1961-1963

MSC: Manned Space Center, the original name of JSC.

MSFC: NASA Marshall Spaceflight Center

MSIS: Man-System Integration Standard, the first two editions of NASA Standard 3000 published 1987, 1995.

NASA: National Aeronautics and Space Administration

Natural log curve: A curve based on a natural logarithm to represent the variance in the dependent variable.

NIST: National Institute of Standards and Technology.

Polynomial curve: A curve based on a polynomial expression to represent the variance in the dependent variable.

Power curve: A curve based on an exponential function to represent the variance in the dependent variable.

Salyut: Series of Russian Space Stations, launched 1971-1981

Shenzhou; Chinese capsule spacecraft derived from the Soyuz template, capable of flying one to three crewmembers.

Soyuz: Russian launch and reentry capsule spacecraft.

Standard Deviation, SD:The square root of the variance; a measure of spread.

$$SD = \sqrt{VAR(X)}$$

STS: Space Transportation System, the Space Shuttle.

Type 1 Error: Find no effect when there is one, e.g. a false negative

Type 2 Error: Find an effect where there is none, e.g., a false positive.

VARIANCE: A measure of how spread out a distribution of data is; a measure of variability. For a single variable *X* having a distribution *P(X)* with the known population mean μ, the population variance, VAR(x) is:

$$VAR(X) = (X - \mu)^2$$

x: For the R^2 value is the independent or explanatory variable.

y; For the R^2 value is the dependent or explained variable.