Risk Acceptance Personality Paradigm:

How We View What We Don’t Know We Don’t Know

Michael J. Massie*
ARES Corporation, Houston, Texas 77058

A. Terry Morris†
NASA Langley Research Center, Hampton, Virginia 23681

The purpose of integrated hazard analyses, probabilistic risk assessments, failure modes and effects analyses, fault trees and many other similar tools is to give managers of a program some idea of the risks associated with their program. All risk tools establish a set of undesired events and then try to evaluate the risk to the program by assessing the severity of the undesired event and the likelihood of that event occurring. Some tools provide qualitative results, some provide quantitative results and some do both. However, in the end the program manager and his/her team must decide which risks are acceptable and which are not. Even with a wide array of analysis tools available, risk acceptance is often a controversial and difficult decision making process. And yet, today’s space exploration programs are moving toward more “risk based design” approaches. Thus, risk identification and good risk assessment is becoming even more vital to the engineering development process. This paper explores how known and unknown information influences risk-based decisions by looking at how the various parts of our personalities are affected by what they know and what they don’t know. This paper then offers some criteria for consideration when making risk-based decisions.

Nomenclature

CAIB = Columbia Accident Investigation Board
CDR = Critical Design Review
CG = Center of Gravity
CxF = NASA’s Constellation Program
FMEA = Failure Modes and Effects Analysis
FTA = Fault Tree Analysis
IHA = Integrated Hazard Analysis
LOC = Loss of Crew
NASA = National Aeronautics and Space Administration
PDR = Preliminary Design Review
PRA = Probabilistic Risk Assessment
RIDM = Risk Informed Decision Making
SDR = System Definition Review
SRR = System Requirements Review
SSA = Software Safety Analysis

* Lead for NASA’s Constellation Integrated Hazard Analysis, ARES Corporation, Johnson Space Center
† Safety-Critical Avionics Systems Branch, Mail Stop 130, SAVIO IHA Lead, AIAA Lifetime Associate Fellow.
I. Executive Summary

The recent experience of working with risk informed decision making (RIDM) used on the NASA Constellation Program (CxP) presented a unique opportunity to study how the risk informed design process arrives at a decision and what influences that decision. Further study of RIDM took the authors on an unexpected path that led to the very heart of how human beings make any decision. Specifically, humans have an amazing inherent capacity to interpolate, that is, to fill in missing parts of a picture to make sense of their world and to make decisions. That capacity comes in many “flavors” which arise from our personalities. In the context of risks, it is a common belief that if we “understand” our risks then we will make good decisions based on that understanding. However, as the authors investigated how humans “gain understanding of risk,” we found that there is a component to risk-based decision making that is known in the scientific community but is not fully explored and leveraged in current real world techniques for technical risk-based decision making. That component is associated with the unknown unknowns. When some prior NASA accidents were briefly revisited, it was found that the unknown unknowns likely influenced the decisions that resulted in the accidents. Furthermore, when this component was combined with the science that drives our inherent human decision making processes, the unknown unknowns seems to be a significant player in the decisions we make about most anything. Recognizing this, the authors looked at how the unknown unknowns could be used to lower cost of NASA programs. It was found that by acknowledging the existence of unknown unknowns and allowing the existence to influence early risk-based decisions, some program decisions would change while most would most likely not change. The authors found that this approach, where it would result in different decisions, could result in significant savings in program total costs.

To really know everything there is to know about a risk, or anything else, one must understand three things: 1) what they know about a risk, 2) what they know they don’t know about a risk and, 3) what they don’t know they don’t know about a risk. Risk analysis tools are developed as a means of communication to inform program management of the risks that have been identified and serve as management tools to assure risks are assessed, mitigated and resolved where appropriate. Risk analysis tools like probabilistic risk assessment (PRA), hazard analysis, fault tree analysis (FTA), software safety analyses (SSA), failure modes and effects analysis (FMEAs), etc each provide alternative perspectives to characterizing/classifying risks based on severity and likelihood. However, in the early stages of a program all of these tools are challenged when attempting to come up with accurate estimates of likelihood. These tools serve very useful roles in risk assessment but cannot completely capture the risk picture because they all are based on assumptions (both explicit and implicit). However, analyses are still vulnerable to likelihood assessment errors due to the fact that whatever an assessor doesn’t know cannot be included in the model. Margins can and often are added to analyses to accommodate these unknown unknowns, but some programs cannot tolerate the impacts of carrying large margins (eg: mass, center-of-gravity (CG), cost, etc) and so try to limit what they can accommodate in the design to what they know they have to mitigate leaving the unknown unknowns to reveal themselves in due time.

This paper investigates (1) how risk-related decision making is influenced by how the decision-maker responds to the unknown unknown contributors to the risk, and (2) whether the decider’s aggregate personality influences their response. We find that this aggregate personality is influenced by a combination of cognitive, affective, volitional, conscious, sub-conscious, unconscious and environmental factors. Because it is an integral part of how we relate to the world, people who make risk-based decisions are usually unaware of this influence. And because these forces are normally sub-conscious, the influences of the unknown unknowns are manifest in accidents to an unknown but real extent. This means that at least some of the decisions that contribute to an accident are influenced by the human struggle to deal or not deal (consciously or sub-consciously) with the unknowns. It is important to note that this is not a failure of engineering or management. It is the inherent consequence of making risk-based decisions with incomplete data and uncertain knowledge. Since it does not seem realistic that one can know everything there is to know about anything, there will always be gaps in our data. The purpose of this paper is not to judge the events that led to prior accidents. Instead, this paper suggests that becoming aware of potential sources of unknowns can be used to help managers strategically make risk-based decisions that can benefit their overall programs. A methodology for design implementation of safety requirements is proposed that addresses this issue and is believed to provide a more robust design that is subject to less rework later in a program, thus reducing overall program cost.
II. Background

A. Risk Tolerance

1. Public Perceptions

Risk is perceived at all levels of conscious thinking and so is influenced at all levels of thinking. The general population can have its own view of risk that is not necessarily the result of some combination of how any one, group or even the entire group of members of the population views the risk. For example: the population of Houston recently voted to disallow the use of cameras to fine drivers who run through red lights at intersections. However, the city managers decided that the cameras must remain operational until they get an agreement to terminate them from the contractor who operates them. The reason the cameras are there in the first place is to mitigate red light runners, without using police officers. The public clearly supports the use of police officers to mitigate red light runners, but not automated machines. The city managers and the voters both represent the population but there are two different answers to how on to deal with the risk associated with red light runners. In fact, the group with control on how to deal with the red light runners is making a decision based on an influence that has nothing to do with the red light runners themselves. The voters perceive the risk one way and the city managers view the risk another. Each group is influenced by different outside forces (which include other related risks) and are arriving at different conclusions. This dichotomy offers an immediate clue into the nature of risk-related decisions: they are affected by many internal and external influences. Our perception of any given risk at any moment in time can be affected by many influences.

2. Levels of Acceptable Risk

Another important point about risk is that the same or similar risk in a different situation or context can result in a completely different decision on the acceptability of a risk. This is particularly evident in the public’s multi-level perception of NASA risks. On one level, NASA failures that result in loss of crew (LOC) are not considered acceptable risks. However, when compared relative to the gains of humanity’s desire to explore, some of NASA’s failures are considered acceptable as a means of ensuring future progress. This perspective may seem counterintuitive, but the desire to know sometimes overpowers the risks of the endeavor. NASA astronauts knowingly and deliberately strap themselves into a vehicle whose operation is considered very hazardous, and yet a social standard has been set that “failure is not an option.” Accidents resulting in LOC in this situation are not considered acceptable losses.

Contrast this situation with oil company workers who work on or around oil, gasoline, and diesel fuel tanks at a refinery. These employees also deliberately and knowingly go out and put themselves in very hazardous situations every day. However, the public perception of this risk is quite different from the NASA astronaut. NASA needs enormous amounts of justification for their crews to take risk because NASA, from a acceptable risk perspective, is held to a higher standard. The refineries also need some justifications but do not receive near the public scrutiny nor the exposure that NASA does even in the event of an accident. In fact, refinery accidents that result in employee loss occur much more frequently than NASA incidents that result in LOC. Actually, the public cares about every employee but because the employee takes that risk many times without incident there is a great deal more tolerance. Thus, the acceptance of risk also varies greatly with the perception of a risk and even the same risk in different context can yield different acceptability answers.

B. Risk Assessment

There are many risk assessment tools for managers or designers to utilize, each with a unique set of features, benefits and limitations. But when it comes to risk-based decision making, there is a very limited set of dimensions that humans draw upon when making decisions. These dimensions are: 1) what we know about a system or circumstance (known knowns), 2) what we know we don’t know about a system or circumstance (known unknowns) and 3) what we don’t know we don’t know (unknown unknowns).

Identification
Every person, group or population must first perceive a risk in order to evaluate and make a decision on whether to accept, mitigate or avoid it. Hazard analysis is a tool which provides a framework for managers of any task or endeavor to systematically identify and mitigate the risks associated with a design or operational plan. The hazard analysis approach starts with a checklist based on energy sources, operational influences, interfaces and physical properties of any design in order to bring out possible safety risks to the public, to ground or flight crews or to the mission inherent in a design. Once hazards and their causes are identified, the hazard analyst works with the designers to refine the details of an identified safety risk in order to outline features which will mitigate the risk. Hazard analysis is one of a number of engineering based approaches used to identify safety-related risks. It is generally the most prevalent risk identification process because it is the most basic.

Mitigation

Once identified, some risks require mitigation. Mitigation, from this perspective, is a means of preventing or hindering hazards or threats from manifesting themselves in a design, configuration or operation. Stated differently, mitigations are features put into the system that help the people involved avoid the realization of the undesired risk consequences. Mitigations must have three components: 1) they must be real and present when needed, 2) they must be verifiable as a means of mitigating the perceived risk, and 3) they must be capable of operating in whatever environment or operational circumstance they are required. Just as there are varying levels of acceptable risks, there are also varying levels of mitigations.

Residual Risk

Once the risk is mitigated to an acceptable level in any design or operational scenario, some will see the risk as no-longer credible or applicable to their own design. There is a yet a credible risk that the mitigations provided for the hazard identified will not be available when needed or may fail due to unforeseen circumstances (unknown states, interactions, influences or environments). The likelihood of the risk occurring despite the use of the mitigations is the residual risk of the design after the controls have been implemented. Risk-based decisions are based on the perception of the residual risks once mitigations are implemented. This perception space is where most difficulties are encountered. As noted earlier, with various risk-based influences at work, the perception of any specific risk can vary dramatically across decision-makers. It’s important to understand why.

C. Understanding Residual Risk

We assert that perception of residual risk is the real driver that any person will use to make a risk based decision and it’s based on the three dimensions of risk: 1) what we know about a system or circumstance, 2) what we know we don’t know about a system or circumstance, and 3) what we don’t know we don’t know about a system or circumstance. In other words, residual risk is simply the risk after employing a set of mitigations.

1. What we know about a system (known knowns)

Our perception of risk is shaped in part by what we know about a system/circumstance. Known risks (hazards) in the system are mitigated where appropriate. These risks are understood and mitigated to the point that the probability of realizing the risk consequence is at an acceptable level. For instance, NASA has significant knowledge on the strengths of materials (metals or composites) used to carry structural loads in a spacecraft design. Engineers spend many man-hours characterizing the loads that the materials will have to carry and then designing the integrated system to carry those loads plus some margin of safety. They optimize their designs to cover all potential loads cases and still maintain sufficient margins of safety to ensure the risk cannot be realized during the mission profile. Their process and results are based on more than 50 years of experience. Hence, once all of the load cases are identified, the risk of structural failure is believed to be controlled without further mitigations. Engineers have learned that they can count on their designs to perform as expected because it has been demonstrated often. Therefore, what is known about the structural design of a spacecraft is believed to be very high. This can lead some risk evaluators to conclude that they know everything about the system under consideration and to feel confident in their decisions that the risk is sufficiently mitigated or no longer credible.

2. What we know we don’t know about a system (known unknowns)
Consider the driver who is trying to pull out of their neighborhood into an intersection that is near the apex of a curve. The driver is faced with what they know and don’t know about the situation and must make a risk based decision on when to proceed into the intersection. The driver knows that they cannot see far enough around the curve to see a vehicle that may be coming at a higher than expected speed in time to avoid a collision. Here the driver often uses what they do know to limit (or place a boundary on) what they don’t know so that they can arrive at a decision that will result in a safe crossing. For instance, if the driver is very familiar with the intersection, then they may know that the traffic signals for oncoming traffic are timed such that there is a big break in the traffic where they can safely proceed. The driver may also know from previous experience that a rapid acceleration will shorten the time they are exposed to a potential collision. The driver may even be experienced as to estimate what is coming around the bend next but they may feel confident that they have bounded what they don’t know and this allows them to proceed when all the criteria is met. Once a person feels they understand the risks of a situation they will often accept that risk based on what they know and what they know they don’t know.

3. **What we don’t know we don’t know about a system (unknown unknowns)**

However, there is still a third dimension that affects risk outcomes - what we don’t know we don’t know (the unknown unknowns). For instance, in the case of the strength of materials of the Columbia wing leading edge, margins of safety were calculated based on expected environments and known load cases that were thought to cover the worst load cases that the spacecraft would ever see. Load cases, however, arose unexpectedly, in flight, that were outside the capabilities of the models used to derive the margins of safety. Findings per the CAIB report²:

- **F3.8-1** The impact test program demonstrated that foam can cause a wide range of impact damage, from cracks to a 16-by-17-inch hole.
- **F3.8-2** The wing leading edge Reinforced Carbon-Carbon composite material and associated support hardware are remarkably tough and have impact capabilities that far exceed the minimal impact resistance specified in their original design requirements. Nevertheless, these tests demonstrate that this inherent toughness can be exceeded by impacts representative of those that occurred during Columbia’s ascent.
- **F3.8-6** NASA’s current tools, including the Crater model, are inadequate to evaluate Orbiter Thermal Protection System damage from debris impacts during pre-launch, on-orbit, and post-launch activity.”

There are many complex reasons for the decisions that resulted in the Columbia Accident and it is not within the scope of this paper to evaluate them. However, the decision making portion that involved the amount of knowledge the experts thought they knew about the system illustrates that some portion of risk-based decision making is made based on how much we think we know about a system. In fact, per F3.8-2 above, the CAIB testing revealed that the structure of the Reinforced Carbon-Carbon (RCC) panels and wing leading edge was even stronger than experts believed. Thus, because of the beliefs in the strength of the structure and the worst case loads expected from a foam debris impact, it was not originally deemed credible that a hole would be found in the wing leading edge. It turned out that the foam debris environment imparted a much bigger impact load than was initially believed. In other words, the circumstance that they “didn’t know what they didn’t know” about foam transport outside of the models ability to characterize presented a greater risk than was originally assessed.

Similarly, considering the Apollo 1 fire accident and its subsequent investigation, one can also quickly see the unexpected outcome was due to what engineers “didn’t know that they didn’t know.” From the Apollo 1 fire US Senate Investigation³:

> “Senator SMITH. What I am trying to get is, where the error was, where we slipped up in not having or taking every precaution before we had that test. I do not see why we would not have precautions in testing before flight.

³ Columbia Accident Investigation Board Report, Section 3 Findings
Dr. THOMPSON. Well, I guess it is a matter of judgment that was made relative to that flight. Maybe I had better ask Colonel Borman. He was going to fly in a Block I spacecraft and he was prepared to go although knowing right much about this. I think we had better let him comment on that.

Colonel BORMAN. Yes. I think, Senator, we were very aware of the problem of fire in flight and we had adopted procedures primarily of venting the command module to a vacuum to eliminate the fire. We had done an extensive study on this before our Gemini 7 flight. However, I think that none of us were fully aware of the hazard that existed when you combine a pure oxygen atmosphere with ignition, and so this test, as I mentioned briefly during the findings and determinations, was not classified as hazardous. I do not believe that anyone within the test organization or the program office considered it hazardous. And, this is the unfortunate trap through which we fell. 

Here it is also clear that the perception of the risk and the actual risk were not aligned despite the fact that the managers and engineers involved believed they understood the risk. They “didn’t know what they didn’t know” about pure O$_2$ in the test model, so their perception of what they didn’t know was that the risks were acceptable. They later learned that their assessment of the risk was not complete.

Per NASA’s Risk Informed Decision Making Handbook, NASA/SP-2010-576

“Risk: In the context of RIDM, risk is the potential for shortfalls, which may be realized in the future, with respect to achieving explicitly-stated performance commitments. The performance shortfalls may be related to institutional support for mission execution, or related to any one or more of the following mission execution domains: safety, technical, cost, schedule.

Risk is operationally defined as a set of triplets:
- The scenario(s) leading to degraded performance in one or more performance measures,
- The likelihood(s) of those scenarios,
- The consequence(s), impact, or severity of the impact on performance that would result if those scenarios were to occur.
Uncertainties are included in the evaluation of likelihoods and consequences.”

In other words, NASA defines a risk as the potential to get an undesired outcome from a design or circumstance and is measured by the severity and likelihood of the undesired outcome. Thus, when making a risk-based decision one of the key parameters is the likelihood of the outcome. Estimating the likelihood of the outcome is ultimately influenced by perceptions because what we know about the design or circumstance gives us a very solid basis for believing we can predict the outcome, but what we know we don’t know and what we don’t know we don’t know introduce uncertainties in that prediction.

For example: (also from the Senate Investigation of the Apollo 1 Fire) we see that the decision on the acceptability of the risk for entering the spacecraft was dependent on how much they thought they knew:

Senator Percy. Colonel Borman, you mentioned before that you would not have hesitated on this fateful day to enter the spacecraft yourself knowing what you did at that time. I now ask the obvious question. Knowing what you know now, would you have refused to enter the spacecraft on that day?

Colonel BORMAN. Yes, sir

More accurately their original decision to accept the risk of entering the spacecraft for the “plugs out” test was really impacted by what they didn’t know they didn’t know. There is a subtle but important distinction here. We use tools, logic, and most of all experience to help us understand a risk. But saying we “understand a risk” implies we know everything we need to know about that risk. People can only reach a point at which they perceive or estimate they

---

$^\dagger$ Excerpt from Dr Floyd Thompson & Colonel Frank Borman Testimony to US Senate committee Investigating Apollo 1 Fire
know enough to make a decision. The subtle difference here is that when we reach that point where we think we can make a decision, it is because we believe we know enough, and this point is different from when we actually know enough.

In fact, since we cannot know everything there is to know about many risks, it is very difficult to ascertain that we actually know enough. In order to cope, our minds actually interpolate or “fill in the blanks” according to our personalities and experience. We, in fact, unconsciously make assumptions to make our world manageable. Engineering handles the unknown by applying margins to the design or circumstance as a way of “filling in the blanks”, but how much margin is adequate is often debated because even when one adds margin they cannot know what they do not know and so may not know that the margin covered a relevant unknown case. So is it then appropriate to add infinite margin? Certainly not! Fortunately, we do know that when margin becomes excessive you stop being able to perform the primary function. However, it could be valuable to investigate how we approach the unknown unknowns and thus reveal if there are additional ways to manage these unknowns.

Characterizing what we don’t know we don’t know varies greatly among people. A person’s environment, personality, social interactions and even their own belief system can influence how this third dimension is perceived. Because so many influences affect how this dimension is perceived, there is significant variation in acceptability of a given risk. Actually, the influence of the unknown unknowns is so profound that even when trying to define boundaries for what risks should be considered, an evaluator is presented with terms such as “credible”. The Merriam-Webster dictionary defines “credible” as “offering reasonable grounds for being believed”. With subjective terms like “credible”, “uncertainty”, and “believed” as the operating parameters for risk evaluation it becomes easy to see that it’s what we don’t know that shapes our ultimate risk-based decisions. Even the basis of the risk evaluation come from what we “believe to be the case”, not from what we know. Stated another way, if you knew everything about a risk the outcome would be 100% certain. Hence, there would be no risk at all because risk arises from uncertainty and uncertainty arises from what we don’t know, not what we know. For an extreme example of this consider that an astronaut while in a space suit can chose at anytime to “accept the risk” and remove the helmet from their spacesuit while in a space vacuum environment. This is not a real risk because the outcome is 100% certain for that action. It’s the uncertainty in the cases where the astronauts helmet could be inadvertently removed that give rise to a risk to the crewmember from a hazard analysis perspective. Here again the crewman may “accept the risk” and go into the space environment knowing that there is some small chance their helmet could be inadvertently dislodged. It’s the uncertainty that drives the risk and subsequent decision not the certainty of the well understood cases. Knowledge removes uncertainty, how much of that uncertainty is removed is subject to interpretation by the individual and their interpretation is influenced by their perceptions of the uncertainties. Thus, it becomes important to understand how those perceptions are shaped and influenced so that better risk-based decisions can be made.

D. The Role of Verification

The hazard analysis process relies extensively on the philosophy of “trust but verify”. The hazard analysis process is an iterative process that operates concurrently with engineering development cycles to identify, evaluate and mitigate safety risks in a timely and cost effective manner (see Figure 1). In order for the hazard analysis to close a hazard (aka safety risk), it must have evidence of the existence and effectiveness of the design features (hazard controls) that are being utilized to mitigate the hazard. Verification in effect increases what is known about a design or circumstance so that some level of confidence may be gained that the hazard mitigations will work when needed.

Verifications take the form of analyses, tests, demonstrations or inspections. Any of these can be used in any combination in order to establish the presence and effectiveness of the hazard controls. Verification in the field of space systems engineering covers two important processes: qualification and acceptance. These processes establish two things. First, they “establish what we really do know about a system.” Second, they “bound what we don’t really know about a system.” Establishment of what we do know about a system is fairly easy to understand as long as the developer had the forethought to identify the constraints under which the hazard control design is valid. Bounding what we know we don’t know about a system can also be achieved with the proper verification approach. For example: in the prior case of the driver attempting to enter an intersection, the driver would had to have experienced the traffic light further beyond the curve to be sure it was there and working in a way that will give them a safe break in the traffic. Periodic reinforcement is essential to giving users confidence in hazard controls. So periodically that driver would have to pass near or through that traffic light to verify it’s still working effectively to mitigate the
potential for a collision. The knowledge of the traffic response to the traffic light gives the driver some capability to limit what they don’t know before proceeding into the intersection. Thus, verification removes some unknowns and establishes confidence in our perceptions of a risk. However, unknown unknowns still exist and even the size of the set of unknown unknowns is not known. In the worst case the set of unknowns is infinite.

III. Risk-Based Decision Making for Hazard Related Risks

A. Personality Influences

Review of Literature

Psychologists today have identified what’s called “the Big 5” personality traits that form the basic traits of our personalities: Openness to Experience, Conscientiousness, Extraversion, Agreeableness and Neuroticism. Openness to Experience means a willingness to experience new ideas, concepts, and experiences and involves imagination and insight. Conscientiousness means to be organized, disciplined, dedicated to goals and thoughtful of others and mindful of details. Extraversion means a high degree of socialization, talkativeness, emotional expression and assertiveness. Agreeableness means to have trust, altruism, affection, and kindness. Neuroticism means to worry, experience anxiety, not be emotionally stable, as well as being unhappy and irritable. Biais and Weber found a strong correlation between risk perception and risk taking via the use of a modified DOSPERT Scale. They were able to measure the correlation between the perception of greater risk and the decreasing willingness to take risks. Furthermore, they assert that risk attitude is a personality trait that can be correlated to risk taking. Nicholson, Fenton-O’Creevy, Soane, and Willman found that increased risk taking is patterned on personality traits of “sensation seeking” and “value openness” especially when combined with influences of low neuroticism and agreeableness which reduce our concern for negative consequences. Moreover, they also found that personalities with high conscientiousness tended to strive for benefit through rigorous disciplined processes where as personalities with low conscientiousness tended toward “get rich, quick” schemes. Hung and Tangpong also

Figure 2. The Johari Window

American Institute of Aeronautics and Astronautics
found that openness is an influence for risk propensity and that another factor known as ambiguity tolerance influences our willingness to accept risk. The findings from the research indicate that risk-related decision making is influenced by our personalities.

NASA, in the context of developing its managers and leaders, has utilized the Johari Window in its training programs for several years. The Johari Window is a cognitive psychological tool, created by Joseph Luft and Harry Ingham in 1955, that was developed to show how our communications and interpersonal interactions are influenced by what we know and what we don’t know. The Johari window identifies four areas that are influenced by knowledge and experience. The Open/Free area represents information we know about ourselves that is also known by others. The Hidden area represents information we know about ourselves but keep from others. The blind area (blind spots) represents information about ourselves of which we are unaware, but it is known to others. The unknown area represents information that is unknown to the self and to others. When generalized for the context of risk, the authors believe that the Hidden (the known unknowns), the Blind (the unknown knowns) and Unknown (the unknown unknowns) areas all represent various components of personality that affect risk-related decisions. Through the mechanisms of feedback, vulnerability and courage, the self can begin to journey into these areas with the express purpose of shrinking them and moving their contents into the Open/Free area.

A Total Risk Model

The Johari window concept can be used as a basis to form a total risk model. Figure 3 gives a simplified view of factors that influence a risk-based decision. One can look at the total risk picture as the sum of the triangular wedges of a wagon wheel. The total risk picture is known if an assessor understands the total sum of what they know, what they know they don’t know and what they don’t know they don’t know. This model can be used to illustrate how the three factors that influence our risk-based decision are perceived. This will be done by coloring the wedges of the wagon wheel in accordance with how different parts of our personalities view a risk and showing what the wagon wheel might look like when a particular part considers a risk acceptable. (Note: the Johari window considers another group called “Unknown-Knowns”. These are facts that we know but suppress and are not considered in this discussion as they are a conscious or unconscious deliberate attempt to cover or suppress information. Engineering analyses are based on openness of data and is invalidated regardless of how one views any other type of information.)

Given that different parts of our personality respond to risks differently, the next section will present a brief summary of the perception tendencies for different parts of our personalities (aka thinking styles) that influence our risk acceptance decisions. For each figure, “Red” indicates unknown facts about a risk, “Green” represents what is known about a risk and “yellow” indicates what we know we don’t know. Graphics are drawn in a spectrum of colors as there is very little hard and fast delineation between how personalities view what lies between the wagon wheel pieces.

The parts of our Decision Making Personality

The “Big 5” personality traits give rise to thinking behaviors that can be labeled as ways we think about things. The personality traits can be divided and re-combined in a myriad of ways to gain perspective. But once established each resultant personality (thinking style) can be evaluated for tendencies toward a particular response given a
particular situation. It is important to realize that each part of our personality has tendencies that can be labeled for the purposes of understanding influences on decisions but no label will describe a single personality part accurately. All people are made-up of a spectrum of the personality parts which vary in degree of emphasis from moment to moment. This paper is based on the premise that we are each made up all of the personality parts and we utilize the results of each part in varying proportions depending our own higher level tendencies and how we view the risk to arrive at a final decision. This paper will discuss some of the more common personality traits and look at how we can use the knowledge of these influences to make better risk-based decisions. The seven proposed traits that will follow differ from the broad all-encompassing “Big 5” personality traits discussed earlier. The authors set of seven risk-related traits are the result of observation of decision makers in real world decision making processes who make risk-based decisions frequently. It is not important that the labels themselves are exactly correct or that the observed tendencies for a particular label are exactly correct. This proposed paradigm and the resulting tendencies are a construct to help the readers see how risk-based decision making is influenced. Once decision making is influenced and understood, recommendations can be made on how to make more informed risk-based decisions.

1.0 Analytical

![Analytical](Figure 3.1)

The truly analytical part of our personality wants to study each factor going into a decision. This part of our personality is immensely logical and helps us by working to build a sound chain of observations and conclusions that result in a well considered decision. This part of our personality continuously seeks more information to close any potential gaps in our knowledge that might represent an increase in likelihood of realizing the undesired consequences of a risk. Its nature is to see the unknown as a potential large risk and thus finds risk acceptable when it feels it knows all it can about a subject. Figure 3.1 shows that the analytical portion of our personality wants to make a decision when nearly all unknowns have become knowns.

2.0 Rule Based

Another part of our personality seeks to compare a risk to some already accepted standard. If something or someone else has made the same or similar risk-based decision before, it gives us a point of reference from which to judge whether a particular risk is acceptable or not. Risk-based decisions then become criteria based via rules. This approach to decision making tends to give a more “black and white” view of the risk. So if it meets the criteria, the risk is considered acceptable and if it does not then the risk is not acceptable. Once satisfied that the risk meets the criteria there is generally minimal perception of additional (unknown) risk. Conversely until the criteria are met, a risk is considered undecided or unacceptable by this part of our personality.

3.0 Negotiator

![Negotiator](Figure 3.3)

The negotiator part of our personality seeks expertise and other social sources for advice on the best decision. The negotiator will weigh information and the credibility of the sources and will generally seek to find the norm of information. The negotiator will try to find the common ground between different views and then build a platform upon which consensus can be built. The negotiator can and will change its decision based on new information. Thus, this part of our personality acknowledges the presence of unknown unknowns and attempts to get the consensus of the group to agree on the implications of those unknowns. Once consensus is achieved the negotiator part of us perceives the risk adequately mitigated.

4.0 Emotional

The emotional part of our risk-based decision making derives information from internal cues on how we feel about a risk. Our prior experiences and observations plus our state of mind at the time of the risk-based decision have a lot of influence on how we will respond to the risk. Football players spend a lot of time and energy “working themselves up.” This heightened emotional approach gives them an edge to be more aggressive when taking risks.
with the game and their bodies. This allows the small wide receiver to run into an open field full of much larger men and catch the ball knowing that they will be hit by several big opponents. Very few people would do this on a normal day, but the emotional decider will use the emotion to help them accept the risk and do things they probably wouldn’t do in other circumstances. Those in military or battlefield conflicts also utilize this trait to a large extent. This trait generally will accept the risk regardless of what it doesn’t know, based mostly on the perceived reward. Conversely, they may reject the risk on the same basis. This part of our risk-based decision making can also influence how we ultimately integrate the other parts of our decision making. For instance, if we do not feel confident about our analytical assessment of a risk we may feel much better utilizing our negotiation capabilities and thus give more weight to that assessment.

5.0 Common Sense

As briefly touched on in the prior section, aspects of our personality also affect each other when making a risk-based decision. The emotional example is the easiest to understand. However, each trait can influence the other. How they interact gives rise to the hybrid portions of our decision making. For instance, the common sense decision maker is a hybrid of the traits above. This trait utilizes external information in the form of rules, data, and general social responses with their own personal emotional responses to arrive at a decision that is practical and immediately implementable. This is the trait that often says “if it walks like a duck and quacks like a duck then it’s a duck”. This trait uses very basic information, assesses the information, utilizes social and emotional input to arrive at what they believe is the best decision. This trait usually does have a strong notion of right and wrong and draws strength from these rules. The major difference between this trait and the rule follower is the rule follower will have numerous rules to apply to every situation and the more intuitive “common sense” trait will have only a few general rules and will use other clues to guide their decisions.

6.0 Reasoner

The reasoner trait will expend energy in the time available to gain an understanding of why one course of action will be better than another. This part of our personality will use all the tools available to find a reasonable answer to a risk mitigation concern. They include: rule following, analysis, emotional cues, expert and social input to arrive at a decision. This personality trait includes a sense of timing. This trait will allow analysis to go on within the time available. This trait will also make a very quick decision when time is short. The true reasoning trait of our personality has little or no fear of being wrong. All the previous discussed decision traits rely on internal or external cues to measure the rightness or wrongness of the decision. Those traits arrive at what they believe is THE decision. The reasoner trait however is often characterized as seeing the world in shades of gray as opposed to seeing absolute right to absolute wrong decisions. The reasoner trait sees decisions in terms of the conditions that exist at the time the decision is required and how those decisions will affect future goals and objectives.

7.0 Directive

The directive part of our personality is goal-oriented and uses our other personality traits to makes decisions relative to the ability to accomplish a goal. The directive trait will dismiss or minimize other factors that do not seem to contribute to that goal. The directive trait tends to be seen as cold, bossy, or heartless as it often sees emotional factors as excess baggage that hinders efficient decisions. However, our directive trait is also seen as the “get the job done” trait. Our directive trait tends to take a big picture view of the situation and will make assumptions and sometimes generalizations to keep things moving toward a goal.
It should be noted that each person is made up of a mix of the above traits that is spread across a personality spectrum. These traits are not necessarily independent. These traits work together or sometimes separately to influence our risk-based decisions. It’s important to note again that no person is confined by nature to purely one trait. The authors believe that all our traits are influenced by our genetics, our meta-genetics, our volitionary natures and the environment. Over time, these influences strengthen some personality traits over others. This strengthening, in time, may become more “hard-wired” so that some particular traits subsequently develop into behavioral tendencies which if repeated again and again will manifest in the form of character. It is true that our individual way of thinking often favors one or two of the traits. In general, we usually associate a person with their dominant personality trait (and perhaps their secondary trait) when, in reality, each person is comprised of multiple traits that manifest themselves differently over time depending on the internal/external influences.

B. Personality Impacts on Risk Based Decisions

Meetings are notorious for being long and boring. The reason is most people are not active participants of any single conversation topic and thus, don’t want to “speak out of turn”. But in a decision making meeting all inputs are needed, especially in a risk-based decision meeting. Each part of our personality has limitations so naturally our resultant total personality will have blind spots due, in part, to how we think. Most often it is difficult for us to see our own shortcomings accurately. David Dunning offers many insights into why we don’t perceive our own weaknesses well and through a series of experiments shows that we have a tendency to overestimate our own capabilities but are fairly accurate in our assessment of the capabilities of others11. These blind spots and biases play a role in risk-based decision making, because it’s our personalities that influence how we “fill in the blanks”.

NASA meetings often end in requests for more information before a decision is made. This is a natural outcome due to the makeup of various NASA teams. Engineers form the vast bulk of the NASA workforce and by and large they are analytical in nature. This is a very good trait to have when designing one of a kind hardware that is used in hazardous operations. Some of the more social personalities at NASA, who learn to balance analytical traits with other traits, find themselves in leadership positions. They are often good negotiators who find “sensible” solutions when confronted with many options. However, many development meetings do end in “indecisions” with subsequent requests for more information. Based on the above, the dynamics of these situations are easy to understand. A large group of analytical thinkers come together with a few negotiators to make a decision. The analytical types are reluctant to commit to a decision without further information. The negotiators are trying to find common ground and often resort to issuing action items for more data so that the team can focus its efforts and move forward. However, when the group is reassembled it is often the case that new requests for information are made and more time and cost are expended on the questions. This process is often repeated many times. For times when the question is not safety-critical, there’s often a missing personality to help the team move forward even in the midst of uncertain or missing information. This missing personality type is the directive person. The directive person would likely have a tendency to make sure a question is explored but when it becomes repetitive will often make a decision themselves and move the team onto new problems. The addition of another personality type can often change the tone of the room and the level of engagement of the personnel. Similarly, the reasoner type can also contribute as they will try to assimilate the new information and make a time based judgment on what should be done at that point in time.

Thus, changing the makeup of the participants in the meeting can greatly influence the thoroughness of the assessment and the time it takes to make a proper risk-based decision. For example: consider a room full of people of primarily directive type traits. Each is self directive, self reliant and confident in their choices for a best solution to a given problem. Each will utilize their other personality trait in varying degrees as observations to “fill in the gaps” for what they don’t know and in the absence of further direction or information when faced with a deadline will make a decision and help his or her team move forward. Except, unless a very unusual situation arises, they will each have different backgrounds and biases and each different perspectives and so are likely to reach different conclusions about a risk. They will each defend those conclusions ardently. With no negotiator to help the decision making process this aggressive group is unlikely to reach a consensus. It’s the combination of analytical, reasoner, emotional, and directive traits that result in a productive meeting.

NASA is very good at encouraging diversity but the diversity used in technical meetings is mostly “did we talk to the right people?” or “Did we get someone from every engineering discipline to look at this risk and give us an input?” In other words, NASA sometimes utilizes functional diversity and thus asks for inputs from more analytical
types. The consequence remains that NASA seldom get inputs by deliberately considering diverse personality types. However, it’s not reasonable to expect NASA or any developer to build a workforce of completely diverse personalities. This is because they must hire engineers into engineering positions and engineers coming out of school or industry are highly analytical by nature. In addition, NASA managers have to find a candidate that will fit the job description. Then there are the legal aspects like the Genetic Information Nondiscrimination Act (GINA) of 2008 that makes job consideration based on traits that come to us via genetics illegal. Our total personality can be attributed to our genetic makeup and so making board or panel selections based on personality criteria could be a precarious proposition at best for a NASA manager.

C. Risk Informed Design

Like many industries, NASA adheres to internal policies and practices regarding risk assessment. This includes tools and frequent reviews to stay aware of programmatic risks. However, risk-based decisions are often very slow in coming as is evidenced by the lengthy durations many risk items remain open. Even after making risk-based decisions, the ultimate outcome of the risk can be very different than the risk initially forecasted. This is often due to an inherent inability to assess the true likelihood of an undesired event.

The Constellation program tried to deal with risk using a new paradigm. The program decided that instead of establishing an essentially rule based risk mitigation process where the severity of the risk alone dictated how much design redundancy was required, the total level of the risk itself would be used to drive the robustness of the design. The concept was called Risk Informed Design Making process (see Figure 4). The premise of this process allows each risk to be judged on its own merits and mitigations are developed based on those merits. NASA’s use of consultation and deliberation are keys in this processes. However, a challenge arose within Constellation due to this approach. In short, while highly successful as a concept, implementation was difficult. Risk is a quantity defined to be the result a combination of likelihood and consequence. But in new one-of-a-kind designs there’s no real demonstrated history to quantify likelihood. So the program may choose a “one failure tolerant” solution to a risk because they believe that it is not high, but there is no concrete data to support that assessment until after the
verification program begins. And the verification program cannot begin until a design is selected and hardware built to test. The verification can show that the original judgments were very wrong or very right, but in the early stages of design and development there simply is no way to know.

This introduces a new problem. When there are essentially two or three personality types available in a NASA environment, the community can become polarized around particular risks. The analytical types will view the unknowns as a significant threat and the directive or emotional folks will sometimes view the unknowns as trivial. There is no data to support either view and so risk items get identified but cannot close, hazard causes get opened but no one can agree on what is sufficient mitigation for them and the community becomes unable to proceed efficiently through design discussions. Engineering decisions become much more reliant on management judgment calls and not on development testing that characterized engineering decades ago. Risk Informed Decision Making is actually a sound and practical design approach but it has limitation when it comes to unique one-of-a-kind hardware operating in unique environments.

D. Risk Quantification

Entire industries have been built up based on risk quantification. Modeling techniques involving very sophisticated combinations of parameters and risk evaluation approaches have been built to gain a better understanding of risk for a particular industry or function. Increasing knowledge about a risk does increase reliability of a risk prediction being correct, but the reliability of that assessment decreases rapidly as the amount of uncertainty increases. Blind adherence to a model can lead to erroneous conclusions and so the limitations of a model must be understood. In other words, if you don’t know what you don’t know then how can you know the model limitations are not affecting the outcome? Our perceptions about the model and what we don’t know we don’t know can lead us to different conclusions based on our backgrounds and experiences.

This is not to say risk likelihood modeling has no value. On the contrary, especially if used to predict the outcome of two very similar designs it can be very successful at determining which design is more reliable. This is because when comparing two very similar designs used in the same way in the same environment you have dramatically limited the cases that add variation. Because you have a high correlation between the two designs you also have a high correlation of unknown unknowns that could influence the outcomes and so significantly reduce the number of unknown unknowns that could impact the outcome of the comparison. Thus, the reliability of the predictions increases dramatically. You may not understand the total reliability of either system but you can much more confidently see the difference in reliability between the systems.

Risk quantification is so inherently difficult to get right that a major disaster like the Columbia accident could not be accurately forecasted. Debris generation and transport models already existed, the hazard analysis already identified that debris impact on the Orbiter thermal protection system could be a catastrophic hazard, but the likelihood of a catastrophic hazard due to foam strike was considered as “infrequent” which was the third highest of the four likelihood categories for Shuttle hazard reports. Today, it is considered the most likely cause to a potentially catastrophic event because we have more data. Whether either answer was right or wrong is not important to this paper, what’s important is that no matter how you mathematically characterize a risk it’s still our perception of the residual risk and our perceptions of the applicability of the model and its uncertainties that give rise to our decisions.

E. Recommendations for Addressing Risk

NASA has been in the risk business for decades and has reaped the benefits and suffered the consequences of risk-based decision making. Since perception of the risk associated with what we don’t know we don’t know influences our risk-based decisions, we now need to establish how this knowledge can be used to the benefit of a program. There are two ways this knowledge can be used: 1) diversify the set of personality traits involved in making the risk-based decision and 2) strategic safety margin management.

1. Diversify the Set of Personality Traits

Since risk assessment comes down to estimating likelihood and data needed to make an effective likelihood assessment, it follows that good risk-based decision making relies on getting the fullest picture you can of the likelihood of the risk. Also, since the risk is really our perception of the risk, and different personalities will see the
same risk differently, it becomes important to get multiple personality type perspectives. This is very similar to how NASA creates boards with multiple discipline perspectives for making decisions (functional diversity). The Rogers Commission Report on Challenger recommended that astronauts be included in the safety decision process. Our concept would take this one step further and the major risk-based decision boards and panels could be evaluated for their personality mixes. This proposed personality mixing should achieve a balance so that the manager can utilize multiple perspectives to reach a better decision. This is more subtle than just saying safety should be there and speak, and design should be there, and operations, etc. It’s saying that the board or panel needs to include as many diverse decision making types as is reasonable to get a more comprehensive risk picture.

Analyzing this proposal, one must determine what is practical. Per the GINA act of 2008, using personality considerations to select board members could be considered illegal. Furthermore, it is not practical to expect to derive a board of decision makers that represent every personality trait AND every necessary technical basis for a decision. However, it is important to mitigate the natural “blind spots” we have in our individual ways of thinking. To resolve these competing goals the manager could at least become aware of what personality traits their thinking most resembles and consciously take steps to assure their blind spots are addressed. For example, if a team/board or decision maker is made up primarily of directive tendencies then making sure extra time is taken to ensure some negotiation between the different supporting analytical groups has occurred and finding out what conditions or reservations those groups had regarding the decision could be enlightening. This is very consistent with the CAIB board findings. This is not easily done by a group who wants to make a decision and move on but the benefit could be avoidance of a downstream problem. Simple steps like requiring assumptions to be shown with any analysis results summary, or stating boundary conditions for any analysis, and bringing forward minor concerns would go a long way to towards keeping eyes open to potential blind spots and thus potential unknown-unknowns. The majority of engineering time and effort goes into demonstrating how well we know what we know. This is not wasted effort but characterizing the boundaries of what you don’t know about a risk will give a lot of additional value. Analysis efforts and tools need to be matured to better capture this aspect of risk.

2. Strategic Margin Management

A second process enhancement mitigates the challenges of early risk-informed design process limitations as well as addresses holes created by the structure of the analysis. The development of new, one-of-a-kind hardware takes away some of the benefits of Risk Informed Design and adds cost in the early stages of a program due to irresolvable debates about what level of risk the design actually presents. In addition, it puts the program at a great degree of “risk” for late redesigns to major systems when risk assumptions are invalidated by actual test and analyses. Since validated likelihood data is not available (and will not be for years) during the early program stages, an alternate way to approach program safety risk is to start with an adaptive rule based approach. Blind adherence to a standard is not effective in engineering a robust but high performance space vehicle, but, relying solely on model data that has not been validated is not wise either. This is because there is considerable risk that some assumptions will be proven wrong much later in the program life cycle and design changes will be required. Instead, by acknowledging that unknown unknowns exist, a program can take a different approach to managing those same risks. As an example, in the initial design stages all functional hazards deemed catastrophic could be designed to be “two-failure tolerant” without exception. Then, as the design details mature and likelihood of the risk gets some validation, the necessary levels of failure tolerance or other mitigations may be removed where warranted. Uncertainty is removed through knowledge and testing, demonstrations and analyses can impart additional knowledge that is sufficient to show that hazard controls can be effective. Once validations are completed there is hard evidence to justify lowering levels of failure tolerance, because the expected performance of the controlling equipment has been shown with some degree of confidence. Effects of this approach are shown in Figures 5 and 6.

Figure 5 shows the relationships when risk is managed by Risk Informed Design. Because large-scale integration programs are one-of-a-kind unique endeavors, there doesn’t exist a pool of real flight vehicles from which to assess the risks associate with a new vehicle design concept. Reliability and probabilistic risk assessment (PRA) engineers rely on two processes to help them estimate the actual undesired event likelihood. They either use data from prior programs to simulate today’s performance or conduct Bayesian type analyses to use performance data from other industries to get a likelihood of the particular concern. Using Apollo, Shuttle and Space Station data where appropriate is the best NASA can achieve, but the number of flights across all these programs together is relatively small. Thus, the set of unknown unknowns may not be well characterized. As shown in Figure 5, as the verification program begins, the amount of uncertainty decreases and the models become more refined and are later used to
Figure 5: Risk Profile - Risk Informed Design Decision Making.

Figure 6: Potential Risk Posture when unknown-unknowns are taken into account.
validate the actual total risk margin on the program. However, since the amount of unknown unknowns is not understood and there is much complexity in large scale integration, it becomes apparent that this path has a significant vulnerability to unexpected changes later in the program. This is especially true for cases where the final verification finds that an assumption about the design was not valid.

Figure 5 also illustrates one of the goals of the analytical tools is to show that the program meets some minimum level of acceptable risk for the integrated design. Since it’s trying to measure relative to the minimum it has the unintended consequence of pushing the design toward to the point where it mitigates the minimum known of hazards/risks of a design. This approach literally pushes the design away from higher margins throughout the program life cycle. The approach creates a hidden assumption that the design is always near compliance with the minimum acceptable level or risk. It does not acknowledge that there is risk in the realm beyond its boundary plus its margin, and thus constrains what the program will consider for hazard controls. However in the early stages of the program this hidden assumption cannot be validated.

Figure 6 shows an alternative approach that provides more robustness against later program changes. This is done by assuming the set of unknown unknowns to be much larger than the prior example. In this case the designers start with the assumption that large margin is needed to cover all cases. If nearly simultaneously the designers also begin testing and test based analysis that validates how the design of a system/subsystem will perform. If that performance is shown to be above the minimum level of acceptable risk line, the design for that system or subsystem can safely be reduced to one failure tolerance. And if additional test based verification shows that the one failure tolerant design is still well above the minimum acceptable risk line, the design concept can safely be reduced to zero failure tolerance. Managing Risk Informed Design via incremental risk reduction this way helps to mitigate the potential costs of modeling a system and then later finding out an assumption was erroneous and having to do a redesign. This is also known as “build a little test a little” risk reduction. The impact of program/project retrofit on program costs is available in the Seminar Series course “Extreme Software Cost Estimating”13. Course results shows that the effect of retrofit is so significant that even when work is spread out longer with lower staffing, the total cost due to rework is unacceptably high and the accompanying schedule delays are significant. Furthermore the authors show that our ability to estimate true program costs is historically (for Government programs) off by factors of two to four times the original estimates due to rework costs. Said simply, using an approach where risk buy-down must be validated before a design can have margins reduced (Figure 6) will save programs large amounts of rework costs that will inevitably be realized when the formal test program gets underway. This strategy appears to represent an advantage for reducing overall costs because it tackles unique, one-of-a-kind, large scale development risks with a strategy that formally acknowledges the existence of unknown unknowns.

IV. Conclusion

NASA has a long history with respect to risk-based decisions and has experienced the consequences of risk-based decision errors and enjoyed the success of correct risk-based decisions. No risk process is foolproof but these suggested modifications to NASA’s already robust risk-based decision processes should effectively increase the knowledge and experience base for future risk-based decisions and as those grow so can the quality of the risk-based decisions. Today’s space exploration programs are moving toward more “risk based design” approaches. Thus, risk identification and good risk assessment is becoming even more vital to the engineering development process. This paper explores the factors of known risk mitigations/verifications and unknown risks and ascertains how various personalities of risk-based decision makers are affected by what they know and what they don’t know. Giving consideration for mitigating some of the unknown unknowns can have very positive benefits by 1) exposing blind spots in our thinking that can lead to erroneous conclusions and by 2) including them in the margin assessments that give rise to high performing designs.

References


1 APOLLO ACCIDENT HEARINGS BEFORE THE COMMITTEE ON AERONAUTICAL AND SPACE SCIENCES UNITED STATES SENATE NINETIETH CONGRESS FIRST SESSION; TO HEAR MEMBERS OF THE APOLLO 204 REVIEW BOARD AND TO DISCUSS THE BOARD’S FINDINGS, DETERMINATIONS, AND RECOMMENDATIONS ON THEIR FINAL REPORT OF INVESTIGATION, PART 3, WASHINGTON, D.C., APRIL 11, 1967.


