Lessons from the Gulf Oil Spill: Black Swan or Black Sheep?

Dr. Adam Borison, Senior Vice President, NERA Economic Consulting
Dr. Gregory Hamm, Senior Economist, NERA Economic Consulting

Introduction

The recent spill from BP's Macondo well in the Gulf is generating a flurry of discussion about risk and risk management. Not surprisingly, this discussion reveals some very divergent opinions. Much of the discussion is dominated by two views. One view emphasizes the nature of the event itself; the spill was essentially a "black swan" – an unprecedented and unforeseeable tail event. Another view emphasizes the nature of the key protagonists; BP and/or the Minerals Management Service (MMS) were essentially "black sheep" – uniquely incompetent or even negligent organizations. These two views are very different, but they share a common implication. The Gulf oil spill was unique. We can point our collective fingers at nature or at BP and MMS, but there are really few if any lessons for the rest of us.

The facts behind the Gulf oil spill are not known at this time and will likely only emerge over months or years through investigations, disputes and the like. But even in these early days, a third view should be considered - that neither the risk nor the risk management in this case was unique. The event was not some black swan beyond the realm of predictability - unusual but not unprecedented. The protagonists were not black sheep behaving beyond the norm - idiosyncratic but basically following common practices.

This third view is both more constructive and more discomfiting. What we have here is neither a rare black swan nor a rare black sheep, but instead is a common seagull covered in crude oil. As such, the Gulf oil spill can serve as a valuable lesson for all of us involved in risk management. This paper uses the Gulf oil spill to highlight important limitations in current practice and makes specific recommendations for improving that practice by addressing those limitations.

Risk Assessment: Difficulty Dealing with Unusual Events

> *"The heart and soul of the whole issue is that this is a Black Swan event," Nansen Saleri, CEO of Quantum Reservoir Impact and a former executive at Saudi Aramco, told CNBC Wednesday. (Lodge, 2010)*

The U.S. OCS platform spill rate [for spills larger than 10,000 bbls] for the last 15 years, 1985-1999, was less than 0.13 spills per Bbbl handled. (Anderson and LaBelle, 2000)

Good risk management begins with good risk assessment; that is, understanding the likelihood and magnitude of uncertain events. Unfortunately, our experience indicates that current risk management practice has remarkable difficulty assessing unusual events. By unusual, we mean events that occur infrequently, even when looking broadly across time and space. This limitation is unfortunate because unusual but potentially significant events are precisely the kind that require the most attention, and where appropriate attention can produce the greatest benefit.

Current practice is dominated by the view that risk can only be formally assessed – that is, quantified - if there is a large amount of repetitive historical frequency data behind it. This results in a clear dichotomy. Where there is sufficient data, risks are quantified. Where there is insufficient data, risks are treated qualitatively. Typically, this means that everyday risks are quantified and unusual risks are not. For these unusual risks, risk assessment is limited to broad qualitative statements. This dichotomy is illustrated very clearly in financial risk management. Everyday risks are quantified extensively using statistical approaches such as VaR. Unusual risks are addressed qualitatively with heuristics such as stress testing. For more on financial risks, see Borison and Hamm (2010).

How is this difficulty revealed in the Gulf oil spill? Significant oil spills are fortunately not everyday occurrences, but they are certainly not unprecedented. Most lists of oil spills over the past several decades include dozens or more. Most are considerably smaller in size than the Gulf oil spill, many comparable in size, and some even larger. Perhaps most notable is the 1979 Ixtoc well blowout in the Gulf of Mexico that resulted in an oil spill not dissimilar in size to the recent event. Based on this history, the Gulf oil spill appears to fit our working definition of an unusual event very well. It is not a black swan, an event so rare that "nothing in the past can convincingly point to its possibility" (Taleb, p. xvii). On the other hand, it is not an everyday event with a wealth of relevant historical frequency data. As Eschenbach et. al. (2010) note: "If the data is used directly, then the point estimates [of large spills]…are unrealistically equal to 0." However, larger spills can clearly occur…"

Looking back, how was the risk of this unusual event assessed? MMS has a system for dealing with oil spills, including formal methods for estimating the likelihood and impact of spills (Anderson and LaBelle, 2000). However, these methods are built on repetitive historical data and are limited largely to everyday events. We find no evidence that MMS quantified the risk of a very large spill from the Macondo well, and MMS regulations do not require such quantification. Instead, when it comes to such unusual events, MMS instead relies on qualitative thinking. They require that developers specify a "worst-case" scenario and provide detailed instructions on how to construct it. See MMS (2007) for the instructions and Replogle (2009) for BP's submission. Consistent with standard practice, they do not require that this scenario be associated with a probability. Based on its "worst-case" submission, BP was aware of the possibility of a spill of tens of thousands of barrels a day or more. But we are unable to find evidence that BP formally assigned probabilities to spills of this size as part of the oversight process. Some have argued that protagonists in the Gulf oil spill underestimated the risk. The view here is not so much that they

underestimated it, but that they adopted standard practice in dealing with unusual events and did not really estimate it at all.

The downside of this qualitative approach to unusual events is considerable. First, with only a qualitative assessment, risk communication is significantly impaired. When we are talking about a "worst-case" event, do we mean 1/100, 1/10000 or even 1/1000000? Without quantification, confusion and misunderstanding can result. Furthermore, it is difficult to evaluate different preparations and responses effectively. For example, the appropriate preparation for a 1/100 "worst case" may be quite different from the appropriate response to a 1/1000000 "worst case." – a distinction that is lost when the assessment is qualitative. In summary, while the intent of this qualitative "worst-case" assessment is to minimize the chance of a disaster, we believe that it can have the opposite effect.

To address this shortcoming, we recommend taking a different – Bayesian - view. With a Bayesian view, risks are assessed by combining available data with expert judgment. Both data and judgment are used to quantify risk, and a statistically significant data set is not required. This Bayesian view is quite common in strategic and operational settings. It has penetrated risk management in spots but is far from the standard practice.

Figure 1 shows an example of the Bayesian view applied to the likelihood and magnitude of an oil import disruption, another unusual but not unprecedented oil-related event of great importance (Beccue and Huntington, 2005). This approach seamlessly combines both data and judgment in a formal quantitative probability and impact assessment. Each of the numbered green "bubbles" represents an individual risk factor that contributes to the possibility of an oil disruption. Some risk factors, such as oil production, are data rich. Other risk factors, such as Middle East conflict, are dominated by judgment. The arrows among these risk factors indicate correlation or dependence. The blue rounded rectangles indicate how these risk factors ultimately affect the likelihood and magnitude of the oil disruption. Such a rigorous risk assessment builds on the available historical data and the available expert judgment to produce a quantitative assessment of an important unusual event.

Current practice could be significantly enhanced by adopting such a Bayesian view in dealing with unusual risks. It is impossible to tell for sure, but we believe that events such as those that transpired in the Gulf could have evolved differently under these circumstances. Probabilities could have been explicitly assigned to oil spills of different magnitudes using both data and judgment. The ensuing discussion could have led to a clearer view of the risk, different risk management and communication activities, and ultimately to better outcomes. See Stewart and Leschine for one of the first discussions of the importance of judgment in oil spill risk assessment (1985) See Winkler, 2003 for a more in-depth discussion of the Bayesian view.

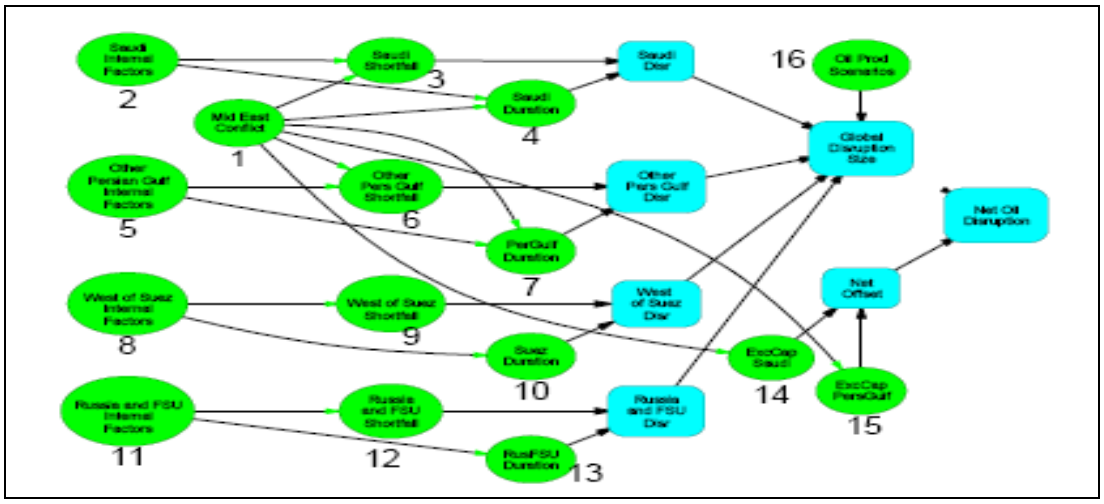

Figure 1. Bayesian Assessment of US Oil Disruption Likelihood and Magnitude

Risk Management: Difficulty Matching Responses to Impacts

> *"'Each of the oil companies' oil spill response plans are practically identical to the tragically flawed BP oil spill response plan,' the lawmakers wrote…No oil company appears to be better prepared for a disastrous oil spill than BP was." (Committee on Energy and Commerce, 2010)*

> *"Embarrassingly we found ourselves having to improvise on prime-time TV and slap bang in the middle of the glare of the global media." (Macalister, 2010)*

Risk management is only as good as the individual preparations and responses that are adopted. Unfortunately, our experience indicates that current practice does not devote sufficient time to elucidating the risks fully, uncovering the breadth of potential impacts, and crafting responses to match those impacts. Instead, current practice relies heavily on off-the-shelf alternatives for addressing everyday risks and is ill-suited to unusual events.

Off-the-shelf preparations and responses are widely available and generally effective for many of the everyday risks faced by organizations today. For financial risks, there are popular market hedges. For hazard risks, there are common insurance policies. For operational risks, there are accepted compliance guidelines. Even for strategic risks, there are industry norms. In current risk management practice, much of the effort is devoted simply to assigning these off-the-shelf alternatives to risks based on what is acceptable, prudent or required. In contrast, there is no systematic process for fully identifying potential impacts and crafting alternatives for those specific impacts, and many risk management alternatives are missed as a result.

How is this off-the-shelf orientation revealed in the Gulf oil spill? The quotes above suggest that, whatever the particular characteristics of the Macondo well, projected impacts and planned responses were largely off-the-shelf. Based on available documents, it appears that the Gulf of Mexico oil spill response plans for all the major firms were similar or identical in many respects

and applied broadly to all wells in the region. As events unfolded, responses had to be tailored to the unique characteristics of the spill on the fly.

In many circumstances, a standardized approach to risk management may be fine. Everyday risks that have common features across time and space often call for such off-the-shelf alternatives. For example, most of us would find an off-the-shelf insurance policy acceptable or even preferable. For unusual risks, however, this one-size-fits-all philosophy is likely to lead to problems and missed opportunities. The downside can be excessive cost, excessive risk or both.

In dealing with unusual risks, we recommend taking a different - decision analysis - approach to generating and evaluating customized alternatives to match the impacts in question. With this approach, rigor is applied not just to the assessment aspect of risk management but to the actual management aspect – the decisions that are actually made to prepare for or respond to risk. This decision analysis approach has become quite common in strategic and operational circles. It has not penetrated corporate risk management particularly well.

Figure 2 shows an example of this decision analysis approach in the context of another oil related risk issue - tanker safety. (Merrick, 2005) In this example, customized risk management alternatives are generated based on their ability to achieve the specific objectives of stakeholders.

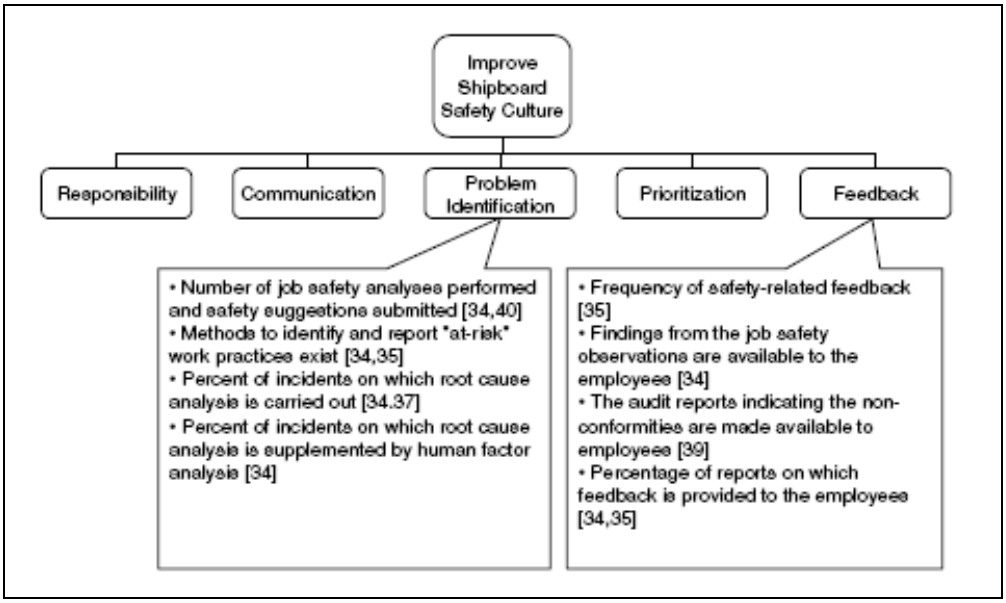

Figure 2. Decision Analysis Approach to Oil Tanker Safety Alternatives

Current practice could be significantly enhanced by adopting such a decision analysis approach to preparation and response alternatives. In our experience, added effort devoted to developing a more complete understanding of the specific impacts and to creating a robust collection of customized preparation and response alternatives has a high payoff. While it is impossible to tell for sure, we believe that the Gulf situation…and similar situations…could have turned out quite differently if protagonists had taken this path. For a more in-depth description of this decision

analysis approach to generating alternatives, see Keeney 1994. For a more general overview of decision analysis, see Clemen 1997.

Risk Communication: Lack of Systematic Learning

> *[Robert Bea, UC Berkeley Center for Catastrophic Risk Management]… laid blame for the Deepwater accident squarely on BP, and he says the company appears to have ignored a study he did for it eight years ago examining problems in the business' refineries." (Baker, 2010)*

> *Last fall BP's chief executive officer, Tony Hayward, and 25 senior staff launched the three-day Operations Academy Executive Program, designed to educate top management on how to lead BP to operations excellence in an enterprise that employs over 96,000 people in over 100 countries. (MITnews, 2008)*

Like many management activities, risk management takes place in cycles where increased understanding, better performance and higher quality can result over time. Unfortunately, current practice is largely static and there is little systematic learning. By systematic learning, we mean a structured process for explicitly comparing what was anticipated to what was realized, and correcting both assessments and actions as a result.

What does the Gulf oil spill reveal about this lack of systematic learning? BP has been cited over decades as a "learning organization" (Donegan, 1990), and even established a joint program with MIT in 2008 focusing on operations safety "designed to enhance the culture of continuous improvement at BP." (MITnews, 2008). BP has experienced its share of oil spills, refinery fires and related unfortunate events over this period. But as we have learned in many organizations, there seems to be a disconnect between the enthusiasm for learning and the actual improvement in key aspects of risk management. While the concept of learning seems particularly relevant when it comes to risk assessment and risk management, the great bulk of the literature in this field has very little to say on the topic, and no concrete recommendations on how to improve, refine and adapt risk assessment and management over time. There are only one or two works on the topic of learning from (other people's) disasters or catastrophes (Toft and Reynolds, 2006; Kunreuther and Useem, 2010), and these works provide little advice on organizational processes for improving with experience.

The downside is that the quality of risk management does not improve significantly over time. Each cycle of risk assessment and management starts largely from square one. Risk assessments can remain uncalibrated, and over or underestimates of risks remain uncorrected. Mistakes in risk responses can also be repeated, with little focus on what worked or didn't work in the past. Our experience indicates that many risk managers and senior executives are concerned not just that risk management stagnates but that it can actually decline over time as a result of fatigue. The net result is that risk management is considerably less efficient and effective than it can be.

We think that the keys to improving learning in risk management are to better define what we mean by quality in risk assessment and risk management and emphasize the importance of

systematic learning in improving this quality over time. Fortunately, there are signs of emerging interest in the application of learning explicitly to risk management. In his seminal book on the learning organization, Senge (2006) includes specific examples where firms have learned effectively from disasters. His work supports the idea that risk management can incorporate systematic learning by providing personnel with the appropriate resources – particularly training – and by adopting appropriate incentives to encourage learning behaviour over time.

Figure 3 shows how a clear definition of quality and a focus on systematic learning can provide substantial benefits in the challenging context of climate change. In this actual application at a water supply firm in Australia, the key risk of concern is the impact of atmospheric CO2 on water supplies. The quality of the assessment of this risk is represented formally in the 10-90 confidence interval regarding this impact. The figure shows the quality of this risk assessment in the 1960's was quite poor. At this time, there was considerable uncertainty over whether rainfall will increase or decrease (or remain unchanged) due to CO_2. With careful observation and analysis of rainfall data over time, the quality of this assessment improves. By the year 2000, the 10-90 confidence interval has been reduced by 80%. This is systematic learning at work. It is accomplished not just by adding new data, but by formally comparing what was anticipated to what was realized and using the difference to refine underlying assessments and models. The narrowing over time has greatly improved the firm's ability to manage and communicate this risk.

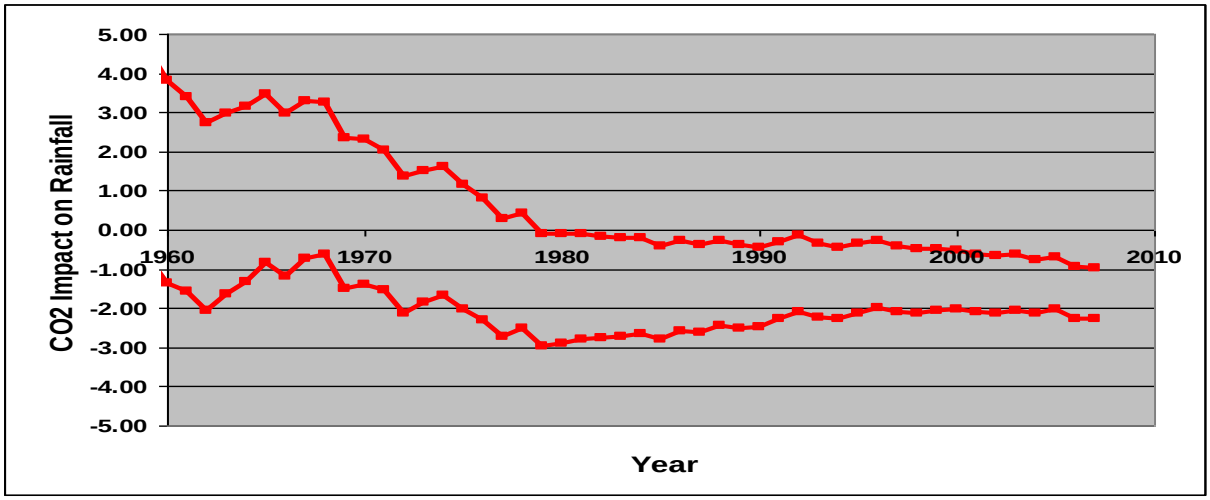

Figure 3. Systematic Learning regarding Climate Risk

Current practice could be significantly improved by incorporating systematic learning into risk management. Again, it is impossible to tell for sure, but it is certainly plausible that events in the Gulf would have transpired differently if protagonists had included a formal learning effort in their risk management activities. Careful comparison of past surprises and past risk assessments, in particular, could have illustrated more clearly that the level of understanding of oil spill risks and impacts was imperfect and reduced the possibility of overconfidence.

Conclusion

Black swan, black sheep or seagull covered in crude oil? This paper argues that the third view reveals three important lessons for anyone involved in risk management. First, the assessment of unusual events -- such as the Gulf oil spill -- can and should be made more rigorous. Rather than relying purely on qualitative treatment. Second, more formal attention should be devoted to developing and evaluating customized preparations and responses. Rather than emphasizing off-the-shelf alternatives. Third, systematic learning should be incorporated to ensure that risk management improves over time. Rather than remaining static or even declining. With such improvements, risks will be addressed more effectively and the chances and impacts of future disasters will be reduced.

References

Cheryl M. Anderson and Robert P. LaBelle, *Update of Comparative Occurrence Rates for Offshore Oil Spills*, Spill Science & Technology Bulletin, Volume 6, Number 5/6, pp. 303-321, 2000).

Alan Bailey, *BP: Learning from Oil Spill Lessons*, Petroleum News, Volume 11, Number 20, May 14, 2006.

David R. Baker, *Engineer Robert Bea a student of disaster*, San Francisco Chronicle, June 10, 2010.

Philip C. Beccue and Hillard G. Huntington, An Assessment of Oil Market Disruption Risks, Stanford University Energy Modeling Forum, October, 3, 2005.

Adam Borison and Gregory Hamm, *How to Manage Risk (After Risk Management Has Failed)*, Sloan Management Review, Volume 52, Number 1, Fall 2010.

Robert Clemen, Making Hard Decisions: An Introduction to Decision Analysis (2nd edition), South-Western College Publishers, 1997.

John Donegan, The Learning Organization: Lessons from British Petroleum, European Management Journal, September 1990, p. 302-312.

Ted G. Eschenbach et. al., *Estimating Oil Spill Occurrence Rates: A Case Study for Outer Continental Shelf Areas of Gulf of Mexico*, Journal of Environmental Statistics, February 2010.

Ralph L. Keeney, Value-Focused Thinking: A Path to Creative Decisionmaking, Harvard University Press, 1992.

Howard Kunreuther and Michael Useem, editors, Learning from Catastrophes: Strategies for Reaction and Response, Wharton School Publishing, Saddle River, New Jersey, 2010.

Melissa Lodge, *Gulf Spill Is a 'Black Swan' Event: Industry Insider*, CNBC.com, June 9, 2010.

Terry Macalister, *Tony Hayward: Public Saw Us as 'Fumbling and Incompetent'*, The Guardian, November 11, 2010.

Jason R. W. Merrick et. al., *Understanding Organizational Safety Using Value-Focused Thinking*, Risk Analysis, Volume 25, Number 4, 2005.

Minerals Management Service, 30 CFR Chapter 8, Section 254.47, July 1, 2007.

Dan R. Replogle, authority, Regional Oil Spill Response Plan, BP, June 30, 2009 revision.

Peter Senge, The Fifth Discipline: The Art and Practice of the Learning Organization, Doubleday, New York, 2006.

Thomas R. Stewart and Thomas M. Leschine, *Judgment and Analysis in Oil Spill Risk Assessment*, Risk Analysis, Volume 6, Number 1, 1986.

Nassim Nicholas Taleb, The Black Swan, Random House, New York, 2007.

Brian Toft and Simon Reynolds, Learning from Disasters: A Management Approach, Palgrave MacMillan, 2006.

Robert L. Winkler, An Introduction to Bayesian Inference and Decision (2nd Edition), Probabilistic Publishing, 2003.

-----*BP-MIT Program Focuses on Operations Safety*, MITnews, April 10, 2008.

Committee on Energy and Commerce, *Waxman, Markey, Stupak Question Oil Companies on Spill Response Plans*, June 28, 2010.