Measuring Safety

A Call for a New Approach

By Philip W. Hurst and Quincey Jones

Inspirational quotes by management gurus abound. These quotes are designed to distill the essence of profound management knowledge into a single, stand-alone statement or slogan that will motivate and influence corporate leaders to make wise decisions and employees to take action. As the information age of society has advanced and computer power has evolved, the famous expression, "What gets measured gets managed," or its cousin statement, "What gets measured gets done" (both quotes often ambiguously attributed to various people, including Peter Drucker, Tom Peters, W. Edwards Deming and Lord Kelvin), has a special appeal to allow data to be our guide. Intuitively, one would assume that the leader who measures the most would, therefore, know the most and can lead the way.

As a result, sophisticated dashboards and real-time data are becoming standard board room fixtures for corporate leaders to huddle around and make their strategic plans, road maps and tactical objectives to produce high-performing organizations. Advancements in quantitative analysis (numerical, objective, repeatable data) have allowed for the measurement of things once seen as too soft to understand. For example, we can now measure critical factors associated with organizational culture and leadership transformational skills (Denison, Hooijberg, Lane, et al., 2012; Kotter & Heskett, 1992; Schein, 2010). But before a leader buys into the quantitative approach wholeheartedly and bets his/her credibility on it, serious questions should be asked about the assumptions we create when we measure.

1) Are Safety Data Believable?

Historically, OSH professionals have used metrics focused on lagging indicators to judge safety performance (e.g., recordable injury rate). More recently, however, the need to examine leading indicators has emerged. Hinze, Thurman and Wehle (2013) further suggest that leading indicators may be seen as passive indicators (e.g., number or percent of subcontractors selected, in part, on the basis of satisfying specific historical safety criterion prior to being awarded the subcontract) or active indicators

IN BRIEF
- **Sophisticated dashboards and real-time data are standard in the field of safety, but safety data may be conceptually flawed. Quantitative analysis has provided critical data to make key safety decisions, but it may not be telling the whole story.**
- **Although safety professionals are collecting more data, measurement systems may not inspire questioning and curiosity.**
- **The future will merge qualitative and quantitative data, which will make for a more robust measurement system.**

Philip W. Hurst, Ph.D., is an executive consultant for Hurst Behavioral Group Inc. He has more than 20 years' experience as an executive coach, organizational consultant, speaker, trainer and facilitator. He works internationally with organizational leaders from a range of industries such as nuclear power, manufacturing, chemical, aviation, petroleum, food and beverage, distribution, recruiting and legal. Previously, Hurst managed and developed internal consultants and provided behavioral consulting for a behavioral services group of a multibillion-dollar corporation. Hurst holds a Ph.D. from Auburn University.

Quincey Jones, M.B.A., is senior vice president at Optimum Performance Solutions. He has helped companies identify breakthrough change possibilities with successful implementation at all levels of their organization resulting in high-performing employees, teams and organizations. Prior to this, Jones worked for Continuous Learning Group, Aubrey Daniels International, Litton Industries and NCR. He holds a B.A. in Economics from Columbia University, and an M.B.A. in Finance/Organizational Development from Bowling Green State University; he is a member of the NACOMS Honor Society.

Regardless of the type of metric used, there remains the possibility of conceptual confusion about the definition of key terms. Definitional variability is a measurement killer, and must be discussed frequently.

(e.g., number of close calls reported per 200,000 hours of worker exposure).

Regardless of the type of metric used, there remains the possibility of conceptual confusion about the definition of key terms. While working with a mining company in Nevada, safety data were reviewed during a supervisory meeting. Injury rates had declined substantially and corporate leaders were proud to show off success. These data were clear and their safety initiatives had a marked impact. No doubt, some leaders received a handsome bonus for this achievement and the company at large received the message that this organization values safety.

But, when supervisors were asked one simple question, the mounds of data soon appeared weak. That question: "What is an injury to you?" Immediately, the vice president of HSE pointed out that the definition is written in the company safety manual and began to quote it from memory. He was stopped midstream. Again, supervisors were asked, "What is an injury to you?"

This question set the occasion for an important conceptual analysis of what constitutes an injury and whether an individual's definition of an injury is the same as what s/he is willing to report. Interestingly, the consensus among supervisors was that an injury occurs if a person cannot walk out of the mine on his/her own volition. One person told a story of how he had seen a miner get cut on the forearm, duct taped it closed and continued working. He did not feel compelled to report his injury.

One can cite many possible explanations for this individual's failure to report what appears to be obvious. Perhaps a "cowboy" mentality sometimes found in mining cultures overrode what should have been reported. Perhaps he did not want to go through the hassle of a root-cause analysis and safety investigation. Perhaps he did not feel enough pain to call it an injury. Or, perhaps he perceived this as something that happens to miners in day-to-day operations. Regardless of the underlying reasons, the miner failed to report the event and it suggests a failure to communicate what constitutes an injury. When this occurs, variability in reporting may increase and all the quantitative data collected becomes suspect.

While many people expect this kind of definitional variability with employees operating heavy equipment, it appears to be a common issue across settings and for other key concepts such as *near-miss*. For example, while working at the headquarters of a major oil company, employees often thought that safety efforts directed toward their white-collar jobs were a waste of time. However, the incident rate was too high and injuries covered a broad range of severity.

Take the example of an office assistant carrying too many boxes, thereby obstructing her view (eyes were not on path). As a result, she fell from the top of a stairwell and broke her shoulder. An injury of this nature is easily identifiable by all (no variability). When employees are asked whether they would categorize this as a near-miss if the worker had fallen and was not physically hurt, they say they would (again, no variability). However, when asked to consider, if a box fell off at the top of the stairs, but the worker did not fall, some say this is a near-miss, others are not so sure. If she had almost slipped but did not fall and did not drop anything, even more people will voice that they are not so sure.

Definitional variability of this kind is a measurement killer, and must be discussed frequently, highlighting examples to anchor a clear understanding of key concepts. Literature is full of examples of definitional variability for near-miss reporting between companies and between industry sectors (Marks, Teizer & Hinze, 2014).

To have reliable data, clear definitions must be operational and understood by the population contributing to the measurement system. As Voltaire notes, "If you wish to converse with me define your terms." And as anyone who has taken a basic philosophy course knows, to determine whether a tomato is a fruit or a vegetable, one must first know what constitutes a fruit and a vegetable, and whether one is discussing it as a scientist or as a cook. Without painstakingly clarifying what is being measured, data will be of little help.

This call for conceptually clarifying terms is important for any company striving for a zero incident vision. Zwetsloot, Aaltonen, Wybo, et al. (2013), note that the idea of zero incident vision is a logical extension derived from a host of zero visions that have been in existence for decades (e.g., zero defects, zero emissions, zero traffic accidents, zero waste, zero economic waste). However, zero incident vision will remain viewed as unrealistic and unachievable when the foundation on which it is built (i.e., the definition of injury) remains unclear within an organization.

To counter this definitional problem, safety leaders must incorporate daily discussions with employees, taking this time to show what defines an injury and what one is not, and give individual examples relevant to company goals. In the authors' experience, it is not sufficient to simply define injury in an all-inclusive manner such as "any physical or mental harm while on the job." Often employees see this all-inclusive terminology as an administrative nightmare with little value. Consequently, this definition can act as a barrier to enhanced incident reporting.

2) Is the Measurement System Telling the Whole Story?

One major sign that the safety field is evolving from measuring lagging indicators and moving toward including leading indicators happens when organizations determine thresholds for employee incentive programs similar to the method proposed by Sparer and Dennerlein (2013).

Another natural progression of moving from lagging to leading would be to seek what sets the occasion for strong leading indicators. Performance is tracked to highlight successes as well as opportunities for improvement. But simply providing numbers and showing charts do not create a measurement system. A system is a group of related parts that move or work together, as defined by Merriam-Webster's dictionary. Providing incident rates is only a snapshot of a measurement system's outcome. It is important to understand the whole system and how the parts move together. Thus, future data will include a chain or pattern of inputs (e.g., testing assumptions and leading indicators) that affect the lagging indicator.

A safety measurement system must create a clear, cognitive understanding of the system's interdependence. This can be the first step to motivating performers into "want to" action. Too often a performer in the field does not value some of the applications the organization wants its employees to use. For example, a petroleum company introduced to the workforce a field-level hazard assessment (FLHA) that was intended to have employees pause prior to performing a task and collectively identify exposure.

Asked whether they valued using the FLHA, employees frequently answered no. Some saw it as another opportunity for management to blame employees who got injured. Some saw it as tool filled out by going through the motions. When asked how they performed the FLHA, these individuals often described filling it out in the truck, then having the team sign off on it. This process voided any real value that could have been achieved if applied properly.

The real purpose for the document was to generate rich collaborative discussions within the team about exposure and to ensure that people did not become desensitized to the job hazards. It was also designed in the spirit of "brother's keeper," that is, to make sure a coworker was situationally aware of all potential dangers prior to performing a task. The assumption that employees would understand this intent was invalid and as a result the measurement system failed to reach its full potential until it was addressed.

Testing assumptions is critical for a measurement system to have significant impact. For example, most leaders assume that if an employee saw someone performing an at-risk behavior, that employee would naturally speak up and take action. This assumption must be examined.

Working with a railroad, it was noted that the workforce had been stagnant for several decades. Few new employees were hired. As the older generation approached retirement, new hires became a significant priority for the organization. As a result, a bimodal distribution emerged in the workforce and the organizational cultural theme became "the old heads versus the new heads." Conflict was a high probability if a new head approached an old head and suggested the person was executing a task unsafely. New heads soon learned to not approach others.

By testing their fundamental assumption regarding approaching others, railroad leaders were able to identify a strategy to overcome this barrier. Employees were brought together to develop best practices for approaching each other when someone perceived that an unsafe act was occurring. These methods for accepted protocol were communicated and practiced. Employees learned to use key language that would act as a cue for discussing exposure and to reduce defensiveness. For example, rather than saying, "Hey, you are doing that wrong," the employee would say, "That makes me nervous, can we talk about it?"

Safety leaders must make a practice of identifying assumptions that underlie safety efforts. By testing assumptions, barriers can be identified and new strategies can emerge. Identify as many assumptions as possible regarding safety, whether implicit or explicit, then critically determine which assumptions have the greatest impact on the measurement system. After this, test how those assumptions are affecting performance by observing in the field and discussing with those in the working interface.

For example, if a leader assumes that peers will approach each other when seeing an unsafe act, ask employees questions such as:

1) When was the last time you saw someone approach a peer about an unsafe act? Describe it.

2) When was the last time you approached someone who was performing an unsafe act? Describe it.

3) How many times this year have you seen a peer approach someone performing an unsafe act?

4) How many times this year have you approached someone who was performing an unsafe act?

5) Describe what you would say to a peer if s/he was doing something unsafe.

6) How do peers approach you when they believe you are doing something unsafe?

3) Does the Measurement System Inspire Questioning, Curiosity?

It has been noted that perception is reality. This reasoning goes back to Plato's famous Allegory of the Cave, in which prisoners are chained in a cave from the start of life and can only see flickering shadows cast on the cave wall that they perceive as real. Only after a prisoner escaped the cave did he learn to discern true forms and patterns. The shadows that are cast as reality for safety data are associated with the person who presents the measurement system. Most organizations communicate their data to make it administratively simple (e.g., each level presents data to the next level). But the presentation of a measurement system that does not convey the curiosity of the leader is one that will ultimately fail.

After a rash of injuries at a chemical plant, management decided to conduct a prejob analysis prior to every task. Rarely did leaders from any level ask those using the prejob analysis to note how the

Before a leader buys into the quantitative approach wholeheartedly and bets his/her credibility on it, serious questions should be asked about the assumptions we create when we measure.

document was helping them or how to improve it for effectiveness. They simply collected the documents and recorded the number conducted. Only when an injury occurred did a leader ask whether a prejob analysis had been completed.

Charts reflecting the number of weekly prejob analyses conducted never compelled interesting discussions. Data were presented as an audit, not as a collective mutual exploration of gathering insights into how to deal with exposure. When the curiosity factor is zero, so goes the measurement system. Leaders need to develop the Socratic skills to ask the right questions with an emphasis on behaviors to make a measurement system reflect the power of curiosity.

An acid test for the curiosity factor of a measurement system is how the following questions are addressed:

- What do people using the documents say that characterizes the measurement system? How does it affect them?
- Do they believe that managers really have curiosity about the numbers and that leaders seek to make changes based on what they hear?
- What actually gets changed or fixed as a result of capturing these data?
- How do leaders make decisions based on the measurement system?
- Can the performers use the system to describe examples of how it helps them be safer?
- What has the performer seen or heard that makes him/her believe that others have bought-in to the measurement system?

These questions should be addressed quarterly by safety leaders. They can do this by holding small focus groups and conducting individual interviews to capture relevant qualitative data. To maximize a measurement system, one must define and align the definitions of the measure being taken; make sure the measurement system tells the whole story and demonstrates the interdependence by testing assumptions; create the appropriate leading indicators and show a relationship to the lagging indicator; and ensure that the measurement system inspires questions and curiosity.

4) Do You Understand Today's Technological Impact on the Measurement System?

As noted, dashboards displaying real-time data are becoming standard features in organizations. This is because civilization has gone through three industrial revolutions:

1) introduction of mechanical production facilities with the help of water and steam power;

2) introduction of division of labor and mass production with the help of electrical energy;

3) use of electronic and IT systems that further automate production.

Oberhaus (2015) notes that society is in the midst of a fourth industrial revolution, which is gaining exponential velocity and eliminating traditional business models. As a scientific or engineering field matures, the development of a unique language increases. Some terminology of the fourth industrial revolution that has become standardized by those leading the way are *machine-to-machine solutions*, *big data*, *predictive analytics*, *digital manufacturing*, *remote monitoring*, *intelligent machines*, *collaborative robots*, *cloud computing*, *augmented reality* and *multimodal interaction*.

Even the name typically associated with the fourth industrial revolution [i.e., Industrial Internet of Things (IIoT)] is reflective of a different way of thinking and talking. IIoT is defined as the network of physical objects (e.g., sensors, devices, scanners) that contain embedded systems that communicate and interact with their internal states or the external environment.

Not to be confused with the consumer Internet of Things (IoT), IIoT has a unique set of data usage requirements, a different measurement system and requires data to be available through open standards. It is being driven by the Industry 4.0 Initiative, which consists of German sponsorship, industry organizations, manufacturers and suppliers that are revolutionizing the manufacturing engineering sector through the establishment of international standards and removal of implementation barriers.

IIoT connects all types of devices to the cloud, which possesses massive data storage, combines it with other data sources in the cloud and performs data analytics. Organizations can use the resulting output of information to gain real-time operational insights that close performance gaps and generate new insights that create new business value.

Numerous articles and books have discussed IIoT, its challenges and opportunities (Greengard, 2015; Mukhopadhyay, 2014; Schroeer, 2015). This fourth industrial revolution represents a paradigm shift for both quantitative and qualitative data analysis. It allows massive amounts of data to be analyzed, then compared to data from smaller sample sizes derived from qualitative analysis that will either confirm or call into question those quantitative numbers.

For example, if a site averages 200 near-misses per month, a company can ask how many of these events actually meet its clearly defined quality standard that detailed information must be included in the form's comment section. By conducting a qualitative assessment (e.g., perception surveys, interviews, focus groups) of key departments and having access to those completed near-miss forms via mobile device, the quality question can be quickly answered by auditing the near-miss database forms, or simply designing a digital dashboard measurement system to display the total number of forms that meet the quality standard.

In addition, technology can facilitate the testing of assumptions by building qualitative databases that capture, track and trend this information. To test the assumption that peers will approach each other when seeing an unsafe act, employees can use mobile devices in the field to capture answers to the six questions referenced previously. This provides an opportunity to question the variance and the reliability of the data, and rethink the original assumptions if necessary.

According to Pew Research Group:

> As the year 2011 began on Jan. 1, the oldest members of the baby boom generation celebrated their 65th birthday. In fact, on that day, today, and for every day for the next 19 years, 10,000 baby boomers will reach age 65. The aging of this huge cohort of Americans (26% of the total U.S. population are baby boomers) will dramatically change the composition of the country. Currently, just 13% of Americans are ages 65 and older. By 2030, when all members of the baby boom generation have reached that age, fully 18% of the nation will be at least that age. (Cohn & Taylor, 2010)

Thus, most organizations will continue to be challenged by this demographic shift creating "the old heads versus the new heads" and must train employees how to comfortably approach each other to discuss unsafe work actions. In addition, many other areas will be affected, including the critical area of knowledge retention.

Against the backdrop of this demographic shift and years of workforce reductions, one chemical client requested assistance to accelerate new employee (i.e., with 1 to 3 years' experience) operator onboarding and training programs. New employees were often found working alone on off shifts (i.e., second and third), without the support of a senior technical operator who possessed the knowledge to answer questions. Numerous incident reports indicated employees were experiencing increased risk exposures by not completely understanding the total work process system, and that they were sometimes unsure of the next steps and how to resolve dangerous situations. Whether in the department or the field, quick access to policies and procedures was paramount. Transitioning all paper-based policies and procedures to a digital format with instant mobile accessibility increased the ability to correctly respond to any situation, while reducing employee risk exposure.

Sophisticated dashboards with real-time information and data are the new norm in many industries. Technological advances allow organizations to receive immediate feedback that either verifies or raises questions about the measurement system and whether leaders are curious enough to ask about it. In addition, mobile technology allows for faster, better decision making throughout the organization and especially at the point of performance. These digital dashboards contain actionable data (e.g., graphs, charts) specifically tailored for key decision makers at each organizational level. The future will merge qualitative and quantitate data, making for a more robust measurement system.

Conclusion

Management gurus in the near future will emphasize qualitative as well as quantitative safety analysis. They will speak to the importance of conceptually analyzing the key concepts of a safety measurement system, such as *injury* or *near-miss*. New techniques will be promulgated to close the definitional variance that exists in organizational cultures. These future gurus will advocate the need to test management assumptions and identify best practices for creating measurement systems that inspire employees and reflect the curiosity of the leaders. Finally, real-time data will be a reality and new technology will change the way we are doing safety. **PS**

References

Cohn, D. & Taylor, P. (2010, Dec. 20). Baby boomers approach 65—glumly. Retrieved from www.pewsocialtrends.org/2010/12/20/baby-boomers-approach-65-glumly

Denison, D., Hooijberg, R., Lane, N., et al. (2012). *Leading culture change in global organizations: Aligning culture and strategy*. San Francisco, CA: Jossey-Bass.

Greengard, S. (2015). *The Internet of things*. Cambridge, MA: MIT Press.

Hinze, J., Thurman, S. & Wehle, A. (2013). Leading indicators of construction safety performance. *Safety Science*, *51*(1), 23-28. doi:10.1016/j.ssci.2012.05.016

Kotter, J.P. & Heskett, J.L. (1992). *Corporate culture and performance*. New York, NY: Free Press.

Marks, E., Teizer, J. & Hinze, J. (2014) Near-miss reporting program to enhance construction worker safety performance. *Construction Research Congress 2014*: pp. 2315-2324. doi:10.1061/9780784413517.235

Mukhopadhyay, S.C. (Ed.). (2014). *Internet of things: Challenges and opportunities*. Geneva, Switzerland: Springer International Publishing.

Oberhaus, D. (2015, April 13). This is what the fourth industrial revolution looks like. *Motherboard*. Retrieved from http://motherboard.vice.com/read/life-after-the-fourth-industrial-revolution

Schein, E.H. (2010). *Organizational culture and leadership*. San Francisco, CA: Jossey-Bass.

Schroeer, T. (2015, May 7). The Internet of things drives new industry 4.0 manufacturing. IBM Big Data and Analytics Hub. Retrieved from www.ibmbigdatahub.com/blog/internet-things-drives-new-industry-40-manufacturing-initiatives

Sparer, E.H. & Dennerlein, J.T. (2013). Determining safety inspection thresholds for employee incentives programs on construction sites. *Safety Science*, *51*(1), 77-84. doi:10.1016/j.ssci.2012.06.009

Zwetsloot, G.I., Aaltonen, M., Wybo, J., et al. (2013, Feb.). The case for research into the zero accident vision. *Safety Science*, *58*, 41-48. doi:10.1016/j.ssci.2013.01.026