

Integrated Reliability Index (IRI) Concepts

Comments and Responses
August 31, 2011

The Operating Committee and Planning Committee approved the Reliability Metrics Working Group (RMWG) to post its draft IRI whitepaper for public comments and to hold a workshop to seek public comments at their June 7-8, 2011 meeting. RMWG posted the whitepaper on July 14 and held a workshop on July 28. A total of 75 comments were received by the end of commenting period (August 14, 2011). RMWG has provided the responses to each comment. The entire comment and response matrix (see Table 1) and additional stakeholder input for the draft IRI whitepaper are included in this document.

For additional information, the draft IRI whitepaper is located at:

http://www.nerc.com/docs/pc/rmwg/Integrated_Reliability_Index_WhitePaper_DRAFT.pdf.

The IRI Workshop's agenda and presentation slides are located at:

http://www.nerc.com/docs/pc/rmwg/IRI_Workshop_Agenda_Slides.pdf.

Comment and RMWG Response Matrix

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
1	Co-relation between all of the 3 indices and their relationship.	<p>The relationship between the 3 Indexes is structurally consistent with the regulatory concept of NERC relative to reliability. Standards are in place to ensure an adequate level of reliability. Compliance with those standards is measured by the SDI. Metrics have been developed to measure performance and a compilation of the results of those metrics has been utilized in the development of the CDI. Finally, the proof of success is in the lack of significant events which is measured by the EDI. Together as a group, the IRI represents a first attempt at harmonizing these three important components into a high level index that will, over time and as it becomes refined, help to communicate industry trends. This will help to ensure that any shifts in the wrong direction can be addressed before they become problematic.</p>

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
2	<p>What is the feedback that will be used to change the weightings in the equation?</p>	<p>The process of changing the equation weightings is currently under RMWG's consideration. A statistical framework for changing the risk factors and weightings is necessary and possible. In fact, such a framework may answer some of the difficult conceptual questions regarding an Integrated Reliability Index, such as how to weigh separate components of the IRI. Statistically based techniques can be established to test the relationship between the index and its independent risk factors. If the risk factors are statistically significant, they will be included as a part of the equation and their significance will influence the weightings. A formal model review and statistical significance calculation will be conducted on an annual basis to adjust the risk factors and their weightings.</p>
3	<p>How will the IRI account for changes in the SDI (changes of the selected standards and CDI (changes of the selected metrics)?</p> <p>Will the SDI and CDI change drastically based on the selected metrics/standards chosen at the time, or will a conscious effort be made to keep the IRI components consistent?</p>	<p>Please see response for comment 2.</p>

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
4	How should the aggregated sum be presented? Does the positive change represent overall industry performance improvement? Do the change trends have any statistical significance? Can these trends be used to predict future outcomes?	<p>Currently, positive change would represent overall industry risk management. Yes, these trends will have statistical significance (see response for comment 2). More data and experience are needed to understand whether these trends can be used to predict future outcomes.</p> <p>Any "significant" changes in performance warrant investigation. Findings may suggest improvement or normal random variation. Data/Conclusions will not be presented unless statistical significance is assured, and, in the future, RMWG desires to develop leading indicators which may have relationships with lagging indicators.</p>
5	Does the current IRI measure the intersection or the union of risks within the bulk power system?	Currently, the IRI is intended to measure overall risk based on statistically significant independent risk factors, it does not measure union or intersection (see response for comment 1). If the components are not factored, this may not be an issue.
6	Are there any consistent methods, beyond industry experience and strategic goal setting, to determine the weighting factors used in EDI, CDI and SDI? If yes, what are they?	Please see response for comment 2. Additionally with time certain correlations between metrics and components may reveal a more scientific weighting method.

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
7	Are there any major risk components missing from current IRI definition? If yes, what are they?	NERC leadership developed the vision and conceptual models that have been shared in the industry. RMWG is working to implement that model. As such, for the risks identified in this model RMWG believes that the IRI covers the major risk components. Investigation will help find any missing factors, for example, whether CIP related factors need to be considered (see response for comment 2). Note that these risks do not include risks outside the bulk electric system, in the form of policies and other actions which could change the fundamental design and foundation of the bulk electric system. As stated above, the risks identified do not include any outside the purview of the bulk electric system. Thus externally-resultant risks, such as those that are a result of policy, are not included within the risk model.
8	How should they be integrated into the IRI and how should the “baseline” performance be refreshed, as metric components change over time?	RMWG is currently considering this idea.
9	ALR1-4 in CDI considers the load loss events. So does EDI. If events have violation, they will be counted in SDI. How does IRI deal with double or triple counting?	The current IRI equation reflects the addition of risks and aims to capture all high risk factors. The risk factors were largely chosen to be independent from each other; minor duplication may exist. RMWG is currently working with Sandia National Lab and DOE's statisticians.

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
10	<p>Can each risk factor be integrated into one outcome measure, given problem specific risk factors? If so, how should this be done?</p> <p>What considerations should be used?</p>	<p>Please see response for comment #2.</p>
11	<p>How should industry weight compliance risks versus historic event risks versus operational metrics that demonstrate the effectiveness of the system?</p>	<p>Please see response for comment 2.</p>
12	<p>How should we be considering operational versus planning aspects in the index?</p>	<p>Please see response for comment 2.</p>
13	<p>The current configuration of the IRI has too much emphasis on historical 'what happened' events focus. In effect, we are measuring events, disturbances, and other adverse incidents on the bulk system.</p>	<p>As more data and experience are gained, other areas can be introduced to the IRI if they prove to be statistically significant.</p>
14	<p>The two problems with this are the undue focus on events per se [some of which are performing exactly as desired, and therefore represent 'good' results of the system design in place], and even more so focus on loss of load as an indicator [when in some cases the consequential load loss is an exact result that has been designed into the protection schemes]. Ultimately, I think we need to re-look at this, to measure only those events that do not occur as designed/planned</p>	<p>RMWG is working to address this issue. More specific data may need to be collected to determine if an event happened "as designed" or "not as designed". However, this does impose an additional reporting burden on registered entities.</p> <p>SRI daily scores do not consider whether the event happened "as designed" or "not as designed". This could be an area to reconsider in the future. However, data collection may limit RMWG's ability to produce meaningful results.</p>
15	<p>Were unplanned outages part of this calculation?</p>	<p>Yes, currently only automatic sustained outages are considered in this calculation (EDI).</p>

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
16	<p>Are there any major risk components missing from current IRI definition? If yes, what are they?</p> <p>How should they be integrated into the IRI and how should the “baseline” performance be refreshed, as metric components change over time?</p>	Please see response to comment 7.
17	Should the indices be weather normalized?	This topic is currently under consideration within RMWG. One index value could include weather related events, and another companion value could exclude weather related events.
18	<p>Should CIP standard violations be included in SDI? If yes, what are the criteria to be used to determine its impact to bulk power system reliability?</p> <p>Should they be weighted the same as the other categories of NERC Reliability Standard violations?</p>	Please see response for comment 7.
19	Should equipment or loss of load that functioned as designed or through predefined operating procedures be treated differently than those events which did not? If yes, how?	Please see response to comment 14.
20	How many years of history are needed to assure an index is ready for measurement?	At least 3 to 5 years of statistically relevant data are needed for trending. The years of history necessary, in many cases, depends on the sampling size and specificity of information.
21	<p>Can the index be used to assess both historical performance and be an indicator of current performance?</p> <p>Can we use the index in a predictive manner?</p>	Currently the index is only used to measure historical performance. More data and experience are needed before it can be used for other purposes.

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
22	<p>Many of the metrics for the CDI are listed as “inconclusive”, “no data”, or “new data”. How will the relative lack of data be accounted for in the CDI in regards to the metrics?</p> <p>Will the CDI jump based on only the amount of data collected, or will it truly measure performance of the system?</p>	<p>RMWG is in the process of collecting missing data. It is expected by the end of 2011 all metric data will be available; nonetheless that will still result in a limited history of data which will be evaluated. How to rebaseline the CDI if new data becomes available is under consideration within RMWG. Sandia National Lab is assisting in development of a rebaseline process.</p>
23	<p>How was randomness considered in this analysis?</p> <p>Random or uncontrolled outages are sometimes a substantial portion of the total outage data count, is this considered?</p>	<p>Please see response to comment 2.</p>
24	<p>How and who determines the level of desired outcome? What if the results are far from the desired level?</p>	<p>RMWG is considering this issue.</p>
25	<p>What does early predictors mean, based on trending, if so, how many years of data to establish a trend?</p>	<p>An early predictor is an index or metric that provides sufficient information to point to a growing problem that left unresolved will ultimately impact reliability. The IRI is still in a conceptual stage, and it is too early to know if the current version of IRI or supporting indices will be early predictors and how much effort or time will be required to achieve such a goal.</p> <p>In general, 3 to 5 years of data is needed for trending. The years of history necessary, in many cases, depends on the sampling size and specificity of information.</p>

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
26	Priority of the equipment from the system point of view. Restoring all equipment does not have the same priority. How does the EDI take that into consideration?	The EDI inferentially determines the effectiveness of the restoration of high priority equipment, assuming that they will drive down the duration of load loss events (the restoration promptness factor gives credit for fast restoration of load).
27	Concern about year to year data changing because of ALR Severity level changes which affects trending.	Please see response to comment 9.
28	A new system will take 4-5 years to implement and achieve consistent and meaningful data.	Agreed.
29	Should EDI include weather-caused events? If yes, why? If not, why not?	Please see response to comment 17.
30	Transmission unavailable is not necessarily a bad state. Planned Outages are allowed on a system that can serve the expected load for that condition, TPL-002 &-003, R1.3.12.	The EDI uses only automatic (forced) outages.

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
31	Will require more resources for utilities that are current overwhelmed with compliance/standards programs that requires funding while keeping rates low	<p>True, but we do not believe it will require an overwhelming amount. Further, utilities should also look at the benefits to this benchmarking effort. Once these benchmarks are on a sound footing, entities, whatever the role within the industry, will be able to assess their relative position with respect to these benchmarks and make informed decisions. Meanwhile, if the metrics enable an operator to operate the system more effectively because they have good performance indicators, it should be worth the cost.</p> <p>The industry is always concerned with respect to undue reporting, and has taken this into account in the development of the process and associated value.</p>
32	Glad that we are looking at developing a measurement to tell us where the industry is today from a reliability performance viewpoint. It would have been helpful to have done this as one of NERC’s first missions.	It is important to know where you start from to figure out where you need to go; it took some time getting minds aligned to develop and that continues to be an ongoing effort.
33	The system “must” be simple to use, understand, and be intuitive.	Indeed, keeping measurements simple and easy to understand has been one of RMWG’s objectives since the metric development started in 2008.

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
34	<p>Consider how Reliability Indices might be tied to costs. IF/AS we improve reliability will we pass a point of diminishing returns?</p>	<p>It is a valid point. However, how the IRI and other metrics are used to resolve any issues uncovered by these tools is not within RMWG's scope. Indexes and metrics are the indicators allowing one to understand the costs and benefits of any given project or measure.</p> <p>Perhaps, stakeholders and industry will be in a position to have this conversation as the metrics demonstrate system performance.</p>
35	<p>Are we addressing Reliability/Loss consistently across areas or are we spending large dollars for minimal benefits in one area, while ignoring low lying fruit in another?</p>	<p>Please see response to comment 34.</p>
36	<p>Could we rank the type of improvements/process that gives the most bang for the buck in improving reliability/capacity of the system?</p> <p>Should indices be able to support such work?</p>	<p>As the indices and metrics are fully developed and trends are identified, there may be an ability to develop inferences in this area.</p> <p>However, this is not currently part of this project's scope.</p>
37	<p>It is still not clear to me who is pushing for an Integrated Index and why.</p>	<p>NERC believes that it is important to understand the state of reliability and to understand if reliability is improving, staying constant or degrading. NERC also wants to provide the industry with metrics that will help identify common issues, potential areas for further study, standard development etc. An IRI is one way that we are considering to do this although we need to better understand the viability and appropriateness of this approach over time.</p>

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
38	I worry that the IRI effort will create a distraction for operators of the bulk electric system.	Indices will not provide real-time trends. IRI and metric trends will be calculated on a quarterly basis using data from the past quarter. Insights provided will assist planners in evaluating what is working and where the risks are presently. As such, we see no situation where this effort could distract a system operator.
39	How will the guidance be determined and how will it be conveyed to the industry?	<p>The metrics sit between the ALR objectives and the requirements of the standards. Hopefully, the metrics will be accurate enough to provide a meaningful assessment of where an entity is with respect to reliability (benchmarking) to answer the question:</p> <p>"What is the margin between an entity's current reliable state and an unreliable state?"</p>
40	Since CDI, EDI, and SDI cover major risk factors to BPS reliability, is it more appropriate to rename the Integrated Reliability Index (IRI) as the Reliability Risk Index (RRI)?	RMWG is currently considering this idea.
41	Are there reasonable example models (such as credit scoring or actuarial models) which could be used to enhance the understanding of integrated risks?	<p>In theory, this is a promising approach. But there needs to be much work done on identifying the proper components of an integrated risk assessment approach.</p> <p>Even in models such as credit scoring, the importance of proper application is important. As people have heard, the improper use of a blended metric is a big concern (credit scores as an indicator of a good employee.)</p>

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
42	Have regulators in other industries used the integrated weighted sums to measure performance?	Yes, NRC, FAA and OSHA have all used integrated indices to measure performance.
43	This document lacks two essential ingredients needed for attribution: (a) a date, and (b) an author. Another referenced document has this same problem which should be corrected: Integrated Bulk Power System Risk Assessment Concepts. It is important for reviewer to know when a document was prepared and by whom.	A cover page will be added to state the date and author.
44	While the entire frame is conceptual, the formulas and weights used are completely heuristic. This fact that should be acknowledged at the outset. Experience will likely change the respective formulas over time.	More clarifications will be added in next IRI whitepaper revision.
45	In my opinion, it is far too early to combine three heuristic indices into one combined index – see line 43. Rather, the three indices should be kept apart so that the information they communicate is not lost by combination.	RMWG plans to track the three indices separately regardless of whether a combined IRI is used or not.
46	All or the data that supports each calculation of an index should be made available so a greater understanding of the strengths and weaknesses of the proposed indices can be understood. If the data is CEII, then it could be “sanitized” if need be so that it cannot be attributed to a particular entity.	All data aggregated at the NERC, regional and interconnection levels will be made available at the NERC website.

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
47	<p>Lines 56-69. It is unclear how overlapping data in the three indices is addressed. For example, an event could create data in both the EDI and the SDI index if it had an associated standards violation. Is data compiled in each index for the same event? What Figure 2 is trying to communicate is unclear. One would be more interested in events that did NOT have associated standards violations. Those would be more indicative of potential reliability gaps.</p>	<p>Please see response to comment 9.</p>
48	<p>Figure 3 – See comment above regarding supporting data. It appears that certain known events are listed, so one would assume CEII was not an issue with them.</p>	<p>Please see response to comment 46.</p>
49	<p>Figure 3 – definition of SRI and Event Categories 1-5. The Event Categories 1-5 defined in the document in footnote 7. Since the SRI index ranges assigned to an Event Category, that analysis should be provided.</p>	<p>The analysis and calculation will be added in next IRI revision.</p>
50	<p>Lines 109-110. Were only GADS outages and TADS outages that met an Event Category criterion used? That should be clarified.</p>	<p>All GADS and TADS forced/automatic outages were used in the calculation.</p>
51	<p>Line 114. The formula for SRI in the footnoted document is as follows:</p> $SRI_{event} = 50\%*(MWL) + 30\%*(NT) + 10\%*(NG) + 5\% (HD) + 5\% (NE)$ <p>The terms MWL, etc. are described as “normalized values.” That calculation should be provided to understand whether the normalized values change from year to year.</p>	<p>The normalized values could change from year to year. The detailed description is available at http://www.nerc.com/docs/pc/rmwg/SRI_Equation_Refinement_May6_2011.pdf.</p>

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
52	Line 116. The justification for EDI is unclear. SRI is not denominated in days, but is subtracted from “days” in the denominator. EDI as presented looks like it’s the percentage of availability of some part of the BPS.	The SRI's units are % loss*day. The top part where "Duration in days" is specified should instead be called "100% availability"*number of days. The bottom instance of "Duration in days" should have the units of days. This will lead to the correct units of %. This will be clarified in next IRI revision.
53	Line 121. Not sure that’s true given the comment above.	The issue has been corrected.
54	Line 144 – “Date” should be “Data.”	The issue has been corrected.

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
55	<p>Table 2 - The CDI starts with 18 metrics which have SMART scores, but does not use the SMART scores. Instead, per Table 2, only two SMART attributes (Relevant and Tangible) are used (lines 130-132), resulting is all having similar weighting factors (WFs) of 4, 5, or 6. Some of the metrics use TADS data, but it’s unclear if different TADS data is used in the CDI data than is the EDI data. Six of the 18 metrics have a zero trend weighting. Therefore, in line 146, those have no metric contribution. As data is added over time for the missing six metrics, the CDI will decrease since the metric contributions for these six metrics will increase. This will falsely indicate a decline in reliability. Also, equation (6) on line 150 uses each metric’s SMART rating (all five) while the metric contribution only uses two of the five SMART ratings. For consistency, the SMART rating should only us those used in the metric contribution calculation in equation (5) on line 146 or the WFs should be used in equation (6) on line 150. Finally, equation (6) on line 150 should be re-examined from another perspective – a metric with trends that improve (TR = 5 or 4) would lower the CDI versus a trend that deteriorates (TR= 2 or 1). This doesn’t make sense if increasing CDI is desirable.</p>	<p>More details of the CDI calculation will be made available in next IRI revision.</p>

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
56	<p>The computation in equation (7) on line 180 treats all registered entities as having the same exposure to violations by equally weighting them in the denominator. For example, consider the PRC standards. Larger entities with more protection equipment will have more exposure to violations than smaller entities. Consider two entities, one with 500 protective relays versus one with 20 protective relays. If each entity has 5 misoperations to report, the total is 10 for two entities, but the entities are obviously not equal. The same applies to other standards such as FAC-003. All violations per standard need to be normalized for exposure to better measure reliability.</p>	<p>Currently, IRI is still in conceptual stage. More details of the CDI calculation will be made available in next IRI revision.</p>
57	<p>Event Driven Index (EDI) Calculation – The document states that the daily generation and AC transmission circuit outages were obtained from GADS (Generating Availability Data System) and TADS (Transmission Availability Data System), respectively. GADS reporting is currently not mandatory. Even with reporting becomes mandatory in 2012, not all generators will be required to report. The result is that this calculation may not adequately represent the industry or provide an adequate measure.</p>	<p>Currently we have 72% of generating availability data. When GADS becomes mandatory in 2012, this will become 92%.</p>
58	<p>Condition Driven Index (CDI) Calculation – The attachment fails to provide enough detail to explain how the values are arrived at that will ultimately be used in the CDI formula.</p>	<p>Currently, IRI is still in conceptual stage. More details of the CDI calculation will be made available in next IRI revision.</p>

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
59	Standards/Statute Driven (SDI) - The document fails to provide enough detail to explain how the values are arrived at that will ultimately be used in the SDI formula. Additionally, the relationship between Regional Entity identified Reliability Impact Statement (RIS) levels and VRFs is not clear.	More details of the SDI calculation, including RIS and VRF, will be made available in next IRI revision.
60	It is difficult to compare performance over time when the system configuration is not remaining constant.	This is a correct statement if the measurement were not at a macro level. However, by measuring in a more broad scale many of the complexities of the measurement become more straightforward. RMWG believes given the risk model it was presented it has created a broad level set of indices which may accomplish this goal. Much more work is required for them to be stress-tested. The task of RMWG is to determine the metrics which preserve the holistic view. If the system configuration changes such that a metric is not measuring reliability "performance", then that metric will be changed.
61	It is difficult to compare performance over time when the system is being operated differently.	Since success rates are the proposed measurement method that should translate independent of the current configuration of the system.
62	The nuclear model works well for a set of assets that are built and operated similarly. Also they share a set of common operating criteria and procedures.	Agreed.
63	When evaluating performance against the IRI, is it simply the trend?	Besides the IRI trend, all components that form the IRI will be tracked and evaluated as well.
64	Who will determine the "acceptable level" of performance?	Please see response to comment 24.

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
65	Combining the indices is a good way to evaluate disparate values with a single index; however, it can mask one or another index that is dragging the system down.	Please see response to comment 10.
66	The real key is to determine the drivers the affect performance.	Please see response to comment 2.
67	Experience with analysis of voltage or frequency measurements indicates that an understanding of the specific reasons for significant deviations or variations in the norm is essential to understanding whether they indicate a risk or are normal variations.	Agreed. Thank you for your comments.
68	Much of the performance variability relates to items over which the power system stakeholders have no control, such as weather patterns, fires, even economic conditions. An IRI itself has little value unless reasons for variations are understood.	As discussed, previously, there will be further investigation into "operated as designed" and "weather initiated events".
69	Equation 3 is missing	The issue has been corrected.
70	Footnote 5 and 7 hyperlinks do not appear to connect properly ("Page not found") Suggest http://www.nerc.com/fileUploads/File/Disturbance%20Reports/NERC_Std_EOP-004-1_01-2007.pdf for footnote 5 Suggest http://www.nerc.com/page.php?cid=5 365 for footnote 7	The issue has been corrected.

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
71	Footnote 9 hyperlink does not appear to connect properly (“Page not found”) Suggest http://www.nerc.com/docs/pc/rmwg/RMWG_Metric_Report-09-08-09.pdf (which is strange since I see that to be the hyperlink---maybe a local issue here???) Suggest verify the footnotes to see if similar issue exists with you)	The issue has been corrected.
72	The Trend Rating (TR) discussion discusses 5 ratings from 1 to 5 then includes a sixth (0-5) in the list. References should be to six TRs, correct? “To integrate individual metrics with differing units of measure, the six common trend ratings (TR) were identified to quantify each metric performance level, shown in Table 2. These six ratings range from 0 to 5.”	This issue is currently being considered by RMWG.
73	Sentence above Figure 5 says “Figure 5 provides the quarterly SDI trends for past three years” then sentence below Figure 5 says “Figure 5 provides the quarterly SDI trends for past three years...”. Suggest striking that part of sentence below Figure 5 and replacing with “Figure 5 indicates.....”	RMWG is considering this issue.
74	Consider the way “equation” is referenced in document- Sometimes it is capitalized while in other places, using similar sentence structure, it is not. Sometimes in parentheses while in other locations, using similar sentence structure, it is not.	The issue has been corrected.
75	Table 3 references on Page 10 and Figure 5 reference on Page 11 are bolded while other Table and Figure references are not.	The issue has been corrected.

Table 1: Integrated Reliability Index Comment and Response Matrix

Number	Comment	Response
76	On page two, line 1, do we really believe the EDI provides a basis for prioritization or should it be categorization?	RMWG is considering the issue.
77	On page two, line three under CDI, do we know that the ALR metrics are “early predictors”, or are we assuming they are early predictors?	RMWG is considering the issue.
78	On page 9 I’m not sure I understand equation (6) correctly or there is a problem with the trend ratings. As I understand the formula, trend improvement should reflect a higher CDI, therefore, the Metric Contribution should go down with improving trends (smaller number subtracted from 100), but the trend rating factors are in the opposite direction. Sample calculations for CDI would be a big help in this section.	RMWG is considering the issue.
79	You have an event driven, condition driven and standards driven. I think we are missing one – security driven. Maybe you could give that some thought – how do we include security in the index – particularly, cyber security.	Statistically based techniques can be established to test the relationship between the index and its independent risk factors. If the risk factors are statistically significant, they will be included as a part of the equation and their significance will influence the weightings. A formal model review and statistical significance calculation will be conducted on an annual basis to adjust the risk factors and their weightings.
80	I think we should do these indexes by interconnection.	RMWG is considering the issue.

Additional Input

Kansas City Power & Light Company

Kansas City Power & Light Company (KCP&L) provides these comments on the Integrated Reliability Index (IRI) whitepaper. Comments are provided in order by topic as they appear in the whitepaper.

General comments

It is not clear how the IRI, as it has been proposed, assists the industry's understanding of reliability risk absent further clarity regarding the details of the IRI. It would seem that applying an individual IRI for each of the identified Standards and Requirements for the Standards Driven Index (SDI) component would be more useful in understanding the areas of risk in the industry. It is also not clear in what manner the IRI would be used by NERC and how the IRI may be intended to modify or change the behaviors or actions of Registered Entities.

Integrated Reliability Index (IRI) concept

KCP&L is supportive of the concept represented of an IRI for assisting with the common definition of adequate reliability. The weighting factors that have been proposed require further analysis and may be considered unsupported. We recommend these components are reviewed in a test environment with conditions of sufficient duration to allow for appropriate adjustments to be made.

The IRI components

The Condition Driven Index (CDI) is intended to be a measure of how far reliability performance is from desired outcome, and if the performance is heading or trending in the preferred direction. However, the way in which the CDI calculation is described is more a measure of the ability to measure reliability performance and how accurately this can be measured rather than a measure of actual performance. For any desirable reliability performance indicator, if it is deemed performance metrics are insufficient, then improvements to those metrics should be considered before utilization as a component of a reliability performance indicator. Until a reliability performance indicator attains a "SMART" criteria level of acceptance, that reliability performance indicator should not be considered in determining reliability risk.

Furthermore, the CDI as it described is confusing. As examples, the Specific, Measurable & Attainable components of the "SMART" table are not considered in Table 2. On page 9, the whitepaper claims that "Metric Contribution" will range from 0 to 100, however, unless the "Metric Contribution" equation has a denominator of the sum of the weighting factors this calculation is not mathematically accurate.

The CDI requires further consideration or a more detailed explanation of the computations. The idea of the CDI is clear, but the method for obtaining the index is unclear.

Bonneville Power Administration

Comments for:

NERC: Integrated Reliability Index (IRI) Measuring Bulk Power System Reliability Performance

These are the combined metrics:

Event Driven Index (EDI)

Indicates Risk from Major System Events

Condition Driven

Index (CDI)

Indicates Risk from Key Reliability Indicators Standards

Statute Driven Index (SDI)

Indicates Risks from Severe Impact Standard Violations

BPA recognizes that this single metric appears to be intrinsically deficient in that it combines three important but dissimilar metrics. Meaning, an entity can lose substantial information about the whole grid if any of the metrics have a combination of higher bad and good values, leaving it untrustworthy to give a meaningful assessment. BPA would recommend instead, considering the use of the three source metrics directly in assessing the grid. BPA feels that this could be highly useful to the industry. If pursuit of a single metric continues, BPA believes it should be on a trial basis over a several year period to see if it does, in fact, present credible and useful information.

BPA suggests that perhaps the most efficient type of CDI indicator of an, overall system robustness will focus on:

- EEA's
- Reserve Margins
- DCS non-recovery
- IROL/SOL exceedances
- Misops
- Human Error

BPA observes the following:

- 1) On NERC's posted Concept document the copy did not have the equation 3 at the bottom of page 4, [just the (3) bracket], the next page (5) was the Appendix A.
- 2) What is happening to the standard indices, SAIDI, CAIDI, CAIFI? Are these replacing those?

3) Regarding SAIDI, etc., many utilities do not count major disturbances (storms, tornadoes, etc.) in their indices, and utilities interpret "major" differently. This reduces the ability to compare apples to apples. Will major natural disasters be counted in these indices?

PacifiCorp

PacifiCorp welcomes the opportunity to engage with NERC and industry stakeholders about measuring levels of bulk power system (“BPS”) reliability. The Company believes that the development of meaningful measures of existing levels of reliability, which can appreciate the costs to society and the benefits society experiences from any measurable changes, is a valuable conversation in which all stakeholders should participate. As such, PacifiCorp offers the following comments:

Marginal Improvements Resulting from Substantial Resources

- While the Integrated Reliability Index (“IRI”) project is designed to reflect currently submitted data, PacifiCorp is concerned that NERC’s ongoing quest to obtain more consistent data and an increasing number of data points will undoubtedly drive improvements in equipment and technology. This will again require the industry to invest time, effort and financial resources in a program that has yet to demonstrate any discernable or measureable improvement in bulk system reliability. As demonstrated in Figure 4 of the IRI White Paper, the industry’s Event Driven Index has ranged from 99.835 to 99.87 between 2008 and 2010. The paper recognizes that “No significant EDI trend changes are observed from 2008 to 2010.” There is limited quantitative support to warrant the significant time, resources, and equipment upgrades contemplated by this effort in the absence of more information.
- A cost-benefit analysis must be performed before any improvements can be expected. In addition, PacifiCorp is concerned that NERC has failed to consider the costs and effects to ratepayers of a .01% improvement in reliability. Ratepayers will have their own expectations about what level of electric service is appropriate at which cost and for what value.

Need for an Integrated Reliability Index

- The Company questions whether a need for an IRI truly exists. It is important that the indicators for reliability are well understood and actionable. As entities’ raw reliability data is aggregated, and as more time passes, the industry will be less able to take action to influence those results on a prospective basis. At the same time, it appears that an integrated metric invites an opportunity for external scrutiny of the reliability of the BPS. If the primary purpose of the IRI is to provide a means for this external scrutiny it is even more important that the data reflecting these elements is not aggregated so much that it is rendered meaningless. In addition, each system should be evaluated based on an individual component rather than a single composite number produced by the IRI formula.
- Additionally, the Company identifies several potential refinements or enhancements of the IRI structure:

- a. An annual approach might be considered as an alternative. Any expectation that this data could eventually be submitted on a “real-time” basis demonstrates a serious lack of understanding about the time and effort required to fully investigate and resolve “events.”
- b. A complete review of this entire approach should be planned for 2-3 years after its initial implementation. Technology advances, industry use, verification of data consistency and collection techniques must be anticipated and evaluated as part of the implementation process.
- c. Any index must support basic component modifications and technology upgrades that may result in differing data sets as a function of time.
- d. Any calculations must be reviewed periodically and, at minimum, annually.

Intended Use of an Integrated Reliability Index

- The Company is concerned that NERC and some stakeholders may not fully appreciate the limitations that exist in the application of such an integrated reliability indicator. It would be a gross disservice to the industry if it were to develop the IRI as a means of creating meaningful changes in BPS design, maintenance and operation but find it being applied in a punitive way. This is particularly noteworthy in that BPS reliability, especially in the changing utility industry environment, is heavily impacted by policies set outside the purview of the industry. Resource mixes, variable frequency generation, and Smart Grid all will result in changes to the reliability of the system, but such variables are not fully (if at all) accounted for in the Venn diagram shown in the IRI White Paper.
- Furthermore, the Company has concerns about the following consequences of the IRI to the extent they are intended:
 - a. Rather than use “trend analysis” to minimize the criticality of some of the standards that prove to have minimal or no impact on reliability, the industry should support an increase in the number of standards covered under NERC’s administrative citation approach.
 - b. It would be inappropriate to establish penalty levels based on event size, without regard to the identified violations of the standards and the actual impact on BPS reliability.
 - c. It is also inappropriate to establish a national average for a level of reliability and then penalize individual utilities for performing “below average” without concrete identified violations. Likewise, it would not be proper to establish levels of performance for individual utilities and then penalize them for a reduction in performance.

- d. In short, the IRI should not be used as a means in which NERC can utilize individual or aggregate performance data to justify penalties, sanctions and/or additional enforcement processes (such as investigations).
- e. Any index should measure the duration of time in which a system event is mitigated and service is restored, to incentivize behavior among utilities to resolve issues expeditiously.

“Operated as Designed” versus Unintended System Operations

- PacifiCorp believes that it is important that NERC recognize that any currently-proposed measurements of levels of reliability include both the intended consequences of safe system operation (operated as designed) as well as unintended system operations. Notably, the Standards/Statute Driven Index and the Event Driven Index are measures which reflect a world in which the system is functioning exactly as planned. As these measures evolve they should be modified to reflect measurements of unintended system operations as an appropriate metric for “unreliability.” If the intended system operations were performed day in and day out, the system should be deemed perfectly reliable, whether or not transmission lines, generators or load were out of service for discrete amounts of time.
- PacifiCorp offers the following suggestions to improve the measures to reflect unintended system operations
 - a. Outages occur for a variety of reasons, including environmental and economic factors. Including these elements into any index is critical to adequately reflect real performance.
 - b. Whether or not a system functioned “as designed” versus “not as designed” must be incorporated in any index, particularly when considering the potential intended uses of the IRI.
 - c. The index currently does not incorporate weather. There is a vast difference between reporting outages due to weather between the Eastern and Western Interconnections, for example. In addition, the index fails to recognize other critical national and regional differences attributable to geography, an area’s urbanization, utility service area, customer demographics, and population. Issues like this must be investigated and made consistent before the data used in this project can be deemed consistent enough to add value.

Period of History to Identify Substantial Trends

The Company believes that it is important to refer to other industries as a model for setting an expectation of the length of time and number of industry resources required to develop rational methods of measuring system performance. Notably, the nuclear industry needed about two decades to develop resilient and comprehensive data sets. Thus, it is critical that the electric utility industry be

afforded the opportunity to distill critical data, understand history for each of these sets, and determine causal relationships before any punitive use is contemplated. Each index should be thoroughly tested and the consistency of data should be verified before they are combined into an integrated index.

American Transmission Company (ATC)

ATC appreciates the opportunity to provide comments in response to the *Integrated Reliability Index (IRI) Concepts* whitepaper and thanks the Reliability Metrics Working Group (RMWG) for its efforts.

ATC's assessment of the proposed IRI is that it is difficult to discern what the metric is portraying as the IRI is complicated mix of reliability performance metrics all rolled into one, single number. The IRI calculation is the combined weighted product of an Event Driven Index (EDI), Condition Driven Index (CDI) and Standards / Statute Driven Index (SDI), which are each likewise derived from weighted products of other indices.

For example, the EDI appears to be a composite based on the number of system events that occurred during that year, weighted by severity. The CDI is a composite of 18 sub-metrics with a subjective weighting factor for each. The SDI is a composite score of industry violations on 26 Standards with a weighting factor for each.

This leaves the reader with the following questions:

- What does the IRI mean?
- How can I tell a "good" from a "bad" IRI score?
- What should I do if the IRI score is bad?

Although the paper provides illustrative examples of the EDI and SDI calculation, neither of the factors appear to offer any additional *meaningful* insight to the industry.

In Appendix C, the SDI is noted as having a *slight improvement in reliability performance of 0.11%* between 3Q2009 to 4Q2010, although the magnitude of the change is so small it is difficult to discern whether the change is relevant.

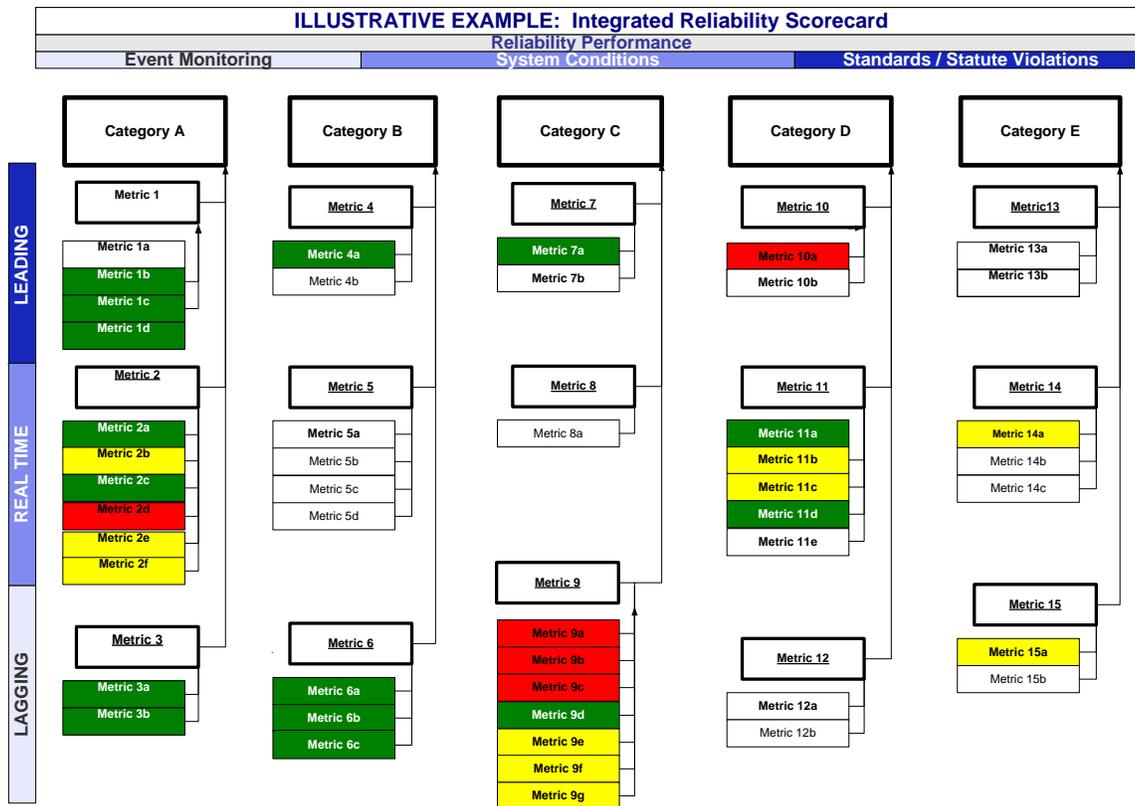
In addition, ATC questions whether the list of Standards utilized to calculate the SDI are the best reflection of risk to the system. Recall the 2003 Eastern Interconnection black-out. The three leading causes have been referred to in the industry as the 3 Ts: Trees, Training and Tools. The list of Standards in the SDI does include a reference to FAC-003 ("Trees"); and PER-002 ("Training"); however, there are current "Tools" specific Standards that ATC is aware of. Therefore, this may be some additional food for thought in terms of how to develop an adequate IRI.

This leaves the reader with the following question:

- With so many factors being brought together, what events or trends could possibly "move the dial" on the IRI?

Currently, the paper does not offer actual values for either the CDI or the composite IRI score. Therefore, it is difficult to assess how the IRI would meet its stated goals. ATC proposes that the IRI be calculated for the 2008 through 2010 period so that this assessment can be made.

One possible suggestion for the RMWG to consider is a Scorecard approach, which would provide a means to consider a multitude of reliability performance metrics and group them as a means of looking at similar metrics (perhaps over differing time horizons; i.e. leading, real-time and lagging) while still providing an easy and meaningful way for the reader to understand drivers of “good” and “bad” reliability performance with colorization to indicate where to focus improvements.



Whatever approach is taken, ATC recommends that the RMWG use historic events (i.e. such as those cited in Appendix A of the paper) as a benchmark to evaluate whether or not the approach used meets the stated goals; i.e. to provide the industry with:

- **Meaningful trends of the bulk power system performance**
- **Guidance on how [to] improve reliability**
- **Support [for] risk-informed decision-making**