**Balanced Scorecarding and Dynamic Evaluation**

**By: Dr. Tom Bigda-Peyton, assisted by Ani Koulian**

## I. Background and Context

This paper will review the potential of combining Balanced Scorecarding (BSC) with Dynamic Evaluation (DE) to create a new, more robust approach to performance measurement in healthcare. Case studies from the three Health Transformation Learning Partnership (HTLP) organizations: the South East Community Care Access Centre (SE CCAC), York Central Hospital (YCH), and North York General Hospital (NYGH) will be used.

Since the early 1990's, Balanced Scorecarding has been the best-known effort to create performance metrics beyond "just numbers" (or cost), and has been associated with Kaplan and Norton and the Balanced Scorecard Collaborative. Leading up to the official start of the HTLP project, during a 2008 conference, in a presentation entitled "Can Local Governance Survive?" Project team member Ted Ball suggested that Ontario was strong in strategy development but weak in strategy execution and deployment. He pointed to Kaplan and Norton's newest book, *The Execution Premium,* which made a similar argument and provided several case studies to support it, such as the experience of the Canadian Blood Services.

This argument was valid. However, it is important to note the difference between strategy execution from a 1st Curve (technical change) and 2nd Curve (adaptive change) perspective. The notion of a 2nd Curve health system was developed by Dr. Martin Merry, based on the work of Ian Morrison and Alvin Toffler.

Writing in 2008, Ball adopted Merry's argument and argued that we need to "change our way of thinking" in order to transform our healthcare system. As a result of the HTLP project, we are now in a position to address this issue by adding three dimensions to Ball's original argument:

- An enhanced perspective on strategy execution;
- New methods for promoting adaptive learning and change; and
- Ways of targeting the model to new challenges, which have emerged, or become more prominent, since 2008 (i.e. Access to Care, Emergency Department/Alternate Level of Care, and preventable hospital readmission).

These three dimensions enable us to take into account shifts in the environment since the HTLP project began (i.e. the adoption of the *Excellent Care for All Act (2010)*, and to aim the resulting model at the challenges that will continue to be faced by the Government of Ontario. We believe those challenges can best be described as making the transition to a transformed, 2nd Curve healthcare system in Ontario. This article attempts to chart a course for how to frame, evaluate, and measure progress through that

transition, and is thus part of an emerging program we call the Health Transition Learning Partnership.

## II. Using Dynamic Evaluation to Solve Emerging Problems in Measurement

Drawing on the disciplines of action science, reflective practice, frame reflection, and systems dynamics (Argyris and Schon; Forrester; Senge), the author has developed Dynamic Evaluation as a response to challenging problems of measurement and evaluation in large system change initiatives. These questions include:

- How can we measure innovation, such as the adoption and spread of best practices?;
- How do we develop reliable systemic indicators?; and
- How can we embrace and fully utilize quantitative as well as qualitative information?

Dynamic Evaluation posits a framework for data-gathering and analysis which implies response to these and other challenges in healthcare measurement and evaluation. Called "When the Stories Meet the Numbers" (see appended slide), this framework is based on several related assumptions:

- When something has happened only once, the data we have is an anecdote, or "story." When it has happened three times, we have a pattern (which we can recognize by comparing and contrasting across instances). When it has happened seven times, we reach an initial threshold of statistical significance because the phenomena has occurred often enough that we have recognizable frequency for the first time.
- The methods which are useful for data analysis vary according to the type, and rarity, of the data. For instance, we use story analysis for anecdotes and narrative (n's of 1), pattern analysis for n's of 3, and numeric analysis for n's of 7).
- Story analysis yields maps that reflect the logic of action which produced desirable or undesirable outcomes. These maps are the basis for anticipating the likely trajectory of change in the next phase, and for designing action experiments that attempt to replicate success models and formulas in the future.
- Dynamic Evaluation thus enables an organization to learn better from its own experience, spot emerging trends, and intervene during a change in order to recalibrate, make needed adjustments, and foster desired changes sooner and more effectively than is possible by using other methods of evaluation.

## III. Dynamic Evaluation: Case Examples from the HTLP Project

As of 2007, the South East CCAC was launched as an amalgamation of three former CCACs which had served southeastern Ontario. One immediate challenge the new

management team faced was how to evaluate performance and measure the impact of change, given that they were a new organization and lacked appropriate data.

As one of the HTLP organizations, the South East CCAC was committed to implementing a Strategy Management System, including a Balanced Scorecard, Office of Strategy Management, and Mutual Accountability Agreements for the Senior team, Directors, and Managers. However, the team faced multiple implementation challenges, including the lack of an organization-wide information technology system and ways of monitoring performance across the region. As they embarked on the journey, they struggled with a key measurement problem: *how do we measure and monitor performance when we lack reliable, consistent data?*

Case #1: South East CCAC

In August 2008, the author attended a meeting of the South East CCAC Strategy team with HTLP Team Member Ken Moore. The team had read *The Execution Premium* and with Moore facilitating, reviewed the Balanced Scorecard and Process Improvement Program (PIP), a performance management program requested by the South East Local Health Integration Network (SE LHIN).

In the process, they also reviewed anecdotes and numbers related to key performance targets, as well as patterns which could be detected from the above. Moore and Bigda-Peyton introduced key frameworks, models, and tools from their respective methodologies (Balanced Scorecarding and Dynamic Evaluation). For instance, they looked at "When the Stories Meet the Numbers" an integrating framework for linking anecdotal evidence (i.e., "storytelling") with "hard," numeric evidence (i.e., choosing metrics for the Balanced Scorecard).

Strategy Team members began identifying three metrics (per Strategy theme) that they could use to measure progress on PIP, and on their Balanced Scorecard, over the next few months. In parallel, several examples of emerging linkages between the BSC, PIP, and Dynamic Evaluation (as represented by using both stories and patterns, as well as numbers, as "data" for monitoring performance) were identified.

For instance, one Finance metric is "cost per case." The theory behind Dynamic Evaluation predicts that qualitative information can reveal otherwise invisible organizational dynamics that affect this metric. In the case of the SE CCAC, another advantage of the method was more compelling in the near-term: at the time, there was no reliable way to identify cost per case, because the quantitative information simply was not available. Questions such as the following were asked:

- Could Dynamic Evaluation assist the SE CCAC in identifying the underlying dynamics affecting cost per case, while usefully incorporating qualitative as well as quantitative information?

- Could this happen in a tight timeframe? (The SE CCAC had to report to their LHIN Board every two months on the PIP initiative, starting soon after the meeting in late September).
- Could the findings be generalized to help other healthcare organizations solve similar problems, such as evaluating improvements in patient satisfaction, quality, and safety without reliable quantitative data?

In particular, it was felt that the most useful application of Dynamic Evaluation would be to evaluate the PIP Program. As of September 2008, there were 5-6 PIP initiatives, all of which were designed to help reduce "cost per case." It was necessary to establish a qualitative baseline and progressively demonstrate progress toward the goal. Could this be done? If so, how?

One project involved implementing service delays, which was later reconstituted as "planned service delivery." As the project began, emerging evidence of the impact of this initiative, using both quantitative and qualitative analysis was gathered. For instance, at a teleconference of CCAC client services managers, it was reported that:

*"We are on Day 52 of the service delay approach. We are hard at work. We have assessed 552 cases, looking at admits and exceptions. We admitted 197; 158 were priority 1; there were 39 exceptions. We have put 338 on the waitlist."*

From this data alone, it would be hard to tell what the impact on cost per case would be (since no average, or aggregate, "cost per case" figure was available). It would be possible to use qualitative data to help, though, by gathering "stories" from key middle managers who were running the case management assessment process, using typical assessment categories (i.e., "admits," "exceptions," "priority 1," and "waitlist") as "filters."

For example, a potential best practice indicator from the qualitative information gathered from calls with the middle managers was gathered. One of the managers noted that she had been working with new staff members to determine delays in homecare. When implementing such a delay, she noted, the practice for the caseworker was to *run it by your manager, use a rating tool, plead your case with the client, then put them on a waitlist*. She was working with staff to *determine how they came up with their rating, is the data correct?* This was worthwhile, she said, but it was time-intensive and layered on top of other everyday duties. She also mentioned having done this with *19 new staff in a year*.

This anecdote provides a mid-level, or even ground level, snapshot of how service delays are implemented, and what the work of managing service delays can entail. It also seems to link back to "cost per case." The guidance and coaching provided by this manager is probably reducing cost per case in one way but adding to it in another. By helping new staff to learn how to implement service delays, the manager appears to be reducing cost per case; but by apparently doing it one staff member at a time, she is inadvertently using more of her own time than she needs to and is thus increasing cost per case. If the 19

new staff could be divided in half or thirds, and helped to learn and practice this standard of practice by others aside from their manager, it was predicted, the process could unfold in a similar way with less time-intensive effort by the manager herself.

This pattern is reflective of a "craft model" way of functioning and has been evaluated (both quantitatively and qualitatively) for system-level performance in other areas of healthcare; thus we can make predictions on cost-per-case impact from outside experiences. *(Specifically, it is know from the work of Dr. Merry and Patankar, Brown, et. al. that a system based on a craft model cannot achieve better than a 4 Sigma error rate (about 6000 errors per million cases), because things will routinely "fall between the cracks." In contrast, a system based on 2nd Curve design will achieve a 6 Sigma error rate (3.4 errors per million cases) because the system is aligned, coordinated, and designed such that things do not fall between the cracks).*

With reference to the South East CCAC, it was suggested that this anecdote should be treated as a leading indicator and should be investigated further. A working forum of case managers, client service managers, and finance-operations personnel could identify ways of obtaining the benefits this manager is achieving (and perhaps doing so more rapidly) while also reducing her personal workload. The effect of such a plan would be to enhance and maintain a high-quality practice while also reducing cost per case.

Another emphasis of Dynamic Evaluation is impact analysis, i.e. identifying unintended consequences of current decisions and actions and designing remedies for them in advance. This is sometimes referred to as identifying "leading indicators." For instance, during a call with the South East CCAC middle managers, the observation that *"an adverse event is becoming likely,"* was surfaced. They felt the organization was at growing risk of an adverse event because *"We have had close near-misses, missed service, intravenous therapy, referral delays, and a provincial Red Cross strike."*

It was suggested that this observation should be explored further. It led to a concern that cost-cutting may have reduced the CCAC's "safety envelope" to an overly risky extent. However, without further inquiry, the extent of the risk could not be determined. From this observation, some of the key metrics used, in practice, by experienced middle managers was identified. According to the consultants' observations at the time, the risk could be reduced simply by examining it, by discussing it with the management and leadership team, and by identifying the proxy metrics being used for wider use and review going forward. In turn, this would reduce cost per case (or at least not increase it), since adverse events (and delays, wasted effort, and re-work in general) are costly. This would positively impact both the PIP report-out and the Balanced Scorecard (which identifies "reducing the cost of adverse events" as one strategic objective).

This kind of story-based analysis is best done iteratively and across levels and areas of the organization, because it highlights leverage points, common practices, and "organizational DNA." It surfaces the logic of action in the organization, thus exposing the "drivers" of the results identified in the PIP and the Balanced Scorecard. By doing

so, it makes the landscape far clearer and more amenable to constructive intervention and change than alternative methods.

At this point, it was felt that combining Dynamic Evaluation with Balanced Scorecarding, the Strategy Management System, and the Leadership and Learning Journey would make each of the component parts more valuable. In the case of Dynamic Evaluation, this is because the consultants were not solely reliant on anecdotes to surface important areas to examine - this could be done through the management forums, Learning Organization tools and methods, and observations generated by the Balanced Scorecarding/ SMS process. When these avenues are used (i.e. when "cost per case" has gone up), the consultants can work backwards to the qualitative evidence underneath that observation. In short, by using Dynamic Evaluation one can get at "drivers" of organizational performance through quantitative, qualitative, and anecdotal pathways.

As a final dimension of the work with the SE CCAC, the consulting team suggested that the organization needed to identify, and agree on, a framework and measurement system for system-level improvement and performance measurement. The Senior Team and Board indicated interest in moving toward "Second Curve" healthcare as such a framework.

This framework involved moving from a Current Reality that is characterized by a craft model and a focus on client loyalty, to a Future Vision that is driven by conscious system design and both client and organization focus. In 2008, the consultants planned to select a few additional qualitative (and quantitative) markers of changes in beliefs, operating assumptions, and organizational culture at the South East CCAC. In 2009, the "Tower of Power" model was introduced as a way of helping the organization move through the "learning dip," or transition, implied by the movement from late First Curve to early Second Curve functioning.

In 2010, the consultants worked with senior management to begin framing an action learning project for the Executive Team and Management and Leadership Team. In 2011, the consultants are implementing that project with the new senior team and the directors and mangers for the Access Team (Client Services).

In brief, the SE CCAC is moving "up" the Second Curve by using the Action Learning Systems methodology for navigating a system-wide transition toward health system transformation. Dynamic Evaluation is a necessary, but not sufficient condition, to make this transition. Given current beliefs and assumptions about valuation and measurement, something like a Strategy Management System is also needed. Beyond that, however, a system is needed, a way of thinking, and a critical mass of leaders with the skills to act as navigators rather than repeat the familiar command-and-control approach, or to go to the other side of the pendulum and believe that innovation will emerge from grassroots initiative entirely. The skills to enact collaborative design and live up to it in day-to-day practice in healthcare, is lacking. Thus, not just the vision, but also the skills implied by the Second Curve, appreciative inquiry, and Dynamic Evaluation in order to create, maintain, and sustain excellence for all in healthcare, ought to be embraced.

Case #2: York Central Hospital

As York Central Hospital developed its Wildly Important Goals (WIGs) in 2010 (see HTLP article, "Enhancing Strategy Execution at York Central Hospital" for further details), the organization also experimented with adding a qualitative dimension to strategy execution.

In the summer of 2010, the author began formally introducing the discipline of storytelling to York Central in a workshop for the Senior Team, Directors, and Managers. This meeting introduced the basic ideas of storytelling (see HTLP Project Book) and applied them to the 100-Day Action Plan program to balance the budget that was upcoming.

Working with David Stolte (Chief Strategy Officer) and Jo-Anne Marr (then Interim CEO), three success stories to emphasize Agency Reduction, Code of Conduct, and Diversity were chosen. All three were campaigns led by Marr in which elements of an emerging discipline of strategy execution could be discerned, recognized, and reinforced. For instance, in 2009, a task force had been set out with a goal of reducing the use of outside agency nurses to zero. This would be an important step toward managing the budget, reducing necessary nursing expense, and working better together across areas, departments, and disciplines. The team had done a good job and had made significant steps toward the original goal, but had also learned some key lessons:

- It took longer than they thought to achieve the budget reduction targets, because they had to invest in building new skills and capabilities first;
- There were unintended consequences of the campaign - sometimes costs went up before they could even out or go down; and
- The numbers did not tell the whole "story." Indeed, the use of the story analysis method helped the group surface the observation that agency use was starting to go back up after the campaign. This meant that the campaign was not "over," but instead required persistence and vigilance to maintain the gains.

The "agency" story turned out to be an important prelude to the 100-Day campaign which started soon after. For one thing, it showed the value of working well together as a multi-disciplinary task force and team. Further, it revealed the benefit of focusing on a few goals, meeting regularly, using project management approaches, "keeping score," celebrating wins, and keeping up the momentum through frequent communication, visual management, and engaging the front lines in coming up with ideas to solve the problem.

In addition, the June 2010 workshop highlighted two other internal YCH "stories": Diversity and Code of Conduct. Both showed that at the heart of budget reduction was connecting to values, mission, purpose, and commitment. The code of conduct project came from a sense that staff were not treating each other with respect. With the help of a task team, Marr engaged each department or unit in reflection on appropriate and needed behavior. This led to a hospital-wide declaration of workplace values and culminated in

a “signing ceremony” in each area to lend personal commitment to the new statement of values, mission, and expected conduct. With respect to diversity, the management fostered discussion both of race and gender, and of opinions, backgrounds, disciplines, and orientation. This created a climate of discussion, inquiry, and greater acceptance of differences. These elements were incorporated as well into the planning for the 100-Day campaign, for instance, by balancing the budget recovery WIG with a WIG on patient safety. This was building on the idea that the WIGs needed to be not only about budget recovery, but also about staff and patient satisfaction.

York Central Hospital also used storytelling in other ways to generate qualitative information to enhance strategy execution, specifically by working with the author to do a story-based debriefing of the 100-Day Action Plan. The idea was to do story-gathering with managers, directors, and Executive Leadership Team (ELT) members, then conduct a feedback session with those interviewed, in order to develop shared learnings and use them to guide and inform next-stage WIGs. In December, the author conducted story-based interviews and focus groups with all ELT members, directors, and managers who were involved in the 100-Day campaign. In March, after consultation with David Stolte, the author facilitated a story-based After Action Review with the same participants. The combination led to several discoveries about the 100 Days, using the qualitative dimension of strategy execution.

Story analysis is a key component of Dynamic Evaluation. It involves looking closely at anecdotes or “stories” to identify outcomes, patterns of practice which are linked to those outcomes, and operating assumptions that drive those patterns of practice. Story analysis relies on the use of several tools, one of which is story mapping. In the 100 Day After-Action Review at York Central Hospital, the author introduced one kind of story map, called a Situation Map™. This was the format for organizing the information gleaned from the focus groups and interviews into a pattern, or patterns, that the group could validate, challenge, add to, and learn from. In brief, it was the mechanism for converting “stories” into “patterns.”

Since the 100-Day campaign for budget recovery at York Central has been discussed elsewhere in the articles and books produced by the HTLP program, we will not re-tell the story. However, the mapping process drives toward a shared interpretation of the story by the group reviewing the map (in this case, the ELT). This will be the focus of our discussion here. This was relevant in the York Central context both for increasing the odds that lessons learned from the 100-Days would be re-used in subsequent campaigns (and in executing strategy in general), but also because David Stolte had mentioned before the review that although there had been many ad hoc conversations about the 100-Day campaign, his sense was that there had not been a coherent agreement among the leadership on the implications of the 100-Days for subsequent work. We set out to reach such agreement while also surfacing areas of divergence for further discussion and review.

The ELT discussion affirmed much of the map but also added to it and changed it in important ways. For instance, the group immediately questioned and challenged the

author's inclusion in the title of the map a quote from the late Gordon Cheesbrough which had become a key aphorism in the HTLP Program: "The First Step Toward Innovation is a Balanced Budget." York Central's ELT felt that this did not represent their experience:

- *It (the 100- Days) was not about that. It was about sustainability and financial health. To "Be the Best in Canada" (a York Central vision), we had to tackle this.*
- *We felt we had to balance the budget first before we could do patient satisfaction. We had to get our house in order.*
- *The situation was dire; we had a burning platform. Our viability was threatened. We felt "this needs to be dealt with."*
- *Our run rate was not sustainable; we had to tackle it in order to stop the "yo-yo" stuff. The focus was on sustained financial health so we can do what we need to do.*
- *It was a major first step toward stability and financial health. The beginning of a journey.*

This kind of discussion is healthy because it helps add to, and re-shape, the map. In addition, since the purpose of the map is reuse in analogous situations in the future, it is important that the group "own" the naming and language of the map as well as its implications for them. In this case, the author acknowledged the need to remove the quote from the title and to add the comments to the "Assumptions" column of the map.

A second area of discovery was the identification of key actions that created momentum early on in the campaign:

*We felt like we had a really large mountain to climb; a daunting task. But three things happened right away that sparked momentum:*

- *The supply chain manager came back and said "I captured $1 million." The rest of us said, "Wow - why not us?"*
- *The bed management task group had a VRE outbreak but made a breakthrough anyway*
- *We had been thinking about implementing a Patient Care Coordinator for years, and when the campaign came we said "let's do it."*

*These all happened within a couple of weeks; they were really significant wins. The senior team has had problems with execution here for a while; finally, we were executing. This built momentum.*

The discussion of the story map led to the identification of the key actions and agreement among the ELT on their significance.

A third area of discovery from the review of the story map was the set of key actions that drove results. These actions were the implementation of daily management techniques

(including visual management), the development of a new, more actionable unit of measurement (UPP), and the introduction of monthly collaborative program review. These three innovations, taken together, provided the needed tools, practice, and reinforcement needed for unit managers to become more proactive and accountable for managing their budgets. Here are excerpts of the discussion on visual management and new metrics for daily work:

- *A plus was the use of visual management.*
- *(Bigda-Peyton): The spread of visual management happened "by surprise" because you were doing 100-Days and Lean concurrently. This gave a compelling context and purpose for using visual management.*
- *And the move from HPPD (Hours per Patient-Producing Day) to UPP (Unit-Producing Professionals) really helped with daily management.*
- *It was an enabling process. An outside consulting firm that had done an Operations Review prior to the 100-Days' campaign had used the language of HPPD; with UPP, you see it every day.*
- *Another unintended positive was that for the first time, managers became aware of the budget on a daily basis.*
- *Managers did not believe the numbers represented what was going on in their areas. How did the numbers play out? UPP cleaned things up and created both understanding and accountability.*

Monthly Program Review was another key piece of the budget recovery toolkit. It was conducted in 30-60 minute roundtable discussions with each unit manager and the full ELT team:

- *Managers had to deal with staff and consequences. The outside firm identified 52 recommendations that were focused on specific areas. What was missing was Program Review and templates.*
- *The underlying assumption was that everyone was clear about their budget target. This was NOT true.*
- *They didn't believe it. We said, "What do we do now?"*
- *Due to Program Review, we changed the focus from "budget target" to "run rate".*
- *This helped each manager focus on their role.*

As with any change, though, implementation was not easy:

- *The first meetings regarding HPPD were challenging.*
- *So were the first meetings of Program Review, there was a sense of fear and trepidation going into them.*
- *The monthly meetings drive accountability. Here, there is shared understanding and shared ownership. It's not like that everywhere. Those meetings are about problem-solving and how to help each other out; how do we figure it out? From identifying variances to collective solutions. A culture of shared accountability.*
- *Somehow we made the revenue issue disappear.*

- *We combined funding, volume, and cost per activity. Understanding leads to clarity and confidence.*

Since the ELT conducted the Program Review sessions in a spirit of collaborative problem-solving rather than blaming, the sessions were productive and soon gained a positive reputation.

In summary, the ELT's review of the initial story map for the 100-Day Action Plan yielded new, actionable insights which could be used in the next cycle of strategy execution. From story-gathering in interviews and focus groups, it was evident that all of the things in the "Actions" column played a role in the success of the 100-Day campaign. What emerged out of the session was more about the ELT's "Beliefs and Operations Assumptions" and about their "Thoughts, Feelings, and Intuitions"- in particular, their sense of a threat to their viability and a resulting feeling that they faced a daunting task. Given that, the standout elements of a new version of the map became:

- The three early wins;
- The development of a culture of shared accountability by how the Program Review was handled (that is, by a culture of reporting and collaborative problem-solving rather than one of blame and shame), and
- The development of a translational, transitional metric that had meaning for staff and could be used to get traction with daily work.

The consultants' updated hypothesis is that these are the key elements of creating and building momentum for other WIGs, in addition to the elements noted on the original map.

## Case #3: North York General Hospital

North York General Hospital used Story Analysis and Situation Mapping in 2006 to make the transition from Strategy Development to Strategy Execution. The hospital had gone through the capacity-building journey and had done a great deal of work with their Strategy Map and Balanced Scorecard. They had a two-year history of becoming a Learning Organization and had become familiar with the use of Personalysis as an individual and team tool for assessing personality differences in ways of processing information. However, they had not included the front lines, or operations improvement, in the journey, in any significant way.

At this juncture, the author was invited to conduct a two-day visit on Storytelling and Dynamic Evaluation at North York General Hospital. This included an orientation to the methodology for the senior team on the first day, followed by an introduction to Storytelling and Story Analysis. On the second day, the author led a session for the Management and Leadership Team and Facilitators' Team. Attendees included the top vice-presidents, directors, and managers of the organization. Since North York General had been exposed to storytelling through the author's coaching and advisement, these

sessions focused on the use of story analysis, including Situation Mapping and the Wheel Model, as the next phase of skill acquisition and capacity-building in the Action Learning Systems methodology.

After an introduction to story analysis and the related tools, the author asked the group to identify a story (of frustration or success) that the group could work on together as a way of learning the methods by applying them to an important problem. They chose the Patient Flow project. They worked through this story and created a map.

The sense of concern, and even urgency, about the "flow" problem was widely shared. Indeed, one of the most powerful standouts was a collective "aha" moment which happened when the group was asked, before proceeding, how many of them were affected by this problem. Easily 75-80% of the hands in the room went up; in and of itself, this seemed to be a surprise to the group. Later on, a senior vice-president referred to this as a turning point, a positive sentinel moment.

From that moment onward in the meeting, she felt, there was agreement that "we're all in this together," in contrast with the invisible, though widely shared belief, before this workshop, that it was up to the Emergency Department to "fix" the patient flow problem. This standout did not show up on the map, but it was sparked by the exercise of working through the map.

A project team had been convened to look at policies, and potential guidelines, for the "use of flex beds when the Emerge is full." Using the Situation Mapping method, it quickly became apparent that this statement of the problem was not useful, because (as the group confirmed) "Emerge is always full." Thus, the focus of the inquiry was changed, early on, to the guidelines for the use of flex beds "when there is gridlock." (The term "gridlock" came up as a useful description, or metaphor, for the situation that the group could recognize and use in action.) It was concluded that, for the most part, the use of flex beds helps in such situations. Thus the first "finding" was a revised assumption: *"When in gridlock, flex up."*

The group had to go further, though, because of the wide variation in practice between units on the use of flex beds. This variation created a variety of unintended consequences, most notably an inability to coordinate flow strategies across units when there was a surge. Thus, the group began to explore a variety of options for approaching the problem. These ranged from adjusting the practice for bed alerts, to coordinating communications across units, to examining interconnections with the Province on after-care.

In order to prioritize possible strategies for improvement, the group focused on intended and unintended consequences, and on current actions. While the use of flex beds had worked to some extent, the overall situation still yielded significant unintended consequences, many of which directly impacted patient safety.

Indeed, one key aspect of the map is that it demonstrates that patient flow is an important patient safety issue. Across all clinical areas, it was discovered that the greater the sense of "gridlock," the greater the sense of frustration among staff and patients, the more "acting-out" one saw from patients, and the greater the risk of adverse events which would not have happened otherwise.

By constructing the map, the group could see that the organization had been making efforts to deal with the issue, but that these efforts were localized by unit and there was significant variation among the various unit-level strategies. The dialogue revealed a major opportunity for cross-unit learning, both to reduce unwanted variation in practice and to increase the spread of useful local innovations.

When the group looked for major leverage points to shift the picture from "gridlock" to a smooth "flow" of patients through the hospital, the areas of greatest impact had less to do with specific actions, such as revised policies, procedures, structures, and processes. Rather, the areas of greatest leverage were found in the operating assumptions that were held as obvious, and which can best be surfaced through this kind of story-mapping.

The key operating assumptions in the NYGH Flow Map were:

1. Doctors can only release patients when they are 100% ready;
2. Emerge is full all the time - it operates at 97% capacity; and
3. The hospital operates on a 9 to 5, five day a week schedule (while patients operate on a 7-day per week, 24-hour per day schedule).

These First Curve assumptions were deeply embedded in the organization's habits, practices, structures, policies, and procedures. They were hard to change because they were so taken for granted, and because they were reinforced every day by standard operating procedures. Yet, changing them represented the best chance for breakthrough to a Second Curve system and performance level.

This application of the Situation Mapping approach to improving flow from the Emergency Department to General Internal Medicine led to a conclusion by the Vice-President of Quality, supported by the CEO, to invest heavily in Lean improvement methods, such as Kaizen events and Value Stream Mapping. NYGH went forward with this initiative and became a provincial leader in process improvement during the next two years after this session.

The work with story analysis and Dynamic Evaluation was pivotal in helping NYGH managers and leaders arrive at an "aha" moment ("we're all in this together") and recognize that the next transition they needed to make was toward reducing variation in practice. At the same time, some of the work of changing basic operating assumptions from First Curve norms (such as operating on a 9 to 5, five day a week schedule) to Second Curve routines (such as operating on a 7-day a week, 24-hour per day schedule) was left unfinished during this era.

## IV. Authors' Observations

It is important to note that the partners were not trained in Dynamic Evaluation as part of the HTLP program, nor was the methodology used except by the author as a facilitation approach.

The author encouraged the SE CCAC to adopt the approach and to integrate it with the accepted approaches to measurement and evaluation, specifically, gap analysis and numeric reporting of statistics.

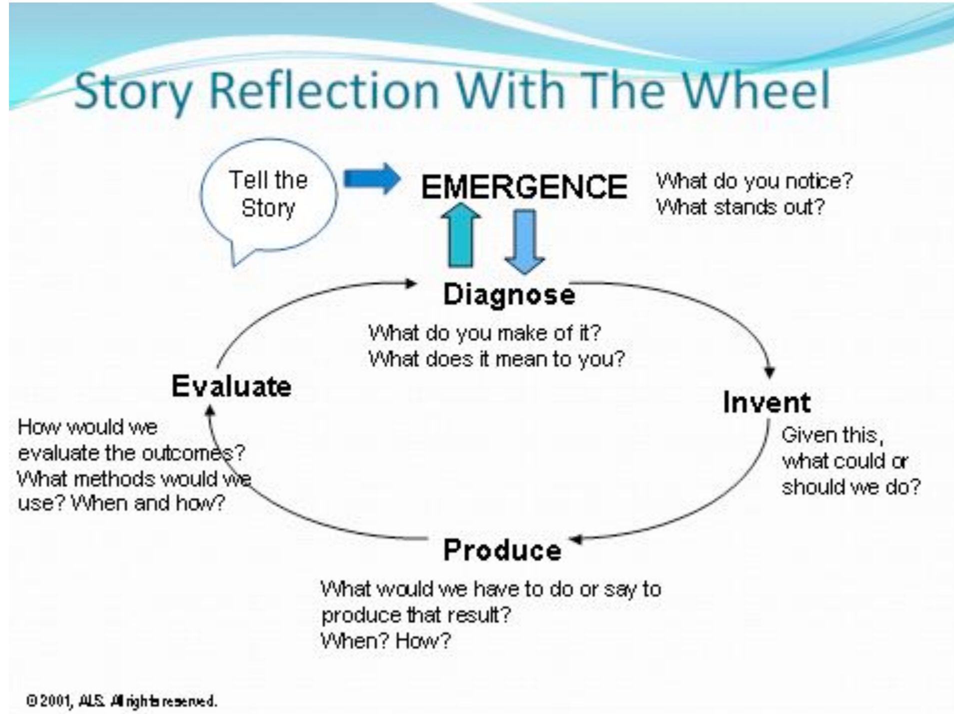
York Central Hospital came closest to using the approach during the project, by including a "qualitative dimension" in their emerging formula for enhancing strategy execution.

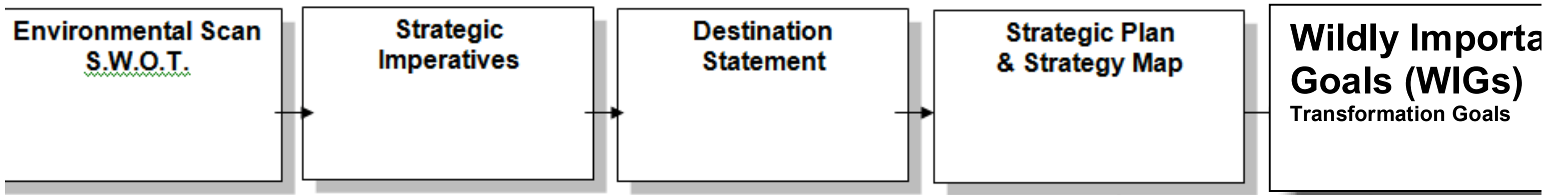
At North York General Hospital, the methodology was introduced as "storytelling." The next step would use Dynamic Evaluation as the methodology for measurement and evaluation.

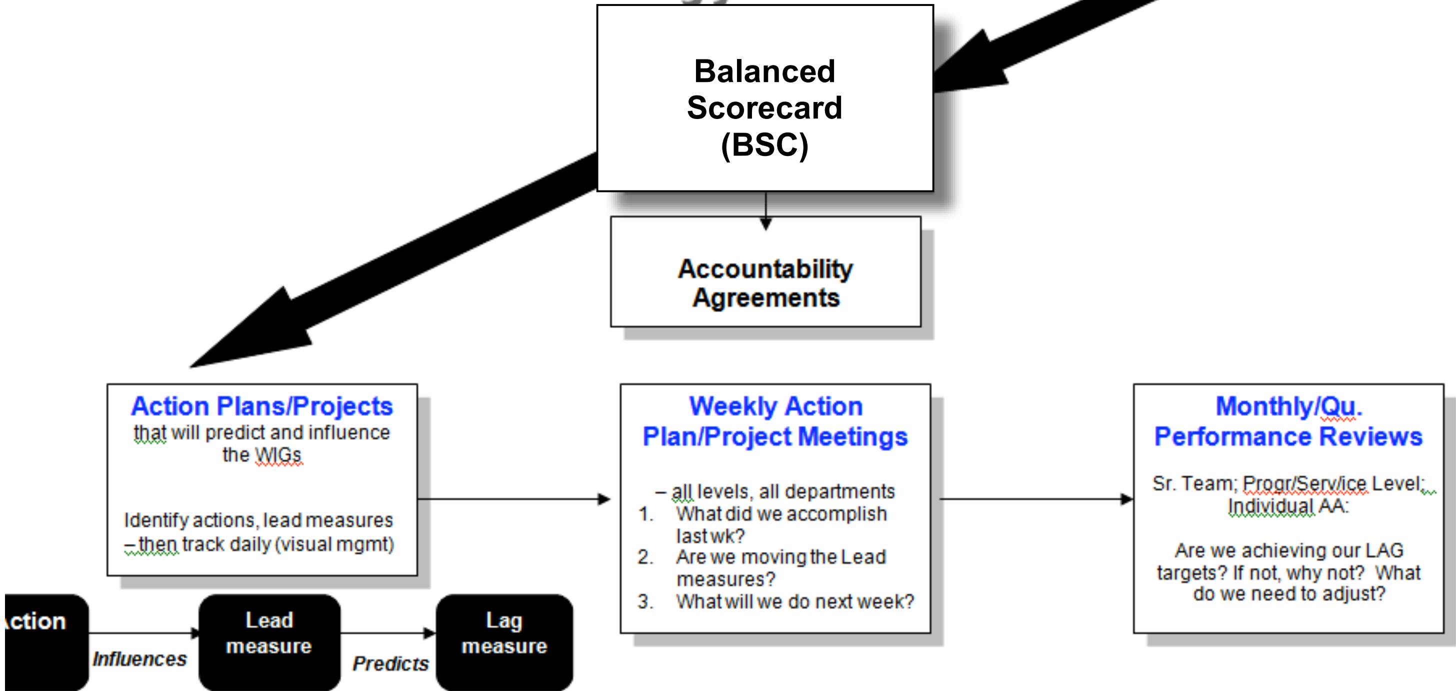
## V. Concluding Points

- Storytelling and Dynamic Evaluation is a practical guide in assisting an organization in identifying areas where execution is weak. It helps to move an organization from strategy development to strategy execution.
- STDE is an important component of executing strategy successfully.
- STDE enables people to see strategy execution in a concrete and attainable manner.
- The STDE methodology proved successful in all three HTLP partner organizations.

**Tools**




Strategy Development
Environmental Scan
S.W.O.T.
Strategic
Imperatives
Destination
Statement
Strategic Plan
& Strategy Map
Wildly Importa
Goals (WIGs)
Transformation Goals


Strategy Execution
Balanced
Scorecard
(BSC)
Accountability
Agreements
Action Plans/Projects
that will predict and influence
the WIGs
Identify actions, lead measures
– then track daily (visual mgmt)
Weekly Action
Plan/Project Meetings
– all levels, all departments
1. What did we accomplish last wk?
2. Are we moving the Lead measures?
3. What will we do next week?
Monthly/Qu.
Performance Reviews
Sr. Team; Progr/Service Level; Individual AA:
Are we achieving our LAG targets? If not, why not? What do we need to adjust?
ction
Influences
Lead
measure
Predicts
Lag
measure

# When the Stories Meet the Numbers

Nature of the Task

| | Ambiguous /New | ⟷ | Routine /Familiar |
|---|---|---|---|
| **Phases & Functions** | ***I Observation*** | ***II Assessment*** | ***III Measurement*** |
| **Data Format** | Stories | Patterns | Numbers |
| **Sample Size** | n= 1 | n = 3 | n = 7 |
| **Features /Characteristic** | Emergent | Diagnostic | Statistical |
| **Descriptors** | "Anecdotal" | "Qualitative" | "Quantitative" |