PM3: A Performance Measurement, Monitoring, and Management System
for Local Children’s Aid Societies

Robert J. Flynn
School of Psychology & Centre for Research on Community Services
University of Ottawa

Raymond Lemay
Services to Children and Adults of Prescott-Russell

Hayat Ghazal
Centre for Research on Community Services, University of Ottawa
& Services to Children and Adults of Prescott-Russell

Stéphane Hébert
Services to Children and Adults of Prescott-Russell

Introduction

People are constantly engaged in measuring, monitoring, and managing their own performance. The driver of a car, for example, uses a speedometer to measure the speed of the car, glances frequently at the speedometer in order to monitor how fast he or she is going, and then takes corrective action to manage his or her driving behaviour (e.g., by slowing down to avoid a speeding ticket or speeding up to be on time for an important meeting). Organizations also engage in these three intimately connected processes to optimize their performance. However, unlike the driver of a car, organizations often have only intuitive and subjective measuring tools at their disposal, such that monitoring their own progress may be slow and imprecise and managing the attainment of their key goals difficult and uncertain.

Despite these challenges, considerable progress has been made over the last decade in the field of organizational measurement, monitoring, and management. The purpose of the present paper is to describe the evolution, current functioning, and future development of a system that we refer to as PM3, to convey succinctly that it is concerned with promoting high-quality performance (P) through the judicious use of measurement, monitoring, and management procedures (M3) in child welfare organizations such as Children’s Aid Societies.

In the literature, performance measurement is often used as a convenient short-form that subsumes all three terms, performance measurement, performance monitoring, and performance management. We will adopt this convention here. Crediting Joseph Wholey with being the main architect of the theory and practice of performance measurement (Wholey, 1983; Wholey & Hatry, 1992; Wholey, Hatry, & Newcomer, 1994), Newcomer offers several descriptive definitions of the term: “an institutionalized, routine process whereby the assessment of programmatic results can be undertaken in an ongoing fashion” (Newcomer, 1997a, p. 1); “the label typically given many efforts undertaken within governments and in the nonprofit sector to meet the new demand for documentation of results” (Newcomer, 1997b, p. 5); and “a fairly inclusive term that may refer to the routine measurement of program inputs, outputs, intermediate outcomes, or end outcomes” (Newcomer, 1997b, p. 7). In program evaluation terms, performance measurement is essentially a form of internal (vs. external) and continuous (vs. episodic) program monitoring (Rossi, Freeman, & Lipsey, 1999).

Performance measurement, monitoring, and management systems are designed to permit organizations to evaluate the outputs and outcomes of their programs on a regular, ongoing basis, in order to improve program management, effectiveness, and efficiency (Rossi et al., 1999). Outputs are the products, services, or activities that a program delivers to its clients. In a Children’s Aid Society (CAS), such outputs might include the number of children or youths served per year or the annual number of child-protection visits made by agency staff. Outcomes, on the other hand, are the benefits that clients experience during or after their participation in a program, such as improvements in children’s safety, mental health, school performance, or self-care skills. Performance measurement, monitoring, and management have become highly influential within the federal governments of the United States and Canada and have begun to appear in the nonprofit sector as well.

Plantz, Greenway, and Hendricks (1997) assert that managers in nonprofit agencies who have implemented outcome-oriented performance measurement systems report a number of organizational
benefits. These include a clear focus on intended program outcomes, the establishment of baseline levels of achievement that can guide future targets, a means of monitoring whether corrective action has, in fact, led to program improvement, a way of motivating staff by allowing them to see the progress of their clients in a more visible and objective manner, a mechanism for promoting one’s agency as a successful contributor to the community, and a way of increasing financial resources and volunteer support.

**Performance Measurement in the Public and Nonprofit Sectors, Including Child Welfare**

In the United States, the Government Performance and Results Act (GPRA) of 1993 gave considerable impetus to what has now become an international movement by requiring that all federal agencies implement operational performance measurement systems by the year 2000 as a condition of requesting and receiving continued funding. The basic assumption underlying the GPRA legislation was that performance measurement would enable government managers to “improve the confidence of the American people in the capability of the federal government by systematically holding federal agencies accountable for achieving program results” and that it would “improve federal program effectiveness and public accountability by promoting a new focus on results, service quality and customer satisfaction” (cited in Newcomer, 1997b, pp. 8-9).

In Canada, performance measurement has also become influential. In the ongoing federal-provincial discussion and debate concerning the financing of health services, for example, the federal government would like to tie future increases in funding to the provinces to the establishment of provincial performance measurement systems that would allow inter-provincial comparisons of the effectiveness, efficiency, and quality of health services. The interest of the federal government in such approaches reflects its recent commitment to performance measurement within its own departments. Treasury Board, the management board since 1997 for the entire government of Canada, now holds public service managers responsible for a focus on service to citizens, results, and value for money in the design, delivery, evaluation, and reporting of all federal programs (Treasury Board Secretariat of Canada, 2000a). Treasury Board is also disseminating an operational guide to the elaboration of evaluation and accountability frameworks that are intended to help managers improve their assessment of and reporting on program performance (Treasury Board Secretariat of Canada, 2000b).

In child welfare in Canada, the timeliness and relevance of performance measurement have been heightened by recent policy reforms that several provincial governments have either initiated or are planning. In Ontario, the province that we know best, the Ministry of Community and Social Services (MCSS) launched a new Child Welfare Reform Agenda in 1998, partly as a response to reports from six coroner’s inquests and the Ontario Child Mortality Task Force (1998) into the deaths of children served by the child welfare system. Phase I of this reform was accomplished in 1998 by the introduction of a risk-assessment system into the 52 local CASs in the province. Phase II (2000-2003) will see the implementation in Ontario CASs of a revised and “Canadianized” version (Flynn & Ghazal, 2001) of the Assessment and Action Record (AAR) from *Looking After Children* (Ward, 1995, 1996).

The use of the AAR across Ontario is intended to encourage the emergence of a consistent approach to planning, monitoring, and managing services for children in care, based on the *Looking After Children* philosophy of good outcomes through effective parenting. The AAR is well suited to performance measurement in that it assesses important outputs (services received) and outcomes (benefits realized) in seven major developmental domains: health, education, identity, family and social relationships, social presentation, emotional and behavioural development, and self-care skills. Use of the
new AAR will provide provincial bodies such as MCSS and the Ontario Association of Children’s Aid Societies (OACAS) with a comprehensive portrait of children’s placement histories and developmental status, one that can be updated each year. It will also allow local CASs to participate in a common information system and “benchmark” their own performance against that of CASs elsewhere of comparable size and type.

The newly revised AAR (Flynn & Ghazal, 2001) incorporates standardized measures from the National Longitudinal Survey of Children and Youth (NLSCY), a 20-year longitudinal survey at 2-year intervals of a nationally representative sample of more than 22,000 Canadian children and youths (Statistics Canada & Human Resources Development Canada, 1995). Begun in 1994-95, the NLSCY has quickly become the standard conceptual and empirical framework within which to view child and adolescent development in Canada. Because it incorporates many NLSCY measures, the new AAR will allow, for the first time, systematic comparisons between the developmental trajectories of children and youths in out-of-home care and those of their gender and age peers in the general Canadian population. This strategy, which offers the major advantage of “situating” information on children and youths in care within the broad context of child and adolescent development in Canada, has previously been used on a more limited basis in our own child welfare research (Flynn & Biro, 1998) and in that of Kufeldt, Simard, and Vachon (2000). Researchers outside of child welfare, such as Willms and Sloat (1998) in the Partners for Youth project in New Brunswick, and Peters, Arnold, Petrunka, Angus, Brophy, Burke, et al. (2000), in the Better Beginnings, Better Futures project in Ontario, have also made good use of this comparative strategy for understanding the development of at-risk children and youths.

Establishing a Performance Measurement System

Two steps are usually followed in setting up a performance measurement system for a particular program. First, a logic model is conceptualized to identify the key elements of the program and describe its underlying theory. Then, the logic model is translated into an operational performance measurement system.

Program logic models. Logic models describe how programs are thought to work to solve specific problems. For McLaughlin and Jordan (1999), they are also tools for telling the “performance story” of a program. The approach to logic models used at the Centre for Research on Community Services at the University of Ottawa and at Services to Children and Adults of Prescott-Russell (hereafter referred to simply as “Prescott-Russell”) is based on the United Way model described by Hatry, van Houten, Plantz, and Greenway (1996). Our logic models (such as the one developed for a new educational tutoring program for Prescott-Russell children in long-term care who are experiencing substantial educational difficulties) specify, in order, the needs (problems) of the children/youths that the program intends to address, the inputs (resources) thought to be necessary to implement the program successfully, the activities (interventions) planned to meet the children or youths’ needs, the outputs (services and “units of service”) that the program is expected to provide, and, most important of all, the intended shorter-term and longer-term outcomes (client benefits) that the program aims to achieve through its inputs, activities, and outputs.

Translation into a performance measurement system. Using the program logic model as a guide, a team translates each conceptual output and outcome into an operational, quantitative performance indicator. The team also specifies how the data will be collected (i.e., by whom, from whom, and at what frequency), how each indicator will be monitored, and how corrective action will be
taken to manage the problem as well as possible. In such endeavours, practical guides such as the United Way manual on outcome measurement (Hatry et al., 1996) provide helpful advice on how to proceed.

The Performance Measurement System at Services to Children and Adults of Prescott-Russell

Development of the System

Several contextual factors and events led to the development, beginning in 1995, of the performance measurement approach currently used at Prescott-Russell:

- A longstanding interest on the part of a former agency Board President and of the current Executive Director in outcome-oriented program evaluation
- A seminar given by Roy Parker in Canada in 1992 on the development of Looking After Children in the UK (Parker, Ward, Jackson, Aldgate, & Wedge, 1991)
- Piloting of the AAR and NLSCY in 1994 by a Prescott-Russell supervisor
- A research and development grant from MCSS (1995-1998) that allowed Prescott-Russell and the University of Ottawa to adapt and translate into French the Assessment and Action Record (AAR) from Looking After Children, which, along with the NLSCY questionnaires, was to form the heart of a new internal evaluation system at the agency
- Collaboration on the final translation and adaptation of the AAR with colleagues from Memorial and Laval universities, who, with funding from Human Resources Development Canada, were beginning to launch, in 1996, a national, six-province Looking After Children project (Kufeldt, Simard, & Vachon, 2000)
- Annual presentations to Prescott-Russell managers and staff of findings from the AAR and NLSCY, which led to efforts to improve services

Current Functioning of the System

Overall program quality. PASSING (Wolfensberger & Thomas, 1983) is an external program evaluation tool that has been used internationally to evaluate the quality of community services (for a review, see Flynn, 1999). The instrument is based on the principle of Social Role Valorization (Lemay, 1996; Thomas & Wolfensberger, 1999). Early in our project, in 1996, PASSING teams assessed the quality of the residential and family services programs at Prescott-Russell. We hope to be able to fund a second PASSING evaluation within the next couple of years, to find out whether, after an extended period of implementation of Looking After Children, overall program quality has indeed improved (in the baseline assessment in 1996, it was found to be within the “fair to acceptable” range).

Twice-monthly monitoring of service outputs and outcomes: Quality Control Committee. In 1997, following a lengthy strike in 1995 and a subsequent operational review of the agency by MCSS, the Prescott-Russell Board of Directors asked the Executive Director to conduct a thorough management review of the agency and come up with new ways of operating. Besides visiting 6 CASs in Ontario, some of which were known for their innovative services, the Executive Director also undertook an extensive literature review. He found Martin’s (1993) book on quality management in human service organizations to be particularly relevant and asked agency managers to study it. In addition, during this same period a person who had been extensively involved in the development of an innovative quality control and performance measurement system at Canada Post became the President of the Prescott-Russell Board of Directors.
Directors. The Executive Director observed at firsthand the Canada Post quality control system, which served as the model for a new Quality Control Committee (QCC) at Prescott-Russell.

Initiated in 1997, the QCC continues to evolve. It currently meets twice a month for an hour or more, as a management subcommittee. The first meeting of the month is devoted to a review of recent quantitative performance indicators that assess mainly outputs (e.g., the volume of cases served) but also, to some extent, outcomes (e.g., graduation from high school). The second meeting of the month is focused on a largely qualitative examination of the performance issues facing a particular service division or program within the agency. Future plans call for the development, over the next several years, of logic models for all agency programs, the translation of these logic models into operational performance measurement terms, an agency-wide focus on client outcomes, and, most important of all, an emphasis on the clinical and administrative corrective actions that will produce ongoing improvements in these outcomes.

During QCC meetings, the electronic presentation of performance data allows trends to be tracked over time. Figures 1 and 2 illustrate the kind of information (here, on outputs) that is presented to the management team. Figure 1 shows that the number of Crown Wards at Prescott-Russell has increased over the last several years. As of the end of October 2000, the number of Crown Wards had reached 202, compared with 174 in 1999, 157 in 1998, and 143 in 1997. Similarly, as of the end of the fiscal year (March 31), the number had reached 184 in 2000, compared with 177 in 1999 and 134 in 1998. This growth reflects, in part, the adoption of risk assessment procedures in Ontario CASs in 1998 and changes in Ontario’s child welfare legislation in 2000. It has placed added pressure on direct-service staff, in the form of increased case loads, and on the staff responsible for recruiting foster families, whose numbers have not kept pace with the number of children in long-term care.

Figure 2 shows the number of cases opened or reopened each month in the family services program over the last several years. Output data such as these allow the prediction of when peak periods of service demand are likely to occur each year (e.g., in September, with the start of the school year) and to foresee when time is most likely to be available for staff training or the updating of files (e.g., in January or February).

Although the QCC is a management subcommittee, its meetings are open to all staff, foster families, board members, or other persons associated with Prescott-Russell. The electronic display of performance indicator data usually confirms—but sometimes disconfirms—senior and middle managers’ impressions about what has recently happened. It allows supervisors to focus quickly on service problems, such as placement breakdowns, hospitalizations, out-of-jurisdiction placements, child abuse allegations in open child protection files, serious occurrences, or case re-openings within 90 days. Supervisors report on the data and problems presented and use the occasion to highlight some of the difficulties in providing good, timely services.


Figure 1. The number of Crown Wards in care at the end of each month at Services to Children and Adults of Prescott-Russell for the years 1997-98, 1998-99, 1999-2000, and, for 2000-2001, to October only.

An illustrative clinical case. The potential of the QCC process for improving services can be seen in the following clinical example. The case involved a young girl who had been abused, despite the fact that a child protection file had recently been opened on her at what was then the Children’s Aid Society (CAS) of Prescott-Russell. (After this case had occurred, an amalgamation took place, involving the CAS, a children’s mental health centre, and an organization serving persons with developmental disabilities, that created the present agency, Services to Children and Adults of Prescott-Russell.) The CAS supervisor responsible for the girl’s case reviewed the file and found that her parents had referred themselves to the CAS three times because of conflicts with their daughter. Each time, the family had indeed required services but, because these would be voluntary in nature, the CAS had referred the family to the local children’s mental health centre. It turned out, however, that over a three-year period the family had often been assessed but had never received any services. The pattern became clear: when things got out of hand, the family would call the CAS which, in turn, would refer the family for mental
Figure 2. The number of cases opened or reopened each month in the family services program at Services to Children and Adults of Prescott-Russell for the years 1997-1998, 1998-1999, 1999-2000, and, for 2000-2001, to October only.

health services. Unfortunately, the CAS would fail to monitor whether or not the family actually received services. After this had happened several times, the father eventually hit his daughter. The discussion of the case at the QCC meeting led the management team to make two immediate decisions: first, the family would receive direct assistance from the CAS, and second, the intake team would monitor the way in which future referrals to the mental health centre were handled.

Placement breakdowns in foster care also generate a lot of discussion at the QCC meetings. Invariably, staff ask themselves whether the child in question has experienced previous placement breakdowns and how further breakdowns can be prevented. A decision to conduct follow-up monitoring is usually made in such cases.

A number of important organizational benefits flow from the QCC meetings. First, senior management is kept well informed about service trends and about the experiences of staff in providing
services to clients. Problem areas are identified and discussions held about agency policies. Second, the degree of public accountability of middle managers (i.e., supervisors) is increased by virtue of their having to report regularly on service problems and successes in their areas of responsibility. Third, the meetings are open to all, such that staff are engaged in an ongoing dialogue about what is most important in a CAS, namely, services to clients (outputs) and client benefits (outcomes). Fourth, direct-service staff have occasions to discuss issues with middle and senior management, thereby exercising an important influence on policy questions and making clear how complex the delivery of services can sometimes be. Finally, frontline staff feel supported by the fact that data on outputs (e.g., volume statistics), which indicate how busy they are, are presented first, before data on outcomes. This practice is consistent with the emphasis placed by the balanced scorecard (Ulrich, Zenger, & Smallwood, 1999) on monitoring data that are important to all stakeholders, not just to those at the top of the organization.

Annual monitoring of service outputs and outcomes: Reports on aggregated AAR and NLSCY data. At Prescott-Russell, as in other agencies using the Looking After Children approach, the AAR is administered once a year to each child or youth in long-term care by his or her child protection worker. The NLSCY questionnaire was also administered once a year to the same children and youths on an experimental basis during 1997-1999. At the end of each calendar year, the person in the agency responsible for research and evaluation aggregates this individual-level AAR and NLSCY data and produces a report on the outcomes observed in the entire group of children and youths served in long-term care. The purpose of this annual report is to inform managerial and direct-service staff about the children and youths’ current developmental status and alert them to areas in which outcomes need improvement. These annual reports are thus complementary to the QCC meetings: they focus mainly on shorter or longer-term client outcomes (with some attention to outputs), whereas the QCC attends principally to immediate service outputs (with some emphasis on outcomes).

In early 1998, the Prescott-Russell management team got its first look at the AAR and NLSCY data on the children and youth in long-term care on whom assessments had been conducted during 1997. The team met for two days, reviewing more than 100 performance indicators. These data provided the first detailed portrait that the agency had ever had of the life conditions, experiences, and benefits that the agency was affording the children and youths in its care. Based on this review, the management team attempted to improve several problem areas, including vaccination records, school suspensions, academic results, children and youths’ engagement in extracurricular activities, and placement stability.

To illustrate the kind of year-by-year monitoring of outcomes that the annual AAR and NLSCY assessments make possible, Table 1 presents data on children and youths’ participation in healthy extracurricular or spare-time activities, such as playing team sports, learning a musical instrument, or being a member of a social or youth club. Taking part in such activities enhances youths’ physical competencies, self-esteem, and interpersonal skills, and reduces the risk of poor mental health or delinquent behaviour (Jackson & Kilroe, 1995). Based on the year 1 data, the management team felt that foster parents and protection workers should encourage participation in such activities and therefore instituted a policy requiring that each child or youth participate in at least one healthy activity per year. After monitoring the data on these indicators for year 2, the management team further required that the objective of participation in at least one such activity per year be written into each child or youth’s plan of care and that foster parents be reimbursed for any costs incurred in such activities.

As can be seen from Table 1, the new policy met with some, albeit mixed, success. On the one hand, the development of skill or interest in a wide range of healthy activities increased over the 3-year
period, from 68% to 85%. On the other hand, participation in team sports declined somewhat over the three years, although it should be kept in mind that many of the children and youths in the sample were making the transition from early to middle adolescence, a period during which the abandonment of team sports is common. Finally, the percentage of children or youths who played a musical instrument or were members of a social or youth club was relatively stable.

Table 1

Results on AAR Indicators of Participation or Skill/Interest in Extracurricular Activities, for Prescott-Russell Children and Youths in Long-Term Care on Whom Data Were Available in 1997, 1998, and 1999

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Participation in team sports</td>
<td>60%</td>
<td>50%</td>
<td>47%</td>
</tr>
<tr>
<td>Playing a musical instrument</td>
<td>30%</td>
<td>27%</td>
<td>30%</td>
</tr>
<tr>
<td>Member of social or youth club</td>
<td>30%</td>
<td>37%</td>
<td>27%</td>
</tr>
<tr>
<td>Has skill/interest in extracurricular activities (e.g., art, theatre, cooking, photography, sewing, dance)</td>
<td>68%</td>
<td>80%</td>
<td>85%</td>
</tr>
</tbody>
</table>

*Note. In the case of the first three indicators, the same 30 children/youths were present in all 3 years. For the fourth indicator, the same 34 children/youths were present each year.*

Table 2 presents data on indicators of positive identity or of support thereof. These three indicators showed mild fluctuations across the 3-year period, within a 7-10% range, with two increasing and one decreasing. Most children or youths reported that their foster parents often praised them and showed interest in their activities. Only about half, however, were rated by their protection workers as usually having a good self-image and a positive perception of their own abilities. The management team incorporated these findings into its foster parent training activities.

Future Development of the Performance Measurement System at Prescott-Russell

PM3 is an evolving system that continues to undergo refinement and improvement. It is noteworthy that two and half full-time positions are currently dedicated to the performance measurement, monitoring, and management function in the expanded Prescott-Russell agency. As a result, logic models in all service fields and for all programs will have been completed in the near future. Translation of their outputs and outcomes into performance indicators is planned over the next few years. In addition, the respective contributions to the process of improving services of the QCC meetings (focused especially on outputs) and the annual reports on aggregated AAR and NLSCY data (oriented particularly to shorter and longer-term outcomes) will be sharpened and enhanced.
Table 2

Results on AAR Indicators of Positive Identity (or Support Thereof), for Prescott-Russell Children and Youths in Long-Term Care on Whom Data Were Available in 1997, 1998, and 1999

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Child/youth reports that he/she often receives praise from foster parents</td>
<td>89%</td>
<td>78%</td>
<td>81%</td>
</tr>
<tr>
<td>Child/youth reports that foster parents often seem interested in his/her activities</td>
<td>78%</td>
<td>78%</td>
<td>85%</td>
</tr>
<tr>
<td>CAS protection worker rates child/youth as usually having a good self-image and a positive perception of his/her own abilities</td>
<td>44%</td>
<td>49%</td>
<td>54%</td>
</tr>
</tbody>
</table>

Note. In the case of the first two indicators, the same 27 children/youths were present in all 3 years. For the third indicator, the same 39 children/youths were present each year.

In these efforts to improve PM3, the new AAR (Flynn & Ghazal, 2001) promises, for several reasons, to be especially useful to staff and other stakeholders responsible for children and youths in long-term care. The computer-scannable nature of the revised AAR and its close alignment with the individualized plan of care will permit more frequent and relevant feedback to managers, direct-care and supervisory staff, foster parents, children and youths, and board members. Also, as was mentioned earlier, the incorporation of many NLSCY measures into the AAR will encourage comparisons between the developmental outcomes experienced by Prescott-Russell and other CAS children and youths in care and those of their age and gender peers in the general Canadian population. In addition, the new AAR has a section for the recording of the type and amount of each service received by a child or youth in care during the preceding 12 months. This will allow, for the first time, the cost of individual service “packages” to be calculated on a regular basis. Finally, the revised instrument will enable the construction of matched samples composed of children and youths in care and those with similar characteristics in the Canadian population. Matched samples will allow evaluators to conduct special quasi-experimental studies that go beyond outcome monitoring and permit at least the tentative attribution of causal effects to the program in question. Such evaluation designs will provide a stronger logical warrant for inferring that it is indeed our child welfare interventions, and not merely extraneous factors, that are responsible for progress made by children and youths in care.
References


