The Toronto International Leadership Centre for Financial Sector Supervision held its inaugural program in Toronto, Canada, from July 19 - 24, 1998. The Toronto Centre was recently created by the World Bank and the Canadian Government, in cooperation with York University's Schulich School of Business, the International Monetary Fund and the Bank for International Settlements. Its objective is to strengthen financial markets by enhancing the leadership capabilities of public sector executives. The target group for the Toronto Centre's programs is comprised of the most senior executives whose primary responsibility is banking supervision. Individuals from 14 countries participated in the inaugural program session. Of these, 3 (21.4%) were women. This inaugural program offering was considered a "pilot" effort from which valuable lessons could be learned for future program offerings. A number of pedagogical methods were used in this program.

The EDI Evaluation Unit (EDIES) designed and supported the evaluation. The evaluation consisted of several components. A rapid response form was used to assess individual Program Leaders on the content of their work sessions and on their presentation ability. This short form contained 10 questions that used a 5-point Likert type scale along with general, open-ended questions. An evaluation form designed to assess general satisfaction with the program and to assess learning was also used in the evaluation. This form was administered at the end of the course and also used a 5-point Likert type scale. One part of this final evaluation form asked questions about the overall performance of the program, including accommodations and other means of support. The second part consisted of questions designed to assess the perceived level of information gain by participants. This method involves asking participants to judge their level of knowledge or information on a set of course related topics before and after the course. These measures are taken at the end of the course in order to provide a common baseline among participants for making these judgments. Comparing the before and after results provides a valid indication of perceived growth in knowledge/information among participants. Data from the evaluation were collected and processed by EDIRP staff. The analysis of these data were conducted by the EDI Evaluation Unit. Following is a summary of the evaluation results.

- The rapid response forms showed some variation among Program Leaders on both their presentations and the content of their cases. While this was worth noting to the individual leaders, the amount of variation was not great. Most Program Leaders scored in the high range on both presentation and content.

- Combining the results of the rapid response forms shows little variation in the scores for the Program Leaders' presentations and Case content. All items received a score of over 4.0 on a 5-point scale. Within the domain of presentation, the item which received the highest score (4.61) was the view that the course, "imparted useful..."
supervisory advice and skills." The item receiving the lowest score in this area was the view that the course, "treated issues in sufficient depth." The score for this item was 4.34, which still is relatively high on a 5-point scale.

- For content of the Case presentations, 2 items received the lowest score. These were the degree to which the course, "covered topics relevant to your country needs," (4.22) and whether the course, "provided enough material to serve as a useful basis for discussion but not too much to overwhelm" (4.23). The highest score, 4.52, was for, "use of English which was easy to understand." It should be noted that these scores were on a 5-point scale and none were low enough to indicate a problem.

- The end of program questionnaire consisted of 21 questions that asked about various features of the program, including issues related to the substance of the training and the logistical operations for the event. The first 12 questions used a 5-point scale that measured "intensity" on an issue and ranged from 1=minimum to 5=maximum. On these questions, the average range of scores for all respondents was 4.23 - 4.69 out of 5.0. Overall, these were very high, or positive, scores. The item with the lowest score, 4.23, was the usefulness of the Risk Factor/Core Principles Questionnaire.

- The remaining 9 questions on the end of program questionnaire used a 5-point scale that ranged from 1=too little to 5=too much. The mid-point, 3=just right, was considered the ideal score. The number and percentage of score above and below 3.0 were also reported as indicators of direction away from the ideal score. All of the scores were 3.0 or higher (3.0 to 3.62), indicating a general inclination for respondents to find the course item "too much" rather than "too little." This is confirmed by observing the percentages above and below 3.0. On only one item, the amount of reading sent in advance, did the percentage responding that it was "too little" exceed the percentage responding "too much." For all other items, there was the general inclination that there was "too much" offered.

- Results from the pre/post self-assessment showed a substantial gain in information or skills in all areas. The item which showed the highest gain was an increase in the ability to access an international peer network as a resource. This item showed an increase from pre to post of over 71% (71.15%). Since one of the major objectives of the program was to promote this type of networking, this is an important finding. The area which showed the least change was in the ability to, "critically analyze your country environment." This item increased nearly 28% (27.78%), the smallest increase of all items. It appears that respondents may have felt comfortable with this information prior to entering the course.

- Some valuable information was obtained from the open-ended questions. There were 3 questions which asked for 1) the most useful part of the program, 2) the least useful and 3) other advice/suggestions. The most useful items noted were the analytical tools/conceptual approach, interface with program leaders and the Case Studies. This was a mixture of pedagogical methods and content. Found least useful were the amount of time devoted to the Group/Team work and the repetitiveness of some information. A number of respondents felt that time could be saved by more efficient handling of the group presentations. Of the 14 respondents to the question of whether they would recommend this course to others, all responded in the affirmative.

- The evaluation, as well as the program, were pilots and intended to provide information that could be used to make changes for later sessions. Based upon this information, there are several limitations to the evaluation that should be considered in future efforts. One major objective is the use of self-assessment as an indicator of knowledge gain. This should be supplemented later with more reliable measures,
such as those obtained from using cognitively based questions. More information could also be gained by linking the questionnaire responses. An anonymous method could be employed to link the end of program questionnaire with the rapid response questionnaires. This would provide more extensive data on course topics. Finally, some effort should begin on tracking respondents so that information on course impact can be collected.