

**Impact Evaluation of
Staff Learning in East
Asia & Pacific and
South Asia Regions**

Phase II Report

Chaoying Liu

Yermal Shetty

Victoria Gunnarsson

WBI Evaluation Studies
No. EG05-94

The World Bank Institute
The World Bank
Washington, D.C.

August, 2004

ACKNOWLEDGEMENTS

This evaluation report was prepared for the Learning Board under the overall guidance of Marlaine Lockheed, Manager, Evaluation Group. This report was prepared by Chaoying Liu, Yermal Shetty, and Victoria Gunnarsson.

A number of people made various contributions to this evaluation study and the report, including: Narendra Sharma, Jennifer Ross, and David Colbert from the Asia Learning Group. Many thanks also go to Elina Manjieva who worked diligently to collect and analyze the data and help with the statistical work, and to Humberto Diaz for the excellent graphics preparation and document production. This report also benefited greatly from expert review comments provided by Elaine Patterson, Regional Learning Coordinator (LAC) and Sukai Prom-Jackson (HDN). Finally, the authors wish to extend thanks to all participants, especially the Operations and regional managers who provided for and shared with us valuable information and experience for the purpose of this study.

The report was discussed at a meeting with the Asia Learning Group chaired by Phyllis Pomerantz on July 26, 2004.

WBI Evaluation Studies are produced by the Institute Evaluation Group (IEG) to report evaluation results for staff, client, and joint learning events. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and do not necessarily represent the view of the World Bank Group. WBI Evaluation Studies are available on line at:

<http://info.worldbank.org/etools/WBIEG/publications/index.cfm?pg=getPubs&category=Publications&Intro=yes&instructions=no&showDetails=no&ID=>

Vice President, World Bank Institute	Ms. Frannie Léautier
Chief Learning Officer	Ms. Phyllis Pomerantz
Manager, Institute Evaluation Group	Ms. Marlaine E. Lockheed
Task Team Leader	Ms. Chaoying Liu

ACRONYMS

ACS	Administrative and Client Support
AO	Advanced Operations
CAS	Country Assistance Strategy
CD	Communication and Dialogue
CPPR	Country Portfolio Performance Reviews
EAP	East Asia & Pacific Region
ESW	Economic and Sector Work
FY	Fiscal Year
GE+	Staff in grade E and above
GF+	Staff in grade F and above
GG	Staff in grade G
GH	Staff in grade H
HQ	Headquarters
HR	Human Resources
IEG	Institute Evaluation Group
ILP	Individual Learning Plan
OPCS	Operations Policy and Country Services
OPE	Overall Performance Evaluation
PRSP	Poverty Reduction Strategy Paper
PSR	Project Supervision Reports
QAG	Quality Assurance Group
SAR	South Asia Region
SLC	Strategic Learning Center
TD	Team Dynamics
TL	Team Leadership
TTL	Task Team Leader
VPU	Vice-Presidential Unit
WB	World Bank

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
ACRONYMS	iii
TABLE OF CONTENTS	v
EXECUTIVE SUMMARY	vii
1. INTRODUCTION AND BACKGROUND.....	1
Background	1
Evaluation Methodology	3
2. IMPACT OF THE REGIONAL CORE COURSES	5
Methods	5
Evaluation Participation.....	6
Overall Results.....	8
Evaluation Results of the Advanced Operations Course	12
Evaluation Results of the Foundations of Team Leadership Course.....	16
Summary and Discussion	19
3. STAFF LEARNING AND PERFORMANCE.....	21
Measurement Issues	21
Hypothesis Examined and Methods	22
Results.....	23
4. STAFF LEARNING AND UNIT PERFORMANCE.....	25
Methods	25
Unit Performance: A Framework and Indicators.....	26
Review Results of Staff Learning for the Selected Units	28
Analysis Results of Performance for the Selected Units	32
Summary.....	35
5. CONCLUSIONS AND IMPLICATIONS	37
ANNEXES	40
Annex 1A: Evaluation Questionnaires – Advanced Operations Course	42
Annex 1B: Evaluation Questionnaires – Foundations of Team Leadership	47

Annex 2: Statistical Analysis Results – Advanced Operations	52
Annex 3: Statistical Analysis Results – Foundations for Team Leadership.....	53
Annex 4: Analysis of the Learning Impact on OPE.....	55
Annex 5: Manager Interview Procedures, Questions and Results.....	65
Annex 6: Unit Performance Framework and Indicators	67
Annex 7: Staff Learning Information Highlights	68
Annex 8: Learning and Development: 2003 Staff Survey Results	69
Annex 9A: Delivery of Products, FY02-03*	70
Annex 9B: Riskiness of Projects And Key Performance Ratios for FY02 and FY03.....	71
Annex 10: Unit Learning and Operational Performance Indicators	77

EXECUTIVE SUMMARY

This report presents the results from the second phase of the impact evaluation of staff learning programs in the East Asia & Pacific (EAP) and South Asia (SAR) regions. Building upon results from the phase I evaluation, the main objective of the phase II evaluation is to estimate the extent to which regional staff learning has impacted regional staff performance and regional Operations. Specifically, the evaluation was designed to address the following three-layer impact questions:

- Did the regional staff-learning program help staff build the knowledge and skills that they need?
- Did staff learning enhance individual staff performance?
- Did an investment in learning/training bring better performance to organizational units?

The evaluation approached these three questions with separate and distinct methodologies:

First, using two core courses—Advanced Operations and Foundations of Team Leadership—as examples of regionally designed learning programs in FY03, the evaluation applied a quasi-experimental design (waiting list control), to compare perception, knowledge/skills, and utilization levels of past course participants with prospective course participants, thus determining the extent to which these courses successfully helped staff acquire the knowledge and skills needed for their work.

Second, using longitudinal Human Resources (HR) data, the evaluation explored the relationship between staff learning participation and staff performance, as measured by Overall Performance Evaluation (OPE) assessments.

Third, the study employed an experimental analysis to explore the possible learning effect on organizational outcomes. An analysis of unit staff learning indicators and unit performance was carried out for eight selected regional units (six operations units and two support units). Due to the nature of available information gathered, the analysis carried out is largely in descriptive nature.

Three key findings pertaining to the evaluation questions are derived from these evaluation activities:

- Both regional courses (Advanced Operations and Foundations of Team Leadership) successfully achieved their objectives. Compared to non-participants (controls), participants gained an increase of 30 percent in their knowledge related to operations and leadership as a result of attending the learning courses. Apart from the knowledge gain, there is evidence that the courses helped participants work more efficiently. Participants also increased

their utilization of some of the content of the Team Leadership course. These changes are clear indications of courses impact on individuals.

- There is only weak evidence, however, of a relationship between staff learning participation and staff performance as measured by OPE scores. This may be due in part to the OPE assessment, which is a relatively poor proxy measure for staff performance.
- The study employed three staff learning indicators¹ and eight operational indicators² to investigate the relationship between staff learning and unit performance. Overall, evidence that learning investment yields positive returns at the organizational level is weak. However, units that have a good record on learning tend to have comparatively high Proactivity and Realism Ratios. As learning may enhance staff's capacity to find solutions and move projects away from the problem project category, learning may also contribute to the unit's business performance as measured by a high Proactivity Ratio and improved Realism Ratio.

Thus, the study findings support the general notion that staff learning is important to help staff build knowledge and skills. However, the lack of positive relationship between staff learning and organizational performance does not support the view that learning is the sole and primary source of change in individuals and organizations, at least in the short run.

The evaluation results have some implications for regional staff learning and its evaluation.

First, as revealed in the phase I evaluation, Headquarter (HQ) staff received more learning than staff stationed in country offices. With the trend towards decentralization of the Bank's Operations through country offices, it is increasingly important to ensure that adequate learning opportunities are available to all country office staff. The Asia Learning Group has put a great deal of effort into ensuring that front line staff members have the knowledge and skills needed to carry out their work. In addition to the Advanced Operations and Team Leadership courses, the regional staff learning program has designed and promoted two other core courses: Development and Ethics and Client Engagement and Communications. What are the likely impacts of these core courses and regional programs? The results of the evaluation of the two core courses indicate great promises for the success of the regional staff learning programs. If a substantial number of regional professional staff and Task Team Leaders (TTLs) were to attend these core-learning courses, and if all of them achieved a 30 percent increase in their knowledge and skills with the ability to use this knowledge more effectively, we would expect to see positive changes in the Regional Operations.

¹ Share of staff in unit having ILPs, number of learning days per staff, and learning budget per staff.

² Projects at risk ratio, commitment at risk ratio, disbursement ratio, realism ratio, proactivity ratio, percentage of projects facing financial management problems, percentage of projects facing financial performance problems, and percentage of projects facing procurement problems.

Second, the evaluation highlights the influence on staff learning of the organizational culture. At the unit level, in addition to developing a unit development plan, efforts need to be made to enable all staff members to have Individual Learning Plans (ILP) in order to encourage and monitor learning. In this regard, the OPE process can be an effective instrument to ensure that staff have ILPs and participate in the core learning required to achieve business results. There seems also to be a need for changing the incentive structure for career development and promotion so that it gives reasonable weight to adequate learning. This approach could provide an incentive for staff to take advantage of learning events and keep their learning records current.

Third, not only unit staff, but also managers need to be equipped with new and updated operations knowledge and skills. Strong leadership is the driving force behind successful unit performance, and it also encourages staff to acquire and utilize new knowledge and skills.

Fourth, the evaluation findings are significantly affected by the comprehensiveness and quality of the data collected. In particular, the time frame, accuracy of available data and validity of outcome measures for these analyses are not ideal. These are common problems that evaluators face in trying to determine the impact of learning. It is recommended that a thorough exploration of the dynamics of outcome measures and exogenous factors be part of future evaluation efforts.

Fifth, it should be acknowledged that the impact of learning takes a long time to become visible, and that the evaluation of learning impact should be a continuous effort. In the context of the current study, evaluating the impact of learning on the third dimension — client relationships — should be a critical item on the future evaluation agenda.

Finally and most importantly, this evaluation confirms that staff learning is not determined simply by an outcome indicator of organizational effectiveness, and that learning also contributes to a unit's business productivity. At the organizational level, the impact of learning may be expected only if a critical mass of staff learning occurs. In addition, the benefits of staff learning may require long lead times to become visible. These points reinforce the value of the commitment to invest in staff learning, to focus on the right topics, to reach and maintain a certain intensity level, and to make learning a long-term strategy. With these commitments, organizations should expect that learning will eventually bring positive returns.

1. INTRODUCTION AND BACKGROUND

1.1 This is the second phase of the impact evaluation of the EAP and SAR regional staff learning program. Several activities were undertaken during this phase in order to assess the impact of the programs. Following the introduction and background, including the objectives and scope of the evaluation, the report is organized according to the following five chapters:

1.2 Chapter II presents the results of impact assessment of two regionally designed core courses: Advance Operations and Foundations of Team Leadership. By employing a quasi-experimental design, the evaluation measured and compared levels of knowledge, skills and utilization of the course content between the past course participants and prospective participants.

1.3 Chapter III seeks to examine the extent to which staff learning has had an impact on staff performance. This was done by analyzing the relationship between staff learning participation and staff performance, as measured by Overall Performance Evaluation (OPE) assessments available from the Human Resources (HR) data set.

1.4 Chapter IV examines how learning impact has contributed to the organization's, as opposed to the individual's, performance. The section discusses the conceptual framework and the indicators used to measure learning impact at the organizational level. It presents analysis results for eight selected regional units, using the framework and indicators identified.

1.5 Chapter V discusses evaluation conclusions and their implications for future regional staff learning and its evaluation.

BACKGROUND

1.6 At the request of the Learning Board, the World Bank Institute Evaluation Group (IEG) initiated an impact evaluation of staff learning programs in the EAP and SAR regions in FY03. Three evaluation activities – a desktop review, a staff survey and manager interviews – were carried out during the first phase of the evaluation³. The phase I evaluation shows that:

- staff learning participation in the EAP and SAR regions during FY99 and FY02 was lower than the WB's standard of 10 days per year, with fewer than a third of regional staff actively involved in regular learning events;
- though both Bank- and Region-sponsored learning events were well-received, Region-sponsored events had the most favorable ratings;

³ Liu and Hanson (2003). Impact Evaluation of Staff Learning in East Asia & Pacific and South Asia Regions – Phase I Report, WBI Evaluation Studies, No. EG04-80, The World Bank Institute.

- as reported by participants, the vast majority of learners utilized the knowledge and skills acquired from the learning events; and
- over half of the learning participants agreed that learning improved their performance and their service to clients.

1.7 During the study period (FY03/04), some important changes occurred in regional staff learning. These include: (a) strengthening the leadership of the Asian staff learning by appointing the first Regional Learning Advisor to oversee the learning program and the formulation of a joint EAP and SAR Learning Group; and (b) establishing a central “Hub” in Bangkok for the delivery of learning courses.

Box 1: Leadership of Asia staff learning

Regional Learning Advisor	Asia Learning Group
<ul style="list-style-type: none"> © A well-seasoned senior Bank staff with more than 30 years experience in the Bank 	<ul style="list-style-type: none"> © Has a small number of staff (3) but dedicated 100% to learning for the two regions
<ul style="list-style-type: none"> © Vast operational experience in multiple regions 	<ul style="list-style-type: none"> © Offers staff one place to go for their learning needs/questions (provides opportunity for greater staff involvement)
<ul style="list-style-type: none"> © A 100% learning position (not a side job). 	<ul style="list-style-type: none"> © Helps unit build learning infrastructure within regions (unit-level)
<ul style="list-style-type: none"> © Accountable directly to both Regional VPs and thereby directly linked to the overall regional strategies and staff issues 	<ul style="list-style-type: none"> © Identifies and institutes core learning programs for staff in both Grades A-D and GE+
<ul style="list-style-type: none"> © Delegated responsibility by Regional VPs to plan, develop and implement VPU Learning Plans 	<ul style="list-style-type: none"> © Customizes learning to meet corporate priorities, regional business needs and staff learning needs
<ul style="list-style-type: none"> © Communication with the Regional HR Manager 	
<ul style="list-style-type: none"> © Greater accountability of staff learning funds 	
<ul style="list-style-type: none"> © Responsibility for Regional initiatives' learning budgets 	

1.8 With these initiatives, regional staff learning has shown significant improvement. Since late FY02, significant progress has been made in regional staff learning. **Strategically**, the Asia Learning Group has assisted regional units in developing Unit Learning Plans that reflect the units’ business priorities and their staff’s learning needs and development plans. The Asia Learning Group has also helped build a learning infrastructure at the unit level that includes unit learning committees and coordination, so that staff are now more involved and are able to contribute to their unit learning plans. Moreover, the Asia Learning Group has set up incentives for staff learning by funding travel and subsistence; therefore, staff need only to request that the manager/unit pays for staff time.

Operationally, the Asia Learning Group has designed, customized and delivered learning courses that meet regional business and staff learning needs and corporate priorities in a timely fashion. The number of topics and learning events offered by the Asia Learning Group has grown substantially (Box 2). The Asia Learning Group also partnered with both the Administrative and Client Support (ACS) Network and Operations Policy and Country Services (OPCS) to roll out pilot operational courses to staff across the Bank. In FY04, based on the experience in FY02 and FY03, the Asia Learning Group refined its core learning program by consolidating a number of important regionally offered courses and by adding the Client Engagement and Communications course in response to the corporate priorities announced by the Managing Director Office of the Bank.

Box 2: Asia learning priorities and offerings

	Priority subject areas	Total offerings
FY02	Corporate strategies Managerial skills Team building Operational skills IT skills	6
FY03	Advanced operational skills Leadership and management skills Team learning Development and Ethics	14
FY04	Advanced operational skills Leadership and management skills Development and Ethics Client Engagement and Communication	27

EVALUATION METHODOLOGY

1.9 By building upon the analysis and findings of the phase I evaluation, the phase II evaluation seeks to assess the extent to which staff learning and regional staff learning programs have bettered regional staff performance better contributed to Regional Operations. Specifically, the evaluation was designed to address the following three impact questions:

- Did the regional staff learning program help staff build the knowledge and skills needed?
- Did staff learning enhance staff performance?
- Did an investment in learning/training improve the performance of organizational units?

1.10 The above evaluation questions were addressed by the three activities summarized in Table 1. More details about the methodology employed for each activity are presented later in their respective sections. Some key elements of each activity are outlined below:

Table 1: Evaluation issues, measures and data sources

Evaluation issues addressed	Outcome/impact measures	Data sources
A) Immediate impact on participants	Participants' knowledge level and utilization	Surveys
B) Individual performance	Individual OPE assessment FY03 (Behavior and Result ratings)	HR data
C) Unit performance	Unit staff learning indicators; Unit operational performance indicators - organizational effectiveness, quality of products, and client relationship	Manager interviews/ Desk reviews

Analysis of the impact of the two regionally offered core courses in FY03

1.11 The first evaluation question was addressed by analyzing information collected through a quasi-experimental design study. That study compared key impact indicators between formal course participants and prospective participants in regionally offered core courses, namely, Advanced Operations and Foundations of Team Leadership for staff in EAP and SAR regions.

Analysis of staff learning and staff performance

1.12 The evaluation addressed the second question by examining the relationship between staff learning data and staff OPE assessments.

Analysis of staff investment and unit performance

1.13 The third evaluation question was addressed by a review and analysis of the relationship between unit staff learning experience and unit performance. This is the first time that the evaluation has used “unit performance” as an indicator of staff learning impact in the context of staff learning evaluation.⁴ Eight regional units (six operations units and two support units) were selected for the review. The review and analysis were undertaken through the following four steps:

- (a) Developing a framework and set of generic indicators that measure unit performance;
- (b) Reviewing a) the unit’s expenditure/investment on staff learning/training over the past two years; b) the unit’s staff learning management, including unit learning plans; and c) staff Individual Learning Plans (ILPs) and staff learning participation;
- (c) Analyzing unit performance for the eight selected regional units; and
- (d) Analyzing the possible effect of learning on unit performance.

⁴ In this case, the unit is the work unit, not the Vice Presidential Unit (VPU). We use eight measures of unit performance, which are discussed in details in Chapter IV.

2. IMPACT OF THE REGIONAL CORE COURSES

2.1 The EAP-SAR Advanced Operations and Foundations of Team Leadership are the two most important learning courses for EAP and SAR staff. These two courses, together with two other courses—Development and Ethics and Client Engagement and Communications—constitute the core courses of the regional learning program.

2.2 The EAP-SAR Advanced Operations (AO) course is designed for Task Team Leaders (TTLs) in grades GE+. The course offers a systemic view of Operations and is results-oriented, using a case-study approach supported by video (interviews of TTLs and managers) to engage participants in real-life problem solving and teamwork. It focuses on emerging corporate priorities such as results and implementation, new products and instruments, and new business approaches. The key objective of the course is to improve the efficiency of operations and business practices.

2.3 The Foundations of Team Leadership (TL) course for staff in EAP and SAR regions is designed to complement the EAP-SAR Advanced Operations course. The Asia Learning Group has customized the traditional Foundations of Team Leadership course (given by Leadership and Organizational Effectiveness) by combining Modules I and II and placing a greater emphasis on the operational content of regional business. The key objective of the course is to improve managerial and leadership capabilities. Similar to the EAP-SAR Advanced Operations course, its primary audience is EAP-SAR TTLs of grades GE+.

2.4 Between December 2002 and May 2004, the Asia Learning Group offered eight Advanced Operations and five Foundations of Team Leadership courses to EAP and SAR regional staff. As of May 2004, 165 staff (25 percent of all GE+ staff in both regions) had attended the Advanced Operations course, and approximately 16 percent of regional TTLs (N = 103) had taken part in the Foundations of Team Leadership course.

2.5 The objective of the impact evaluation is to determine the short and intermediate effects of these two courses. Specifically, the evaluation addresses the following outcome questions:

- Did the regional courses help participants build capacity (knowledge and skills) in Operations and leadership?
- Did participants utilize the knowledge acquired from the courses?
- What factors influenced participants' knowledge utilization?

METHODS

2.6 A quasi-experimental design was employed to estimate the course impact on participants' knowledge and skills. Outcome measures obtained from past course participants (in FY03) were compared to that of prospective course participants (on the schedule for FY04), serving as waiting list controls. It was assumed that past participants and prospective participants would have similar demographic characteristics and motivations to attend the courses. By using prospective course participants as a

comparison group, we eliminate any possible bias due to selection of the learning events, thus measuring the potential effects of the courses on a set of outcome indicators on equal terms. Only two offerings for each course are analyzed, as these were the only offerings available at the time of the evaluation that were offered at the same time of year, allowing a correction for seasonality.

2.7 All past course participants (referred as participants) were surveyed via email and telephone 12 months after attending the FY03 course offerings.⁵ The survey questionnaire asks participants for feedback regarding their overall perception of course relevance, quality, and impact, as well as their knowledge and skill utilization and the factors influencing knowledge utilization. The follow-up questionnaire can be found in Annex 1A.

2.8 A pre-course survey was conducted with all prospective course participants (referred as controls) from three selected offerings in FY04.⁶ These staff were asked to fill out a web-based survey questionnaire about one week before they attended the courses. The survey asks participants a set of similar questions regarding their familiarity with and level of use of course-related knowledge in their work. The pre-course questionnaire can be found in Annex 1B.

2.9 The evaluation employed paired two-tailed t-tests to estimate the size and significance of the difference in relevance, level of knowledge of the course topics and the utilization of the topic areas between participants and controls of the courses. Given the sample size and commonality of the two courses in terms of key outcome measures, i.e., relevance, knowledge, and utilization, we looked first at the overall course effect by pooling the two courses together. Subsequently, because the course content area is different, which may have a potential effect on participants' behaviors, a separate analysis was carried out for each course.

EVALUATION PARTICIPATION

Survey/interview response rate

2.10 Forty-eight staff participated in the Advanced Operations course in FY03 and 38 were in the prospective FY04 Advanced Operations comparison group as controls. Forty-five staff participated in the Team Leader course in FY03 and 19 participated in FY04 Team Leadership pre-course comparison group. Survey response rates exceeded 70 percent for all four groups (Table 2). Thirty-eight of the Advanced Operations participants (79 percent) responded to the follow-up survey/interviews and the same number of prospective course participants (controls) responded to the online pre-course survey (100 percent). Thirty-eight participants responded to the follow-up survey of

⁵ December 2002 and April 2003 offerings for Advanced Operations; February and May 2003 offerings for Foundations of Team Leadership.

⁶ December 2003 and February 2004 offerings for Advanced Operations; February and April 2004 offerings for Foundation of Team Leadership. However, as the May 2004 offering was canceled, the pre-course survey could not be conducted for this offering.

Team Leadership course participants (84 percent) and 14 pre-course survey respondents – controls (74 percent).

Table 2: Number and survey response rate of study participants and controls, by course

Course	Study Group	Total course participants	Total respondents	Response rate
AO	Participants	48	38	79%
	Controls	38	38	100%
TL	Participants	45	38	84%
	Controls	19	14	74%

Demographic characteristics

2.11 Table 3 presents the demographic characteristics of the study participants. The demographic characteristics of the respondents to the Advanced Operations course evaluation are very similar to those of the participant group and control group. The only difference is the number of years in Bank Operations. Participants had nearly a year of experience over the controls. However, the difference is insignificant. Further, regression analysis shows that there is no significant effect of experience on the level of knowledge and skills.⁷

2.12 The characteristics of the participant group and control group for the Team Leadership course are quite different across grade levels. The participant group is heavily concentrated (60 percent, N = 20 in the GG grade level), while there is a smaller share of GG staff among the controls (35 percent, N = 5). The staff in the control group have less operations experience, but more regional experience, than do the participants. However, a regression analysis shows that there is no significant effect of grade level on knowledge level⁸, which indicates that these two groups are equivalent in terms of the outcome measures.

2.13 In comparing the two courses, the Advanced Operations course seems to attract more experienced staff than the Team Leadership course does. It is possible that some senior staff took the Team Leadership course earlier in their careers. It is also possible that the content of the courses reflects the needs for different staff groups.

⁷ β Operational experience = 0.055, $t = 1.56$, β Regional experience = 0.012, $t = 0.40$, $N = 75$, $R^2 = 0.112$.

⁸ β Grade = -0.085, $t = -0.69$, $N = 48$, $R^2 = 0.010$.

Table 3: Demographic characteristics of the respondents

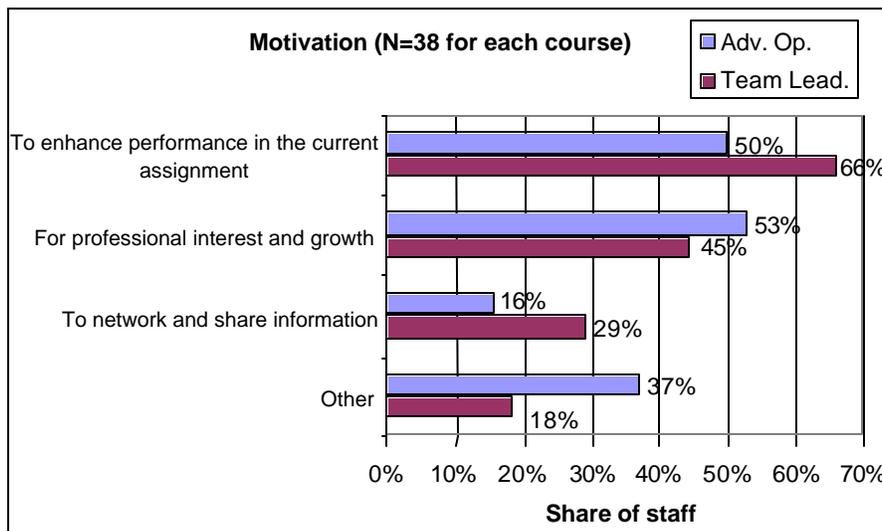
Demographic	Participants	Controls
<u>Advanced Operations</u>	<u>N=38</u>	<u>N=38</u>
Years in EAP/SAR	6.00	6.04
Years in Bank' s Operations	7.05	6.16
Grade level	Average GF/GG	Average GF/GG
<u>Team Leadership</u>	<u>N=38</u>	<u>N=14</u>
Years in EAP/SAR	5.60	6.23
Years in Bank' s Operations	5.83	5.35
Grade level	Average GG	Average GF

OVERALL RESULTS

Motivation for attending the learning

2.14 At the 12-month follow-up, past course participants were asked to reveal their main reasons or motivations for having attended the Advanced Operations and the Team Leadership courses. As shown in Figure 1, the most common reason is to enhance their own work performance, followed by professional interest and growth. Only a few staff mentioned that the reason they attended the learning was that there were no other learning opportunities available for them or that the course attended was simply needed in their work.

Figure 1: Motivating reasons for attending the courses

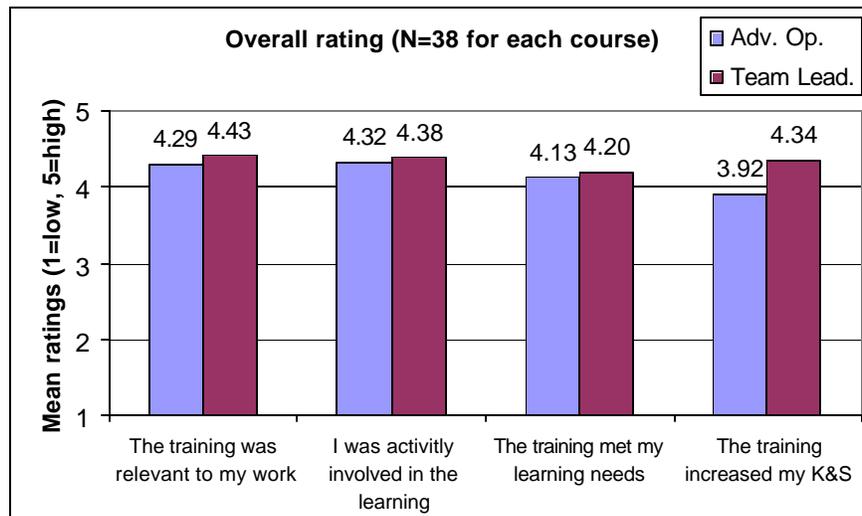


Overall quality of the courses

2.15 The quality of the course was assessed by participants’ ratings of the following four aspects: relevance, participant involvement, meeting the learning needs, and enhancing knowledge and skills. As shown in Figure 2, the average ratings of all four

categories are above four on a scale from 1 to 5 (low to high). Respondents reported that the learning was highly relevant to their work; they were actively involved in the learning, that it met their learning needs and, to a slightly lesser extent, that it increased their existing knowledge and skills in general.

Figure 2: Participants’ overall satisfaction with the courses



Perceived relevance, level of knowledge and utilization

2.16 As indicated earlier, the prospective course participants (controls) have similar demographic background and years of work experience in the same regions, particularly for the Advanced Operations course. Therefore it might be expected that prospective participants’ perception of relevance, level of knowledge and degree of utilization of content was similar to that of the course participants’ just before taking the course.

2.17 Based on this assumption, ttests were used to compare the two groups with regard to mean level of relevance of the course topics, level of knowledge and skills of the course topics, and the use of the course topics in work. Separate overall measures concerning relevance, knowledge and skill, and utilization level were constructed from participants’ and prospective participants’ ratings of the various course topics.⁹

2.18 Figure 3 compares these three measures between the two groups. It can be seen that:

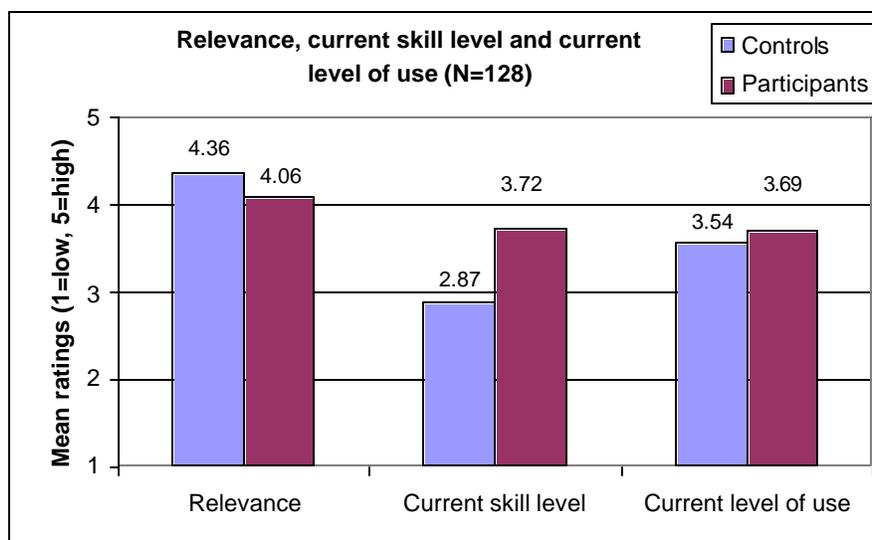
- *Relevance of the topics of the two courses to staff’s work is high* Both participants and controls rated the course topics as highly relevant. However, the rating was significantly statistically higher in the control group than in the participant group ($t = -3.10, p < 0.01$), suggesting that staff find the course topics more relevant to their work before attending the course. This difference might suggest that participants rated relevance from different perspectives. Before attending the course, participants may rate relevance

⁹ The topics of the two courses are discussed separately below.

according to the topics of the course sessions, whereas, after attending the learning, they may rate relevance according to their familiarity with and current knowledge of the content of the courses. This point was also made in another evaluation study carried out by IEG.¹⁰ Another possible explanation is that staff may rate course relevance according to their current work priorities.

- *Level of knowledge of the course topics is modest, but significantly higher for the past participants.* Control group participants rated their level of the knowledge of the topics of the course that they were about to attend below the midpoint 3 on a scale of 1 to 5 (low to high), but participants' ratings were significantly higher, averaging 3.72 on the same scale. The t-test confirms that this nearly 30 percent difference is significant ($t = 8.61, p < 0.00$), and indicates that participants increased their knowledge as a result of attending the learning.
- *Utilization of knowledge and skills in the topic areas also averaged around the midpoint, but there was no significant difference between participants and prospective participants (controls).* Although participants show a higher level of knowledge and skills than controls, there is no evidence of a significant increase in the utilization of acquired knowledge and skills for those attended the course. Average utilization of skills rated modestly above the midpoint for both groups. The t-test confirms that the difference in level of use between the two groups is insignificant ($t = 1.23, p = 0.22$), suggesting participants did not increase the use of the course contents substantially (in quantity).

Figure 3: Comparison of perceived relevance, level of knowledge and utilization between participants and prospective participants (controls)

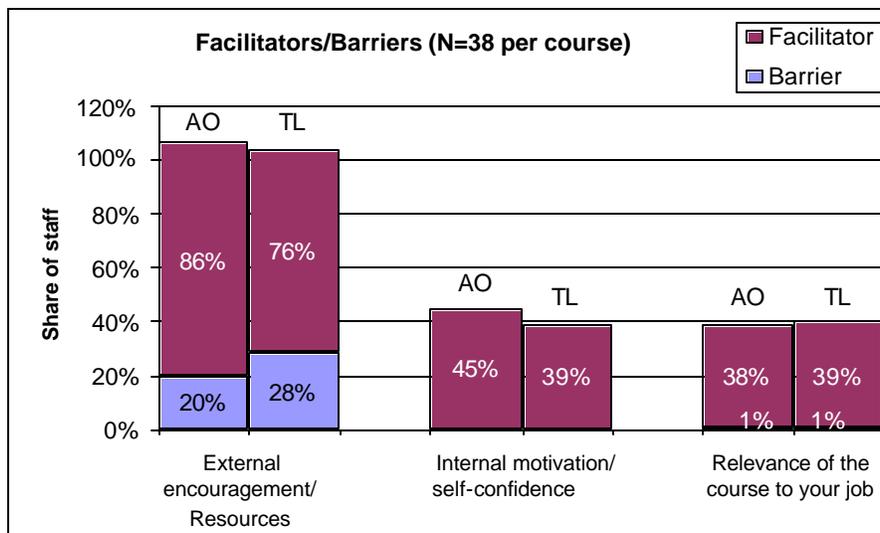


¹⁰ Liu et al. (2004) Evaluation of Human Development Forum 2003. WBI Evaluation Studies, No. EG05-98. The World Bank Institute.

Factors influencing knowledge utilization

2.19 *No major barriers to the utilization of acquired skills are identified.* At the follow up, participants were asked to rate the extent to which factors both internal and external helped or prevented them from using the knowledge gained from the course. To simplify the analysis, motivation and self-confidence are defined as internal factors, while encouragement from managers and peers, institutional and organizational policy and tools, and resources are defined as external factors. As seen in Figure 4, participants did not report any barriers. Rather, over 80 percent believed that one or more external factors facilitated the utilization of new or improved knowledge and skills. This conclusion pertains particularly to participants who attended the Advanced Operations course, which suggests that positive, external factors are critical for utilizing operational knowledge and skills. Fewer than 30 percent in either course found barriers in the external work environment. Moreover, approximately 40 percent believed that internal factors helped them to apply learned skills, and a similar percent of staff stated that the relevance of the course material to their workload was an important facilitator. These findings highlight the importance of a positive environment for facilitating behavioral changes.

Figure 4: Factors influencing utilization of knowledge and skills*



* Note that external facilitators is a group of several factors and since respondents could choose more than one factor the sum of staff may total more than 100 percent.

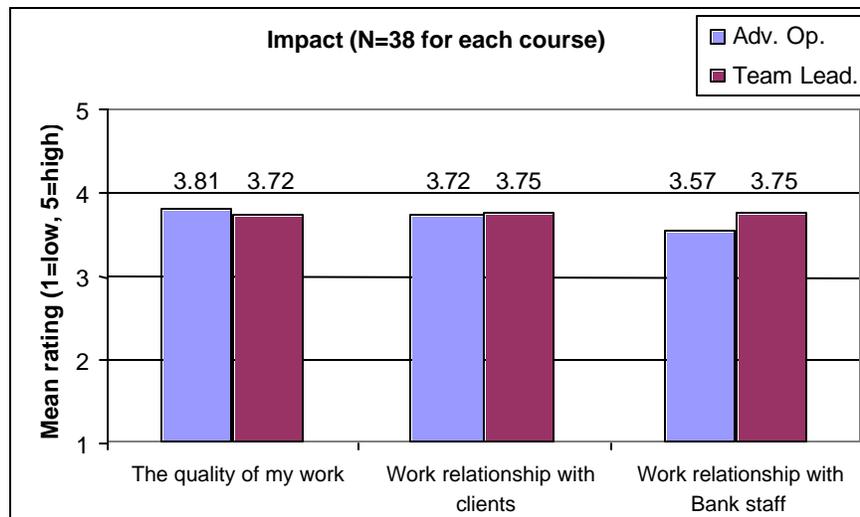
Participants' perceived overall impact

2.20 At the follow-up survey, three questions were used to assess participants' perception of the impact of the course and its subsequent benefits. On a 5-point rating scale (with 1 = low and 5 = high), participants gave an average rating of 3.6–3.8 for the questions assessed (Figure 5), suggesting that participants perceived that the impact and subsequent benefits of the courses were at a modest to somewhat higher level.

2.21 Moreover, some small differences between the two courses are observed in Figure 5. The Advanced Operations course shows a stronger effect on the quality of work.

Adversely, the Team Leadership course has more influence on internal working relationships. These differences may simply reflect the nature of the two courses.

Figure 5: Perceived impact of the courses



Comments and suggestions from participants

2.22 Participants of the course follow-up survey believed that they could benefit more from real-world examples, case studies and more practical rather than theoretical learning, especially for the Team leadership course. Many participants also expressed a need for more follow up and occasions to refresh their knowledge. Some respondents also mentioned that the material presented was quite heavy, that they need more time to digest it and that the length of the event was excessive; hence, consolidating or reducing the length of the courses should be considered.

EVALUATION RESULTS OF THE ADVANCED OPERATIONS COURSE

2.23 This section presents the evaluation results of the Advanced Operations course specifically.

2.24 The Advanced Operations course is designed for EAP-SAR TTLs, at levels of grade GE+. The course covers seven topics:

- Regional strategies and corporate priorities;
- Quality of new products;
- Quality at entry;
- Quality of supervision;
- Results-based project management;
- Quality at completion; and

- Portfolio management and new style Country Portfolio Performance Reviews (CPPR).

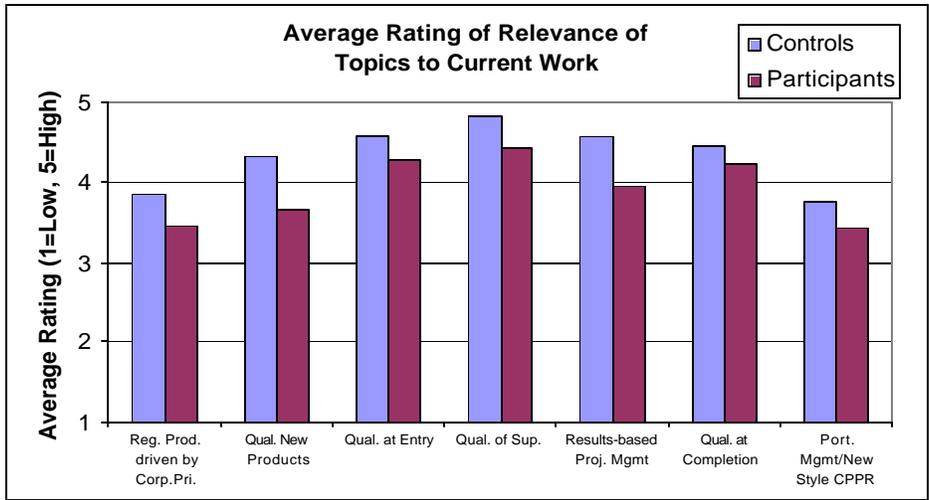
2.25 In the survey, both past and prospective participants were asked to rate each of the seven topics covered in the course on a 5-point scale according to their perceived relevance, their own current knowledge level, and their use of related knowledge and skills.

2.26 The analysis was carried out for each of the three measures (relevance, level of knowledge, and utilization), as well as for an aggregated measure. Detailed results of t-tests for the difference between past and prospective participants (controls) are summarized in Annex 2.

Relevance

2.27 *Prospective participants (controls) rated the relevance of all topics more highly than did participants.* Figure 6 displays the relevance ratings of the Advanced Operations course topics. The highest rated topics of the Advanced Operations course in terms of applicability and relevance are Quality at Entry and Quality at Supervision, which both averaged over 4.0 for both participants and controls. Both groups rated the relevance of Regional Products Driven by Corporate Priorities and Portfolio Management and New Style CPPRs as somewhat relevant to their current work just above the midpoint of the scale. Collectively, the aggregated score of relevance was significantly higher for prospective participants (controls) than for past participants ($t = - 3.56, p = 0.00$). This, again, may be because staff rated the relevance from different perspectives and at different stages in their work programs.

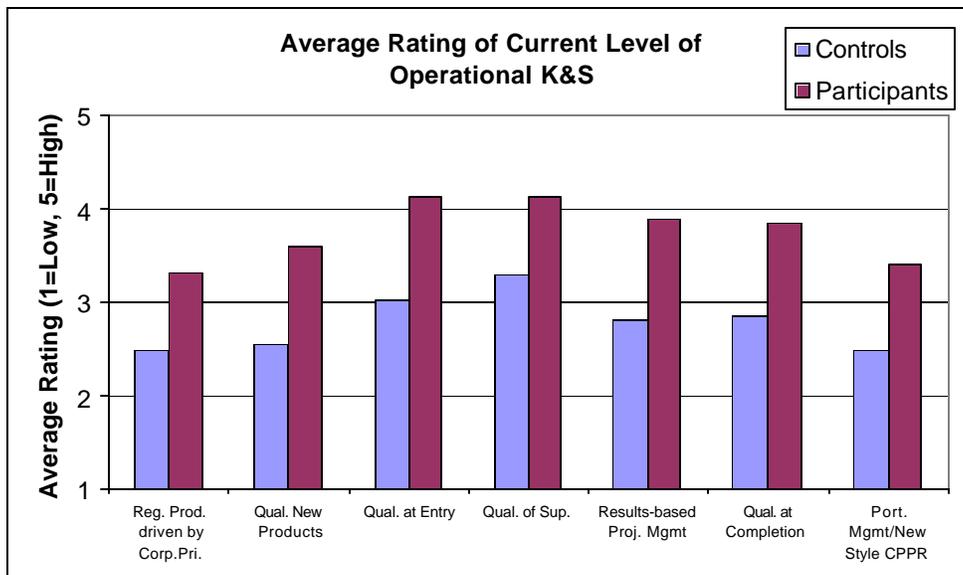
Figure 6: Comparison of perceived relevance of course content between participants and prospective participants (controls)



Current knowledge and skill level

2.28 *Past course participants reported significantly higher knowledge of the course content.* Figure 7 shows that, similar to the results on relevance, Quality and Entry and Quality of Supervision were rated by both the participants and controls as the areas where current level of knowledge and skills were the highest. The knowledge of Regional Products Driven by Corporate Priorities and Portfolio Management and New Style CPPRs was rated very low, averaging around 2.5 for pre-course participants and 3.3 for those who had taken course. More importantly, the aggregated score of knowledge and skills is significantly higher for participants than for controls ($t = 7.23, p = 0.00$), representing a nearly 40 percent difference (increase) on average (from 2.78 to 3.76). Across all topic areas of the course content, participants show between 22 and 41 percent higher current knowledge compared with prospective participants (Annex 2). Without training, past participants should have the same knowledge level as the prospective participants (controls), which means the observed differences in current knowledge level between participants and controls is a measure of the course effect.

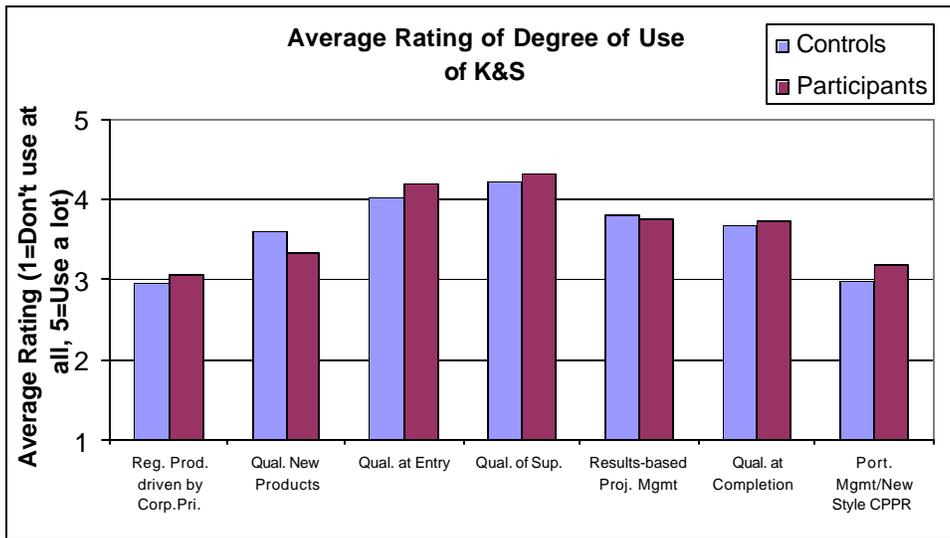
Figure 7: Comparison of perceived current level of knowledge and skills between participants and prospective participants (controls)



Use of knowledge and skills

2.29 *Participants did not increase their use of the course content when they returned to work.* As shown in Figure 8, the level of use of the seven course topics is rated as moderate to high, or between 3 and 4 on a scale of 1 to 5 (low to high), for both participants and prospective participants (controls). No great difference is observed between the two groups across the seven topic areas, suggesting that participants did not increase their use of course content greatly in comparison with the pre-training level (represented by the controls). A possible explanation for this will be discussed later.

Figure 8: Comparison of perceived level of use of knowledge and skills between participants and prospective participants (controls)



How the participants utilized knowledge and skills acquired

2.30 Despite the fact that the statistical analysis did not show any significant difference in terms of the level of utilizing knowledge and skills between the two groups, some participants reported that learning improved their efficiency and quality of their work. As captured by the open-ended questions, participants most often mentioned issues in project preparation and project supervision and explained that, having attended this course, they are more effective in communicating with their clients, have improved their planning and coordination skills across sectors, realize the importance of client-oriented projects, and are more realistic about their clients' capacity. Participants also learned specific ways to improve Quality at Entry reports, Project Supervision Reports (PSR) and CPPRs. According to the participants, having taken the course, they pay attention to more details during project preparation and before going on a supervision mission.

2.31 Specifically, participants were able to mention several occasions during which they had improved their relations with other Bank staff and clients. One follow-up respondent said:

“Before taking the course I couldn't relate well, for example, to procurement because I didn't know what they were doing. Understanding the work of others improves personal/work relationships a lot.”

2.32 Almost all participants agreed that attending the course has in some ways influenced their approaches, for example, in conducting strong quality reviews at different stages. Some participants argued that although the course did not change their approach specifically, it reinforced their existing knowledge or provided them with more focus. Some stated that the course helped them to acquire a broader and more systematic understanding of the Bank's Operations. Understanding of the project cycle and the work of other sectors were mentioned most frequently in this category.

EVALUATION RESULTS OF THE FOUNDATIONS OF TEAM LEADERSHIP COURSE

2.33 The Foundations of Team Leadership course for staff in EAP and SAR regions covers three content areas spanning a total of ten individual topics:

Self-awareness

- Learning from multi-instrument feedback on preferences;
- Thinking styles and behavior;
- Engaging in self discovery and self-reflection; and
- Developing an action plan for self-improvement based on peer feedback.

Team Dynamics

- Generate team synergy;
- Balance people-related factors such as listening, supporting, etc.;
- Balance task-related factors like goal selection, alternatives, implementation; and
- Manage teams under conditions of stress.

Communication and Dialogue

- Understand the impact of mental models on one's thinking and behavior; and
- Increase influence through better communication.

2.34 Similar to the Advanced Operations course, we asked both participants and prospective participants (controls) about their perceived relevance, current knowledge level, and use of the course content.

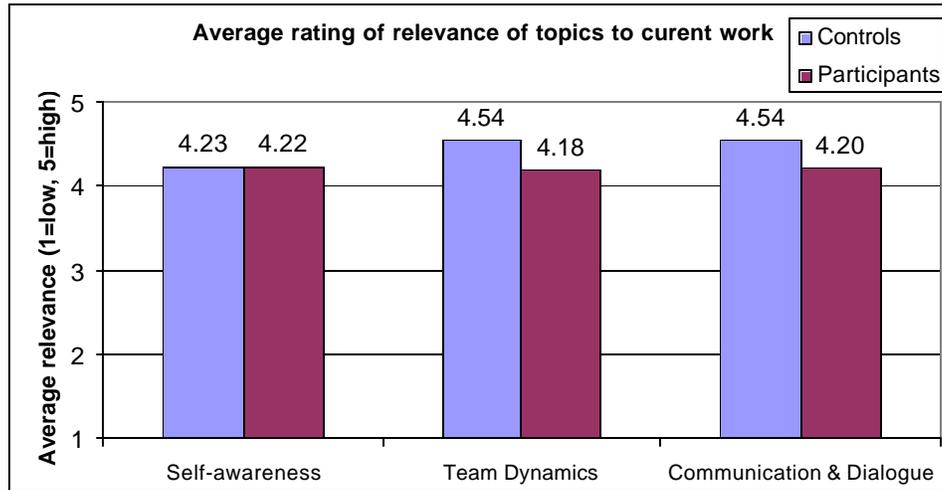
2.35 Again, the analysis was carried out separately for relevance, level of skills and utilization as well as for aggregated measures for the individual topics in each content group.¹¹ Detailed results of t-tests for the mean score difference between the participants and controls are summarized in Annex 3.

Relevance

2.36 *Both participants and prospective participants (controls) rated the learning as highly relevant to their work* (Figure 9). Although prospective participants (controls) rated two topic areas slightly higher than did participants, the difference is insignificant ($p = 0.07$; $p = 0.10$). With respect to the individual topics within the three topic groups, the highest ratings were given to topics such as Engaging in Self-discovery and Self-reflection, Thinking Styles and Behaviors, Generating Team Synergy and Balancing Task Related Factors.

¹¹ Both factor analysis and correlations support the intuitive division of the data into the three groups.

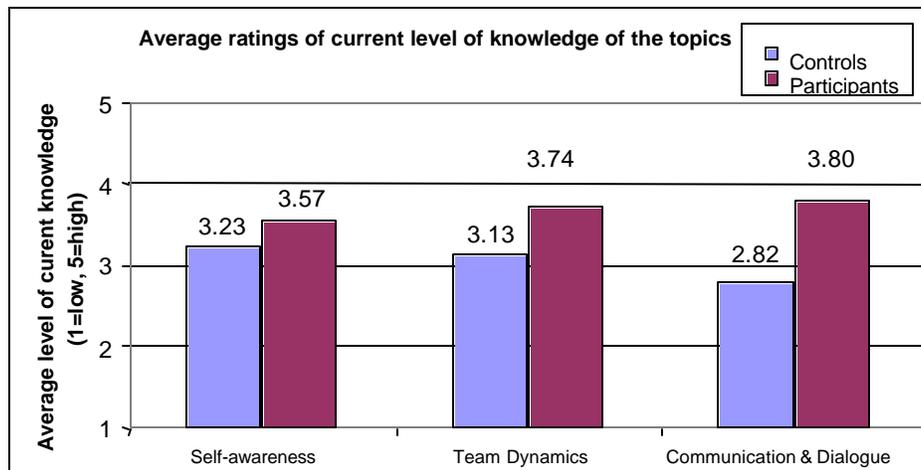
Figure 9: Comparison of perceived relevance of course content between participants and prospective participants (controls)



Current knowledge level

2.37 Past participants reported a higher knowledge level than did prospective participants (controls), and the levels of knowledge are slightly above the midpoint. Figure 10 displays the participants' and controls' mean ratings of current knowledge and skills in the course topics. On a scale of 1 to 5, the average score of participants' knowledge and skill level before the course (as represented by the controls' score) is around 3, while it reaches around 3.7 after the course for an increase of 20 percent. The current scores for the topics area of Team Dynamics and for Communication and Dialogue are significantly higher for past participants than for prospective participants ($t_{TD} = 3.34$, $p = 0.00$) ($t_{CD} = 5.02$, $p = 0.00$). Across individual topics, a significantly higher score for knowledge and skills is seen in six of the ten individual topics (Annex 3). As the knowledge score of the past participants should be the same as that for the prospective participants, the difference (increase) in knowledge level can be attributed to the course effect.

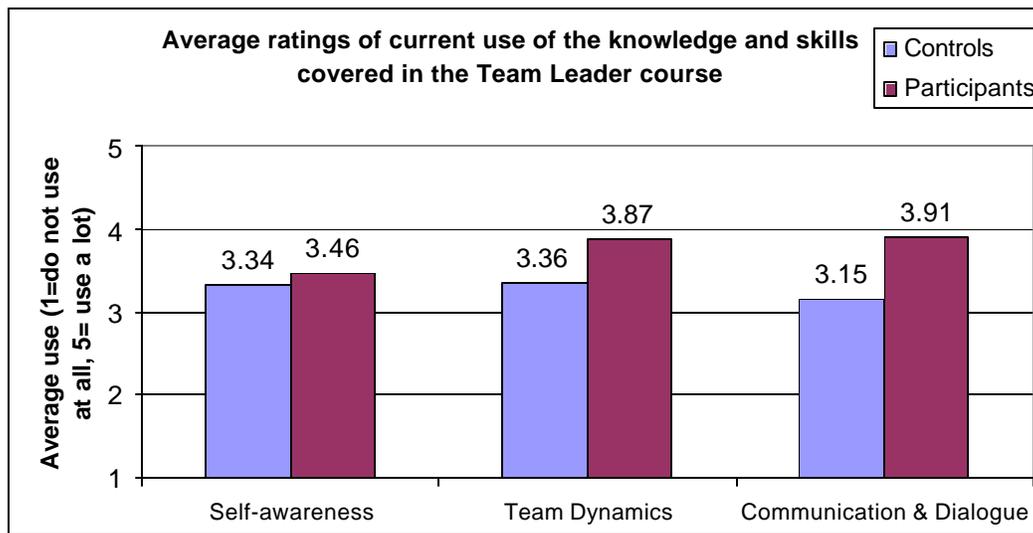
Figure 10: Comparison of perceived current level of knowledge and skills between participants and prospective participants (controls)



Use of knowledge and skills

2.38 *The level of use of knowledge and skills was higher for those who have attended the learning event.* Figure 11 illustrates the level of use of the knowledge and skills associated with the course contents for past course participants and prospective participants (controls). Among the three content areas, the current use of knowledge and skills is significantly higher for Team Dynamics (15 percent, $t = 2.41$, $p = 0.00$) and Communication and Dialogue (24 percent, $t = 3.36$, $p = 0.00$), but not for self-awareness. Topic areas where participants made much more use of knowledge and skills than prospective participants (controls) include Understanding the impact of mental models (28 percent), Generating team synergy (19 percent), Managing teams under stress (17 percent), and Balancing people related factors (16 percent). Again, all these differences (changes) may be due to the learning course.

Figure 11: Comparison of perceived level of use of knowledge and skills between participants and prospective participants (controls)



How the participants utilized knowledge and skills acquired

2.39 A few follow-up participants provided concrete examples of how they had made use of the knowledge and skills they acquired in the Team Leadership course, although some mentioned that they had been able to better deal with particular people who presented difficulties. A few other people said that the use they made of the course came through their action plan or that they were able to be more open to the suggestions, ideas and opinions of other team members.

2.40 One participant answered in this manner when asked to explain how attending the Team Leader course has impacted his work:

“I would say that I was able to communicate better with my manager and negotiate a change in my work with my manager and another colleague,

and they are both very different people. The course also helps me to communicate better with clients in the field.”

2.41 Twenty-two of the participants interviewed (58 percent) also mentioned that the course has changed their approaches to their work. Content analysis reveals that the majority of respondents mentioned that it improved their communication, particularly their listening skills. A few people also added that it helped them to understand people better, realize the strengths and minimize the weaknesses of their teams and become more patient and tolerant with others.

SUMMARY AND DISCUSSION

2.42 One important finding derived from this analysis is that participants have benefited from both courses. It is estimated that participants increased their knowledge and skills by almost a third after attending the course. The finding that there was no increase in the average level of utilizing the knowledge and skills for either course may be due to the fact that both courses target staff who have been working in Regional Operations and have been playing a team leader role for a while at the time they attended the learning courses. The real impact is that both courses have helped participants to work more sufficiently, which is exactly what the courses aim to achieve. Additionally, there is evidence from the content analysis that the majority of course participants perceived that the courses have had a modest impact on their practice with respect to their performance efficiency and relationship with internal and external clients. This analysis suggests that the two regionally designed core courses successfully achieved their objectives.

2.43 A number of differences between the two courses were observed. For the Advanced Operations, all participants (past and prospective) reported higher relevance for the course topics Quality at Entry and Quality of Supervision, than other course topics. In practice, these topics represent areas where participants changed their approaches to their work and made the most effective use of the course knowledge and skills acquired. By contrast, for the Team Leadership course, no topics were rated as more relevant than others. Likewise, the increased knowledge level and knowledge utilization appear less selective and cross cutting across content areas for the Team Leadership course.

2.44 This phenomenon may be explained by the nature of the course. The topics covered in Advanced Operations are technical and operational in nature. The relevance of the course may be perceived differently and the level of utilization could differ for a TTL at a certain point in time related to project cycles. The results obtained, particularly high relevance at the start of the course, improvement in skills but no increase in the level of the application of skills after the course, seem tied to the immediate need of particular operational skills that the Advanced Operations course addresses. The impact brought by the Advanced Operations course may have more likely concentrated on the quality of the work.

2.45 In contrast, materials covered in the Team Leadership course offered a more behavioral and general type of knowledge that could be used throughout a staff jobs. It seems possible that since behavioral knowledge, especially that in Team Dynamics and Communication and Dialogue, is not more relevant at a certain point in time than at others, and the learning can be applied at many occasions, not just in particular tasks. The impact of the Team Leadership course is primarily concentrated on individual behavior for such as the improvement of relationships with Bank staff.

3. STAFF LEARNING AND PERFORMANCE

3.1 Evidence from the previous chapter clearly shows that learning can bring immediate benefits to individuals by enhancing their knowledge and skills. The question that remains is how the immediate and measurable benefits from individual learning events transfer or contribute to overall staff performance. This evaluation seeks the answer by examining the relationship between staff learning and the staff performance data collected from FY03 OPE. A more detailed discussion of the methodology used in this analysis and tables with statistical results are attached in Annex 4. Some key methodological issues and results are discussed below.

MEASUREMENT ISSUES

Outcome measures: the OPE and its limitations

3.2 Bank staff members' work performance is reviewed annually using an HR management tool known as the Overall Performance Evaluation (OPE). For general staff, two sets of staff performance indicators constitute OPE assessment: Results and Behavioral.

3.3 One needs to keep in mind the limitations of OPE scores when using it as a measure to determine concrete results. Although an OPE assessment is different from survey information, and is a record of the evaluation conversation between a staff member and his/her manager, it is still a subjective measure. Thus, some biases associated with differences in managers' personal perception and level of observational competence may be present. Moreover, the experience from two previous IEG studies suggests that variations in OPE scores is relatively modest.^{12,13} It is therefore unlikely that a single learning event, especially one of short duration, would affect staff performance as measured by OPE scores.

3.4 Even with these limitations, in the absence of other valid and direct measures of staff performance, OPE as a proxy measure is seen as a useful complement to self-reported information. Also, because the current evaluation targets a large and diverse staff population (over 1,000 staff in EAP and SAR together), and staff with long learning experience (over four years), it is worthwhile to investigate whether learning participation and diverse learning experience enhances staff performance as measured by OPE.

3.5 It is also important to acknowledge that there is no simple relationship between learning and staff performance. Any observed differences in OPE scores corresponding

¹² Bardini et al. (2003). ACS Network Impact Evaluation: The Building your Skills in a Team Based Environment Course. WBI Evaluation Studies No. EG04-75, The World Bank Institute.

¹³ Eckert et al. (2004). Evaluation of World Bank Sector Fora, FY02. WBI Evaluation Studies No. EG04-83, The World Bank Institute.

to various learning practices are not necessarily attributable to learning entirely. Staff performance is a complex issue, and the influencing factors cannot always be identified and controlled. Furthermore, the ability to predict future performance from these analyses is diminished by the “ceiling” effect (limit to how high the scores can go) of the OPE performance measures.

Input measures: learning participation

3.6 The evaluation extracted longitudinal staff learning data (FY99-FY02) from the HR PeopleSoft. Consistent with the methodology used in the phase I evaluation, learning participation was measured in terms of both intensity and regularity. Intensity refers to how frequently during each fiscal year staff took part in learning, and is measured by the number of activities and days of learning that a staff member attended per fiscal year. Regularity is measured by the number of years during a four-year period (FY99-FY02) in which the staff took part in learning. Regularity is said to be high if the staff participated in some kind of learning at any point of time during each one of the four years.

HYPOTHESIS EXAMINED AND METHODS

Hypothesis

3.7 This evaluation examines the hypothesis that learning participation has a direct effect on staff OPE scores. This hypothesis is based on the assumption that learning participation may lead staff to change their performance in the workplace. As a result, learning participation may influence the managers’ perception of the staff performance, and hence, may ultimately influence a change in the staff OPE scores.

3.8 In addition, based on an observation from the phase I analysis, it is hypothesized that learning participation (both *Intensity* and *Regularity*) can be influenced by the exogenous demographic characteristics of staff members.¹⁴ The exogenous variables analyzed were gender, age, length of time working for the Bank, grade level in FY02, highest level of education obtained and whether or not staff is stationed at HQ. These demographics characteristics may also affect OPE scores indirectly through their effect on participation in learning. In this analysis, we do not distinguish between those who participate in regional learning programs and those who attend courses offered by other learning program providers.

Methodology

- 3.9 Two methods were used to estimate the effects of staff learning on OPE scores:
- Ordinary Least Squares (OLS) and Ordered Probit regressions of OPE scores on participation; and

¹⁴ An additional influence on participation is that managers have to approve activities and may recommend certain training for staff, or not approve sending staff to a specific course.

- Propensity Score Matching of staff's probability of learning participation using t-tests to evaluate the differences in the OPE scores between two groups.

RESULTS

Determinants of learning

3.10 Consistent with the phase I finding, the regression results show that the main determinants of learning days among these variables are gender, number of years in the Bank, grade level and station country. Results show that female staff participate in more learning days than males; the longer a staff member has been working in the Bank, the lower the level of learning participation; higher grade level staff attend more often than lower level staff; and those stationed in HQ (as compared to country offices) attend more learning days per year.

Impact on OPE

3.11 It appears that, among the demographic characteristics of the staff, having been at the Bank for many years and having a high grade level has a negative effect on overall Behavioral assessment OPE scores. However, there is no evidence that there are significant effects on learning *Intensity* or *Regularity* in the overall Behavioral OPE scores. The same general findings hold for the effect of learning *Intensity* and *Regularity* on overall Results OPE scores.

3.12 Because our initial statistical regression techniques were not able to explain the variation in OPE scores, we evaluated the difference in OPE scores between only high-intensity learners and low-intensity/non-learners, as the effects of learning attendance are likely to be most apparent in these groups.

3.13 Figures 12a and 12b display the distribution of high- and low-intensity learning staff in each OPE category. It can be seen that the variation in OPE scores across both groups is rather small in terms of Results and Behavioral ratings.

3.14 It is therefore not surprising that the Propensity Score Matching method generated results consistent with the regression methods: learning participation has no effect on OPE scores. There is no significant difference in the mean of any of the OPE scores between the high-frequency participants of regional staff learning and those who never participate or participate very rarely. The difference in both Behavioral and Results assessment OPE scores for staff who often take part in learning, and staff whom rarely or never do so, is statistically insignificant.¹⁵

¹⁵ We also conducted an analysis that isolates low and high OPE performers (the top and bottom quartile/decile of staff) and compared their level of training participation. Consistent with the main results, this analysis finds no evidence that high OPE performers attend more learning than low OPE performers.

Figure 12a: Distribution of high and low intensity learning staff in each OPE category – Results rating

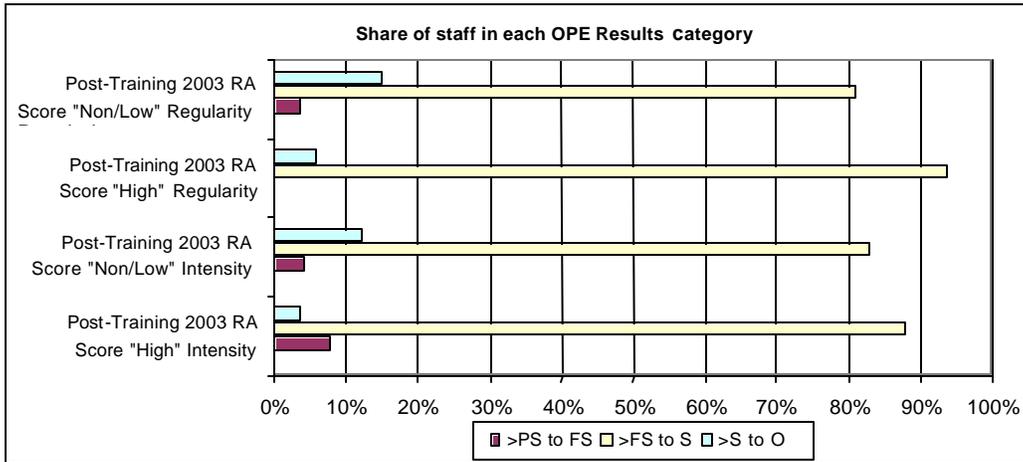
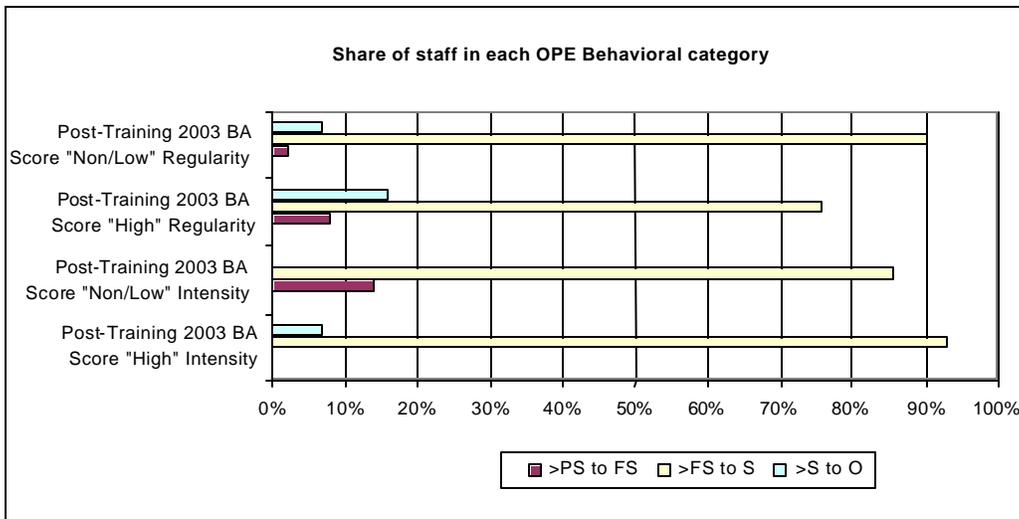


Figure 12b: Distribution of high and low intensity learning staff in each OPE category – Behavioral rating



3.15 The small variation in OPE scores across individuals may preclude finding a statistical relationship between OPE scores and learning. There may also be a substantive reason for this lack of relationship that is undetected by the study. On the other hand, it may be possible that a relationship exists in the opposite direction: OPE on participation. Staff with lower OPE scores may actually be more motivated to engage in learning opportunities than staff with higher OPE scores due to a greater need for improving their performance. However, a further regression analysis did not support this hypothesis.

3.16 Thus, the analysis did not support the relationship between staff learning participation and staff performance. Nevertheless, the experience from this analysis may again highlight the limitations of using OPE scores as an outcome measure for staff learning.

4. STAFF LEARNING AND UNIT PERFORMANCE

4.1 After presenting the analysis of learning impact on individuals, the focus of this chapter is on exploring the links, if any, between staff learning and unit performance.

METHODS

4.2 A number of qualitative methods are employed for data collection and for the analysis, including consultation with the Asia Learning Group, interviews with regional unit managers and a desktop review and analysis. Due to the nature of available information gathered, the analysis carried out is largely descriptive in nature.

Consultation with the Asia Learning Group

4.3 Several consultation meetings took place between the evaluation team and the Asia Learning Group for the development of the framework and to identify indicators for measuring unit performance. Meanwhile, in consultation with the Asia Learning Group, eight regional units (four in EAP and four in SAR) were selected for the review and analysis. They include three operations units and one support unit¹⁶ in each region. These units were selected because of their importance in the lending and operational priorities of the regional business. For confidentiality reasons we do not report the names of the units analyzed, but label them randomly as EAP1 (support unit), EAP2-4 (operations units), SAR1 (support unit) and SAR2-4 (operations units).

Manager interview

4.4 A total of 11 managers from the eight selected regional units and from the OPCS and the Quality Assurance Group (QAG) in the Bank were interviewed. The interview focused on two key issues: an overall picture of unit staff learning, and the use of operational performance indicators to measure unit performance with respect to portfolio quality and outcome. The interview questions and a summary of interviewees' responses are given in Annex 5.

Desktop review

4.5 The unit profile and staff learning participation information for the selected units came from HR. Learning investment/budget for these units were provided by the Asia Learning Group. Furthermore, results of the 2003 Bank Staff Survey with respect to learning and development were also taken into account. The rest of the data pertaining to regional and unit portfolio performances was extracted from the Bank Business Data Warehouse.

¹⁶ These units provide administrative support to regional operation units.

UNIT PERFORMANCE: A FRAMEWORK AND INDICATORS

Framework

4.6 A framework was developed in consultation with the Asia Learning Group for measuring unit performance. The framework represents common practices in the Bank and covers three dimensions:

- (a) organizational effectiveness;
- (b) portfolio quality and outcome; and
- (c) client relationship.

4.7 A set of specific operational indicators was developed within the framework to study unit performance in relation to staff learning. Most of these indicators were developed based on those already used in the monthly operational reports provided by the units to regional management, and on the interviews carried out with managers of the eight selected units and some senior staff from EAP, SAR, QAG, and OPCS. A summary of the framework, desired outcomes and performance indicators is provided in Annex 6. A brief description of the framework and indicators is presented below.

Organizational effectiveness

4.8 Organizational effectiveness is a measure of the ability of an organization to achieve its business goals effectively and efficiently. It is about each individual doing everything he/she knows how to do and doing it well. It is about the capacity of an organization to produce the desired business results with minimum time, money, and human and material resources. It is also an ongoing process of transferring input to outcome. Of all the indicators measuring organizational effectiveness, this study is specifically interested in staff learning-related indicators. The main indicators identified and used for this evaluation are:

- (a) number and percentage of staff having ILPs;
- (b) average number of days of learning per staff; and
- (c) average expenditure on learning per staff.

4.9 In addition, the unit staff learning management and 2003 Bank Staff Survey results for learning and development were also considered when available. These learning indicators are used in this evaluation as outcome indicators not only of organizational effectiveness, but also as input indicators of the units' productivity.

Portfolio quality and outcome

4.10 Initially, about 20 measures were extracted from the Bank Business Warehouse as operational performance indicators for the selected units. However, although all of these indicators are important, some of them are strongly influenced by external factors rather

than by a unit or staff capacity. For example, lending volume, in most cases is affected by a country's condition and needs of a borrowing country. It is difficult for the evaluation to measure the influence of learning on these operational performance indicators. After consulting with the Asia Learning Group, the following eight indicators were selected to measure productivity, portfolio quality and outcome of unit performance in this evaluation. They are:

- (a) projects at risk ratio;
- (b) commitment at risk ratio;
- (c) disbursement ratio;
- (d) realism ratio (the percentage of problem projects in the total number of actual and potential problem projects);
- (e) proactivity ratio (percent of projects rated as problem projects 12 months ago that have since been restructured, closed, upgraded, suspended or partially cancelled);
- (f) percentage of projects facing financial management problems;
- (g) percentage of projects facing financial performance problems; and
- (h) percentage of projects facing procurement problems.

4.11 These indicators have been widely used in the Bank as a measure of unit performance. In the regional context, all but the last three indicators have been used by the units in their monthly operational reports to regional management. The last three indicators are added as they also help evaluate the impact of the two support units selected.¹⁷ The underlying assumption is that if these support units carry out adequate work programs for Bank staff and clients and participate in supervision missions adequately, fewer projects will experience procurement and financial problems.

Client relationship

4.12 Bank management has been emphasizing that the challenge for Bank staff is to work with country clients to develop more result-oriented poverty reduction and other national strategies; undertake more integrated institutional assessment of country readiness; and offer a more strategic menu of tools to strengthen capacity in diverse country circumstances. In this context, to get better results from the Bank's Operations, there is a need to learn from past experiences to improve the results focus of country programs and to align sector projects and programs with country priorities. The following five measures are suggested for the client relationship dimension:

- (a) developing the behavioral and integrative skills of Bank staff;
- (b) involving civil society, including non-governmental organizations (NGOs), in project preparation and implementation, as well as in the preparation of Country Assistance Strategies (CAS) and Poverty Reduction Strategy Papers (PRSP);

¹⁷ Both support units deal with procurement and financial management to the regions.

- (c) reducing the time lag in client response by following a rapid response system and making effective use of information technology;
- (d) client capacity enhancement for better economic policy and project implementation; and
- (e) building country demand for, and capacity to adopt, results-based approaches to development effectiveness by incorporating monitoring and evaluation systems to evaluate outcomes.

4.13 Due to time constraints, this last element is not included in the current phase of the evaluation.

REVIEW RESULTS OF STAFF LEARNING FOR THE SELECTED UNITS

4.14 Detailed statistics of staff learning for the selected units are summarized in Annex 7. Some highlights are discussed below.

Individual Learning Plan (ILP)

4.15 Ideally, every staff member should have an ILP at the time of the annual OPE process. As shown in Table 4, in FY02 the percentage of staff having ILPs for both regions is comparable to the Bank as a whole. However, an increase is seen in FY03 for both units, especially for SAR (a 70 percent increase), whereas the percentage for WB as a whole decreased somewhat.

Table 4: Staff with Individual Learning Plan

FY	WB	EAP	SAR
02	1721 (23.4%)	123 (21.2%)	116 (24.0%)
03	1616 (20.7%)	157 (25.6%)	212 (40.9%)

Source: HR data.

4.16 The picture varies across the eight selected units. Annex 7 shows that in FY02 the percentage of staff having ILPs ranges from 5 to 15 percent in five of the eight units selected. Two units (SAR3 and EAP3) show better results, with ILP percentages of 71 and 72 percent, respectively.¹⁸ Results improved in FY03 in six units with ILP percentages ranging from 11 to 87 percent. However, out of the remaining two units, SAR1 had no ILP for its staff, and EAP2 had ILPs for only four percent of its staff in FY03. As for the EAP and SAR regions as a whole, the ILP percentages were 21 percent and 24 percent respectively in FY02, rising to 26 percent and 41 percent respectively in FY03. Of the eight units reviewed, four showed better performance against the Bank average in FY03, while the remaining three operational units and one support unit (EAP2, SAR1, SAR2 and SAR4) showed lower results than the Bank average. The actual results achieved are fairly satisfactory only for one operational and one support unit (EAP4 and SAR3), which, respectively, had 72 percent and 87 percent staff with ILPs in FY03.

¹⁸ Data for the remaining unit (EAP4) was not available for FY02 as it was reorganized that year.

Staff learning participation

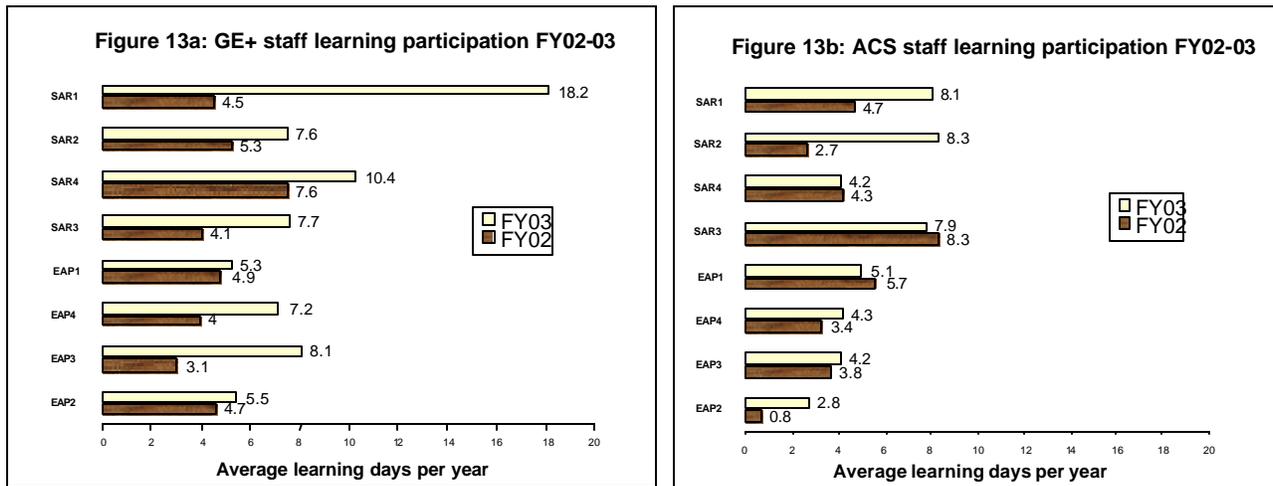
4.17 Staff learning participation is reflected in the average number of learning days per staff per year. Compared with the Bank as whole, as seen in Table 5, the average level of learning participation for both regions is comparable to the Bank's overall level. For the eight units reviewed, Figures 13a and 13b show that the learning days ranged from 3.1 to 7.6 days in FY02 for professional staff, and that there is great variation among ACS staff across the eight units studied, ranging from 0.8 to 8.3 days. However, the situation improved dramatically in FY03 for most units in both the professional staff and for the ACS staff, with a unit average ranging from 5 to 14.2 days for those eight units.

Table 5: Staff learning participation in FY03

	Average learning days		
	GE+	ACS	Average
WB	6.1	6.6	6.3
EAP	6.8	4.5	6.0
SAR	6.7	11.6	8.3

Source: HR data

Figure 13: Staff learning participation FY02-03



Expenditures on staff learning

4.18 Table 6 below summarizes the investment/expenditure on learning per staff of the eight selected units during FY02 and FY03. It can be seen that:

- In FY02, the investment on learning per staff in three of the four EAP units was higher than the EAP average of \$5,064; in the remaining one support unit (EAP1), it was lower. This unit increased its investment per staff in FY03, exceeding the EAP average for that year of \$4,622. In the other three operational units (EAP2, EAP3 and EAP4), the investment per staff decreased substantially in FY03 compared to the previous year. Two of those units

(EAP2 and EAP3) still showed a higher investment per staff than the EAP average in FY03, while one unit (EAP4) showed lower investment per staff. The data shows that there was nearly a nine percent decrease in investments per staff in the EAP region as a whole in FY03. This trend was also observed in three of the four EAP units reviewed.

- The four SAR units reviewed showed considerable increases in investments per staff in FY03. Three of those units had higher investment per staff than the SAR average of \$5,052 in FY02. One unit (SAR3), which had lower investment per staff than the SAR average, increased its investment substantially in FY03, exceeding the SAR average of \$7,262 for that year. Of the remaining three units, one support and one operational unit (SAR1 and SAR4) had higher investment per staff than the SAR average, while one unit (SAR2) had a lower investment even though, it increased that investment over during FY02. For the South Asia Region as a whole, there was a substantial (nearly 44 percent) increase in the average investment in learning per staff in FY03.
- The average investment per staff was comparable for EAP and SAR regions in FY02, but the investment increased considerably in FY03 for SAR, while it declined for EAP that year. This may help explain why the intensity of staff learning was slower in EAP than it was for SAR in FY03.

Table 6: Investment on learning per staff (\$)

Units/Region	FY02	FY03	Diff.
All EAP	5,064	4,622	-442
EAP1	4,685	5,114	429
EAP2	8,756	5,841	-2915
EAP3	6,332	4,821	-1511
EAP4	6,812	4,117	-2695
All SAR	5,052	7,262	2210
SAR1	6,550	8,747	2197
SAR2	5,399	6,049	650
SAR3	4,292	7,843	3551
SAR4	8,279	9,081	802

Source: HR data

Staff learning management

4.19 Some information on staff learning management was collected from the managers interviewed:

- Of the eight units studied, none has a unit development plan. None of the units has a systematic approach to monitoring the completion of ILPs. In some cases, courses taken differ from those included in ILPs as more relevant learning events are announced in the Bank’s Learning Catalog during the year. Overall, staff learning-related indicators for organizational effectiveness are not at present systematically used by the units selected.

- Managers also indicate that the learning record is not seriously taken into account by sector boards and networks in promotional decisions, which are primarily based on performance records.
- The annual budget is generally not a constraint for learning except in learning events involving travel and subsistence costs (as in foreign study tours). Most managers encourage staff to engage in learning. However, work pressure on staff prevents them from attending a satisfactory level of learning events.
- Managers are paying attention to the role of learning in improving client relationships. However, it is difficult to develop indicators to monitor the impact of learning on client relationships. In this context, managers agree that learning to focus on the improvement of staff behavioral skills and client governance is useful. They also feel that learning has an important role in deepening civil society involvement and promoting a participatory approach to project/program development and implementation.
- One interviewee felt that adequate attention to and progress in staff learning depends on improving the incentive structure. Currently, there is no incentive for experienced staff to carry out challenging tasks. Those with extensive Bank experience and adequate learning could be entrusted with the responsibility of dealing with challenging problems (e.g., how to develop a state like Bihar, India, which has serious governance problems), and those persons could mobilize the civil society and devise a strategy based on a participatory approach for solving specific serious problems identified in the field. Their work could provide the foundation for new types of operations.

Results of Bank Staff Survey 2003

4.20 The Bank Staff Survey of 2003 had three main questions relating to staff learning and development. The first question was whether the World Bank is providing continuous opportunities for staff training. The second question was whether the Bank gives the staff adequate opportunity to learn new skills to do his/her job better. The third question was whether the staff takes advantage of the Bank's available learning opportunities. Staff responses from the selected units to these questions are summarized in Annex 8. In summary:

- *Similar to the Bank as a whole, around 80 percent of both EAP and SAR regional staff is satisfied with the continuous learning opportunities provided by the Bank. Of the six units from which response data were received, the favorable response of the staff to the first question ranges from 76 to 91 percent, that is, from somewhat lower to higher than the 80 percent favorable response for the Bank as a whole. The averages for both EAP and SAR regions are 83 percent.*
- *The level of perceived adequate opportunities for learning new skills to do a better job is varied but generally average. The favorable response to the second question ranges from only 51 to 82 percent across the six units, which, again, is both lower and higher than the Bank average of 65 percent and the*

EAP and SAR averages of 62 percent and 63 percent, respectively. In the six units reviewed, two operations units (EAP2 and SAR3) had significantly higher favorable staff response to this question compared to the averages for EAP and SAR as a whole.

- *Similar to the Bank as a whole, around 60 percent of regional staff take advantage of the learning opportunities available to them.* Although a majority of regional staff is satisfied with the learning opportunities provided by the Bank, a substantial portion do not make use of these opportunities. The favorable response to the third question is low (ranging from 63 to 66 percent) for the EAP and SAR regions, which is comparable to the overall Bank rating. However, the favorable response shows great variance across the units reviewed. One operational unit (EAP2) had a 91 percent favorable rating to this question, and another (EAP3) had a 70 percent favorable rating. The other four units (EAP1, EAP4, SAR3 and SAR4) had both much lower and similar favorable ratings (ranging from 46 to 61 percent), compared with 63 percent for the Bank as a whole.

ANALYSIS RESULTS OF PERFORMANCE FOR THE SELECTED UNITS

Possible learning effect on unit performance for the six operations units

4.21 As mentioned earlier, about 20 operational performance indicators were extracted from the Business Warehouse and briefly reviewed for the selected units.

4.22 Detailed statistics for these indicators, together with a technical note describing the performance status across the regions and the six operations units selected, are provided in Annexes 9A and 9B.

4.23 To explore the possible relationship between staff learning and unit performance as measured by portfolio quality and outcome for the six operations units, an analysis was carried out relating five selected operational performance indicators to three core staff learning indicators: percent of staff having ILPs, staff learning days per year and average dollar spending on learning per staff. Annex 10 provides an overview of findings for the learning and operational performance indicators applied to these units.

4.24 To simplify the analysis, the six studied units are grouped into three “learning units” according to findings about their changes on the three learning indicators from FY02 to FY03:

- (a) units that experienced increases in all three learning indicators;
- (b) units that experienced increases in two indicators (learning days and ILPs) but showed decline in learning investment per staff; and
- (c) units that showed increases in learning days per staff, but experienced decline in learning investment per staff as well as ILPs.

4.25 As can be seen from Table 7, all three SAR operations units reviewed fall into the (a) category (reflecting increased attention to all learning parameters); one EAP operations unit (EAP4) falls into the (b) category (having increases in two learning parameters but decline in one); and two EAP operations units (EAP2 and EAP3) fall into the (c) category (showing an increase in one parameter but declines in the other two).

Table 7: Changes in learning indicators for the operational units from FY02 to FY03

	Units	ILPs (%)	Learning days per staff/year	Learning budget per staff (\$)
	SAR2	+6.0	+3.4	+650
(a)	SAR3	+16.3	+2.5	+3551
	SAR4	+10.0	+1.9	+802
(b)	EAP4	+57.5	+2.1	-2695
	EAP2	-8.1	+1.1	-2915
(c)	EAP3	-25.7	+3.4	-1511

4.26 Table 8 below shows the changes in operational performance indicators from FY02 to FY03 for the three groups of learning units. The results appear to be somewhat mixed.

- Compared to “low learning units,” units with better learning indicators (ILPs, learning days and learning spending)—category (a)—learning seems generally to correspond to a high proactivity ratio (which measures percent of projects rated as problem projects 12 months ago but that have since been restructured, closed, upgraded, suspended or partially canceled). The Proactivity Ratio is considered to be an important performance indicator for monitoring improvement in the quality of supervision. As seen in Table 8, another corresponding indicator is possibly Realism Ratio.
- No connection was found between learning and performance indicators for the following: (i) projects at risk; (ii) disbursement; and (iii) commitment at risk. In some cases, study units experiencing declines in learning investment (categories b and c) or in ILPs (category c) still show improved performance.

Table 8: Changes in operation indicators for the study units from FY02 to FY03

Operational indicators	Category (a)			Category (b)	Category (c)	
	SAR2	SAR3	SAR4	EAP4	EAP2	EAP3
Projects at Risk (%)	Risk increased	Risk increased	Risk increased	Risk decreased	Risk decreased	Risk decreased
Commitment at Risk (%)	Risk increased	Risk decreased	Risk increased	Risk decreased	Risk decreased	Risk decreased
Disbursement Ratio (%)	Declined	Declined	Declined	Increased	Increased	Decreased
Realism Ratio (%)	Improved	Improved	Worsened	Remained high	Worsened	Worsened
Proactivity Ratio (%)	Remained at 75%	Remained high at 100%	Remained high at 100%	Remained high at 100%	Worsened	Worsened

Source: Bank Business Warehouse

Possible learning effect on unit performance for two support units

4.27 Two of the eight units selected for review are regional support units; one provides financial management and procurement specialist services to EAP and the other provide procurement services to SAR. With respect to staff learning, both of them increased the number of learning days per staff in FY03 and the learning cost per staff (Annex 10).

4.28 As these two units are not directly involved in the delivery of projects and lending programs, operation indicators for these units have to be different from those for the operations units. Considering the nature and responsibility of EAP1, the following three proxy operational performance indicators are used to study learning effects on unit performance: (i) percentage of projects having the financial management flag in EAP; (ii) percentage of projects having the financial performance flag in EAP; and (iii) percentage of projects having the procurement flag in EAP region. For SAR1, which provides only procurement, procurement is used to measure impact on unit performance. As summarized in Table 10 below, both units present an increase in staff learning parameters from FY02 to FY03, especially for the SAR1 unit. However, the increased learning input in these units was not accompanied by an improvement in their operational performance measures.

Table 10: Changes in learning and operational indicators for the support units from FY02 to FY03

Change in learning and operational indicators	EAP1	SAR1
Staff learning days (mean)/Staff	+0.2	+9.6
Staff learning investment (mean)/Staff (\$)	+429	+2197
Proj. with fin. mgmt. flag in Region (%)	Increased	NA*
Proj. with financial perf. flag in Region (%)	Increased	NA*
Proj. with proc. flag in Region (%)	Increased	Increased

Source: Derived from Bank Business Warehouse data

* SAR1 provides only procurement services to the Region

SUMMARY

4.29 This section presents the review and analysis of unit performance in relation to staff learning.

4.30 A unit performance framework with indicators covering three key dimensions was developed to guide the analysis. Of the three dimensions — organizational effectiveness, product outcome and quality, and client relationship — the current analysis specifically focuses on portfolio outcome and quality.

4.31 Overall, the unit performance in staff learning is uneven across the eight units. The percentage of staff having such annual plans is low in six of the eight units reviewed. Most units increased staff learning days despite the fact that some units cut their learning budget in that year. This may suggest that the learning budget is perhaps not a sensitive measure of unit performance in relation to staff learning, especially with an increasing use of new learning modes, such as online, team learning, etc., which cost is relatively little compared to traditional modes of learning delivery.

4.32 The 2003 Bank Staff Survey results for six of the units show that even though the staff in these units are satisfied with the continuous learning opportunities that the Bank provides, they are not taking full advantage of those learning opportunities, and they feel that learning events do not provide adequate opportunity to learn new skills.

4.33 According to the unit managers' interviews, none of these units has a unit development plan geared to meeting the priority learning needs of the unit. Moreover, these units do not have a systematic approach to monitoring the completion of agreed ILPs. Further, managers feel that, currently, there is no incentive for the staff to take learning events seriously and give priority to learning. Learning records of individuals are not taken into account seriously when career promotion decisions are made. As a result, learning takes place as an add-on activity with the primary focus being on the performance record based on delivery of products and services.

4.34 The study identified and used eight indicators for measuring the units' operational performance with respect to operational portfolio quality and outcome. These indicators have been used by the unit to evaluate and account for performance in the monthly management reports to regional management. These indicators include: (1) projects at risk; (2) commitment at risk; (3) disbursement ratio; (4) realism ratio, and (5) proactivity ratio. In addition, the following three indicators were added for this study: (1) percentage of projects with financial management problems; (2) percentage of projects with financial performance problems; and (3) percentage of projects with procurement problems. These additional indicators help evaluate performance of the two support units in the sample of eight units selected for the evaluation of learning effects on unit performance.

4.35 In analyzing these operational indicators, the six operations units are divided into three groups: (a) the units that had increased ILPs, learning days per staff as well as learning expenditures per staff; (b) units that had increased ILPs and learning days per staff, but showed lower learning expenditures per staff; and (c) units that had increased

learning days but showed decline in ILPs and learning expenditures. Although the results are mixed, overall positive changes in learning parameters are not associated with parallel improvement in the unit operational performance. However, units that have a good learning record tend to have comparatively high proactivity and realism ratios. Learning may motivate staff to focus on finding solutions and move projects from the problem project category. In this context, it is worthwhile to consider that staff learning may have an effect on Proactivity with regard to problem projects and Realism ratios. However, there is no or unclear indication that staff learning has effect on the remaining indicators.

5. CONCLUSIONS AND IMPLICATIONS

5.1 This report presents the results of the second phase of the impact evaluation of staff learning programs in the East Asia and Pacific and the South Asia Regions. Building upon the results from the phase I evaluation, the main objective of the phase II evaluation is to estimate the extent to which the regional learning program has enhanced regional staff competence, performance and the Regional Operations. Several findings pertaining to the evaluation questions are derived from various analyses carried out in this study.

5.2 Both the Advanced Operations and the Foundations of Team Leadership courses have successfully achieved their objectives. By attending the learning, staff achieved more than a 30 percent increase in their knowledge and skills in operations and leadership. Apart from the knowledge gain, there is evidence that courses also helped staff to work more efficiently. Moreover, participants in the Team Leadership course reported increases of between 15 to 24 percent in some course content areas. These changes are clear indications of the course impact.

5.3 There is little evidence, however, to support the relationship between staff learning participation and staff performance as measured by OPE scores. This may mean that the OPE assessment is a poor proxy measure for staff performance.

5.4 The study has selected and applied three key staff learning indicators (ILPs, staff learning days and learning expenditure) and eight operational performance indicators to investigate unit performance with respect to organizational effectiveness and portfolio quality and work performance. Overall, there is little evidence of a positive relationship between learning investment and organizational performance. However, units that have a good record in learning tend to have comparatively high Proactivity and Realism Ratios. Learning may motivate staff to focus on finding solutions and moving projects out of the problem project category. In this context, in the short run, learning may have some impact in terms of helping the units to maintain a high Proactivity ratio and to improve Realism Ratio. For the remaining operational performance indicators, learning dosage does not seem to have an observable effect.

5.5 When interpreting these evaluation findings, it is important to keep in mind some limitations and constraints of the evaluation. First, we know that learning impact is likely to occur in the medium and long term, while the performance analyzed in this study covers a short period of only two years – FY02 and FY03. Learning impact may not become visible due to insufficient time elapsed of learning events; more specifically, many regional initiatives took place after FY02. Second, learning impact becomes clear only when the intensity of learning is strong enough to encourage change. Staff learning impact cannot reach maximum at unit level unless a critical mass of staff learning occurs at the unit level. As of this writing, less than 25 percent of professional EAP and SAR regional staff attended the Advanced Operations course and this course is considered to

be substantially important in terms of improving staff operations skills. A large proportion of regional staff has not yet been brought up to date in terms of knowledge and skills. Finally, as mentioned earlier, many of these unit operational performance indicators are affected by external factors. Without a more solid understanding of the dynamics of these indicators and more information about external influences, it is difficult to determine the learning impact on unit performance.

5.6 Thus, the study findings support the general notion that staff learning is important to help staff build knowledge and skills. However, the lack of relationship between learning and staff and organizational performance is not supportive of the view that training/learning is a sole or primary source of change among individuals and organizations, at least in the short run.

5.7 Some observations from this evaluation have implications for future regional staff learning and its evaluation.

5.8 **First**, as revealed in the phase I evaluation, staff stationed in Headquarters (HQ) received more learning than staff in country offices. With the increasing trend towards decentralization of the Bank's Operations through country offices, country office staff are experiencing practical constraints on adequate learning. These constraints have been addressed in EAP and SAR by introducing a Regional Learning Center in Bangkok, Thailand, to serve the learning needs of country office staff of both regions. Strategically, the Asia Learning Group has developed a set of intensive learning programs in response to corporate and regional priorities. Many additional efforts have been made to ensure that front line staff members have the knowledge and skills needed to carry out their duties. In addition to the Advanced Operations and Team Leadership courses, the regional staff learning program has designed and promoted the two other core courses: Development and Ethics and Client Engagement and Communications, which are tailored to regional business and staff needs, especially for professional staff. What are the likely impacts of these core courses and regional programs? Clearly, the results of the evaluation of the two courses indicate great promise for the success of the regional staff learning programs. There could be considerable impact if, as found for study participants, all regional professional staff and TTLs were to take these core training courses, achieve a 30 percent increase in their knowledge and skills, and develop an ability to use this knowledge more effectively. We would definitely expect to see positive changes in the Regional Operations brought on by well equipped and prepared staff.

5.9 **Second**, the analysis highlights the importance of the influence of the organizational culture on staff learning. At the unit level, in addition to developing a unit development plan, efforts need to be made to enable all staff members to have ILPs. In this regard, the process of OPE can be an effective instrument to ensure that staff have ILPs and take the core learning required to achieve business result. Subsequently, achievements against the ILPs need to be monitored at the time of the annual OPE process. There seems also to be a need for changing the incentive structure for career development so that it gives reasonable weight to adequate learning. The staff could be informed that the learning record of every staff member will be kept up to date and that it

will be taken into account in decisions on promotion. This approach could provide an incentive to the staff to take advantage of learning events and keep their learning records current.

5.10 **Third**, not only the unit staff but also managers need to be equipped with new and updated operations knowledge and skills. Strong leadership is the driving force for unit performance, as it helps to create an environment that encourages staff to utilize new knowledge and skills. The analysis of the impact of the Advanced Operations course presented in Chapter II shows clearly that a positive, supportive environment is crucial for knowledge application.

5.11 **Fourth**, the evaluation findings are significantly affected by the comprehensiveness and quality of the data employed. In particular, the time frame, accuracy of available data and validity of outcome measures for the analyses are not ideal. The lack of an established outcome instrument to measure staff performance is a clear weakness. The attempt to relate staff learning to unit performance has been even less successful, partly reflecting the limitations of the data to examine the process in which staff learning is transferred to business outcomes. Many of the operational performance indicators selected are likely to be influenced over the short term by country situations, management approach, and overall experience of staff than by learning programs. These are familiar problems that evaluators face in evaluating the impact of learning. It is recommended that a thorough exploration of the dynamics of outcome measures and exogenous factors should be part of future evaluation efforts.

5.12 **Fifth**, acknowledging that the impact of learning takes long time to become visible, the evaluation of learning impact should be a continuous effort. There is a need for continuous monitoring and analysis, including continuous refinement of indicators that assess the impact of learning on organizational performance. In the context of the current study, evaluating the impact of learning on the third dimension — client relationship — should be a critical item on the forthcoming evaluation agenda.

5.13 **Finally and most importantly**, this evaluation has served to reiterate that staff learning is not found simply in an outcome indicator of organizational effectiveness, learning is also an input for a unit business outcome. At the organizational level, the impact of learning may be expected only if a critical mass of staff learning occurs. In addition, the benefits of staff learning may require long lead times to be visible. These points reinforce the value of commitments to invest in staff learning, to focus on the right topics, reach and maintain a certain intensity level, and make learning a long-term strategy. With these commitments, organizations should expect that learning will eventually bring positive returns.

ANNEXES

ANNEX 1A: EVALUATION QUESTIONNAIRES – ADVANCED OPERATIONS COURSE

PRE-COURSE QUESTIONNAIRE FOR EAP-SAR ADVANCED OPERATIONS COURSE

This short questionnaire aims to obtain some information on your knowledge and skills in the topics included in the Advanced Operations course that you will be attending in November 2003. The information will be serving as a baseline that enables the WBI Evaluation Group to evaluate the effectiveness of the course. Your answers will be treated as confidential with respect to individuals and will be purely used for the evaluation purpose.

1) The Advanced Operations course is designed to enhance staff operational knowledge and skills relating to the following topics. To what extent is the content of each of the topics directly **APPLICABLE** to your work?

- | | Not at all 1 | Very much 5 | N/A |
|---|--------------|-------------|-----|
| a. Regional Products Driven by Corporate Priorities | | | |
| b. Quality of New Products | | | |
| c. Quality at Entry | | | |
| d. Quality of Supervision | | | |
| e. Results-based Project Management | | | |
| f. Quality at Completion | | | |
| g. Portfolio Management and New Style CPPR | | | |

2) What is your assessment of the **CURRENT LEVEL** of operational knowledge and skills in the following course topics?

- | | Low 1 | High 5 | N/A |
|---|-------|--------|-----|
| a. Regional Products driven by Corporate Priorities | | | |
| b. Quality of New Products | | | |
| c. Quality at Entry | | | |
| d. Quality of Supervision | | | |
| e. Results-based Project Management | | | |
| f. Quality at Completion | | | |
| g. Portfolio Management and New Style CPPR | | | |

3) To what degree do you **USE** the knowledge and skills related to the following topics in your daily work?

- | | Low 1 | High 5 | N/A |
|---|-------|--------|-----|
| a. Regional Products driven by Corporate Priorities | | | |
| b. Quality of New Products | | | |
| c. Quality at Entry | | | |
| d. Quality of Supervision | | | |
| e. Results-based Project Management | | | |
| f. Quality at Completion | | | |
| g. Portfolio Management and New Style CPPR | | | |

4) Have you attended any of the following operation related training courses in the last 12 months? (if yes, check all that apply, if no, go to the next Part)

Yes No

Introduction to Bank Operations
Safeguards

Financial Management
Procurement
Lending Instrument
Monitoring and Evaluation
Project Planning and Preparation
Project Supervision
Project Cycle Management
Trust Fund Learning Accreditation Program
Other (please specify) _____

Before submitting the questionnaire, please tell us about yourself.

5) In all, how long have you worked for the **EAP and SAR** regions (including time as consultant/temp)?

Number of years _____ and number of months _____

6) In all, how long have you worked in the World Bank's **Operations** (including time as consultant/temp)?

Number of years _____ and number of months _____

7) What is your grade level?

? GA-GD ? GE ? GG ? Other, specify: _____
? STC/STT ? GF ? GH or above

**** Many Thanks ****

We greatly appreciate your completing this questionnaire

FOLLOW-UP QUESTIONNAIRE FOR EAP-SAR ADVANCED OPERATIONS COURSE

Name _____

Date of Training _____

Today's Date _____

Please, either mark "x" or give the rating number when necessary in the brackets next to your answer!

1. Of the options below, what was your main reason for taking this course?

-]To enhance performance in current assignment
-]To network and share information
-]For professional interest and growth
-]Other, specify:

2. At the time of training, how would you rate the degree (i.e. 1 is low and 5 is high) to which:

Low 1 High 5 N/A

- (1) The training was relevant to your work []
- (2) The training met your learning needs []
- (3) You were actively involved in learning []
- (4) The training increased your knowledge and skills []

3. Since taking this course, have you attended any of other operation related training courses?

No [] Yes []

If yes, do you recall the name of the course?

- Introduction to Bank Operations []
- Safeguards []
- Financial Management []
- Procurement []
- Lending Instrument []
- Monitoring and Evaluation []
- Project Planning and Preparation []
- Project Supervision []
- Project Cycle Management []
- Trust Fund Learning Accreditation Program []
- Other (please specify) _____

4. Since taking the course, have you changed the type of work that you are performing? (Are you in the same regional unit?)

5. The Advanced Operations course is designed to cover 7 key operations topics. Could you please indicate to what extent (i.e. 1 is not at all and 5 is very much) the content of each of the topics below is directly **APPLICABLE** to your **current** work?

Not at all 1 Very much 5 N/A

- (1) Regional Products driven by Corporate Priorities []

- (2) Quality of New Products []
- (3) Quality at Entry []
- (4) Quality of Supervision []
- (5) Results-based Project Management []
- (6) Quality at Completion []
- (7) Portfolio Management and New Style CPPR []

6. What is your assessment of the **CURRENT LEVEL** (1 is LOW and 5 is HIGH) of operational **knowledge and skills** in the following course topics?

- | | Low 1 | High 5 | N/A |
|--|-------|--------|-----|
| (1) Regional Products driven by Corporate Priorities [] | | | |
| (2) Quality of New Products [] | | | |
| (3) Quality at Entry [] | | | |
| (4) Quality of Supervision [] | | | |
| (5) Results-based Project Management [] | | | |
| (6) Quality at Completion [] | | | |
| (7) Portfolio Management and New Style CPPR [] | | | |

7. Has the course changed your approach to your work? If yes, then in what ways? (Can you give us some examples?)

8. Since the training, have you had an opportunity to use the knowledge and skills you learned from the course?

- No [] => go to Question 10
- Yes []

If yes, could you describe to what degree (1 is low and 5 is high) you **USE** the knowledge and skills related to the following topics in your daily work?

- | | Low 1 | High 5 | N/A |
|--|-------|--------|-----|
| (1) Regional Products driven by Corporate Priorities [] | | | |
| (2) Quality of New Products [] | | | |
| (3) Quality at Entry [] | | | |
| (4) Quality of Supervision [] | | | |
| (5) Results-based Project Management [] | | | |
| (6) Quality at Completion [] | | | |

(7) Portfolio Management and New Style CPPR []

An example of the usage:

9. To what degree has your use of knowledge and skills had an impact on:

No impact at all 1 High impact 5 N/A

(1) The quality of your work []

(2) Work relationship with Bank staff []

(3) Work relationship with clients []

(4) An example of the impact:

10. To what degree did the following factors facilitate your use of the knowledge and skills you learned in the course?

Major barrier 1 Major facilitator 5 N/A

(1) Relevance of the course to your job []

(2) Motivation/self-confidence []

(3) Encouragement from manager/supervisor []

(4) Encouragement from peers/colleagues []

(5) Institutional/organizational policy []

(6) Tools and resources []

(7) Other, specify:_____

11. Do you have any recommendations for improving the course? (list up to three)

12. In all, how long have you worked for the EAP and SAR regions (including time as consultant/temp)?

Number of years____ and number of months____

13. In all, how long have you worked in the World Bank's Operations (including time as consultant/temp)?

Number of years____ and number of months____

14. What is your grade level?

GA-GD[] GE[] GG[] STC/STT [] GF []

GH or above [] Other, specify_____

15. Gender

Male [] Female []

This is the end of the interview and Thank You for your time

ANNEX 1B: EVALUATION QUESTIONNAIRES – FOUNDATIONS OF TEAM LEADERSHIP

PRE-COURSE QUESTIONNAIRE FOR EAP-SAR FOUNDATIONS OF TEAM LEADERSHIP

This short questionnaire aims to obtain some information on your knowledge and skills in the areas included in the Foundations of Team Leadership course that you will be attending in February 2004. The information will be serving as a baseline that enables the WBI Evaluation Group to evaluate the effectiveness of the course. Your answers will be treated as confidential with respect to individuals and will be purely used for the evaluation purpose.

The Foundations of Team Leadership course is designed to help East and South Asia Team Leaders develop managerial and leadership capacities. It focuses on three areas: Self-awareness, Team Dynamics, and Communication and Dialogue.

1) To what degree are the knowledge and/or skills covered in each of the areas **RELEVANT** to your work?

Not at all 1 Very much 5 N/A

Self-awareness

- a. Learn from multi-instrument feedback on preferences
- b. Thinking styles, and behaviors
- c. Engage in self-discovery and self-reflection
- d. Develop an action plan for self-improvement based on peer feedback

Team Dynamics

- a. Generate team synergy
- a. Balance people related factors such as listening, supporting, etc
- b. Balance task related factors like goal selection, alternatives, implementation, etc.
- c. Manage teams under conditions of stress

Communication and Dialogue

- a. Understand the impact of mental models on one's thinking and behavior
- b. Increase influence through better communication

2) What is your **CURRENT LEVEL** of knowledge and/or skills covered in each of the areas?

Not at all 1 Very much 5 N/A

Self-awareness

- a. Learn from multi-instrument feedback on preferences
- b. Thinking styles, and behaviors
- c. Engage in self-discovery and self-reflection
- d. Develop an action plan for self-improvement based on peer feedback

Team Dynamics

- a. Generate team synergy
- b. Balance people related factors such as listening, supporting, etc.
- c. Balance task related factors like goal selection, alternatives, implementation, etc.
- d. Manage teams under conditions of stress

Communication and Dialogue

- a. Understand the impact of mental models on one's thinking and behavior
- b. Increase influence through better communication

3) In your work, to what extent do you currently **USE** the knowledge and/or skills covered in each of the areas?

Not at all 1 Very much 5 N/A

Self-awareness

- a. Learn from multi-instrument feedback on preferences
- b. Thinking styles, and behaviors
- c. Engage in self-discovery and self-reflection

FOLLOW-UP QUESTIONNAIRE FOR EAP-SAR FOUNDATIONS OF TEAM LEADERSHIP

Participant Name _____
Date of Training _____
Today's Date _____

This short questionnaire aims to obtain some information on the knowledge and skills you gained during the course on Foundations of Team Leadership. The WBI Evaluation Group is conducting an impact evaluation of the course and this questionnaire is an integral part of the evaluation process. Your answers will be treated as confidential and used only for the evaluation purposes.

- Of the options below, what was your main reason for taking this course?
 To enhance performance in the current assignment
 To network and share information
 For professional interest and growth
 Other, specify: _____
- At the time of training, how would you rate the degree to which:

Low 1 High 5 N/A

 - The training was relevant to your work
 - The training met your learning needs
 - You were actively involved in learning
 - The training increased your knowledge and skills
- Since taking the course, have you attended the Advanced Operations courses?
Yes No
- Since taking the course, have you changed the type of work that you perform?
Yes No
- The course focuses on three areas: self-awareness, team dynamics, and communication and dialogue. To what degree are the knowledge and skills covered in the following areas relevant to your work?

Not at all 1 Highly relevant 5 N/A

Self-awareness

 - Learning from feedback on preferences
 - Learning from feedback on thinking styles and behaviors
 - Engaging in self-discovery and self-reflection
 - Developing an action plan for self-improvement based on peer feedback

Team dynamics

 - Generating team synergy
 - Balancing people related factors such as listening, supporting etc.
 - Balancing task related factors like goal selection, implementation etc.
 - Managing teams under condition of stress

Communication and dialogue

 - Understanding the impact of mental models on one's thinking and behavior
 - Increasing influence through better communication
- What is your assessment of your current level of knowledge and skills covered in each of the following areas?

Not knowledgeable 1 Very knowledgeable 5 N/A

Self-awareness

 - Learning from feedback on preferences

- d. Encouragement from peers/colleagues []
 - e. Institutional/organizational policy []
 - f. Tools & resources
 - g. Other, specify:_____
10. Do you have any recommendations for improving the course (list up to three)?
- a. _____
 - b. _____
 - c. _____
11. In total, how long have you worked for the EAP and SAR region (including time as a consultant/temp)?
- Number of years_____ and number of months_____
12. In total, how long have you worked in the World Bank's Operations (including time as a consultant/temp)?
- Number of years_____ and number of months_____
13. What is your grade level?
- GA-GD [] STC/STT [] GE [] GG []
- GF [] GH or above [] Other, specify:_____
14. Gender
- Male [] Female []

THIS IS THE END OF THE SURVEY. THANK YOU FOR YOUR TIME!

ANNEX 2: STATISTICAL ANALYSIS RESULTS – ADVANCED OPERATIONS

T-TESTS OF MEAN DIFFERENCE OF RELEVANCE, CURRENT KNOWLEDGE AND UTILIZATION BEFORE AND AFTER REGIONAL LEARNING

Effect	Variable	Group (N)	Mean	Difference	2-tailed T-stat	P-value
Relevance	Overall relevance (Average)	Controls (38)	4.34	-0.40	-3.56	0.000**
		Participants (38)	3.92			
	Regional products	Controls (37)	3.84	-0.39	-1.67	0.098*
		Participants (38)	3.45			
	Quality of new products	Controls (35)	4.31	-0.65	-2.90	0.005**
		Participants (38)	3.66			
	Quality at entry	Controls (37)	4.58	-0.31	-1.84	0.070*
		Participants (38)	4.27			
Quality of supervision	Controls (38)	4.82	-0.61	-2.75	0.008**	
	Participants (38)	4.21				
Res.-based proj. mgmt.	Controls (37)	4.57	-0.62	-3.42	0.001**	
	Participants (38)	3.95				
Quality at completion	Controls (37)	4.46	-0.22	-1.30	0.197	
	Participants (38)	4.24				
Port. mgmt./ new style CPRP	Controls (36)	3.75	-0.32	-1.35	0.182	
	Participants (38)	3.43				
Current Knowledge and Skill Level	Overall current skill level (Average)	Controls (38)	2.78	0.98	7.23	0.000**
		Participants (38)	3.76			
	Regional products	Controls (37)	2.49	0.83	3.98	0.000**
		Participants (38)	3.32			
	Quality of new products	Controls (37)	2.55	1.06	5.92	0.000**
		Participants (38)	3.61			
	Quality at entry	Controls (37)	3.03	1.10	5.82	0.000**
		Participants (38)	4.13			
Quality of supervision	Controls (38)	3.39	0.74	4.75	0.000*	
	Participants (38)	4.13				
Res.-based proj. mgmt.	Controls (37)	2.82	1.07	6.30	0.000**	
	Participants (38)	3.89				
Quality at completion	Controls (37)	2.86	0.98	5.62	0.000**	
	Participants (38)	3.84				
Port. mgmt./ new style CPRP	Controls (38)	2.49	0.93	4.77	0.000**	
	Participants (38)	3.42				
Current Use	Overall current use (Average)	Controls (38)	3.63	0.04	-0.22	0.829
		Participants (35)	3.67			
	Regional products	Controls (37)	2.95	0.11	0.44	0.664
		Participants (32)	3.06			
	Quality of new products	Controls (37)	3.59	-0.27	-1.01	0.276
		Participants (32)	3.32			
	Quality at entry	Controls (37)	4.03	0.17	0.80	0.425
		Participants (35)	4.20			
Quality of supervision	Controls (38)	4.21	0.10	0.46	0.647	
	Participants (35)	4.31				
Res.-based proj. mgmt.	Controls (37)	3.81	-0.05	-0.18	0.857	
	Participants (34)	3.76				
Quality at completion	Controls (38)	3.66	0.08	0.30	0.767	
	Participants (34)	3.74				
Port. mgmt./ new style CPRP	Controls (34)	2.97	0.23	0.87	0.388	
	Participants (35)	3.20				

**statistically significant at the 0.05 level.

*statistically significant at the 0.10 level.

ANNEX 3: STATISTICAL ANALYSIS RESULTS – FOUNDATIONS FOR TEAM LEADERSHIP

T-TESTS OF MEAN DIFFERENCE OF RELEVANCE, CURRENT KNOWLEDGE AND UTILIZATION BEFORE AND AFTER REGIONAL LEARNING

Effect	Variable	Group (N)	Mean	Difference	2-tailed T-stat	P-value
Relevance	Overall relevance (Average)	Controls (14)	4.42	-0.22	-1.23	0.223
		Participants (38)	4.20			
	Self-awareness (Average)	Controls (14)	4.23	-0.01	-0.04	0.969
		Participants (38)	4.22			
	Learn from multi-instrum. Feedback	Controls (14)	4.07	0.03	0.13	0.900
		Participants (38)	4.11			
	Thinking styles, behaviors	Controls (14)	4.50	-0.21	-0.88	0.383
		Participants (38)	4.29			
	Self-discovery and reflection	Controls (12)	4.08	0.42	1.76	0.084*
		Participants (38)	4.50			
	Action plan for self-improvement	Controls (14)	4.29	-0.29	-0.96	0.339
		Participants (38)	4.00			
	Team dynamics (Average)	Controls (14)	4.54	-0.35	-1.83	0.073*
		Participants (38)	4.18			
Generate team synergy	Controls (14)	4.64	-0.40	-1.62	0.111	
	Participants (38)	4.24				
Balance people related factors	Controls (14)	4.43	-0.07	-0.29	0.774	
	Participants (38)	4.37				
Balance task related factors	Controls (14)	4.57	-0.46	-2.03	0.048**	
	Participants (38)	4.11				
Manage teams under stress	Controls (14)	4.50	-0.47	-2.01	0.050*	
	Participants (38)	4.03				
Communication and Dialogue (Average)	Controls (14)	4.54	-0.34	-1.69	0.097*	
	Participants (38)	4.20				
Understand impact of mental models	Controls (14)	4.46	-0.22	-0.86	0.392	
	Participants (38)	4.24				
Increase influence through comm.	Controls (14)	4.57	-0.41	-1.75	0.086*	
	Participants (38)	4.16				
Current Knowledge and Skill Level	Overall current skill level (Average)	Controls (14)	3.12	0.56	3.89	0.000**
		Participants (38)	3.68			
	Self-awareness (Average)	Controls (14)	3.23	0.33	1.86	0.068*
		Participants (38)	3.57			
	Learn from multi-instrum. Feedback	Controls (14)	3.00	0.63	2.76	0.008**
		Participants (38)	3.63			
	Thinking styles, behaviors	Controls (14)	3.21	0.40	1.83	0.074*
		Participants (38)	3.61			
Self-discovery and reflection	Controls (13)	3.38	0.38	1.46	0.150	
	Participants (38)	3.76				
Action plan for self-improvement	Controls (12)	3.33	-0.07	-0.24	0.813	
	Participants (38)	3.26				
Team dynamics (Average)	Controls (14)	3.13	0.61	3.34	0.002**	
	Participants (38)	3.73				
Generate team synergy	Controls (14)	3.00	0.84	3.31	0.002**	
	Participants (38)	3.84				

(Annex 3 continues on next page.)

(Annex 3 continued.)

Effect	Variable	Group (N)	Mean	Difference	2-tailed T-stat	P-value
	Generate team synergy	Controls (14)	3.00	0.84	3.31	0.002**
		Participants (38)	3.84			
	Balance people related factors	Controls (14)	3.21	0.58	2.86	0.006**
		Participants (38)	3.79			
	Balance task related factors	Controls (14)	3.07	0.67	2.93	0.005**
		Participants (38)	3.74			
	Manage teams under stress	Controls (13)	3.23	0.35	1.27	0.209
		Participants (38)	3.58			
	Communication and dialogue (Average)	Controls (14)	2.82	0.98	5.02	0.000**
		Participants (38)	3.80			
	Understand impact of mental models	Controls (13)	2.62	1.17	4.88	0.000**
		Participants (38)	3.79			
	Increase influence through comm.	Controls (13)	3.15	0.67	3.12	0.003**
		Participants (38)	3.82			
Current Use	Overall current use (Average)	Controls (14)	3.31	0.40	2.41	0.020**
		Participants (37)	3.71			
	Self-awareness (Average)	Controls (14)	3.34	0.12	0.59	0.555
		Participants (37)	3.46			
	Learn from multi-instrum. feedback	Controls (14)	3.14	0.27	0.92	0.363
		Participants (37)	3.41			
	Thinking styles, behaviors	Controls (13)	3.38	0.20	0.75	0.453
		Participants (36)	3.58			
	Self-discovery and reflection	Controls (13)	3.46	0.15	0.50	0.618
		Participants (36)	3.61			
	Action plan for self-improvement	Controls (14)	3.43	-0.24	-0.77	0.447
		Participants (36)	3.19			
	Team dynamics (Average)	Controls (14)	3.36	0.51	2.41	0.020**
		Participants (37)	3.87			
	Generate team synergy	Controls (14)	3.29	0.63	2.01	0.050*
	Participants (36)	3.92				
Balance people related factors	Controls (14)	3.57	0.59	2.66	0.010**	
	Participants (37)	4.16				
Balance task related factors	Controls (14)	3.36	0.34	1.54	0.129	
	Participants (37)	3.70				
Manage teams under stress	Controls (13)	3.15	0.55	1.75	0.087*	
	Participants (37)	3.70				
Communication and dialogue (Average)	Controls (13)	3.15	0.75	3.36	0.002**	
	Participants (37)	3.91				
Understand impact of mental models	Controls (12)	3.00	0.84	2.92	0.005**	
	Participants (37)	3.84				
Increase influence through comm.	Controls (13)	3.38	0.59	2.48	0.017**	
	Participants (37)	3.97				

**statistically significant at the 0.05 level.

*statistically significant at the 0.10 level.

ANNEX 4: ANALYSIS OF THE LEARNING IMPACT ON OPE

METHODOLOGY

OPE ratings

For the purpose of this study, OPE Results and Behavior rating scores are converted to a set of numerical values from 1–5, which are itemized as follows: Unsuccessful (1), Partially successful (2), Fully successful (3), Superior (4) and Outstanding/best practices (5). OPE Results and Behavior scores for a staff member are each determined by the sum of the five individual item scores. The higher the score, the better the performance. To equalize the possible differences in the number of result rating items for different years, an average score is calculated for an individual for each year. For example, in FY99, if a staff member had 5 results areas with scores of 4, 3, 5, 3 and 4, his/her raw total results rating score would be 19 (4+3+5+3+4). If, in FY02, he/she had only 3 Results areas but each scored a 5, the raw total score would be 15 (5+5+5). Looking only at the raw total score, his/her performance would appear to have declined. However, if the scores are averaged for each year, it is clear that, controlling for the number of scores per year, the performance in fact improved in 2003, 5 (15/3) being greater than 3.8 (19/5).

OLS and Ordered Probit Models¹⁹

We employ simple regression analysis techniques as one way to evaluate the relationship between training and impact on OPE scores. The analysis is divided into two stages. In the first stage two dimensions of learning participation are regressed on a bundle of exogenous demographic characteristics. OLS estimation is used to predict Intensity. However, because Regularity is a discrete ordered measure ranging from 1–4 we employ an Ordered Probit Model to evaluate what determines Regularity of learning.

In the second stage we evaluate the extent to which learning participation, both in terms of Intensity and Regularity, impact OPE scores, controlling for demographics. OPE scores range from 1–5; however we are interested not in each score separately but in the average of the five results agreement scores, as well as the average of the four behavioral assessment scores. Averaging the components in each category results in continuous composite variables and OLS can be used for the estimation.

Propensity Score Matching

Because of the self-selection problem in learning participation, the most suitable model for this is estimation is Instrumental Variables estimation (IV)²⁰ where Intensity and

¹⁹ Ordered probit regression can be used when the dependent variable is of ordinal nature. Ordered data is a special case of discrete choice data where the outcome is ranked. The measure of training regularity is an ordered index from 0–4 where 0 means no attendance at all during the four years, and 4 means that the staff attended activities each of the four years. The OLS, multinomial logit and multinomial probit models fail to account for ordinal data. One may regard the qualitative responses to the question on how regularly the staff attended training as an ordinal categorical measure of underlying continuous variables. Ordered probit is similar to ordered logit, but uses a different distribution function for calculating probabilities.

Regularity are instrumented by one or several exogenous variables that effect learning participation but do not have any direct effects on OPE scores. Given the data available, no such instruments could be defined. In this case, we applied propensity score matching: comparing the OPE scores of staff with similar probabilities of regional learning participation over the given period of time.

Because learning participation is a continuous variable (number of learning activities or days) and not a dichotomous variable (participated, did not participate) we generate two groups of learners: High-frequency participants (staff in the 4th quartile of the distribution of number of learning activities) and low-frequency or non-participants (staff in the 1st quartile of the distribution of number of learning activities). By employing this technique, we use a dummy variable (1 = high-frequency participation, 0 = low-frequency/non-participation) to measure those factors that determine high-low frequency participation and to establish each person's predicted probability to be a high-frequency participant. For reasons of simplicity, this analysis is only applied to the effect on job performance of one of the two measures of intensity—number of learning activities—and does not employ number of learning days or learning regularity.

RESULTS

OLS and Ordered Probit regressions results

Intensity and Regularity

The results from the first stage estimation, which covers FY99–02, are summarized in Tables A1–A3. A variable description and summary statistics of the variables used in the analysis are presented in Table A4. Examining the results of the regressions on number of learning days over the three different time periods (FY01-02, FY00-02, FY99-02) the main determinants of learning days that stand over the periods are gender, number of years in the bank, grade level in FY02 and station country. Female staff attend learning more days than males; the longer the staff remains in the Bank the lower the frequency of learning days; higher level staff attend more often than lower level staff and being stationed in the U.S. (as compared to in a country office) increases the number of learning days.

The results are consistent also for number of learning activities attended over the three time periods as presented in Table A2.²¹ Women attend more activities than men; new staff tend to attend more activities; higher level staff participate in learning activities more frequently than lower level staff and staff stationed in the U.S. attend a greater number of activities than staff stationed elsewhere.

²⁰ Instrumental variable estimators (IV) are obtained by first regressing Y on X and V and subsequently using the predicted value \hat{Y} , obtained from the first regression, regressing Z on \hat{Y} and X. X are variables, such as demographics, affecting both Y and Z and V are the instruments used to predict Y, though they do not directly affect Z.

²¹ Estimation over FY00–02 and FY01–02 separately were also performed and generated similar results.

Similarly, the same variables that determine Intensity of learning also are the most important determinants of Regularity (Table A3). Education level does not impact either Intensity or Regularity; however, some of the effects of education on participation may be accounted for by grade level, since higher-educated staff tend to occupy positions in high grade levels.²²

Impact on OPE

The simple correlation between staff learning and OPE scores is very low. Of all the possible combinations of learning types (Intensity - days, activities, and Regularity), learning time periods (FY99–02, FY00–02 and FY01–02) and OPE scores (Behavioral - BA and Results - RA) nowhere is the direct relationship between learning and OPE scores higher than 0.1. The correlation does not however take into account other observable and unobservable factors that can possibly have effects on scores. Hence, regression analysis controlling for some (observable and available) of these factors will help explain the conditional effect of learning on OPE scores.

Selected results from the second stage regression analysis of learning effects on OPE behavioral assessment scores are presented in Table A5 and the effects on OPE results assessment scores are presented in Table A6.²³ We choose to present only the results from the analysis over the period FY99–02 because the extended period is the most representative of long term learning on results and behavioral changes.²⁴

It appears that, among the demographic characteristics of the staff, having been at the Bank for many years and having a high grade level decreases overall behavioral assessment OPE scores. The coefficients on these variables are, interchangeably, negative and significant in the regressions on behavioral scores. Furthermore, in contrast to our initial hypothesis that learning improves results and behavior, there is no evidence at all of significant effects of learning Intensity or Regularity on overall behavioral OPE scores. More importantly, these results cannot be given any importance because of the extremely low explanatory power of the regressions ($R^2 = 0.024-0.028$). With the variables that we employ we cannot explain the variation in OPE scores; 97 percent of the variation in scores is left unexplained or can be explained only by unobservable (to the authors) factors.

The same general findings hold true for the effect of learning Intensity and Regularity on overall results assessment OPE scores. Staff with higher education (such as Masters or Doctorate degrees) perform significantly lower on the results assessment part of the OPE scores than staff with lower degrees. Also, staff stationed at Headquarters (HQ) in Washington, D.C., score lower than staff in country offices. Nevertheless, similar to the findings of the behavioral scores, there is no effect of learning Intensity and Regularity on results assessment OPE scores. However, again and more importantly, the explanatory power of these regressions are much too low ($R^2 = 0.015-0.027$) to draw any

²² The correlation between Education and Grade FY02 is 0.55.

²³ Although not presented, activities, days and regularity were also separately included in the regression.

²⁴ The results of the analysis covering FY00–02 and FY01–02 are very similar to those presented in Tables A5–A6.

conclusions on what determines OPE scores, other than that participating in learning over the years preceding the OPEs has no significant effect on results assessment scores.

In addition to the simple correlations and regressions of the effects (direct and conditional) of learning on OPE scores, we study the possibility of effects of interaction of learning together with other variables on staff results and behavior. Although not presented in detail, the study finds no evidence of interaction effects of learning and demographic variables on OPE scores. Consistent with previous results, the explanatory power of these results was much too low to form any kind of conclusions.

The results of the analysis of mean difference in OPE scores have to be interpreted with the quality of the OPE scores in mind. There are a number of limitations in the OPE scores, all of which may contribute to the poor results of the analysis.

- First, there is almost no variation in OPE scores. According to the summary statistics in Table A4, no staff member in our sample had an average score of less than 2.25 on a scale from 1 to 5. Moreover, the variance is very small and, as a matter of fact, 92 percent of the sample have an average score between 3 and 4 on both assessments, making quantitative analysis of any kind very unlikely to detect differences between any sub-groups.
- Secondly, OPE scores may be influenced by the managers' personal opinions of their staff and may not necessarily reflect performance or behavior changes due to learning participation. Often there is a sizeable degree of separation between staff and managers, such that managers may not be aware of learning participation and even less of any results or behavior changes due to such learning.

Propensity Score Matching

In the first stage logit regression on high-frequency participation, we include the demographics and the Bank information as determinants of frequent participation. For a more in-depth description of the variables used in the logit regression, see Table A4. In order to generate robust results, we applied three different matching techniques developed by Becker and Ichino (2002):

- Nearest neighbor matching with replacement and equal weight of forward and backward matches.
- Nearest neighbor matching with replacement and random draw of forward and backwards matches.
- Radius matching.

The first technique performs nearest neighbor matching and allows matching both forwards and backwards. Then equal weight is given to the groups of forward and backward matches. In contrast, the second technique randomly draws either the forward or backward matches. In the third technique, radius matching, each treated unit is matched only with the control units whose propensity score falls in a predefined

neighborhood of the propensity score of the treated unit. Becker and Ichino²⁵ point out that the three approaches each reach different tradeoffs between the quality and quantity of the matches and that none of the approaches is a priori superior to the others.

Logit regression results

An example of results of the logit regression outcomes is presented in Table A7. In addition to the variables used when seeking the determinants of participation, we add pre-learning OPE scores to the logit regression in frequent participation. We hypothesize that a person's pre-learning score may influence the level of learning that the person wants to engage in. Almost identical results to those in Table A7 were obtained when we include the pre-learning behavioral assessment OPE scores instead of the two pre-learning results agreement OPE measures. Pre-learning results agreement is used in the logit regression when it is used as a basis for matching for subsequent testing of the difference in post-learning results agreement OPE scores for high-frequency participants and low-frequency/non-participants. Behavioral OPE scores replace result scores in the logit regression as a basis for matching for subsequent testing of the difference in behavioral OPE scores.

Consistent with the results from the OLS estimation, as displayed in Table A7, the probability of frequent participation decreases if the individual is male ($\beta = -0.718$, $z = -2.71$) and has spent many years in the Bank ($\beta = -0.162$, $z = -7.27$). However, the probability of frequent participation is higher for individuals with a higher FY02 Grade level ($\beta = 0.189$, $z = 2.68$) and for staff that are stationed in the U.S. ($\beta = 0.909$, $z = 3.25$).

Matching and results

The probabilities of being a high-frequency learning participant predicted by the first stage logit regression were then used to match low-frequency/non-participants to the frequent participants.²⁶ The results of the three different matching methods are discussed in the following section.

Nearest neighbor matching

The first two techniques yield almost exactly identical results.²⁷ For the tests of mean difference of the behavioral agreement scores and based on the logit regression presented in Table A7, a total of 313 high-frequency participants were matched with 246 low-frequency/non-participants (control group) when we do not take into account common support. For the results agreement matching and t-test, 317 high-frequency participants were matched with 245 control units. Matches of the two techniques and the means of the matched groups are presented in Tables A8 and A9.

²⁵ Becker, Sascha. and A. Ichino. (2002). "Estimation of Average Treatment Effect Based on Propensity Scores." *The Stata Journal*. 2(4): 358-377.

²⁶ The balancing property of the matches was satisfied and the number of blocks (five) ensures that the mean propensity score was not different for treated and control in each block.

²⁷ We discuss only the results of training over the two most current years, FY01-02. The exact same process and tests were applied to test the difference in OPE scores over longer periods of high-frequency training, FY99-02 and FY00-02, both yielding similar, non-significant differences in work performance.

After performing the matching, we performed a t-test of the mean difference in OPE scores between the matched high-frequency participants and control group. The column in Tables A8 and A9 labeled “Difference” reports what is referred to as the average effect of treatment on the treated, or the mean OPE of the high-frequency participants minus the mean OPE of the matched control group. A negative difference indicates that the low-frequency/non-participants had higher post-learning OPE scores than the high-frequency participants and vice versa. The columns labeled “t” report the t-statistic of the t-test performed.

It is clear from the results reported in Tables A8 and A9 that there is no significant difference in the means of any of the OPE measures between the high-frequency participants of regional staff learning and those who never participate or participate very rarely. T-statistics range in absolute value from 0.198 to 0.403, which is below significance at the five percent significance level.²⁸ Interestingly, the difference in both behavioral and results assessment OPE scores between staff who often take part in learning and staff who do not is negative, meaning that staff with little or no learning participation in FY01-02 actually performed and behaved better than high-frequency learning participants, according to their managers. Nevertheless, the difference is minuscule and therefore highly insignificant.

Accounting for common support means that we only match high-frequency participants and control units whose predicted probability of being in the high-frequency participation group are located on a common range. Although not reported, when applying the common support condition, the number of staff in the sample was reduced. The quantity of those matched is reduced and the quality of the matches may be improved; however, the insignificant mean difference remains.

Radius matching

In addition, the radius matching technique was applied with two different sizes of the radius: 0.1 and 0.01. No common support restriction was added. Similar to the nearest neighbor matching approaches, the radius matching finds no significant differences in the mean of OPE scores between the matched high-frequency participants and the low-frequency/non-participants in the regional staff learning. None of the mean differences are significant at the five percent significance level.

²⁸ At these sample sizes, significance at the 5% level requires a t-statistic of at least 1.96 in absolute value.

Table A1: Outcome of OLS regression on intensity (days) FY99-02

Variable	Coef.	Std. Err.	t	P> t
Male	-3.694**	0.800	-4.62	0.000
Age	0.035	0.054	0.65	0.519
Years Bank	-0.273**	0.058	-4.71	0.000
Grade FY02	1.221**	0.174	7.00	0.000
Education	0.658*	0.400	1.65	0.100
Station	3.977**	0.826	4.81	0.000
Constant	7.596**	2.224	3.42	0.001
R ²	0.160			
N	681			

** statistically significant at the 5% significance level.

* statistically significant at the 10% significance level.

Table A2: Outcome of OLS regression on intensity (activities) FY99-02

Variable	Coef.	Std. Err.	t	P> t
Male	-2.068**	0.485	-4.27	0.000
Age	0.024	0.033	0.73	0.468
Years Bank	-0.150**	0.036	-4.23	0.000
Grade FY02	0.534**	0.107	4.97	0.000
Education	0.204	0.246	0.83	0.406
Station	7.350**	0.502	14.65	0.000
Constant	3.318**	1.349	2.46	0.014
R ²	0.316			
N	746			

** statistically significant at the 5% significance level.

Table A3: Outcome of Ordered Probit regression on regularity FY99-02

Variable	Coef.	Std. Err.	z	P> z
Male	-0.577**	0.088	-6.59	0.000
Age	0.010*	0.006	1.81	0.070
Years Bank	-0.030**	0.006	-4.61	0.000
Grade FY02	0.193**	0.020	9.62	0.000
Education	0.006	0.043	0.14	0.886
Station	0.888**	0.091	9.76	0.000
Pseudo-R ²	0.121			
N	702			

** statistically significant at the 5% significance level.

* statistically significant at the 10% significance level.

Table A4: Variable description and summary statistics

Variable	Description	N	Mean	Std. Dev.	Min	Max
<u>Impact variables</u>						
RA FY03	Average results assessment OPE score FY03	1276	3.672	0.311	2.25	5
BA FY03	Average behavior assessment OPE score FY03	1266	3.651	0.386	2.25	5
<u>Intensity</u>						
Days FY01-02	Number of learning days FY01-02	1192	5.046	5.202	0	20
Days FY00-02	Number of learning days FY00-02	1174	9.180	7.918	0	29.90
Days FY99-02	Number of learning days FY99-02	1155	13.810	10.611	0	40
Activities FY01-02	Number of learning activities FY01-02	1265	3.357	3.772	0	20
Activities FY00-02	Number of learning activities FY00-02	1271	5.685	5.709	0	30
Activities FY99-02	Number of learning activities FY99-02	1272	8.005	7.601	0	40
Quart 4 Act. FY01-02	Dummy if individual is in 4 th quartile of activities FY01-02	633	0.502	0.500	0	1
Quart 4 Act. FY00-02	Dummy if individual is in 4 th quartile of activities FY00-02	589	0.499	0.500	0	1
Quart 4 Act. FY99-02	Dummy if individual is in 4 th quartile of activities FY99-02	576	0.502	0.500	0	1
<u>Regularity</u>						
Regularity	Number of years of learning during FY99-02	1278	2.411	1.378	0	4
<u>Demographics</u>						
Male	Dummy if the individual is male	1278	0.527	0.499	0	1
Age	Age if the individual	1278	44.843	9.074	23.70	66.50
Years Bank	Number of years the individual has been in the Bank	1278	10.683	7.558	0.10	36.57
Grade FY02	Grade level in FY02 (1=GA, lowest, 10=GJ, highest)	1219	5.016	2.320	1	10
Education	Level of education (1=diploma/certificate, 4=Ph.D.)	786	2.810	0.930	1	4
Station	Dummy if the individual is stationed in the U.S.	1278	0.416	0.493	0	1
RA FY99	Average results assessment OPE score FY99	857	3.549	0.503	1.75	5
BA FY99	Average behavior assessment OPE score FY99	860	3.666	0.535	2	5

Table A5: Outcome of OLS regression of intensity (days, activities) and regularity on Behavioral OPE scores FY99-02

Variable	Coef.	Std. Err.	t	P> t
Days FY99-02	-0.000	0.003	0.00	0.996
Activities FY99-02	-0.005	0.004	-1.11	0.267
Regularity FY99-02	0.003	0.020	0.15	0.878
Male	-0.025	0.037	-0.68	0.497
Age	0.002	0.002	0.93	0.353
Years Bank	-0.005*	0.003	-1.68	0.094
Grade FY02	-0.013	0.009	-1.36	0.174
Education	0.006	0.017	0.38	0.707
Station	-0.020	0.042	-0.47	0.638
BA FY99	0.073**	0.033	2.19	0.029
Constant	3.447**	0.159	21.75	0.000
R ²	0.028			
N	541			

** statistically significant at the 5% significance level.

* statistically significant at the 10% significance level.

Table A6: Outcome of OLS regression of intensity (days, activities) and regularity on Results OPE scores FY99-02

Variable	Coef.	Std. Err.	t	P> t
Days FY99-02	-0.001	0.002	-0.27	0.788
Activities FY9902	0.002	0.003	0.45	0.652
Regularity FY99-02	0.001	0.016	0.06	0.955
Male	0.033	0.030	1.11	0.266
Age	0.002	0.002	0.92	0.359
Years Bank	-0.001	0.002	-0.66	0.510
Grade FY02	-0.004	0.007	-0.60	0.551
Education	-0.021	0.014	-1.53	0.126
Station	-0.040	0.033	-1.20	0.232
RA FY99	0.045	0.027	1.64	0.101
Constant	3.531**	0.131	26.97	0.000
R ²	0.027			
N	615			

** statistically significant at the 5% significance level.

Table A7: Outcome of Logit regression on high-frequency intensity (days) FY01-02

Variable	Coef.	Std. Err.	z	P> z
Male	-0.718**	0.265	-2.71	0.007
Age	0.024	0.017	1.41	0.158
Years Bank	-0.162**	0.022	-7.27	0.000
Grade FY02	0.189**	0.071	2.68	0.007
Education	0.207	0.130	1.59	0.112
Station	0.909**	0.279	3.25	0.001
BA FY01	-0.123	0.239	-0.51	0.607
Constant	0.498	1.055	0.47	0.637
Pseudo-R ²	0.156			
N	444			

** statistically significant at the 5% significance level.

Table A8: Nearest neighbor matching with replacement and equal weight of forward and backward matches (activities) FY01-02

N Participants	N Control	Difference	Std. Error	Two-tailed t	P> t
Overall BA					
313 (3.629) ^a	246 (3.639)	-0.010	0.051	-0.198	0.843
Overall RA					
317 (3.646)	245(3.622)	-0.016	0.041	-0.386	0.699

^a Mean values of the OPE scores of the respective group are presented in parenthesis.

Table A9: Nearest neighbor matching with replacement and random draw of forward and backward matches (activities) FY01-02

N Participants	N Control	Difference	Std. Error	Two-tailed t	P> t
Overall BA					
313 (3.627) ^a	246 (3.639)	-0.012	0.055	-0.216	0.829
Overall RA					
317 (3.644)	245 (3.662)	-0.018	0.044	-0.403	0.687

^a Mean values of the OPE scores of the respective group are presented in parenthesis.

ANNEX 5: MANAGER INTERVIEW PROCEDURES, QUESTIONS AND RESULTS

Interview procedures

An email invitation was sent to the managers to schedule a face-to-face interview with the consultant. During the interview, managers were briefed on the main findings of the Phase I study and on the main objectives of the current Phase II study. Following this, the three dimensions developed to evaluate the impact of learning on unit performance were discussed: (i) capacity enhancement for organizational effectiveness; (ii) productivity, portfolio quality and outcome; and (iii) client relationship. Each interview lasted 30 to 40 minutes.

Interview questions

A. An overall picture of unit staff learning.

- (1) Does the Unit have an individual learning plan (ILP) for all staff?
- (2) Does the Unit monitor completion of ILPs?
- (3) Does the Unit have a Unit Learning and Development Plan?
- (4) Does the Unit provide incentives to encourage staff to get adequate learning?
- (5) Is the staff learning record taken into account at the time of career promotion decisions?
- (6) Does the Unit face problems in getting an adequate learning budget?
- (7) Is the Unit following any indicators to monitor the progress of learning?

B. Use of operational indicators to measure the unit's performance in respect with portfolio quality and outcome.

- (1) What indicators are followed to evaluate productivity, portfolio quality and outcome?
- (2) Is the Unit giving priority to impacting client relationships through staff learning?
- (3) How does the Unit view its overall performance in training staff and realizing its impact?

Managers' responses

Response to question (1): Most managers said that the number of staff having ILPs is low. Only two managers said that their units have a high number of staff with ILPs. (Data from the Human Resources Department show that more than 70 percent of their staff had ILPs during FY02-FY03. Other Units had significantly lower percentages.)

Response to question (2): None of the Units have a systematic approach to monitoring the completion of agreed ILPs. One manager mentioned that ILPs could change in the

course of the year as courses come become available and individual staff members volunteer to take those courses. Therefore, the courses actually taken may vary from what is planned for in the ILP. Another manager felt that a computerized system could be developed to monitor planned courses in ILPs vs. the actual courses taken, as well as the duration of the courses taken, but the Unit faced budget constraints in developing such a computerized monitoring system for ILPs.

Response to question (3): None of the units have a Unit Learning and Development Plan. One manager said that he wants to develop one for his unit to intensify staff learning efforts.

Response to question (4): Currently, no Unit is giving incentives to staff to undergo learning.

Response to question (5): The learning record is not seriously taken into account by Sector Boards and Networks in making promotional decisions. Promotional decisions are currently based primarily on performance record.

Response to question (6): One manager felt that there is a constraint in getting enough learning budget. Another manager said that he is not facing much difficulty in getting additional funding for learning, if necessary. Other managers felt that budget is not a constraint on learning, except for when it involves foreign study tours, which have travel and subsistence costs not covered by the learning budget.

Response to question (7): None of the units are currently using specific indicators to monitor closely the progress of learning. When suggested, they felt that indicators such as the number of staff having ILPs, the average number of learning days per staff, and the average learning investment per staff could be used.

Response to question (8): Managers mentioned that they use the commonly used Bank indicators to evaluate performance with respect to delivery, portfolio quality and outcome, and that these indicators are used in their monthly management reports to the Regional management. These indicators include: delivery, volume of lending, projects at risk, commitment at risk, disbursement ratio, realism ratio, and proactivity ratio. (The ratios are included in the generic indicators developed to evaluate the impact of learning on unit performance).

Response to question (9): Managers are giving priority to the role of learning in improving client relationships. In this context, they agree that learning focusing on behavioral skills and client governance improvement are useful. They also agree that learning has an important role in deepening civil society involvement and promoting a participatory approach to project/program development and implementation.

Response to question (10): Some managers mentioned that work pressure on staff prevents them from getting a satisfactory level of learning even though staff learning is encouraged. However, three managers mentioned that they are doing well in ensuring a high level of learning despite the usual work pressure on staff.

ANNEX 6: UNIT PERFORMANCE FRAMEWORK AND INDICATORS

Demission	Desired outcomes	Performance indicators
Organizational effectiveness (Focusing on learning)	<p>Align learning events to achieve better business results and strengthen learning of newcomers in Bank procedures and evaluation methodology, and provide advanced learning to senior staff</p> <p>Every staff to spend adequate time on learning</p> <p>Ensure adequate investment in learning</p>	<p>Percentage of unit staff having Individual Learning Plans (ILPs)</p> <p>Number of annual learning days per staff</p> <p>Amount of annual investment per unit Staff.</p>
Portfolio quality and outcome	<p>Meet delivery commitments in a timely manner</p> <p>Improve financial management, financial performance and procurement management</p> <p>Reduce riskiness of projects</p>	<p>Number of projects delivered to board. Actual lending volume</p> <p>Percent of projects under supervision facing financial management, financial performance and procurement flags</p> <p>Projects at risk (%)</p> <p>Commitment at Risk (%)</p> <p>Disbursement Ratio (%)</p> <p>Realism Ratio (%)</p> <p>Proactivity Ratio (%)</p>
Client relations	<p>Reduce time required to respond to clients in communications by following a rapid response system</p> <p>Ensure the client follows a result-based approach to project implementation and development effectiveness</p> <p>Convince the client to use effective communications with the public who have a stake in the development agenda and have a right to be heard</p> <p>Achieve improvement in governance.</p> <p>Achieve civil society participation in Bank activities</p>	<p>Effective use of information technology</p> <p>Success in working with clients to build country demand for, and capacity to adopt, result-based management</p> <p>Giving particular attention to the development of a satisfactory M&E system by the client to monitor and evaluate project/program implementation and outcomes</p> <p>Using power of persuasion and the technique of consensus -building to achieve this objective</p> <p>Success in making public institutions work closely with civil society to enhance transparency</p> <p>Following a participatory approach for getting the participation of civil society in project preparation and implementation as well as in the preparation of CAS, PRSP, ESW, etc</p>

ANNEX 7: STAFF LEARNING INFORMATION HIGHLIGHTS

	WB	EAP	EAP2	EAP3	EAP4	EAP1	SAR	SAR3	SAR4	SAR2	SAR1
<u>FY02</u>											
Professional Staff (GE+)	5415	395	39	49	28	14	328	49	40	57	12
ACS Staff	<u>1929</u>	<u>184</u>	<u>11</u>	<u>11</u>	<u>6</u>	<u>6</u>	<u>155</u>	<u>23</u>	<u>17</u>	<u>23</u>	<u>6</u>
Total Staff (IBRD, MIGA, GEF)	7344	579	50	60	34	20	483	72	57	80	18
Staff Having ILP	1721	123	6	43	5	NA	116	51	3	4	2
% of Total Staff	23.4	21.2	12.0	71.7	14.7	NA	24.0	70.8	5.3	5.0	11.1
Learning Days/Prof. Staff	NA	NA	4.7	3.1	4.0	4.9	NA	4.1	7.6	5.3	4.5
Learning Days/ACS Staff	NA	NA	0.8	3.8	3.4	5.7	NA	8.3	4.3	2.7	4.7
Ave. Learning Days/Staff	4.7	NA	3.9	3.3	3.8	5.1	NA	5.3	6.6	4.5	4.6
Total Learning Cost(\$'000)	NA	2932.0	437.8	379.9	231.6	93.7	2440.3	309.0	471.9	431.9	117.9
Learning Cost/Staff (\$)	NA	5063.9	8756.0	6331.7	6811.8	4685.0	5052.4	4291.7	8278.9	5398.8	6550.0
<u>FY03</u>											
Professional Staff (GE+)	5891	427	43	52	30	47	363	51	44	59	13
ACS Staff	<u>1933</u>	<u>186</u>	<u>8</u>	<u>11</u>	<u>6</u>	<u>5</u>	<u>155</u>	<u>19</u>	<u>15</u>	<u>23</u>	<u>4</u>
Total Staff (IBRD, MIGA, GEF)	7824	613	51	63	36	52	518	70	59	82	17
Staff Having ILP	1616	157	2	29	26	22	212	61	9	9	0
% of Total Staff	20.7	25.6	3.9	46.0	72.2	42.3	40.9	87.1	15.3	11.0	0.0
Learning Days/Prof. Staff	6.1	6.8	5.5	8.1	7.2	5.3	6.7	7.7	10.4	7.6	18.2
Learning Days/ACS Staff	6.6	4.5	2.8	4.2	4.3	5.1	11.6	7.9	4.2	8.3	8.1
Ave. Learning Days/Staff	6.3	6.0	5.0	6.7	5.9	5.3	8.3	7.8	8.5	7.9	14.2
Total Learning Cost (\$'000)	NA	2,833.0	297.9	303.7	148.2	265.9	3761.5	549.0	535.8	496.0	148.7
Learning Cost/Staff (\$)	NA	4621.5	5841.2	4820.6	4116.7	5113.5	7261.6	7842.9	9081.4	6048.8	8747.1
Source: HR data											

ANNEX 8: LEARNING AND DEVELOPMENT: 2003 STAFF SURVEY RESULTS

Questions/Responses	WB	EAP	EAP2	EAP3	EAP4	EAP1	SAR	SAR3	SAR4
1. The World Bank provides continuous opportunities for learning of its staff.									
Favorable (%)	80	82	91	85	80	77	79	90	76
Neutral (%)	13	13	0	12	14	23	12	5	16
Non-favorable (%)	7	5	9	4	6	0	9	5	8
2. I have adequate opportunity to learn new skills to do my job better									
Favorable (%)	65	62	82	62	51	62	63	71	54
Neutral (%)	21	23	18	25	20	8	22	15	24
Non-favorable (%)	15	15	0	13	29	31	15	15	22
3. I take advantage of the learning opportunities available to me.									
Favorable (%)	63	63	91	70	51	46	66	61	54
Neutral (%)	24	25	9	21	31	46	23	28	26
Unfavorable (%)	13	12	0	9	17	8	11	11	20

Source: 2003 Bank Staff Survey

ANNEX 9A: DELIVERY OF PRODUCTS, FY02-03*

Region/Unit	EAP	EAP2	EAP3	EAP4	SAR	SAR3	SAR4	SAR2
<u>Project Delivery (No.)</u>								
FY02	32	7	8	2	27	2	4	8
FY03	33	3	6	4	31	9	5	7
Increase (Decrease)	1	-4	-2	2	4	7	1	-1
<u>Lending (\$m.)</u>								
FY02	1,805	298	473	130	3674	136	430	1750
FY03	2426	271	200	387	3007	482	409	1289
<u>Lending Increase (%)</u>								
FY03	34.4	-9.1	-57.7	197.7	-18.2	254.4	-4.9	-26.3

* Including special operations and GEF projects.

Source: Bank Data Warehouse

ANNEX 9B: RISKINESS OF PROJECTS AND KEY PERFORMANCE RATIOS FOR FY02 AND FY03

	EAP	EAP2	EAP3	EAP4	SAR	SAR3	SAR4	SAR2
<u>Portfolio (No.)</u>								
FY02	313	63	78	40	167	40	46	45
FY03	296	60	69	41	168	41	39	48
<u>Problem Projects (No.)</u>								
FY02	26	7	7	4	14	4	3	5
FY03	25	6	6	3	20	4	5	9
<u>Potential Problem Projects (No.)</u>								
FY02	14	5	4	0	3	2	1	0
FY03	9	4	2	0	6	3	1	1
<u>Projects at Risk (No.)</u>								
FY02	40	12	11	4	17	6	4	5
FY03	34	10	8	3	26	7	6	10
<u>Projects at Risk (%)</u>								
FY02	12.8	19.0	14.1	10.0	10.2	15.0	8.7	11.1
FY03	11.5	16.7	11.6	7.3	15.5	17.1	15.4	20.8
<u>Projs. With Fin. Mgmt. Flag (No.)</u>								
FY02	7	2	3	1	10	2	4	1
FY03	12	5	5	2	13	3	4	3
<u>Projs. With Fin. Mgmt. Flag (%)</u>								
FY02	2.2	3.2	3.8	2.5	6.0	7.5	8.7	2.2
FY03	4.1	8.3	7.2	4.9	7.7	7.3	10.3	6.3
<u>Projs. With Fin. Perf. Flag (No.)</u>								
FY02	28	5	6	12	8	3	1	3
FY03	27	6	6	11	11	5	0	5
<u>Projs. With Fin. Perf. Flag (%)</u>								
FY02	8.9	7.9	7.7	30.0	4.8	7.5	2.2	6.7
FY03	9.1	10.0	8.7	26.8	6.5	12.2	0.0	10.4
<u>Projs. With Proc. Flag (No.)</u>								
FY02	28	6	7	4	16	6	5	4
FY03	32	11	11	5	21	12	4	4
<u>Projs. With Proc. Flag (%)</u>								
FY02	8.9	9.5	9.0	10.0	9.6	15.0	10.9	8.9
FY03	10.8	18.3	15.9	12.2	12.5	29.3	10.3	8.3
<u>Net Commitment Amount (\$m.)</u>								
FY02	30465.4	3738.0	7627.1	3936.2	21068.3	4975.9	4297.9	8471.2
FY03	27767.1	4043.1	6021.2	3976.5	20159.8	4603.1	3746.1	8965.6
<u>Commitment at Risk (\$m.)</u>								
FY02	2438.5	590.2	505.0	266.1	2242.2	854.7	322.9	648.5
FY03	1536.5	428.7	180.5	159.3	2555.8	650.4	603.5	1245.8

(Annex 9B continues on the next page.)

(Annex 9B continued.)

	EAP	EAP2	EAP3	EAP4	SAR	SAR2	SAR3	SAR4
<u>Commitments at Risk (%)</u>								
FY02	8.0	15.8	6.6	6.8	10.6	17.2	7.5	7.7
FY03	5.5	10.6	3.0	4.0	12.7	14.1	16.1	13.9
<u>Disbursement Ratio (%)</u>								
FY02	29.4	29.0	35.9	23.0	29.4	30.6	32.5	27.6
FY03	31.3	35.4	33.9	23.1	29.4	26.6	28.9	26.7
<u>Realism Ratio (%)</u>								
FY02	65.0	58.3	63.6	100.0	82.4	66.7	75.0	100.0
FY03	73.5	60.0	75.0	100.0	76.9	57.1	83.3	90.0
<u>Proactivity Ratio</u>								
FY02	95.0	100.0	87.5	100.0	90.9	100.0	100.0	75.0
FY03	87.0	60.0	83.3	100.0	90.9	100.0	100.0	75.0

Source: Bank Data Warehouse

TECHNICAL NOTE: UNIT PERFORMANCE

EAP Operations Units

Delivery of Projects

From FY02 through FY03, EAP2 and EAP3 showed a decline in project delivery (from 7 to 3 in one and from 8 to 6 in the other), while EAP4 showed an increase from 2 to 4.

Amount of Lending

During FY02–FY03, the amount of lending declined in two units (by 9.1 percent in EAP2 and 57.7 percent in EAP3), while it increased substantially (by 198 percent) in the third unit (EAP4).

Projects at Risk

The percentage of the projects at risk for the three selected units ranged from 10–19 percent in FY02 and the situation showed improvement in FY03 in all the three units, with the range dropping to 7.3–16.7 percent.

Financial Management Risk

In FY02, the percentage of projects with financial management risk was 2.5 percent in EAP4, while EAP2 and EAP3 had higher risk percentages of 3.2 percent and 3.8 percent, respectively. The situation deteriorated in all the three units in FY03, with the percentage of projects with this risk increasing to 4.9 percent in EAP4, 7.2 percent in EAP3 and 8.3 percent in EAP2.

Financial Performance Risk

In FY02, the percentage of projects with financial performance risk was 7.7 percent for EAP3 while EAP2 and EAP4 had higher risk percentages of 7.9 percent and 30 percent, respectively. In FY03, the percentage declined to 26.8 in EAP4 but increased in EAP3 and EAP2 to 8.7 percent and 10 percent, respectively, showing deterioration in this risk situation.

Procurement Risk

In FY02, the percentage of projects with procurement risk was 9 percent for EAP3 while EAP2 and EAP4 had higher risk percentages of 9.5 percent and 10 percent, respectively. The situation deteriorated in all the three units in FY03, with EAP4 showing 12.2 percent of the projects with procurement risk, EAP3, 15.9 percent, and EAP2, 18.3 percent.

Commitment at Risk Ratio

For EAP3 and EAP4, the commitment-at-risk ratio was lower in FY02 (ranging from 6.6 to 6.8 percent), while it was significantly higher (at 15.8 percent) in EAP2. The ratio

showed significant improvement in all the three units in FY03, ranging from 3 to 10.6 percent.

Disbursement Ratio

The disbursement ratio in FY02 was comparatively low at EAP2 and EAP4 (ranging from 23 to 29 percent), while it was significantly higher (at 35.9 percent) in EAP3. In FY03, the ratio declined marginally to 33.9 percent in EAP3, while it increased significantly in EAP2 from 29 percent to 35.4 percent and the ratio remained at around 23 percent in EAP4.

Realism Ratio

The realism ratio was less than 70 percent for EAP2 and EAP3 in FY02, while it was significantly higher (at 100 percent) than the benchmark in the other unit (EAP4). In FY03, the realism ratio continued to be less than 70 percent in one unit (EAP2), while in the other two units it was more than the QAG recommended benchmark.

Proactivity Ratio

EAP2 and EAP4 showed better performance (with the proactivity ratio for both the units reaching 100 percent) in FY02, while the ratio was lower at 87.5 percent in EAP3. In FY03, two units (EAP2 and EAP3) showed lower performance (with the ratio ranging from 60 to 83.3 percent), while one unit (EAP4) continued to show high proactivity performance at 100 percent.

SAR Operations Units

Delivery of Projects

During FY02-FY03, project delivery increased from 2 to 9 in SAR3 and from 4 to 5 in SAR4, while the delivery showed a slight decline (from 8 to 7) in the other SAR operations unit (SAR2) reviewed.

Amount of Lending

The amount of lending declined in two units (by 4.9 percent in SAR4 and 26.3 percent in SAR2) in FY03, while it increased substantially (by 254 percent) in the other unit (SAR3).

Projects at Risk

The projects-at-risk ratio for the three selected units ranged from 8.7 to 15 percent in FY02 and the situation showed significant deterioration in FY03 in all the three units, with the range increasing by 15.4 to 20.8 percent.

Financial Management Risk

In FY02, the percentage of projects with financial management risk was 2.2 percent for SAR2, while SAR3 and SAR4 had higher risk percentages of 7.5 percent and 8.7 percent, respectively. The percentage declined slightly to 7.3 percent in SAR3, while it increased in SAR2 and SAR4 to 6.3 percent and 10.3 percent, respectively, in FY03.

Financial Performance Risk

In FY02, the percentage of projects with financial performance risk was 2.2 percent for SAR4, while SAR3 and SAR2 had higher risk percentages of 7.5 percent and 6.7 percent, respectively. The percentage declined to 0 in SAR4 but it increased in SAR2 and SAR3 to 10.4 percent and 12.2 percent, respectively, showing deterioration in this risk situation in FY03.

Procurement Risk

In FY02, the percentage of projects with procurement risk was 8.9 percent for SAR2, while SAR4 and SAR3 had higher risk percentages of 10.9 percent and 15 percent, respectively. The percentage declined slightly in SAR2 and SAR4 to 8.3 percent and 10.3 percent, respectively, while it increased significantly in SAR3 to 29.3 percent in FY03.

Commitment at Risk Ratio

In FY02, the commitment at risk ratio ranged from 7.5 to 7.7 percent for SAR4 and SAR2, while it was significantly higher (at 17.2 percent) for SAR3. In FY03, the ratio showed a decrease in SAR3 from 17.2 percent to 14.1 percent, while in the other two units (SAR4 and SAR2), the risk ratio increased significantly (with the range being from 13.9 to 16.1 percent).

Disbursement Ratio

The disbursement ratio was comparatively low (27.6 percent) for SAR2 in FY02, while it was higher (ranging from 30.6 to 32.5 percent) in the other two units (SAR3 and SAR4), showing better performance. The disbursement ratio was lower in all the three units (ranging from 26.6 to 28.9 percent) in FY03.

Realism Ratio

For the three selected SAR operations units, the ratio was less than 70 percent in one unit (SAR3) in FY02, while it was significantly higher (ranging from 75 to 100 percent) than the benchmark in the other two units (SAR4 and SAR2). In FY03, the realism ratio continued to be less than 70 percent in SAR3 (with the ratio at 57.1 percent), while in the other two units (SAR4 and SAR2), it ranged from 83.3 to 90 percent, significantly higher than the QAG recommended benchmark of less than 70 percent.

Proactivity Ratio

SAR3 and SAR4 showed better performance (with the ratio for both the units reaching 100 percent) in FY02, while the ratio was lower SAR2, which had a proactivity ratio of 75 percent. In FY03, both SAR3 and SAR4 maintained high proactivity performance (with the ratio remaining steady at 100 percent), while SAR2 continued to show comparatively low performance with the proactivity ratio remaining at 75 percent.

ANNEX 10: UNIT LEARNING AND OPERATIONAL PERFORMANCE INDICATORS

	EAP2	EAP3	EAP4	EAP1	SAR3	SAR4	SAR2	SAR1
A. Learning Parameters								
1. Staff Having IPLs (%)								
FY02	12	71.1	14.7	NA	70.8	5	4	2
FY03	<u>3.9</u>	<u>46.0</u>	<u>72.2</u>	<u>42.3</u>	<u>87.1</u>	<u>11</u>	<u>9</u>	<u>0</u>
Difference	-8.1	-25.7	57.5	42.3	16.3	6	5	-2
2. Number of learning days per staff								
FY02	3.9	3.3	3.8	5.1	5.3	6.6	4.5	4.6
FY03	<u>5</u>	<u>6.7</u>	<u>5.9</u>	<u>5.3</u>	<u>7.8</u>	<u>8.5</u>	<u>7.9</u>	<u>14.2</u>
Difference	1.1	3.4	2.1	0.2	2.5	1.9	3.4	9.6
3. Learning cost per staff (\$)								
FY02	8756	6332	6812	4685	4292	8279	5399	6550
FY03	<u>5841</u>	<u>4821</u>	<u>4117</u>	<u>5114</u>	<u>7843</u>	<u>9081</u>	<u>6049</u>	<u>8747</u>
Difference	-2915	-1511	-2695	429	3551	802	650	2197
B. Selected Performance Parameters								
1. Project Delivery (Nos.)								
FY02	7	8	2	NA	2	4	8	NA
FY03	3	6	4	NA	9	5	7	NA
Difference (%)	-4	-2	-2		7	1	-1	NA
2. Lending Volume (\$ million)								
FY02	298	473	130	NA	136	430	1750	NA
FY03	271	200	387	NA	482	409	1289	NA
Difference	-9.1	57.7	197.7		254.4	-4.9	-26.3	
3. Projects at risk (%)								
FY02	19.0	14.1	10	NA	15	8.7	11.1	NA
FY03	<u>16.7</u>	<u>11.6</u>	<u>7.3</u>	NA	<u>17.1</u>	<u>15.4</u>	<u>20.8</u>	<u>NA</u>
Difference	-2.3	-2.5	-2.7		2.1	6.7	9.9	
4. Projects with Fin. Mgmt. Flag (%)								
FY02	3.2	3.8	2.5	NA	7.5	8.7	2.2	NA
FY03	<u>8.3</u>	<u>7.2</u>	<u>4.9</u>	NA	<u>7.3</u>	<u>10.3</u>	<u>6.3</u>	<u>NA</u>
Difference	5.1	3.4	2.4		0.2	1.6	4.1	
5. Projects with Fin. Perf. Flag (%)								
FY02	7.9	7.7	30	NA	7.5	2.2	6.7	NA
FY03	10	8.7	26.8	NA	12.2	0	10.4	NA
Difference	2.1	1	-3.2		4.7	-2.2	3.7	
6. Projects with Procurement Flag (%)								
FY02	9.5	9.0	10.0	NA	15.0	10.9	8.9	NA
FY03	<u>18.3</u>	<u>15.9</u>	<u>12.2</u>	NA	<u>29.3</u>	<u>10.3</u>	<u>8.3</u>	<u>NA</u>
Difference	<u>8.8</u>	6.9	2.2	-	<u>14.3</u>	-0.6	0.6	-

(Annex 10 continues on the next page.)

(Annex 10 continued.)

	EAP2	EAP3	EAP4	EAP1	SAR3	SAR4	SAR2	SAR1
<u>7. Commitment at Risk (%)</u>								
FY02	15.8	6.6	6.8	NA	17.2	7.5	7.7	NA
FY03	<u>10.6</u>	<u>3.0</u>	<u>4.0</u>	NA	<u>14.1</u>	<u>16.1</u>	<u>13.9</u>	NA
Difference	-5.2	-3.6	-2.8		-3.1	8.6	6.2	
<u>8. Proactivity Ratio (%)</u>								
FY02	100	87.3	100	NA	100	100	75	NA
FY03	60	83.3	100	NA	100	100	75	NA
Difference	-40	-4	0		0	0	0	
<u>9. Realism Ratio (%)</u>								
FY02	58.3	63.6	100	NA	66.7	75.0	100	NA
FY03	60	75	100	NA	57.1	83.3	90.0	NA
Difference	1.7	11.4	0		0	0	0	0
<u>10. Disbursement Ratio (%)</u>								
FY02	29	35.9	23	NA	30.6	32.5	27.6	NA
FY03	35.4	33.9	23.1	NA	26.6	28.9	26.7	NA
Difference	6.4	2	0.1		-4	-3.6	-0.9	

Source: Bank Data Warehouse