

Evaluation of the Global Development Network, 2007

December 2007

Marc D. Shapiro, Ph.D., Principal Investigator and
Managing Partner, MDS Associates

Savi Mull, Senior Research Associate

Tina Khanna, Research Assistant

EXECUTIVE SUMMARY

I) Methodology and Overview

This purpose of this external evaluation of the Global Development Network (GDN) is to provide evidence on the strength of GDN's organizational processes while focusing on attempting to measure the effect or impact of its activities. Multiple methods including desk reviews, contextual visits of partner and GDN-run workshops, interviews, online surveys of over 1,500 end users or stakeholders, and external reviews of GDN activity outputs were used in collecting data across a wide range of stakeholders.¹ Multiple analytical methods using appropriate statistical tests also were used including cross-sectional analyses, change over time, cursory comparisons with best practice in other organizations, analysis of publishability, analysis of improvements from proposals to papers, and non-experimental attempts to create matching control groups, in addition to professional judgment. Due to space limitations, the evaluation assumes a working knowledge of GDN.

The evaluation distills in the executive summary in mostly qualitative terms the numerous primary findings, conclusions and key recommendations for stakeholders that result from the wide-ranging terms of reference. The executive summary has been lengthened to include key quantitative results and recommendations in response to GDN's request to do so. Additional recommendations may be found in the conclusions and recommendations section. The executive summary is organized roughly by its ordering in the main document to make it easier to know where to look for greater detail.

The evaluation's overall **findings are generally positive indicating some progress** from the previous process evaluation in certain areas **as well as qualified evidence of capacity built and knowledge creation from GDN-funded activities.**

Process Evaluation: In terms of the process evaluation, there are many **areas of progress** including

- broadening the reach (representation) across beneficiary types through GDN's centrally-run activities and across some regionally-run activities;
- decreasing somewhat conference expenditures;
- creating board audit, executive, and program committees;
- forming a Donor Advisory Council;
- showing that its activities are scalable during fiscal crises; and
- planning to deepen or broaden capacity building effects through individualized training activities to match target constituencies' priorities.

Despite a financial downturn during fiscal years 2004-2006, GDN was been able to fund its primary activities to a level deemed at least satisfactory to most key stakeholders. **GDN was able to diversify and increase funding sources** in fiscal years 2007 and probably 2008 en route to meet goals of reducing dependence on its largest donor, the World Bank, despite considerable challenges in broadening its donor base into new areas such as corporate donors. Despite sometimes sharp disagreements among stakeholders over highest value added from the different activities in GDN's portfolio, GDN generally is perceived as providing **mostly unique, relevant, and valuable services.**

¹ Table 1.1 in section 1 provides a summary of the types of questions asked by stakeholder types in the online surveys. Response rates for end users, although typical for online surveys without extensive follow up, are sufficiently low (between 16 and 37 percent) that respondents likely are biased towards those more satisfied with GDN's activities.

Impact Evaluation: In terms of the impact evaluation, grantees suggest **moderate individual-level effects** from GDN-funded capacity building and knowledge creation activities. Given insufficient funding and a relatively brief period to generate broad-ranging long-term impacts, the evidence suggests **broader impacts only in some regions** where GDN's efforts accompany those of other actors with similar goals. The evaluation does not find evidence of policy impacts nor does it find a reason to believe this is an appropriate objective, but it does find **evidence of building a cadre of professionals capable of policy relevant work** and helping build policy relevant research as well as suggestions of some outreach to policy makers. GDN and the annual conference are seen as potential vehicles for providing additional public goods for its targeted beneficiaries. The evaluation found **potential differences across regions in uniformity in quality of capacity building** from GDN-funded regional research competition (RRC) activities. **Grantee satisfaction with GDN's overall portfolio of activities on average is between high and medium** (the second or third highest rating on a five point Likert scale) but closer to high, especially among respondents from Sub-Saharan Africa.

The evaluation also notes many **areas for further improvement**. This list includes

- clarifying disciplinary focus,
- increasing the strength of board governance and management capacity,
- increasing capacity in policy relevance and links with policy makers,
- clarifying the extent to which RNPs are considered and treated as partners versus grantees,
- increasing quality of processes across RRCs and learning across RNPs,
- increased openness to inputs from donors and network partners,
- better monitoring grant fund distribution by categories of researchers and overhead costs, and
- reviewing and strengthening key human resource and management policies.

Areas for improvement on financial issues include

- continuing to reduce income volatility,
- maintaining or increasing staff expertise in research and fundraising as feasible,
- creating new relationships at the World Bank to champion long-term financial support needed to maintain its long-term capacity building mission, and
- working towards further and more refined joint fundraising with RNPs.

II) Reach, Outcomes, Effectiveness, Relevance, Impact, and Cost Effectiveness

A) Objectives and Reach

The evaluation's analysis begins (section 2.1) with a key area of concern raised in the previous evaluation and an important starting point to assess an organization – its objectives. Compared to the prior evaluation, this evaluation finds evidence there probably is **somewhat greater agreement and lower tensions regarding GDN's objectives** and target beneficiaries and how they are manifested. GDN's many and varied objectives and target beneficiaries still generates some clear disagreements due to the large number of objectives it claims and the innate conflicts among some of them such as the balance between targeting individuals and institutions and the implicit tension between emphases on capacity building and supporting only highest quality inputs and outputs. The **role of multidisciplinary remains an area of disagreement**, although no longer quite the source of tension previously. These disagreements complicate evaluating based on a set of clear, limited objectives as well as creating consensus among stakeholders on GDN's success in strategies, plans, and activities' measured against agreed upon goals. The evidence suggests a portfolio approach to meet many different objectives is appropriate. Such approach allows greater leeway in programming by GDN's management, which, while positive generally, also reduces board control somewhat, but it also increases the ability to create new or innovative programming as well as to diversify itself as an organization reducing dependence on the World

Bank. The primary recommendation relevant to GDN's objectives, which should be considered during the current strategic review process, is:

- Clarify further the extent to which GDN generally plans to focus on building capacity broadly across the social sciences or building capacity primarily in economics with other disciplines serving mostly as an instrument to answer questions of interest.

GDN's **board and staffing suggests dominance by economists**. However, the evaluation finds that **for its more centrally run activities, GDN is making concerted efforts to have broad representation**. Women, younger researchers, and increasingly people from a variety of disciplines in addition to economics are somewhat better represented than in the past, and regional representation is broad. GDN's **record in broadening beyond targeting only economists is mixed with delegated activities** – the regional research competitions (RRCs) – with some regions broadening considerably and others not at all.² Similar variation appears **true to a lesser degree in terms of gender representation** by RRC region, although the evaluation has no clear benchmark by region. The evaluation did not have information on reach by higher needs countries or areas within regions. The key priority recommendation in this area is:

- Continue GDN's recent unofficial policy of encouraging the selection of themes in calls for papers that are more naturally interdisciplinary in subject, supported by outreach outside of economics networks.

An additional recommendation includes:

- Begin to develop measures to track higher needs countries or even areas within countries.

B) Outcomes, Effectiveness, Relevance and Cost Effectiveness, by Activity

The evaluation reveals varying evidence regarding the relevance and effects of GDN's primary activities. There is some evidence of individual level effects but evidence of broader impact only for regions in which GDN's activities are considered best run and in which they operate in concert with other complementary organizations. The evaluation considers cost efficiency of GDN's activities but does not emphasize the results, as it lacks benchmarks for comparisons, and activities involve different outputs and outcomes, some not readily quantified nor easily comparable. The portfolio approach of activities allows GDN to legitimately claim relevance of all its activities toward some objective and target group, even if preventing unanimity of agreement about each one's relative value toward the larger mission.

RRCs and Global Research Projects (GRPs): As shown in interviews and survey results, **stakeholder support is strongest and most universal for RRCs**. RRCs suffer, however, from somewhat greater difficulty in attracting new funds and lack of visible regional impact in some regions due to ineffective implementation and lack of other complementary efforts of other organizations. The evaluation's review of survey data suggest increases in capacity built for funded individuals, with RRC grantee respondents indicating an **average value of effectiveness of the RRC process between medium and high** (the second or third highest rating on a five point Likert scale). The evaluation found **differences in levels of support and mentoring for RRCs across regions** as well as perceived value added through the process, with particular concerns in the Asian regions. In terms of effect of the RRC process on individuals' research, the average indicated the average change across the five point Likert scale (rated from zero ["none or very little"] to four ["very substantially"]) was 2.1 to 2.3, or closer

² The three year averages for FY 2004-2006 in terms of proportion of RRC grantees outside of economics ranges by region from 0 to 53 percent, with an overall average of 38 percent.

to “moderately” than “substantially” changed.³ End user survey responses regarding value added were correlated with the amount of feedback and value/quality of review sessions suggesting a need to more consistently encourage feedback and mentoring of RRCs.

The evaluation finds **mixed evidence regarding GRPs with stronger individual-level impacts suggested by grantees than perceived value of outputs by stakeholders**. GRPs are considered a **highly unique activity but also criticized by some stakeholders for insufficient relevance to objectives** due to disagreements regarding the GRP theme selection process and selected GRP themes relative to differently perceived objectives across stakeholders. Some also believe GRPs provide highly variable quality across themes, especially given the higher expectations and resources provided per unit output, as noted in the cost efficiency section. Although the evaluation cannot assess that claim objectively through its data collection activities, it does find **grantee survey responses suggest increases in capacity built equivalent to that found for RRC processes**. In contrast, **little statistical evidence was found of capacity built as measured in improvements from proposals to research paper** from pilot test research for one GRP and one RRC, although limitations of proposals provided may help explain the non-findings. The evaluation also did not find positive effects from RRC and GRP capacity building through a non-experimental technique that matches respondents from the grantee and broad constituency surveys based on similar observable characteristics, although the evaluation notes that this technique cannot always surmount biases it seeks to overcome, especially given the low survey response rates.⁴ Given the most successful fundraising approach for GRPs focusing on different issues and techniques by GRP, they are unlikely to be a vehicle for long-term measurable impact unless GDN changes its approach.

In terms of outputs, or knowledge created, from GDN-funded activities, self-reported survey data indicate a **median of two publication types per grant for RRCs and GRPs**. Taken individually, on average one of every two grantees published in an international journal, two of three in a national journal, two of three as a book chapter, and just over one working paper per person. These figures represent upper bounds on GDN’s effect on knowledge creation, as GDN cannot take credit in every case either for increasing the publishability of the papers nor for helping these grantees conduct research that otherwise definitely would not have been conducted.⁵ **RRCs generally are found to be considerably more cost efficient mechanisms in the strictest sense of the term** (\$19,200 per team or \$12,200 per team member on average) than the GRPs (averaging between \$54,000 and \$135,000 per grant), although theoretically GRPs can lead to greater knowledge added in a concentrated subject area or lead to other follow-on activities besides just publications. Further, most funding for GRPs are earmarked and not fungible for application to RRC projects. The evaluation did not find a satisfactorily objective, comparative way of judging *cost effectiveness* given limited resources and time. The next evaluation ultimately will have to judge the cost efficiency of the otherwise cost inefficient Bridging Research and Policy (BRP) and Health GRPs based on the activities and direct and indirect outputs from them. Key priorities GDN should consider regarding RRCs and GRPs include:

- Giving greater consideration from the outset of future GRPs toward strategies for disseminating research findings to the broader policy and development community.

³ (See section 2.3.1.4 for further details.)

⁴ The matching processes, known generally as propensity score matching, are multivariate techniques to create “treatment” and “control” groups from non-experimental data by matching pairs or groups of respondents based on observable information. Results all were either negative or statistically insignificant across seven outcome measures and using six different matching methods. GDN should have independent researchers, who are less time and resource constrained review this data and undertake further tests using this approach.

⁵ Furthermore, there is an upward bias in this type of self-reported data probably augmented by response rates well below 70 percent where such biases are not noticed.

- Develop better delivery vehicles, such as requiring policy briefs (accompanied by training to do so), for clarifying the policy relevance of papers from RRCs, if not GRPs.
- Continue to follow through with the findings of its “Workshop on Methodology for Comparative Analysis” to be more proactive in promoting a diversity of methodologies in the choice of intended research design through its GRPs and other activities.

Additional, lower priority, recommendations that may be more resource intensive or require other changes include:

- Consider holding some GRP workshops before or after a disciplinary conference chosen appropriately for the theme to emphasize capacity building of attendees.
- Consider individual-level training programs or GDN self-training modules beyond the top priority of methodology and technical issues such as use of Internet resources as modules that could accompany RRC workshops to broaden and deepen their effects.

Annual Conference: The annual conference remains GDN’s **most contentious activity in terms of relevance, uniqueness, quality, and cost, but it received generally strong ratings on average from stakeholders for its overall value** for its broad set of constituents. Concerns appear somewhat more muted or less widespread than in the previous evaluation, as GDN has reduced somewhat earlier extravagances although business meetings and GRP workshops, which perhaps should be considered separately, keep total costs relatively high. Its networking function is seen as at least a moderately useful global event for bringing researchers together from different parts of the world. The grantee survey reflects at least a **moderate breadth, if not depth, of South-South networking within and across regions**, although stakeholders indicated a need to try to foster greater inter-regional interactions. The evaluation did not have time for an independent review regarding concerns of some stakeholders regarding both average quality and variance in quality of papers and presentations and was unable to attend an annual conference as part of writing this report. **Different measures of cost efficiency for the conference could lead to sharply differing conclusions.** For instance, the per paper cost (about \$10,500) is very high given the lack of capacity building outside of separate workshops. Other measures could lead to the conclusion that the net conference cost is relatively reasonable given that most researchers are understandably subsidized with average costs under \$1,300 across all attending researchers, if one excludes the estimated \$4,000 cost per participant in the conference side business meetings and the costs for workshops ascribed by the evaluation to GRPs. High priority recommendations include:

- Consider consistent specific track themes across years, increasing networking among clearer sets of audiences.
- Explore publishing and distributing proceedings and papers in a post-conference (if possible, pre-conference) compact disc and on-line to extend the long-term value of the conference.

More time and resource consuming recommendations that should be considered carefully in coming years include:

- Consider experimenting with alternative scheduling such as an 18 month schedule that would consume fewer resources, allow secretariat and board to focus on additional priorities and activities, potentially increase uniformity of paper quality, and allow scheduling across regions to vary by appropriate season.
- GDN also could experiment one year pairing itself with a relevant academic conference and using surveys across years to compare the added value versus the ABCDE conference.
- If GDN desires to pursue both capacity building and highest-quality knowledge building objectives for the conference, GDN should experiment with purposely focusing on increasing capacity building in at least selected sessions in a more meaningful and effective way while ensuring more consistent quality in other sessions.

GDNNet: Among those who understand GDNNet best as well as targeted users, there is some level of **agreement regarding GDNNet's usefulness as well as the need to expand its awareness among and relevance for developing country researchers.** GDN has been actively pursuing strategies to increase the relevance of its GDNNet offerings, broadening access to journals, although some impediments to doing so are out of its control. GDN should explore through its strategic review and separate evaluation of GDNNet additional potential uses for GDNNet, since it likely will be difficult to differentiate itself as adding value through aggregating development information from the southern perspective. Rapidly emerging uses of the Internet as a networking tool might provide GDNNet additional visibility. Key priority recommendations for GDNNet include:

- Consider ways to more make GDNNet a portal for Southern researchers to leading working paper and database collections worldwide.
- Change grantee contracts so that GDNNet serves as a central location for datasets funded by GDN process so that it can confirm its policy of open access data is followed and to reduce search costs for researchers.
- Continue to pursue ways to extend the reach of journal availability through GDNNet or otherwise, considering the use of key stakeholders and board members to serve as liaisons.

A lower priority recommendation is:

- Assess the extent to which more senior researchers might be aided by serving as a focal point for changing access policies of national datasets by setting up an international model program of data access from national institutions in the Global South, which currently often severely limit access.

Additional recommendations for the upcoming evaluation of GDNNet are provided in the recommendations and conclusions section.

Awards and Medals Competition (AMC) and Most Innovative Development Project (MIDP): These awards are funded through an individual and decreasing earmarked grant that is not fungible across activities. The analysis suggests that **AMC's are rated more highly by AMC grantees than RRC and GRP grantees** for increased career advancement, contact with policy makers, and increased professional visibility. However, they **do not appear a cost effective alternative to GDN's other activities.**⁶ These are an area where GDN's appeal to a broader clientele than just economists appears more successful.

Support of Institutions: In terms of GDN's activities in institution building, **GDN's support for institutions thus far has been narrow and *ad hoc*,** although a pilot BRP-related training activity is planned. Further work with institutions could be highly relevant and increase the cost effectiveness of its BRP but would depend on additional resources and development of an inexpensive, efficient mechanism for delivery and possibly hiring a person with such experience on staff. There is no objective way for the evaluation to determine the extent to which shifting resources from supporting individuals to supporting institutions would increase or decrease effects or impacts, but GDN should consider carefully measures of cost effectiveness as presented in this report. GDN's efforts to support institutions obviously could be much stronger, were additional funding available and agreement reached on the board that more work in this area fits sufficiently within GDN's broad mandate to demand its limited resources. The evaluation recommends consideration of facilitating research partnerships between Northern and Southern researchers and institutes through its current ongoing activities to better fulfill the organization's objective to generate high quality research.

⁶ About \$17,800 per medal, \$61,800 per award and \$70,000 per MIDP.

C) Overall Satisfaction, Highest Marginal Benefit, and Policy Relevance and Impact

Grantee average satisfaction with GDN's overall portfolio of activities is closer to high (second highest rating on a five point Likert scale) than medium with survey respondents from Sub-Saharan Africa in particular indicating higher satisfaction. These findings indicate value added from GDN's funded activities as a whole but with room remaining for improvement and greater success.

In terms of where grantees would prefer GDN spend additional resources, **grantees are most interested in training both researchers and institutes on how to influence policy and policy debates and in funding workshop trainings on technical issues.** The next highest responses were funding additional research grants within their region and increasing access to journal articles or working paper series for those without access. Commendably, **GDN has begun to plan a training of individual researchers that appears to parallel some of the top priorities** of end users and should complement and deepen effects, if not impacts, of its current portfolio.

There is modest evidence that GDN's efforts thus far have resulted in **increased capacity for policy relevant research as well as more modest evidence of outreach to policy makers**, but not always supported by the GDN-funded processes. GDN activities, training, and mentoring generally generate **more academically than policy focused outputs** reflecting the potential conflict perceived by traditionally trained economists between the objective of high policy relevance and impact and the objective to foster highest quality research. The evaluation finds **anecdotal evidence of capacity built informing policy making in certain cases and regions.** There is no compelling evidence that GDN-produced research and efforts themselves have produced a "policy impact," which the evaluation considers an inappropriate indicator for GDN given most of its objectives. Planned training resulting from the BRP research and the Health project and GDNNet as a distribution method could further GDN's ability to make further claims in that direction, however.

III) Overall Process and Performance

A) Financial Sustainability and Risk Diversification

GDN has struggled, with some success, toward increasing its financial sustainability and diversification over the time period analyzed. GDN has been **able to fund its primary activities to a level deemed satisfactory to key stakeholders** even if it is not near meeting the high absorptive capacity of its many potential beneficiaries for its conceivable activities. **GDN failed to meet its over-ambitious financial goals of its 2005-07 strategic plan** and actually suffered through three years of financial decreases through FY 2006 due primarily to reductions from two key donors – the World Bank (declining by half) and a bilateral donor Italy (declining to zero). **GDN has made significant, if qualified, progress in diversifying its donor base for FY 2007 and 2008**, however (see Table 3.1-1). Using several different methods to calculate financial volatility, the evaluation found that the **income volatility has decreased slightly** from the FY 02-04 to FY 05-07 time periods if one includes the decrease in funding from the World Bank, although income volatility across donors besides the Bank has remained the same or increased slightly from the previous time period depending on the measure (see Annex Y). On net, this is positive as long as GDN can find substitutes for its current major non-Bank donor (the Gates Foundation) in the future. Current funding commitments and other prospective funding make **GDN's viability appear strong and potentially promising to at least maintain scale in the near term.** Given limited time and resources, the evaluation did not find evidence in its admittedly limited interactions with donors of continuing reluctance among donors to contribute to GDN due to a "World Bank whiff" noted in the prior evaluation.

As with most grant-dependent organizations, GDN's long-term financial sustainability is uncertain and will remain so unless and until the donor base of sizeable donors is widened and funding becomes less volatile. GDN's mission is one unlikely to connect with a highly diverse and rapidly growing donor base,

especially in the corporate sector, which GDN mostly unsuccessfully attempted to pursue. GDN will have to remain flexible in scaling activities up or down annually response to unavoidable fluctuations in overall funding. This is unfortunate for GDN given that its mission reflects a focus on capacity building changes that require a long-term commitment for success. To a large extent, however, GDN appears able to undertake necessary changes. In this respect, GDN may be ahead of some other similar organizations.

B) Governance, Organization and Interaction with Regional Partners and Stakeholders

The evaluation lists in section 4.2.1 the several ways GDN has responded directly or indirectly to **predominantly governance and process issues and recommendations from the previous evaluation** including

- creating audit, program, and executive committees;
- making concerted, albeit unsuccessful, efforts to raise money from corporate sources;
- reducing somewhat their conference expenditures;
- forming a Donor Advisory Council;
- raising funding levels at least in the past year; and
- showing that its activities are scalable during fiscal crises.

These activities indicate responsiveness donors seek although further work remains, and some innate difficulties and contradictions may not be easily overcome.

The analysis of organizational structure indicates there is some disconnect between GDN management's description of GDN being "an efficient administrative mechanism to conduct research" and a quasi-think tank model assumed by some stakeholders, who suggest greater research capacity in the secretariat. The evaluation does not advocate the latter view but does suggest it as an option worth exploring as a way to increase the extent to which highest quality work is facilitated as an objective. It also notes the value of adding staffing with fundraising expertise, although the cost-benefit ratio for an expensive position such as that is unclear for a research organization and should be considered as a provisional appointment for 18 to 24 months based on success.

The evaluation's analysis of board member responses suggests significantly **lower division in the board than during the previous evaluation** since many members were not around for the challenging key decisions of the past. Key problems appear to be **continuing weak governance from the board, insufficient transparency in the process of choosing board members, and insufficient structures to guarantee institutional memory or improved board management capacity**. With many organizations, board governance is not prioritized until after an organization has to overcome problems a board failed to prevent through proper oversight. Although GDN has been fortunate enough not to have yet encountered such a problem, best practice is to move forward to strengthen governance beforehand. Despite board member interest in working more on strategic issues, the board is unlikely to be able to do so meeting only twice a year. Key priority recommendations in line with best practices are:

- Develop requirements probably by positions for future incoming members to guarantee at least some positions have greater experience in managing organizations or boards and financial background.
- Increase board interactions by extending the number of days of board meetings and/or adding one or two virtual meetings per year via the Internet.
- Increase the number of permanent committees to handle the board's business issues considering such committee types as programmatic and development committees. Committees are where most boards get the real work and oversight accomplished.
- Clarify for prospective incoming members appropriate time expectations for board membership such as through a frequently asked questions document and a nonbinding agreement about board member responsibilities.

- Providing new board members a strong orientation to the organization within their first month on the board.
- Work with nominating organizations to make recruitment of board members more transparent.
- Make sure the board has a conflict of interest policy signed at least once if not annually.

Additional recommendations include:

- Consider an additional donor representative to help groundtruth GDN's ideas from the perspective of funding.

In terms of GDN's interactions with its RNPs, the evaluation does not find a clear, compelling rationale to change models. Insufficient clarity exists however regarding the extent to which it is a network of networks – that is, more of a partnership or more of a grantor/grantee relationship. The evaluation supports GDN's new efforts toward creating dedicated work plans with joint fundraising with RNPs, which could prove financially rewarding and decrease perceptions of potential competition for fund raising and increase funds that can be leveraged regionally. There are some concerns about insufficient monitoring of grant fund distribution by categories of researchers and overhead costs. GDN also should be careful not to reduce overhead costs paid to RNPs too much despite shrinking budgets. The evaluation recommends a few priority recommendations regarding GDN's relationship with its RNPs:

- Clarify the extent to which GDN is a network of networks versus a centralized organization with networks as grantees, increasing contacts if more of a partnership approach is decided upon.
- Widen the new pilot idea of working with RNPs in designing a coordinated fundraising plan toward a dedicated work plan with joint fundraising after clarifying roles more fully.
- Provide greater institutional support for RNPs that rotate administrative centers to increase continuity during transitions. Look for partnerships and other ways to generate greater impact in the Asia region in building research capacity.
- Improve learning across regional networks through an additional, dedicated real or virtual meeting outside of the annual conference.

Additional recommendations include:

- Augment input in decision making process from RNPs on GRP themes or candidate selection.
- Consider more detailed reporting to monitor the grant fund distribution to categories of researchers and overhead costs. Consider adopting the common reporting framework used by many other donors.

In terms of the ways in which GDN interacts with its target constituencies and establishes policies for inclusion of stakeholders in programmatic activities, there is some call among stakeholders for seeking greater participation of targeted beneficiaries in planning activities and more actively participating in other existing networks and building further strategic alliances in its field. As with the prior evaluation, this evaluation does not see a workable alternative model for administering GRPs given their size and subject matters, and it does not have the impression that GDN overly centralizes intellectual leadership within GDN currently. The evaluation team also notes improved branding and self-marketing as an area for additional review and improvement.

C) Management Processes

The evaluation considered transparency and disclosure as well as democracy as two key dimensions of organization policy. The evaluation finds among RNP heads and board members **a desire for a more participatory approach in selection of GRP and annual conference themes**. GDN recently has shown **increased openness to donor inputs**, although it is unclear that donors consider the input meaningful and

attended to. In terms of GDN's financial policies, GDN conducts annual audits and conducted a recent review by the World Bank's Development Grant Fund unit of its financial policies. Although the evaluation did not receive a copy of the report from GDN, the team's understanding of the review as interpreted by GDN itself suggests **GDN has reasonable to strong systems given its age and size** and that GDN appears to be incorporating the review's suggestions into its financial policies. Stakeholder survey results and interviews suggest that stakeholders feel GDN has **some feedback learning processes built into the system but that they could be expanded**.

Key priority recommendations in this area include:

- Being attentive to remaining as non-hierarchical as possible given its small size and need to have partners feel ownership and involvement to remain engaged as partners.
- GDN should hire a consultant to work with the organization to develop a coherent branding strategy and advise it regarding a communications plan.

Additional recommendations include:

- Consider hiring a staff member with deep research experience if GDN wishes to focus more on providing intellectual direction to new activities and oversight for conference activities.
- Alternatively, or perhaps additionally, consider hiring a fundraiser as a fulltime staff member or consultant based in the US or Europe.
- Look for ways to create additional strategic partnerships and linkages with additional consortia and participation in other conferences.
- Engage more individuals to participate more fully in planning processes.

Regarding management and administrative systems, while GDN has undertaken advisable external reviews of its financial and accounting systems, the evaluation's understanding is that GDN has **yet to undertake external reviews of its human resources practices and policies**. The report notes several issues likely to be raised in such a review and also points out **changing legal implications now that GDN has acquired international organization status**. The previous president's use of working groups for recurring activities or long-term issues is seen as a helpful organizing tool. There is a common theme that staff appear overburdened during the months leading up to the conference. Key priorities for attention in alignment with organizational best practice include:

- Undertaking a more complete external review of human resource policies.
- Conduct a legal review of its situation and options given its newly acquired international organization status in reference to liabilities, taxes, etc.
- Formalize how staff would handle problems with supervisors or top management including naming an ombudsperson to handle serious or personal concerns with management or supervisors.

Additional, lower priority recommendations include:

- Establish clear and implemented policies on selection and supervision of family or friends
- Establish a policy on and training regarding sexual harassment.
- Review application of travel rules for consistency and cost effectiveness in line with other similar NGOs or smaller international organizations and not the World Bank.
- Improve and more regularly update GDN's internal database and better maintain it for consistency of information across staff.
- Undertake a market survey of similar organizations in India and internationally to check on rough comparability of local and international salary rates.

- Examine whether it is feasible to add at least temporary additional local staff if the conference remains annual.

Finally, this evaluation provides several recommendations for the next evaluation and some minor technical recommendations in Annex AB. A list of interviewees is provided in Annex A.

TABLE OF CONTENTS

INTRODUCTION	1-5
1.1 Introduction	
1.2 Study Design and methodology	
1.3 Description of Survey Samples	
1.4 Structure of Report	
1.5 Limitations	
FINDINGS: RELEVANCE, REACH, OUTCOMES, EFFECTIVENESS, IMPACT, AND COST EFFECTIVENESS	6-40
2.1 Understanding GDN's Objectives in Order to Measure Relevance, Reach, Outcomes, and	
2.2 Relevance – Summary	
2.2.1 Relevance to Targeted Beneficiaries	
2.2.2 Uniqueness	
2.3 Outcomes, Effects, and Impact	
2.3.1 <i>Outcomes of RRC and GRP Capacity Building Activities</i>	
2.3.1.1 <i>Grantee Survey Self-Reports</i>	
2.3.1.2 <i>Propensity Score Matching of RRC versus General Survey Self-reported Output</i>	
2.3.1.3 <i>External Review of Research: Final Paper Quality, and Proposal to Paper</i>	
2.3.1.4 <i>Effectiveness of RRCs across Regions, between RRCs and GRPs, and Impact</i>	
2.3.2 <i>Annual Conference</i>	
2.3.2.1.1 <i>Annual Conference – Perception on Overall Value Added and Paper</i>	
2.3.2.2 <i>Annual Conference – Perceptions on Networking</i>	
2.3.2.3 <i>Annual Conference – Other Objectives</i>	
2.3.3 <i>GDNet</i>	
2.3.4 <i>Policy Impact</i>	
2.3.5 <i>Effect on Careers</i>	
2.4 Reach and Multidisciplinarity	
2.5 Cost Efficiency and Cost Effectiveness	
2.5.1 <i>General Questions</i>	
2.5.2 <i>Annual Conference</i>	
2.5.3 <i>GRPs</i>	
2.5.4 <i>RRCs</i>	
2.5.5 <i>AMC and MIDP</i>	
2.5.6 <i>GDNet</i>	
2.6 Overall Success	
2.6.1 <i>Overall Success Aside from Cost</i>	
2.6.2 <i>Relative Valuations</i>	
FINDINGS: OVERALL PROCESS AND PERFORMANCE	41-54
3.1 Financial Sustainability and Risk Diversification	
3.1.1 <i>How well has GDN been able to mobilize funds over past three years to implement its</i>	
3.1.2 <i>Has GDN met its own financial targets?</i>	
3.1.3 <i>Is there a sustainable portfolio of donor support with greater diversification and smaller</i>	

- 3.2 Governance, Organization and Interaction with Regional Partners and Stakeholders
 - 3.2.1 *Organizational Structure*
 - 3.2.2 *Board Composition*
 - 3.2.3 *Interaction with RNPs*
 - 3.2.4 *Interaction with Target Groups and Other Organizations and Communications*

3.3 Management Processes

- 3.3.1 *Organizational Policies – Transparency/Disclosure and Democracy*
- 3.3.2 *Management and Feedback Processes*
- 3.3.3 *Secretariat and Daily Management*

CONCLUSIONS AND RECOMMENDATIONS

55-63

4.1 Objectives, Activities, and Effects

- 4.1.1 *Objectives and Strategies*
- 4.1.2 *Relevance*
- 4.1.3 *Outcomes, Effects and Impacts*
 - 4.1.3.1 *RRCs and GRPs*
 - 4.1.3.2 *Annual Conference*
 - 4.1.3.3 *GNet and Upcoming GNet Evaluation*
 - 4.1.3.4 *Institutional Support*
 - 4.1.3.5 *Policy Impact*
 - 4.1.3.6 *Reach*

4.2 Governance, Organization and Processes

- 4.2.1 *GDN's Responses to Prior Evaluation, Including Governance*
- 4.2.2 *Relationships with RNPs*
- 4.2.3 *Organizational and Management Processes*
- 4.2.4 *Secretariat and Daily Management*
- 4.2.5 *Recommendations for Future Evaluations and Other Technical Issues*

REFERENCES

64

ANNEXES

65-91

Acronyms

ABCDE	Annual Bank Conference on Development Economics
AERC	African Economic Research Consortium (RNP – Sub-Saharan Africa)
AMC	Awards and medals competition
BRP	Bridging Research and Policy
CEA	Cost Effectiveness Analysis
CEE	Central and Eastern Europe
CERGE-EI	Center for Economic Research and Graduate Education – Economics Institute (RNP – Eastern and Central Europe)
CGIAR	Consultative Group On International Agricultural Research
CIS	Commonwealth of Independent States
DFID [UK]	Department for International Development UK
EADN	East Asian Development Network (RNP – East Asia)
EERC	Economics Education and Research Consortium (RNP – CIS)
ERF	Economic Research Forum for the Arab Countries, Iran, and Turkey (RNP – Middle East and North Africa)
FY	Fiscal year
GDN	Global Development Network
GDNNet	The umbrella title for GDN's online information initiatives (Web site: www.gdnet.org)
GFY	Grant fiscal year
GRP	Global Research Project
IDEAS	Internet Documents in Economics Access Service
LACEA	Latin American and Caribbean Economic Association (RNP – Latin America and Caribbean)
LAC	Latin American and Caribbean Economic Association (RNP – Latin America and Caribbean)
MENA	Middle East and North Africa
MIDP	Most Innovative Development Project
NGO	Non-Governmental Organization
ODN	Oceania Development Network (RNP – South Pacific)
PM	Prime Minister
PtP	Paper to Proposal
RNP	Regional network partner
RRC	Regional Research Competition
SANEI	South Asia Network of Economic Research Institutes (RNP – South Asia)
SIDA	Swedish International Development Cooperation Agency
TOR	Terms of reference
U.S.	United States

1.0 INTRODUCTION

1.1 Introduction and Background

GDN commissions external evaluations of its activities on a periodic basis, the last of which was conducted in March 2004 as just a process evaluation, followed by an internal evaluation in 2005 focused on outcomes. The current evaluation combines both types of analyses examining the stated role of GDN and the extent to which it has positioned itself to achieve its objectives given criteria established by GDN itself. The evaluation attempts to

- Measure the outcomes, relevance and effectiveness of GDN's capacity-building activities in generating, sharing and applying knowledge.
- Assess, in limited fashion, the cost-effectiveness and sustainability of GDN.
- Examine GDN's operational effectiveness as a global network of research and policy institutes, including its internal organizational performance and its interaction with its regional network partners (RNPs). Governance structure is examined in greater detail than day-to-day management processes.

The terms of reference for the external evaluation is attached as Annex B including background on GDN as an organization. Due to space limitations, the evaluation assumes either familiarity with GDN's processes or that the reader has read at least summary information about the organization.

1.2 Study Design and methodology

Discussions with GDN at the start of the evaluation made it clear that the evaluation would update the previous evaluations and go beyond them. The evaluation conducted a similar process from the internal evaluation in collecting information from end-users and externally reviewing GDN commissioned research papers, moving beyond it in several key areas and adding the process evaluation of the previous external evaluation – attempting to do so within the same budget and less time.

The qualitative collection of data included conference and grantee workshop visits and visits to GDN's two offices to provide the contextual background. Data collection included these visits as well as qualitative and quantitative methods including both secondary and primary data analysis to allow triangulation of results. Secondary data analysis consisted of reviewing internal process materials and written documentation. The evaluation gathered information on GDN's management processes related to human resources, in house database and financial policies. A qualitative review of the policies and monitoring and feedback systems relative to the needs and standard industry practices also was undertaken. Special attention was paid to issues of concern in the previous external evaluation.

The primary data collection techniques relied on interviews, focus groups, and surveys of RNP coordinators, GRP coordinators, staff, interns, current and former (since 2006) board members, donors, and end users. The evaluation team also spoke with selected mentors and resource experts at RRC and GRP workshops but did not include them in surveys, as it believed that only those few with perspectives across many years would make appropriate targets. Online surveys (pilot tested and available for up to three weeks with emails and follow up calls) also were conducted to assess the impact of activities on end-users including GRP, RRC, and AMC grantees and researchers registered on GDN's Knowledgebase (sometimes referred to as the "broad constituency"). The evaluation team created an online stakeholder survey of current and former Board members, RNP heads, staff and donor contacts part of which was adapted from the Swedish International Development Cooperation Agency's (SIDA) Octagon Model. This part of the survey includes 14 questions covering issues of organizational performance dimensions involving identity and structure, activities, organizational capacity, and relationships. The stakeholder survey also asks about value of activities and objectives, a key issue in the previous evaluation. One example each of the end user and stakeholder surveys – the survey of RRC grantees and the stakeholder survey of board members – are provided in the annexes (Annex

C and D), but all surveys are available from GDN or the principal investigator on request. Donors were not asked to answer questions across all dimensions due to a narrower perspective and an expectation of less time available.

Survey data was analyzed using appropriate methodologies. The analysis relies primarily on comparisons across two groups using difference of means (t-tests) when data is cardinal (numeric) and Wilcoxon signed-rank test when data is ordinal (rank ordering known, but not distance between the options). For general changes in categorical data, chi-square, tau-beta, or Fisher exact tests are used. Other techniques used are described in the analysis.

The report generally presents survey results for all respondents in a given category. In a few cases, results for the grant fiscal year 2003-2005 are presented when significantly different from the rest of the sample to allow cohort comparisons for future evaluations, although in most cases this group was relatively small and did not provide statistically significantly different responses from other respondents.⁷

The evaluation also collects measures that reflect on GDN's capacity building. There are a number of ways to assess GDN's capacity building activities based on research output due to the multiple potential objectives of the RRC and GRP activities and multiple definitions of or levels of capacity building implied. To the extent that GDN is intending to add to the level of knowledge on development issues from the perspective of the Global South or do so while providing opportunities for cross-regional knowledge fertilization among those with strong skills already, volume and quality of research output generation itself could be an appropriate measure. This assumes, however, it can be determined that GDN is funding papers that otherwise would not be undertaken. To the extent that GDN is intending to build capacity among those in regions where capacity is lacking, better measures would be increase in the quality of research output. Adding to the level of knowledge on development issues from the perspective of the Global South is measured through responses to survey questions about amount of research output generated⁸ as well as external reviews of publishability of the papers. The caveat on this information is that the evaluation does not have evidence on the extent to which these papers might have been undertaken regardless of GDN funding. Thus, these measures should be seen as upper bounds on the extent of knowledge generation by GDN. Building capacity is measured through responses to survey questions about improvement in skills among grantees as well as through a longitudinal review of the change in quality from proposals to final papers among grantees. The methodology for this paper to proposal review is described further in section 2.3.1.3. The evaluation considers this a pilot test since it is the first time attempted and only uses a partial sample of capacity building activities.

The evaluation also examines effects from GDN's RRC and GRP capacity building through a two-step non-experimental multivariate statistical method called propensity score matching. This relatively new technique in the evaluation field allows for unbiased estimates comparing between "treatment" groups (grantees) and "control" groups (broad constituency survey respondents), who share similar objective characteristics. This technique can overcome large biases across groups when the average respondent may not be similar to those found in the treatment group and where selection into treatment is measurable by observable characteristics. In

⁷ The analysis was planned to focus on those for whom grants were provided in 2003-2005 (GFY03-05). This time period was chosen to allow time (no less than two years) for the grantees to complete their work, get published, and have measurable career advancement and to have clear, non-overlapping cohorts for each evaluation. Analyses of this cohort versus all other respondents, emphasizing use of a Wilcoxon signed-rank test that assumes only ordinality of data, resulted in no significant differences across nearly all questions of interest. For this reason and because the cohort size was relatively small, the evaluation primarily reports results for all respondents.

⁸ The evaluation was unable within the time and resources available to perform an analysis comparing curriculum vitae of grantees and matched non-grantees.

fact, this likely bias is the reason GDN indicates they did not attempt to make such comparisons in its past internal evaluation. Some bias remains if selection is based on unobservable characteristics that are not measured in the dataset.⁹ The evaluation does not emphasize the results from this approach given the difficulty in dealing with potentially different selection processes across regions and the potential for selection based on unobservable characteristics. It suggests more detailed use of the technique under less constrained time limits.

Although only basic descriptive tables are provided in the annexes (See Annex E: Basic End User Survey Tables), all anonymous data sets and data output runs are available through GDN or the primary investigator.

1.3 Description of Survey Samples

In order to assess GDN's organizational performance and effects quantitatively and cost effectively, the evaluation collected online survey responses from key stakeholders and end users based on GDN's previous success in being able to do so. The table below shows by respondent categories the survey question issue areas related to GDN's core activities asked of each respondent category. The questions on organizational assessment adapted from the SIDA Octagon Model were restricted to the stakeholder survey.

Table 1.1: Sample Sizes and Issue Areas Asked of Respondent Categories

Categories of Sample	# of Respondents / Response Rate	Question Category					
		Organizational performance	Global research projects	Regional research competitions	GDNNet and newsletters	Annual conference	Awards and medals competition
Key stakeholders							
Board members ¹	14 / 64%	√	√	√	√	√	√
RNP heads ²	9 / 41%	√	√	√	√	√	√
GRP + Regional window coordinators	3 (GRP) + 4 (regional windows) / 78%	√	√	√	√	√	√
GDN staff and interns	31 / 100%	√	√	√	√	√	√
Donors ³	12 / NA ⁴	√ ⁵	√	√	√	√	√
End Users							
RRC grantees	142 / 35%	No	√	√	√	√	√
GRP grantees	54 / 37%	No	√	No	√	√	No
AMC awardees	65 / 28%	No	No	No	√	√	√
Broad Constituency	1224 / 16%	No	No	No	√	√	No

⁹ This technique matches treatment cases from the set of grantees with controls from the broad constituency survey respondents. It matches people based on observable characteristics that generate a single score measuring the propensity, or likelihood, that any given individual would be selected for treatment. Cases from the control group that have a similar propensity score are matched with those in the treatment group. In the evaluation's analysis, the matching is made using position, experience, age, discipline, and region. See discussion in Rosenbaum and Rubin (1983) and Heckman, Ichimura, and Todd (1998), among others for motivation and greater detail.

Notes:

¹ 14 board members responded to the stakeholder survey and three who did not fill the survey were interviewed. Of the 14 board members surveyed, six were also interviewed. The board members include both current and members who left the board between 2005 and 2007. (See Annex A for list of interviewees.)

² Nine RNP heads completed the stakeholder survey, and three were interviewed additionally. (See Annex A for list of interviewees.)

³ Eleven donors responded to the stakeholder survey and one was interviewed and responses entered.

⁴ The evaluation surveyed a small, non-random sample of all active donors and so does not consider the 50 percent response rate to those surveys to represent a response rate relevant to the entire universe of past, present, and potential donors.

⁵ Due to expected time limitations, donors were asked only half of the organizational assessment questions but were asked additional questions about donor relationships with GDN.

The evaluation collected quantitative data from 73 stakeholders that included board members, RNP heads, GDN staff and interns (in New Delhi, Cairo and Washington D.C.), donors, GRP coordinators, and RNP window coordinators. Response rates among most stakeholder groups were reasonably high. To avoid very small respondent categories from which responses might be identifiable and to maintain the integrity of the clear sets of board and RNP heads and because there were no other clear categories in which they fit, the evaluation team combined GRP coordinators and regional windows coordinators together. The responses from this category most often were not statistically significantly different from those provided by staff and interns. Annex F includes a justification of the categories used in the stakeholder survey and the rationale for considering stakeholder perspectives in addition to those of end users in response to questions about doing so from GDN. Annex G provides tables of responses to closed-ended questions by stakeholder group.

Since the GRP grantees group that responded to the online survey is relatively small (54), they sometimes are combined with RRC grantees to create a larger sample size regarding GDN capacity building activities if no statistically significant differences in responses is found. The total number of researchers registered online on GDNNet who completed the survey was 1,224. Response rates for end users are typical for online surveys without extensive follow up especially given its long length to capture many indicators. They are sufficiently low however – between 16 and 37 percent – that there are strong reasons to believe respondents are biased towards those most satisfied with GDN's activities. (See Table 1.1.) Response rates are much higher for stakeholder surveys.

Grantee survey respondents are roughly three quarters male and economists, while the broad constituency survey respondents are only 42 percent economists. Most grantee respondents have their Ph.D., consider themselves senior level, and mostly are primarily affiliated with academic universities, although again each of these figures are slightly lower for the broad constituency survey. Supporting tables with these basic descriptive data regarding the end user samples are provided in Annex #. The evaluation generally reports differences only when statistically significant at the $p \leq 0.05$ level, although it notes noteworthy differences that are statistically significant up to the $p \leq 0.10$ level.

1.4 Structure of Report

Due to space limitations, the evaluation does not describe GDN and its components in any depth, assuming readers have some familiarity with the organization or have read summary information provided in the TOR or elsewhere. Each section or sub-section in the report begins with the section's point and key findings.

This section provides the background and methodology of the evaluation and defines the sample selection as well as outlining limitations of the study. This evaluation's findings begin in section 2, which discusses relevance followed by outcomes, effects and impact; reach and multidisciplinary; and cost efficiency or cost effectiveness. This is followed by a section on overall success. Section 2.3, "Outcomes, Effectiveness, and Impact," is the longest in section 2 and analyzes findings by activity. Section 3 discusses findings on process

and performance of GDN related to financial and risk diversification and on governance and management. The fourth section provides conclusions of the evaluation and highlights recommendations for GDN.

1.5 Limitations

The evaluation study covered most methodological aspects outlined in the Terms of Reference but was altered due to tight time constraints and unexpected hurdles that impeded progress, adding some analyses not requested and forgoing others. These adjustments, as well as other limitations of the analysis are discussed in Annex H in order to conserve space.

2.0 FINDINGS: RELEVANCE, REACH, OUTCOMES, EFFECTIVENESS, IMPACT, AND COST EFFECTIVENESS

In this section, the evaluation examines issues of relevance and reach; outcomes, effectiveness and impact; and cost effectiveness of GDN's funded activities. The evaluation focuses more on outcomes and effects than impact given that the relatively young age and modest scope of GDN's activities are likely to make only modest impacts in targeted regions and fields that the evaluation is not tasked to measure. The evaluation, however, encourages GDN to think carefully about donor desires to see tangible, if not measurable, broader impact.

2.1 Understanding GDN's Objectives in Order to Measure Relevance, Reach, Outcomes, and Impact

The evaluation's TOR relies on the *Statement on the Governance of the Global Development Network* to define GDN's "core objective" as "building capacity to *generate, share, and apply to policy multidisciplinary knowledge for the purpose of development.*" However, there remains a multiplicity of GDN mottos and statements about its objectives, as discussed in the previous evaluation. Further, lack of clarity about objectives was a key theme in the previous external evaluation, and current donors are interested in understanding progress regarding key concerns of that evaluation. Thus, an important first step in evaluating relevance and outcomes and to answer the overall focal question for the evaluation ("Has GDN achieved its stated objective?") is to clarify GDN's objectives and target beneficiaries in practice.

The evaluation finds evidence there probably is somewhat greater agreement regarding objectives and target beneficiaries and how they are manifested than during the period previously evaluated. Interviews with some key stakeholders indicate the lack of clarity along some of GDN's objectives created considerable tension and internal problems for board members, particularly toward the beginning of the time period analyzed. The interviews also suggest that over time and with changes on the board, these tensions have become less pronounced. In the stakeholder survey, GDN is rated on average (assuming cardinal responses) as "good," or a 3 on a scale of 1 (non-existent) to 7 (excellent),¹⁰ in terms of formulation of its objectives and mission indicating both generally positive responses and continued room for improvement.

GDN's portfolio approach towards its many and varied objectives and target beneficiaries generates some disagreements, as discussed in particular in the section on Reach and Multidisciplinarity. Based on interviews and desk research, prior to analysis of survey results, the evaluation had decided the most likely best operationalization of GDN's primary objective for evaluation purposes was "building methodologically sophisticated policy-relevant social science research capacity and facilitating exchange of knowledge across the Global South." This definition does not reference individuals or institutions, nor on the level of capacity, as those foci are unclear.

Results of the stakeholder survey questions about specific objectives suggest relatively strong agreement that GDN in practice emphasizes developing country researchers, generation of development-focused research, and providing a global forum for the exchange of ideas. These three can be categorized as being rated on average (assuming cardinal responses) that they "mostly" "represent GDN's objectives in practice" which still is slightly lower than expected given their centrality to GDN's mission. The objective ranked lowest by stakeholders on average are ranked around "about half-way." (Summary tables are shown in annex I.)

The stakeholder survey results suggest GDN emphasizes supporting individuals more than institutions and facilitating policy relevance more than policy impacts or links with policy makers. There is an implicit tension

¹⁰ All stakeholder survey questions derived from the SIDA model are seven point Likert scales rated, in order, as Excellent, Very good, Good, Reasonable, Weak, Very weak, and Non-existent. Don't Know is provided as an option as well.

between the often stated emphasis on capacity building¹¹ and its often stated sub-goal of supporting only the highest quality research, researchers, and institutions. The ratings suggest more agreement that in practice GDN builds research capacity at many levels of quality than that it supports only the highest quality research, researchers, and institutions. Stakeholders also believe GDN in practice facilitates cross-fertilization of research more within than across regions and both of those less than generally facilitating knowledge sharing of developing country research.

The average rating given by key stakeholders is strongly correlated with the extent to which stakeholders believe change is necessary ideally. For instance, both facilitating cross-fertilization of research across regions and emphasizing linking between policy makers and research are rated nearly at the bottom of all objectives listed. Both also receive the highest proportion of respondents indicating a change is needed at 69 to 73 percent – predominantly indicating an increase is needed. Some internal contradictions and lack of attention to fiscal constraints are notable since 42 percent indicated a need to increase the extent to which GDN supports only the highest quality research (with another 49 percent not indicating a need for change), but 60 percent desire an increase in the extent to which GDN builds research capacity at many levels of quality.

Due to these disagreements, respondents on average rated GDN's "Formulation of Objectives and Mission" as "good" (or third from the highest rating between 1 (non existent) and 7 (excellent) on a seven point Likert scale), with a quarter rating GDN here as "reasonable" or lower. This suggests room for improvement, but the results are arguably higher than might have been predicted from the last evaluation and suggests GDN may not be as weak as some organizations along these measures. Related to this dimension, the evaluation asked stakeholders how well GDN has formulated realistic strategies, courses of action, and sub-goals linked to its objectives and integrated them into its planning processes. The average response is much the same as formulation of objectives, and 29 percent rated GDN here as "reasonable" or a four on the same scale or lower. Although the evaluation has observed that GDN has clear work plans and routines for much of its work, the linking of strategies to objectives by management remains an area of perceived need. This perception is not surprising given its linkages with the formulation of objectives and mission and weaknesses noted there.

2.2 Relevance – Summary

The evaluation next considers relevance and effectiveness of GDN activities. It focuses on the extent to which current activities meet the needs, type if not in magnitude¹², of the intended beneficiaries. It also examines the extent to which the specific activities run by GDN are competitive with efforts by other organizations, complementary to them, or unique.

Among GDN stakeholders, if not secretariat, there is modest disagreement on the balance between targeting individuals and institutions as well as types of individuals and institutions that should be primary beneficiaries. GDN attempts to solve these debates by targeting multiple beneficiaries through a portfolio approach of activities. The evaluation finds that GDN's use of a portfolio of activities is an appropriate approach to meet various needs across and within regions and across funding constraints by donor, to deal with multiple objectives that are not agreed upon uniformly, and to diversify and sustain itself as an organization reducing somewhat its dependence on its founding donor.

¹¹ GDN has responded to the evaluation team in writing noting that neither terminology are found in the mission statement. As but two examples, capacity building is listed as its third bullet point on its web page as of December 2007 under "what we do" "**Build research capacity** to advance development and alleviate poverty." (Emphasis in original.) In a briefing book for the incoming president dated May 2007, GRPs are described in the following way: "They balance GDN's twin goals of generating high-quality research and building research capacity...."

¹² This is discussed instead in the section on outcomes, effectiveness, and impact.

Due to GDN's plentiful objectives and beneficiaries, all parts of the portfolio conceivably have relevance to different targeted beneficiaries. RRCs are seen as the most relevant and complementary, if not entirely unique activity, but it also is the most difficult for which to raise funds. Those stakeholders who know most about GDN believe it is relevant and those who do not note suggestions that parallel some of GDN's initial and long-term goals for the program. Multiple stakeholders indicate greater needs for improved training of individual researchers, and GDN has recently begun to assess how to meet these needs through a training program to increase relevance and effects of its activity mix. GDN's support for institutions thus far has been narrow and *ad hoc*, although a pilot project on bridging research and policy is planned. Further work with institutions could be highly relevant but would depend on additional resources and development of an inexpensive, efficient mechanism for delivery. Somewhat reduced costs and greater institutionalization of the annual conference has led to stronger support now for the conference among many stakeholders, although the conference remains the most contentious activity with suggestions that GDN should associate itself with conferences other than the Bank's ABCDE to better complement capacity building objectives. GRPs are considered a highly unique activity but also the next most frequently discussed activity as being insufficiently relevant probably due to selected GRP themes and the theme selection process rather than the concept of the GRP itself.¹³ The MIDP and medals competitions also are perceived as low relevance, but their small size and earmarked funding does not much negative attention.

2.2.1 Relevance to Targeted Beneficiaries

The definition of intended beneficiaries is not straightforward. Within GDN, there is modest tension if not disagreement on the balance between targeting individuals and institutions as well as types of individuals and institutions that should be primary beneficiaries. One such tension is regarding the balance between targeting top researchers – a goal stated by a few board members as well as mentors/resource persons as well as those familiar with GDN's first GRP, "Explaining Growth" – and targeting researchers who meet some region-specific threshold of quality but who still need some form of capacity building – potentially quite a bit.¹⁴ GDN attempts to solve these debates by targeting multiple beneficiaries through a portfolio approach of activities. The evaluation's analysis assumes that both categories therefore represent targeted beneficiaries, and that emphasis of GDN's objectives target individuals. It acknowledges but does not attempt to resolve a concern of some stakeholders that by donors targeting money primarily at individuals, the strength of their associated institutions could be undermined in some regions by allowing researchers to act as free agents if sufficient funding is available. This may be a significant concern only in Africa currently.

GDN has undertaken surveys of researchers regarding their perceived research needs to see how target beneficiaries' needs align with activities. Nonetheless, there is not uniform agreement among stakeholders regarding the largest needs of either individuals or institutions, even assuming that this choice could be made

¹³ GDN asks the evaluation to note here that GRP themes are selected through a consultation process and approved by the board, although that consultation process is not considered entirely transparent to many – an issue raised in the prior evaluation as well.

¹⁴ GDN's initial written comments contested the evaluation's representation of this tension stating that GDN does not target the 'best researchers' but rather the ones who demonstrate talent and potential and are still in need of capacity building." In addition to the previous footnote citation, the evaluation team notes that its understanding of GDN's first GRP – Explaining Growth – was that most RNPs chose the best experts they could find on the relevant issues rather than expressly looking for people who in some way required capacity building. Further, some RNPs indicated that some funds were allocated toward researchers who they knew needed no capacity building – only to have some time bought to generate good research.

entirely independently of donor/funding demands.¹⁵ More importantly, there is disagreement regarding the best ways to reach GDN's various objectives with no conclusive research evidence regarding the optimal mix of approaches required across different regions. As one stakeholder noted, this means GDN's activities represent large, uncontrolled experiments with modest time series and comparisons possible.

Despite the assertions of some stakeholders to the contrary, GDN's general contention that a mix, or "portfolio," of activities best meets its objectives has merit based on the concept that the ideal needs may vary across and within regions and across funding constraints by donor. A mix of services allows the potential for follow-up support for researchers whose capacity is built and provides a variety of funding options for donors who only earmark. It is also a key way that GDN management attempts to solve the otherwise unsolvable conundrum attempting to meet these varying needs across multiple objectives that are not uniformly agreed upon.¹⁶ This approach also permits GDN to diversify and sustain itself as an organization through different donor channels so that it can meet its commitment to the World Bank (discussed later) in reducing dependence on its founding donor. The question then becomes which parts of the portfolio may be less relevant to targeted beneficiaries or what components need to be added. Despite GDN's request to do otherwise, the evaluation organizes the relevance section by discussing activities found to be more contentious rather than those with the strongest agreement. The rationale is that the highest value of an evaluation is in looking at areas of disagreement rather than agreement, but all are considered ultimately.

The most contentious activity among stakeholders interviewed in terms of relevance and uniqueness is the annual conference. The concern about the annual conference remains from the prior evaluation although it appears to be somewhat lower than in the previous evaluation. This may be due to several factors including somewhat decreased costs of the conference over time and the possibility it has established itself on the development community agenda and has permitted stakeholders and donors to network successfully. Some stakeholders as vehemently support the conference as a relevant activity to stakeholders (and as a necessary to GDN for reputational and other purposes). Interestingly, GRPs are the next most frequently discussed activity as being insufficiently relevant. Interviews suggest the reason for this are selected GRP themes rather than the concept of the GRP itself, as all interviewees seem to identify at least one GRP they thought was extremely relevant, even though their preferences sometimes were diametrically opposed.

The stakeholder survey asked respondents to rank the value, and thus implicitly the relevance, of its main activities "on average for targeted participants" using a five point Likert scale.¹⁷ Responses varied widely and are discussed comparatively in a later section. It should be noted that since the evaluation did not ask a separate question about effectiveness, it is possible that respondents considered both relevance and effectiveness in their responses.

Considering responses of stakeholders as groups, all stakeholder groups gave RRCs an average rating between 1.2 (close to 1.0, or "very valuable") and 2.0 ("Between very and moderately valuable") with only a few

¹⁵ In response to this statement, GDN notes that they "conducted surveys to explore the needs of the researchers in developing countries (both in 1999, 2005 and 2007 [sic]) and thus there is no need to ask (often less informed) stakeholders what the needs of the target beneficiaries are when the direct responses and assessments are available." The evaluation team believes stakeholder understanding of target beneficiary needs is relevant since boards ultimately are responsible for assuring that activities meet needs. Further, survey responses do not necessarily always lead to straightforward interpretations of priorities.

¹⁶ A quantitative metaphor could be of trying to solve a large series of equations with disagreement over constants, number of variables, and perhaps arithmetic expressions, resulting in widely varying solutions by person.

¹⁷ This scale is "Very valuable," "Between very and moderately valuable," "Moderately valuable," "A little valuable," and "Not at all valuable."

indicating “A little valuable” or “not at all valuable.”¹⁸ This, combined with discussions appears to represent a strong consensus on the relevance and probably effect of RRCs. The consensus is less strong concerning the relevance and effect of GRPs. In part, this appears to be due to less clarity on the objectives of the activity in relation to both researcher needs and to GDN’s mission and to a lesser extent a sense that they are less effectively run than some of the RRCs. The average rating among staff and GRP coordinators for GRPs is 1.6 to 1.7¹⁹, while board members and RNP representatives rate it 2.4 to 2.9 on average, the latter essentially equal to “Moderately valuable.”

Board members on average found GRPs and the Awards and Medals and MIDP competitions least relevant to targeted beneficiaries, while RNP representatives indicated GDNNet as least relevant. The primary question about the MIDP is its relevance to a research-focused organization and the long term value of the medals competition for researchers. Given that these programs constitute small amounts of earmarked funding and have had a long-term presence in the portfolio, they appear tolerated by almost all stakeholders even if they question their value or quality of awardees’ work.

Some of the respondents who rated GDNNet least relevant did note suggestions, however, about how to increase relevance. These parallel some of GDN’s initial goals for the program, where they have encountered resistance from providers – to use GDNNet’s regional windows or regional partner websites as a portal to make as many recent journal articles as possible available to registered developing and transition country researchers without access through their institutions. A similar suggestion is to make GDNNet a portal to leading working paper and database collections worldwide. Other suggestions to increase relevance include serving as a focal point for changing access policies of national datasets by setting up an international model program of data access from national institutions in the Global South, which currently often severely limit access. The evaluation has been told that GDN now requires all grant recipients in their contracts to make public the access to their data – a policy that provides a public good that individual researchers not used to such a policy may want to resist. The policy does not indicate where the data will be made available and so the evaluation cannot assess the extent to which this policy is being followed, so stakeholders have suggested GDNNet serve as a central location for all researchers to find the datasets and to confirm its policy is followed.

Multiple stakeholders, including resource people working with grantees, have indicated greater needs for improved training of individual researchers in most developing regions with varying solutions to this problem accompanied by widely varying implied costs and depth and breadth of impact. GDN conducted its own limited training needs survey for 30 non-randomly selected individuals across each region, with a total of 167 respondents. Because the respondent sample comes from the networks’ constituents, it probably is dominated by economists and not necessarily broadly representative of the needs of individuals within each region across disciplines or even necessarily within disciplines. Nonetheless, results suggest differing needs by regional respondent groups, with the desire for training in quantitative techniques and within specific research fields most highly uniform, but generally strong interest in all trainings suggested on average (See Annex J “Regional Training Survey: Preliminary Results”). GDN has indicated recently that it will attempt to put a separate training program on its long-term agenda to increase relevance of its activity mix. The evaluation notes, as have some stakeholders, that scheduling individual training to correspond with RRC workshops and the beginning of RRC or GRP processes could reduce total costs and augment the effect of each.

In terms of institutional support, GDN efforts thus far have been narrow and *ad hoc*. GDN does provide institutional suggestions to their network partners on an as needed basis and provides a short annual meeting of RNP heads at the annual conference for several purposes including comparing notes. Stakeholders indicate that this frequency and length of meeting, amidst the conference’s other distractions, is insufficient for real

¹⁸ These averages assume cardinal ratings.

¹⁹ GDN asked that regional window coordinators be excluded here, although doing so makes no difference in results.

institutional learning. GDN also has begun to advance the concept of institutional training related to the findings of its Bridging Research and Policy GRP with a pilot project with a newly formed policy research institute in the Caribbean. As with the rest of GDN's model of capacity building, it draws on resources of external resource people rather than serving as a source of training itself. A small number of stakeholders suggest that for GDN to have a larger impact on policy, it should branch out into working with more institutions including policy research institutes and government agencies. Some other stakeholders have suggested that for GDN to have a larger impact on capacity, particularly in areas with few strong institutions, GDN should branch out into working directly or through RNPs with more research institutes or networks to build capacity as institutions.²⁰ The evaluation recognizes that working with institutions would be highly relevant and could expand GDN's effects, and is indeed an area GDN is beginning to pursue. However, given limited resources, staffing expertise, and evidence regarding best ways to have an impact, it is unclear the extent to which GDN should pursue institutional support as a focal area barring considerable funding and development of an inexpensive mechanism for delivery. Funding success and trial-and-error likely will end up being the way forward.

2.2.2 Uniqueness

Another aspect of relevance is the extent to which the activities run by GDN are unique, complementary to efforts by other organizations, or competitive with them. It appears universally agreed that RRCs are considered needed and a highly complementary activity for all RNPs (except LACEA) if not entirely unique. In Latin America, the annual conference funded by GDN is mostly unique in scope and appear considered by stakeholders to be highly complementary to the activities and needs of the economics-focused network. The journal also sponsored by GDN funds is considered to be unique within the region and language in terms of accessibility and outreach to policy makers as much as academics. LACEA representatives indicated an interest in beginning an internship program to increase its currently limited efforts in increasing capacity in weaker countries and areas.

GRPs are considered by all to be a highly unique activity in its multi-regional scope and focus on perspectives from the Global South. Their structure of thematic selection and governance is not, however, considered to be set up to be complementary to RNP efforts and are sometimes seen as competitive for funding and allegiance of researchers. The implied model for the structure of GRPs, with the exception of the first GRP, is clearly one of a centralized international organization rather than a network of networks, which again creates confusion among some regarding which model is closer to GDN's mission. GDNet generally is considered unique in terms of its focus on researchers in developing and transition countries and complementary with RNP websites as well as those of at least one portal, Eldis, with which it partners in delivering access to Southern researchers.

The annual conference is considered by some stakeholders to be the least unique in that many other organizations and disciplines hold annual research conferences, although the GDN event is not perceived as directly competitive with other conferences. Other stakeholders believe that it is unique in respect to bringing together researchers from across multiple regions, offering at least the potential of South-South as well as North-South cross fertilization. In recent years, the GDN conference has been held alongside the World Bank's ABCDE conference. This itself suggests complementarity with that conference, although traditionally it has considerably less focus on the perspective of the developing and transition world, and stakeholders did not offer teaming with that conference as great value added (although they were not questioned directly about it). A few stakeholders indicate that in terms of meeting researcher needs, GDN's conference might complement better regional partner annual conferences or provide greater value added to junior researchers if paired (perhaps in rotation) with other disciplinary conferences.

²⁰ Stakeholders generally support GDN's general approach of funding researchers grants relatively directly with the individual to prevent "mediocre institutions from rent seeking," to quote one stakeholder.

2.3 Outcomes, Effects, and Impact

2.3.1 Outcomes of RRC and GRP Capacity Building Activities

The evaluation's review of outcomes of GDN's RRC and GRP capacity building activities suggest increases in knowledge created and capacity built for individuals funded across most regions. The evaluation found only modest variation in publishing productivity across regions in part given the small sample size. It did find, however, differences in levels of support and mentoring for RRCs across regions as well as perceived value added through the process, with particular concerns in the Asian regions. The outcomes and effectiveness of the different GRP activities receive mixed reviews from stakeholders, especially given the higher expectations and resources provided per unit output. Grantee responses regarding value added was correlated with the amount of feedback and value/quality of review sessions suggesting a need to more consistently encourage feedback and mentoring of RRCs. Little evidence was found of capacity built from the GRP and RRC processes as measured in improvements from proposals to research paper from pilot test research, although limitations of proposals may explain the non-findings.

Although all GDN activities have some element of capacity building, this section focuses on the outcomes and effectiveness of RRCs and GRPs as the two activities most frequently described as having a more direct objective of capacity building.²¹ This section also discusses the AMC, where similar measures are relevant although only the awards portion is seen as building capacity. As discussed in the introduction, increase in knowledge on development issues from the perspective of the Global South generally is measured through responses to survey questions about amount of research output generated as well as external reviews of publishability of the final papers. Capacity building is measured through responses to survey questions about improvement in skills among grantees as well as through a longitudinal review of the change in quality from proposals to final papers among grantees.

The evaluation also examines effects from GDN's RRC and GRP capacity building through a two-step non-experimental multivariate method called propensity score matching. This is a relatively new technique in the evaluation field that creates matching "control" cases or groups relative to "treatment" cases matching people on observable characteristics that generate a propensity score that any given individual would be selected. Lacking a proper control, such a technique is the best method to find comparable observations that could be used to simulate a control group. The evaluation does not emphasize the results from this approach given the difficulty in dealing with potentially different selection processes across regions and the potential for selection based on unobservable characteristics as well as concerns in the research literature about biases that might remain in several situations such as when there are low response rates, as found here.

The evaluation also considers responses of key stakeholders to a survey question on the value of the RRC and GRP activities on average for targeted users. In addition, the evaluation analyzes survey responses of grantees and interviews of key stakeholders to assess effectiveness of these activities. Qualitative feedback regarding how to extend and improve effectiveness of GRPs include consideration of ways GDN could better facilitate partnerships between universities and research institutes represented by its resource people and those in developing countries. For instance, the India Development Foundation, GDN's partner in coordinating the Health GRP, recognized the opportunity to bring together the resource people with research expertise on health issues to themselves work on publishing with researchers from the South both remotely and in person. They also arranged "reverse" fellowships with young Northern researchers getting the opportunity to work together

²¹ Both activities also have networking and knowledge generation objectives as well, discussed elsewhere.

with and learn more from the Southern perspective. This “collateral damage” from the Health project generates strong mentoring opportunities that GDN cannot claim directly, but it might be well advised to figure out how to capture that intentionally.

2.3.1.1 Grantee Survey Self-Reports

To get an idea of total output on average from RRCs and GRPs, the evaluation reviews responses from grantees without considering any weighting for representativeness. Combining the four primary publication types together for each respondent, the median is two publication types per grant, and an average of 2.9 due to the 10 percent of prodigious individuals claiming to publish between six and 12 total publications as a result of the grant and only 10 percent indicated publishing nothing (see Annex K). No significant difference was found between RRC and GRP types. Taken individually, on average one of every two grantees claims to have published in an international journal, two of three in a national journal, two of three as a book chapter, and just over one working paper per person.²² The evaluation team urges caution in relying on the results given potential upward or positive biases given the response rate. Amount of publications was correlated (at the $p \leq 0.01$ level) with the number of review sessions and frequency of feedback between workshops as well as the perceived value of the workshop review sessions of research reports and was not correlated with seniority at time of grant. This research-based evidence complements the informal suggestions from grant interviewees of the value of more concerted efforts at improving the quality of workshop sessions and more consistent mentoring across all RNPs for those in need.

Analysis by region indicated higher overall research productivity²³ only for RRC grantees from CEE (although also for CIS at the $p=0.11$ statistical significance level) and lower productivity for grantees from Sub-Saharan Africa. Although each region varies greatly in the training of those being chosen and their research environments, the evaluation notes that the RNPs with highest productive output both provide designated mentors (sometimes paid) for those judged in need and take pains to make sure the workshops are well attended and include a variety of inputs. CIS respondents reported producing twice as many policy briefs as the others, an effort championed by the former RNP coordinator, while those in Sub-Saharan Africa produce fewer.

2.3.1.2 Propensity Score Matching of RRC versus General Survey Self-reported Output

In this section, the evaluation examines effects from RRC and GRP capacity building activities through a two-step non-experimental multivariate method called propensity score matching. This technique matches “treatment” cases (grantees) with “controls” (broad constituency survey respondents) based on a single score measuring the likelihood any individual would be selected for treatment. The matching is made using whether academic discipline of training is economics and region. It also includes measures at a set reference year position level (junior, mid-level, senior), years of experience in research, and age. The evaluation does not find positive effects from RRC and GRP capacity building.

²² Some caution must be taken in interpreting results on the one hand because people may exaggerate their productivity and, on the other hand, because the evaluation has been told that some GRP participants might hold back their best work as outputs so that they can be published in refereed journals rather than as GDN publications. Although the GDN requirement for publishing in a GDN volume for some GRPs also could lead to higher publications among GRP respondents as book chapters, the difference was not statistically significant, although GRP respondents did claim to publish fewer articles in national journals.

²³ Performed by adding together the first four measures of productivity and using region of birth as the regional indicator. Differences for CEE were due to publishing working papers and international publications and for CIS for working papers and national publications. There were too few respondents from LAC to analyze.

The first step of the process is a multivariate probit (binary choice) regression that was found to be statistically significant and explain 23 percent of the variation in the dataset.²⁴ This is an acceptable level given the sample size to continue with the matching process. This level might suggest, however, that some selection is occurring due to unobservable (or at least unmeasured) factors – a conclusion which the evaluation team itself believes is likely given its own discussions with selection committee members in several regions. It is difficult for the evaluation to measure likelihood a person has publishing potential and yet needs capacity built without using the variables intended as outcomes – publication record. For this reason, the evaluation uses seven different methods of matching to look for results that are robust across these different methods. The evaluation also ran all analyses using both the GRP and RRC grantees combined and then just the RRC grantees alone.

The next step in the methodology is to match cases from the treatment and control groups and compare average effects of the treatment in terms of outcomes. The evaluation uses seven different measures of outcomes relative to a reference year (for the grantees, the year they first received a grant). The outcomes include receiving a promotion, received other research grants, and five indices of publications that range from just having published in national or international journals to including all types of publications and presentations. The survey was designed originally to exclude publications sponsored by GDN in the analysis, but the evaluation team later included both GDN-sponsored and other publications in two of the outcome indices. The second stage of matching and analyzing estimated treatment effect was run first for all regions. Next, it was run for just two regions identified as having similar criteria for selection (CEE and CIS) to see if results differ isolating individual regions. Finally, it was run just for those respondents using reference years 2003 or earlier to provide the longest period for differences in outcomes.

In all cases, the analysis found either negative (the control group showed higher publishing productivity or other outcome measures) or statistically insignificant results. The results that generally came closest to that expected (positive, or the treatment group showing higher productivity) is when GDN-sponsored papers were included and when outcomes included all types of publications and presentations included in the questionnaire including working papers, book chapters, policy briefs, articles for popular media, presentations in academic events, policy briefings, and press conferences. The evaluation does not suggest from this analysis that GDN's activities actually have a regressive effect but that greater thought be considered by GDN for creating control groups in at least some of its work in the future if it wishes to find valid evidence of its capacity building. GDN should have independent analysts undertake additional, more careful analysis that is less constrained in terms of resources and time using the same data and types of matching techniques to explore carefully the matches made and whether changes to the analysis might lead to superior matches.

2.3.1.3 External Review of Research: Final Paper Quality, and Proposal to Paper Review

Final Paper Quality

A weak measure of development knowledge built from people in developing and transition countries – regardless of level of capacity built – is the quality of research output generated, as measured through external reviews of at what level final papers are judged publishable.²⁵ This type of analysis is deficient in that no measure is collected of the probability that papers funded by GDN would not have been published anyway nor what GDN's contribution has been in terms of building capacity. GDN previously conducted a similar review, selecting outputs from across all of their activities for FY05. In each case, just under half were judged to be publishable in some type of journal without revisions and between two-thirds and four-fifths with revisions and

²⁴ This figure is slightly higher when using just 2003 and earlier data and much lower (8 percent) using just two regions, as discussed later.

²⁵ This analysis focuses only on one RRC and one GRP because the primary interest was using a review process to measure capacity built using a longitudinal review of proposals to papers, discussed below.

sometimes at higher levels.²⁶ (See Table 2.3-1.) Few papers were found unworthy of at least working paper status, admittedly a low threshold but a reasonable one to consider given the weak capacity level in certain areas within all regions and across most areas in some regions. Given the modest sample sizes and differences in judgments provided for each set of papers, no statistically significant difference was found by set of papers. (See Table 2.3-2.) This non-finding is mildly surprising given that the pool of RRC papers draws from only one region with smaller grants averaging around \$11,000 (and generally smaller teams and potentially more junior talent) and, in theory, less intense mentoring than GRP processes. The evaluation expects somewhat lower figures would be likely if using some other RRCs.

Table 2.3-1: External Review of Publishability of GDN-funded Activity Outputs, One RRC, and One GRP

	Across All Activities FY05*		CERGE FY03-05		Impact Project	
	Current Form	With Revisions	Current Form	With Revisions	Current Form	With Revisions
International Refereed Journal	24%	48%	8%	42%	4%	28%
National Journal	24%	30%	37%	32%	40%	32%
Chapter in volume	18%	3%	26%	21%	12%	20%
Working Paper	18%	9%	18%	0%	28%	16%
Not publishable	15%	9%	11%	5%	16%	4%
Total	100%	100%	100%	100%	100%	100%
N	33	33	38	38	25	25

Note: Provided through GDN 2004 internal evaluation.

Table 2.3-2: T-test statistics for External Review of Publishability of GDN-funded Activity Outputs

	Current form			Revised Form		
	All Activities vs CERGE	All Activities vs Impact	CERGE vs Impact	All Activities vs CERGE	All Activities vs Impact	CERGE vs Impact
T-test probabilities	0.72	0.30	0.42	0.86	0.28	0.17
T-test statistics	0.36	1.04	1.04	0.18	1.09	1.09

GDN has suggested the evaluation team apply the publishability proportions found in Table 2.3-1 across all of its GRP and RRC publications. The evaluation team is reluctant to do so however. As the report discusses earlier, the evaluation team expects that the results from the CERGE RRC might be representative of some, but not all, regions, as supported by stakeholder comments discussed elsewhere. The same is true of the Impact evaluation representing all GRPs. The only broad measure across activities is that generated previously by GDN, which it could use for such purposes regardless of the evaluation.

The evaluation does not directly disagree with GDN's characterization of the previous results as indicating that GDN's processes, given the dual aim of capacity building discussed below and limited coverage of some activities and regions, generally results in value added to the base of knowledge in development from the Southern perspective albeit diverse and of variable quality. It only disputes the extent to which GDN can claim that all of GDN-funded papers represent knowledge added. The next evaluation should attempt to include a question to help measure the likelihood the team would have produced the document absent the funding, and GDN itself can assist by attempting to collect such information directly through its contracts or some other means. In those cases where the team would have generated the research anyway, GDN's potential contribution

²⁶ One reviewer noted that this is generous in that most papers to any journal require some revisions before publication and that greater clarification should be provided reviewers about extensiveness of revisions "in current form."

would be capacity built in most, although not all, cases. In using the above figures correctly for promotional materials, GDN should indicate that these represent upper bounds of the intellectual contribution added via GDN funding.

Proposal to Paper Review

The evaluation team agreed upon a pilot longitudinal review of proposals to papers (PtP) analysis using grants for fiscal year 2003-05²⁷ for one RRC (CERGE) and one GRP (the Impact Project) as an initial test toward how to measure increase in paper quality, as a proxy for increase in capacity built through the process of receiving feedback on proposals, interim presentations of findings, and final reports, and any other mentoring processes. The review process asked established researchers to judge first the proposals and then the papers along five criteria²⁸: originality, clarity, methodological appropriateness, policy relevance, and overall quality. Proposals and papers were readily available for one RRC (CERGE) that has a highly routinized process managed by the same organizers and involving many of the same volunteers since 2001 and includes an assigned, remunerated mentoring process aimed at grantees judged in highest need given available funding. For this reason, results were expected to represent toward the top of the distribution of capacity building potential among RRCs.²⁹ All papers were selected from the Impact project and 38 selected from CERGE, suggesting around 0.8 power for each at $\alpha=0.05$ for a medium effect size ($\delta=0.5$).³⁰

The key problems encountered in the pilot test involved lack of specificity in the proposals, especially for RRC researchers. The evaluation team believes RRC researchers in their proposals promised more than they could or did deliver substantively and were insufficiently specific regarding methodology leading to high expectations among reviewers regarding methodology to be used. In these cases, reviewers – lacking information on the grantee's training and capacity building needs and not having been asked to account for differences from Northern training – generally provided *lower* assessments for final papers than proposals along several criteria.

On average, the respondents indicated a small decrease in quality from the proposals to the final papers along four of five dimensions for the RRC papers with an average increase only in terms of clarity. (See table below.) The respondents indicated small increases on average along four of five dimensions for the Impact project papers. These small differences were statistically significant, however, only in terms of a decrease in policy relevance for CERGE at the 0.05 level and for increases in overall quality for Impact papers at the 0.10 level. To consider the hypothesis that capacity building would most likely be found among only those requiring significant capacity building, the evaluation examined proposals across both sets of papers rated below 3.5 as a proxy for need for capacity built. This analysis resulted in small average increases judged in quality along four criteria and decreases for one, again policy relevance, although the differences were statistically significant only for clarity and overall quality.

²⁷ The selection of years was intended to match the FY 2003-2005 cohort planned for analysis in the surveys to allow this group to complete their projects and report on publication and effect on careers.

²⁸ The review template defines 1 and 5 in each case and defines 3 as well where terminology for a mid point could be agreed upon. Subjective differences in interpretations across reviewers is a potential problem, but since the most important measure of concern is change between proposal and paper, differences in levels across reviewers is somewhat ameliorated.

²⁹ On the other hand, the selection process included all proposals and papers from CERGE in those years, which included some from the education field funded separately and were expected by resource people to be of lower quality. The evaluation tests for this and did not find that group to explain the negative changes in perceived quality.

³⁰ The sampling included 13 from each year excluding anyone funded in a previous year by CERGE to minimize attenuation of effects. One paper was dropped for technical reasons.

Table 2.3-3: Proposal to Paper Review, by Criterion

	Criterion				
	Originality	Methods	Clarity	Policy Relevance	Overall
Ave difference proposal to paper, CERGE	-0.16	-0.11	0.29	-0.75	-0.09
T-test probability, CERGE	0.28	0.56	0.11	0.00	0.50
T-statistic, CERGE	1.11	0.59	1.66	3.32	0.68
Ave difference proposal to paper, Impact	0.12	0.16	0.16	-0.16	0.24
T-test probability, Impact	0.33	0.33	0.44	0.21	0.08
T-statistic, Impact	1.00	1.00	0.78	1.28	1.81
Ave difference proposals to paper, All proposals rated <3.5	0.09	0.19	0.34	-0.14	0.27
T-test probability, All proposals rated <3.5	0.62	0.29	0.01	0.61	0.04
T-test statistic, All proposals rated <3.5	0.49	1.07	2.58	0.52	2.06
Difference in Averages Impact minus CERGE	0.67	0.54	0.66	-0.08	0.44
T-test probability, Impact vs CERGE final paper judgements	0.01	0.05	0.02	0.79	0.09
T-statistic, Impact vs CERGE	2.67	2.08	2.51	0.27	1.77

Notes: * Statistically significant at the $p \leq 0.10$ level.

** Statistically significant at the $p \leq 0.05$ level.

Statistically significant calculations are shown in bold type except for difference between Impact and CERGE.

Although the analysis is not centered around measuring differences between outcomes between RRCs and GRPs, it recognizes the implicit comparison in the lack of evidence of greater value added from the GRP process and the explicit comparisons stakeholders make about outcomes given money spent (assuming some fungibility of funds). The review forms contain a proxy for the challenging question of value added to the field through the criterion of originality of the paper, on which the Impact project (certainly as a whole) should have greater effect. The evaluation is cautious about these comparisons, as the research design emphasizes changes within not across papers and because results may not generalize to all GRPs and RRCs. Nonetheless, there is some evidence that the GRP papers may show greater originality, clarity, and better use of methods than the RRC papers. This suggests real added knowledge to the broad subject of impact of rich countries on poor countries from the Southern/transition perspective.

The general approach of this research could be refined and replicated across all RRCs and GRPs. Assuming RNPs can request greater specificity in proposals³¹, especially methodologically, such an approach could provide GDN and donors quantitative evidence of capacity building. This could be facilitated through requesting RNPs maintain proposals and papers electronically. Findings could help encourage RNPs to identify ways to improve delivery. Another idea to consider is use of citation indices to understand the extent to which GDN-

³¹ The evaluation decided against using second drafts of proposals or first drafts of papers under the assumption that this would attenuate measurement of capacity built.

sponsored research is being cited by other authors, but the evaluation team is unaware of how well these could be applied in a Southern context.

2.3.1.4 Effectiveness of RRCs across Regions, between RRCs and GRPs, and Impact

Through interviews, the evaluation found that key stakeholders are less uniform in their assessments regarding the effectiveness of the RRCs across regions than the general relevance and effectiveness (value to targeted users) of RRC activities generally, as discussed previously in the section on relevance. There was particular concern even among donors about the level of value added and mentorship especially for SANEI, ODN, and EADN. Stakeholders closest to the work of the GRPs generally are more positive about the effectiveness of the GRPs (although not necessarily regarding outputs), but still provided suggestions for improvement including additional training (especially regarding research design and choice of quantitative methods) earlier in the process. Those somewhat further removed had stronger critiques and varying ideas for improving effectiveness. The suggestions from these stakeholders range from providing greater research leadership within the secretariat or greater RNP input on helping suggest or vet participants to restructuring the entire process of GRP theme selection to increase transparency and ranking of intellectual and development contribution of the project.³² Shorter timelines and GRP project size also has been suggested.

Grantees themselves provided responses via the survey as well as through focus groups conducted by GDN and informal discussions with evaluation team members. A small percentage of RRC grantees indicated they received no review sessions on their proposals and a larger percentage receiving no feedback between sessions with a large plurality indicating they received one review session and received feedback once or twice between sessions.³³ (See Annex E for basic end user response tables.) An equivalent plurality of GRP grantees indicated receiving two review sessions and received feedback two to three times between workshops suggesting more frequent mentoring as well as a longer time period of receiving feedback.

Despite the higher resource input for GRP than RRC grantees, no significant difference was found between these respondents.³⁴ In terms of effectiveness of the RRC process, a small percentage of grantees indicated very low value of review sessions for their proposals or reports or of feedback between workshops, but most rated the value of each between medium and high. Grantees interviewed from South Asia in particular noted a lack of guidance, when desirable, which was mirrored in survey responses by region.³⁵ There was a relatively strong correlation (0.24 to 0.51) between the amount of feedback and perceived value of review sessions suggesting that GDN and RNPs should consider ways to encourage feedback and mentoring more consistently. Some resource persons interviewed noted that these interactions were entirely up to the initiative of the grantees. At the same time, in interviews some grantees indicated lack of clarity regarding whether they were allowed or supposed to initiate contact unless it came from the resource persons – especially for cultures where open, horizontal relationships are not traditional.

In terms of effect of the RRC process on the outcomes of individuals' research, grantees on average provided similar answers for all items on the extent of change in four elements of their research: quality of proposal, research design and methodology, substantive direction, and publishability with slightly lower responses for

³² This would shift GDN's role more toward acting as an intermediary between donors and researchers and less as project entrepreneurs.

³³ No statistically significant difference was found between the 2003-05 cohort of RRC grantees and others.

³⁴ In addition, no significant difference was found between the 2003-05 RRC cohort and all others.

³⁵ Mentoring was rated higher than average among those in Sub-Saharan Africa and CIS and lower in South Asia. There were few respondents born in the Pacific region and LAC.

change in the publishability of the research.³⁶ In all cases, the average change rating across the five point Likert scale (rated from zero [“none or very little”] to four [“very substantially”]) was 2.1 to 2.3, or closer to “moderately” than “substantially” changed.³⁷ The average valuation of the grant process among GRP and RRC grantees was between 2.7 and 2.8, or closer to “substantially” than “moderately” changed, for the following: effect of the grant process on technical skills, subject knowledge, and academic output. The exception is lower ratings among GRP grantees for the effect of the grant process on technical skills (2.2) and academic output (2.4, significant at the $p \leq 0.06$ level).³⁸ These responses from grantees suggest moderate direct effects at the individual level of GDN-funded capacity building. The responses, in concert with feedback from interviews, also suggests room for improvement, especially in terms of technical training early in the GRP process and ways to encourage feedback and mentoring of grantees in need in the RRC process more consistently.

A qualitative note should be made in reference to the issue of regional impact, as opposed to individual-level outcomes from the RRCs and GRPs emphasized in this analysis. The impact of RRCs and GRPs, and of GDN more broadly, must be taken in perspective of the relatively modest scale of operations given the vast potential scope of researchers to target, especially if GDN is truly attempting in its mission to serve all social science disciplines, many of which are vastly weaker in many regions of the Global South.³⁹

Nonetheless, many stakeholders, including some donors, note disquiet that there remain limited indications of broader impact in capacity built within certain regions – particularly all those in Asia. Those interviewed were more positive about impacts of GDN and similar efforts in areas within the former communist country regions, although the effect even among most CIS countries was seen as modest to minimal outside of key cities. GDN is noted as a small player in Africa, where combined efforts regionally are seen as uphill battles against larger factors. Some interviewees noted the problem with the GRP model of one-off projects including researchers from many regions that are unlikely to have noticeable regional impacts without replication along similar themes or methodologies.

The evaluation acknowledges these concerns about limited impact. It notes that even at 10 years, however, given the magnitude of GDN’s budget relative to the large numbers of inadequately trained researchers in many regions, it would be ambitious to expect regional impacts from GDN’s efforts alone. GDN appears most successful where it is working in concert with others, especially where institutions are being formed with similar missions or externally trained researchers are returning to and remaining in their regions or both. Unfortunately for GDN, its need for long-term focus is insufficient for the funding timelines of most donors, which suggests it should consider additional strategies to broaden its impact – and the perception of progress towards impact – in the weakest regions. This perception represents some threat to supporting GDN’s work in these regions. GDN generally does not work directly on setting up or strengthening institutions and only recently has begun work on stand alone individual training, but to the extent that it might consider areas to do so, Asia might be a focal point.

³⁶ Statistically significant at the $p \leq 0.05$ level for most comparisons including between the 2003-05 cohort and others.

³⁷ Analysis by region indicates that respondents from Sub-Saharan Africa consistently indicated greater change as a result of the workshops across all four measures, while respondents from other regions noted slightly less (South Asia, East Asia, or CEE) or more (CIS) change than the other regions only in one or two categories at the $p \leq 0.01$ level.

³⁸ The average valuations for the FY03-05 cohort was slightly lower (significant at the 0.10 level) regarding the value of the grant process on subject knowledge. Grantees from Sub-Saharan Africa rated the value of the RRC process more highly than grantees from other regions, with those from East Asia and possibly the Pacific (significant at nearly the $p = 0.10$ level, but an n of only 6) regions rating theirs lower.

³⁹ The low ability to have a measurable impact across many fields given GDN’s scale is, in fact, a reason frequently cited by those who believe GDN should focus on its “competitive advantage” in the economics field.

2.3.2 *Annual Conference*

The evaluation next examines the extent to which GDN's annual conference succeeds in the objective of knowledge generation and "cross-fertilization" as demonstrated by perceived value of the conference and networking not only within but across regions⁴⁰. It also explores to a lesser degree the extent to which GDN's capacity-building activities are justified by the volume and quality of the resulting research output. The effectiveness, quality, and value added of the conference given its costs has been a long-standing concern to stakeholders.⁴¹ Interview and survey results indicate that the annual conference is seen as between moderate and high value by both key stakeholders and attendees, especially in terms of networking. Younger researchers and those from Sub-Saharan Africa are more likely to ascribe higher value to the conference. The evaluation finds evidence that, despite strong concerns from several stakeholders, at least some post-conference networking does occur and that it includes at least some cross-regional networking in addition to intra-regional networking although stakeholders suggest that efforts to encourage cross-regional networking could be improved. Since the survey measures are not fine tuned to the quality and depth of such networking and the evaluation is unable to attend the conference before completing the report, it assumes some veracity in key stakeholder concerns about the need for additional attention on these issues.

The evaluation is provided no objective measurement for success of the conference, since one would expect some networking within and across regions regardless of how well it is run or integrated with GDN's other activities. It therefore relies on subjective analysis of qualitative and quantitative responses from stakeholders and the grantee survey on overall value added through capacity building and generation of knowledge as well as of networking that occurs – the primary rationales related to end users provided by GDN for holding the conference. It later considers "cost effectiveness" of the activity. A short description of the conference and profiles of the 2005 to 2007 conferences are found in annex L.

2.3.2.1 *Annual Conference – Perception on Overall Value Added and Paper Quality*

Despite strong disagreements found during interviews with key stakeholders, on average in the stakeholder survey, almost all key stakeholders rated the annual conference for GDN's constituencies to be between moderately and very valuable (or about 2.0 on a scale of 1 to 7) for three target groups (researchers; GDN itself; and board members, RNP coordinators, and donors). The conference is viewed as at least a moderately useful global event for bringing researchers together from different parts of the world and gives them opportunities for increasing their visibility. The evaluation did not have time for an independent review regarding concerns of some stakeholders regarding both average quality and variance in quality of papers and presentations.

The disagreement is stronger among stakeholders on the quality of the conferences and value given limited resources. A couple of stakeholders note that opulent World Bank-style extravagances from earlier conferences have visibly decreased, although others believe the conferences remain too showy and expensive given the value added. While some stakeholders believe the quality of papers and presentations has improved over the past three years, others believe strongly that the conference provides low intellectual value, that paper quality within and between sessions varies too greatly, and that the quality can be vastly improved. The evaluation notes subjectively and anecdotally that the quality of another multidisciplinary conference not focused on development issues and with a completely Northern perspective, the Association for Public Policy Analysis and Management, took many years to become more consistently of strong quality with a consistent set of

⁴⁰ The evaluation did not explore quantitatively the issue of cross-disciplinary networking, which should be included in future post-conference surveys and evaluation surveys, nor of other aspects of knowledge cross-fertilization outside of qualitative discussions.

⁴¹ Other concerns are discussed in relevant sections.

participants across time and that there is great disagreement regarding the consistency of quality of some other well established conferences, especially with large conferences where the proportion of less established researchers or less rigorously trained researchers is larger. Indeed, several stakeholders interviewed had concerns about the effect of large conference size on quality and GDN resources to coordinate large conferences. Some stakeholders also suggested specific track themes found at most other regular conferences to attract more similar papers over time, increasing networking among clearer sets of audiences and overall conference value.

The 334 respondents to the end-user surveys who had attended GDN's annual conferences on average rated it to have between moderate value added to high value added for, in descending rank, forum for making presentations (half-way between moderate and high value), enhancing research skills and knowledge, feedback on research, and increasing policy relevance of the conducted research. (See Annex M.) The respondents attributed slightly higher value added on average for the networking aspects of the conference, especially in networking across and within developing regions. Sixty-five percent of respondents rate either high or very high value for opportunities for networking and making new contacts from other developing and transition countries at the conferences.

There was considerable variation among conference goers, with attendees from LAC and East Asia (at $p \leq 0.06$) rating the conference lower than others with respondents from Sub-Saharan Africa rating it higher⁴² potentially a function of fewer opportunities for visibility and networking especially for non-economists in the resource deficit region. The latter finding is driven predominantly, but not entirely by the non-economist respondents from Africa, although the evaluation did not find a systematic difference between disciplinary training and value perception of the conference's value. Interestingly, GRP and AMC (at $p \leq 0.08$) participants rate the conference (exclusive of pre/post-conference GRP workshops) lower than the other beneficiaries. As suggested by several stakeholders, evidence was found that younger participants rated the conference more highly⁴³, although this was not attributable to level of seniority of a researcher. Female participants rated the conference's value slightly higher than male participants (at the $p \leq 0.10$ level).

Some stakeholders indicate that the real value added from conferences such as GDN's is when the same researchers meet multiple times over years and that a onetime event would provide low networking value, although it does not appear that GDN purposefully aims to have many junior researchers attend multiple times. The survey results do indicate a small correlation (0.10, $p \leq 0.08$) between frequency of conference attendance and the value attributed to it across all items asked (although not mirrored in any individual items).⁴⁴ However, it also finds a strong positive correlation between both age (0.20) and seniority (0.15) with number of times one attends a GDN conference.

These measures are designed to purposely exclude pre- and post-conference workshops, where GDN management indicates that some of the highest value added occurs, since workshops for GRP are covered under that activity. The evaluation notes that some other one- to two-day workshops also are held by GDN and other organizations, which augment the value of the conference. It also excludes any added value gained by participants by attending any conferences planned in coordination with the annual conference, such as the World Bank ABCDE conference in past years. Any value added is captured only indirectly if key stakeholders or grantee respondents have attended them and considered them in their valuations, and so the overall value of the conference plus workshops is somewhat under-estimated.

⁴² Using Wilcoxon rank sum tests.

⁴³ Rho = -0.13, tau = -0.07.

⁴⁴ One must also note here self-selection bias, which could explain this difference entirely.

2.3.2.2 Annual Conference – Perceptions on Networking

While GDN expects the benefits of the conference to include opportunities for global level discussions and collaboration among researchers across regions, there is disagreement among stakeholders on the amount of such collaboration that results from the conferences. Overall, the grantee survey findings seem to reflect at least a moderate amount of South-South networking within and across regions supporting GDN's contention of one of the primary rationales of the conference.

Opinions expressed by several key stakeholders reflect opinions similar to that the previous independent evaluation regarding insufficient value-added in holding a GDN conference annually or more value added in holding more frequent but smaller multiple-region conferences to lower costs (conference costs are discussed further later) and increase opportunities for cross-regional collaboration and cooperation in regions closer and more similar to each other. One of the primary alternatives proposed is an 18 month schedule that would consume fewer resources, allow secretariat and board to focus on other issues as well, potentially increase uniformity of paper quality within sessions as well as generally, and allow scheduling across regions to vary by appropriate season as well. Even among some of the supporters of an annual conference, the opportunities for collaboration across regions were viewed to be diminished by the presence of cliques or natural tendencies for intra-regional networking that were not mitigated by purposeful GDN programming.

In contrast, some key stakeholders and grantee survey respondents from the Sub-Saharan African region view GDN's annual conference as a "lively forum" with opportunities for interregional interaction and GDN as one of their most important affiliations. There was a need expressed to use some sessions, workshops, and social events purposely to increase inter-regional collaboration with caution that regional pairings do not always work in terms of consistent paper quality if training levels are vastly different.

Almost all (91 percent) participants claimed to remain in contact with at least one person met at the conference for the first time and 65 percent remaining in contact with three or more persons. More of the participants remain in touch with persons from either their own regions or other developing or transition countries than northern/non-transition countries. The latter is not surprising since the majority of the participants come from developing countries. In contrast to concerns from several stakeholders that they notice little cross-regional interaction or networking occurring at the conference itself, respondents claim to stay in touch with people from other regions at least as much or slightly more than people from just their own regions although the survey cannot judge the intensity of such contacts. Participants from Latin America and East Asia appear to be somewhat more insular than those from other regions, maintaining post-conference ties more within their own region than other regions, and those from CIS region were twice as likely to remain in contact with people from developed countries.

Some stakeholders mentioned or alluded to the positive outcomes of networking between donors and regions or others at the conference. GDN should try to take stock better of such outcomes of the conferences, if anecdotally, and invest resources in ensuring the publication or distribution of the proceedings and papers.

One criticism among some key stakeholders is that capacity building of participants and cost efficiency might be increased by focusing less on showcasing and bringing "big names" for plenary sessions. The capacity building of researchers can be conducted in a more meaningful and effective way than traditional presentations and discussants. On the other hand, a focus on capacity building may not attract the highest quality research – another concern of stakeholders. This conflict demonstrates the challenging trade-offs in effect implicit in GDN trying to pursue multiple objectives for the conference.

2.3.2.3 Annual Conference – Other Objectives

Despite the fact that the TOR does not specify judging activities such as the annual conference on additional criterion, GDN management note that the annual conference indeed is intended to meet other objectives for GDN. These include serving as a relatively cost efficient mechanism to conduct needed annual meetings of RNP coordinators and of donors and semi-annual meetings of the board. Additionally, it provides these stakeholders an opportunity to observe GDN's work up close, network with each other, and for these stakeholders to serve as resource persons. The evaluation's interviews and survey responses generally support this contention, as discussed further in Annex N, although with multiple goals may come a further decrease in clarity on the conference's goals and thus effectiveness judged on any one of them. The evaluation's cost effectiveness analysis section also points out that combining these purposes together may make the annual conference seem less cost efficient than it otherwise might with its components separated.

2.3.3 *GNet*

The GNet portal, managed from GDN's Cairo office, is one of GDN's two key activities aimed primarily at increasing information sharing among developing country researchers and between researchers and policy makers. Envisioned as an online tool for enabling developing country researchers and research institutes, the GNet portal serves as a package of services that allows researchers to register individual profiles and access databases, datasets and selected journal sets that otherwise require paid subscription and access to newsletters related to research and funding opportunities.

Due to limited time and resources and because DFID will be funding a more in-depth evaluation of GNet specifically, this evaluation does not collect thorough information for a separate evaluation of GNet as a website. It did, however, collect stakeholder and user opinions regarding GNet that provide an idea of its usefulness to targeted beneficiaries. The TOR for the evaluation provides no clear objective for measurement suggesting a criterion such as "does GNet generate any networking and cross-fertilization?" The evaluation team believes this represents too low a bar and suggests for the GNet evaluation including cost effectiveness and comparisons of outreach and value relative to similar websites as well as relative value among GNet's different elements and prospects for the future to encourage experimentation.

In the stakeholder survey results, on average the board members rated GNet approximately half way between only moderately valuable and the next higher category up (between very and moderately valuable) and RNP coordinators rated GNet as only moderately valuable. Board and RNP heads do not represent the targeted beneficiary users of GNet, and several indicated they were largely unfamiliar with GNet's services. Among those who understand it best as well as targeted users, there is some level of agreement regarding its usefulness as well as the need to expand its awareness among developing country researchers, as exhibited for example in one RNP workshop, where RRC grantees all showed surprise about its existence as a tool. GDN should increase the visibility of GNet through systematic efforts through its RNPs, partner organizations, as well as through workshops and conferences it holds.

The evaluation focuses its attention next on survey responses for those registered on GDN's user base.⁴⁵ One measure of the value of GNet to users is that implied by the frequency of visits to GNet. Two thirds of registered researchers visit the GNet portal at least every three months, with 36 percent doing so between once a month to once a week, the latter of which suggests a relatively high implied valuation at least among the small proportion who responded to the survey (and the associated positive biases that result). Users on average rate GNet most highly for increasing their knowledge of their subject area (halfway between medium and high) followed by outreach and building contacts with other researchers and the rest, including dissemination to policy

⁴⁵ The drawback is that, to a large extent, this represents a self-selected sample of those who find GNet useful.

makers, rated medium.⁴⁶ The comments of some interviewees suggests that the last finding is surprising only in how high the perceived outreach is to policymakers even at a “medium” rating given that the most likely audience for most of the research is other researchers rather than policymakers and that researchers they are unlikely to know whether policymakers are accessing their research unless contacted. The evaluation of GDNNet should attempt more direct measures of outreach to policymakers as a baseline for future reference and a way to understand how to improve outreach in this area.

In order to increase the website traffic, it has been suggested by both board members and RNP coordinators that the portal should aim at a more user-friendly interface and re-designing the webpage layout in addition to providing access to national datasets. This represents an area for research in the separate GDNNet evaluation. In supporting researchers to engage in meaningful literature search, GDNNet has attempted, with mixed success, to provide free access to notable journals and datasets through J-STOR and with greater success to ELDIS. GDN has been trying to broaden access to journals in concert with stakeholder needs and requests. Of the registered researchers, 64% found high value in accessing research related resources through GDNNet, and this was supported by views of stakeholders, who hoped journal access could be increased.

GDNNet’s implicit competition in terms of its purpose for existing is other websites offering, aggregating, or linking to knowledge about development. Although this comparison requires more careful exploration, the evaluation briefly compared survey responses for several potential “competitors.” GDNNet currently was rated behind search engines and researchers’ access to journals through their own institutions followed by the World Bank data and research website.⁴⁷ GDNNet ranked above IDEAS (although not when considering only economists by training) and above Google Scholar at just above the $p \leq 0.10$ level. The finding on the World Bank and IDEAS may suggest that respondents did not consider carefully the instructions specifying “development-related research produced *by researchers in developing or transition countries*” (emphasis added). Nonetheless, if GDNNet wants to continue to differentiate itself as adding value through aggregating information on development information from the southern perspective in ways not available otherwise, it has a difficult task ahead to do so.

Over the past two years GDNNet has developed regional windows on the portal to facilitate dissemination of regional researchers and research institutes as well as announcements made within a region.⁴⁸ Thirty-six percent of registered researchers found the windows to be of high or very high value with 24 percent rating them at moderate value.

Besides the portal, the GDNNet team coordinates the dissemination of Research in Focus and Funding Opportunities newsletters, which users generally rate highly⁴⁹ and provides valuable communication about

⁴⁶ Views vary highly in each case a standard deviation of at least 1 (assuming ratings are cardinal values), with 120 respondents consistently rating GDNNet below medium and 271 consistently above medium.

⁴⁷ Based on a Wilcoxon signed-rank test that assumes only that the data are ordinal, and excluding no opinions.

⁴⁸ As discussed previously, it also represents a way to be more multidisciplinary in outreach.

⁴⁹ Of the registered users, 56 percent find the Research in Focus newsletter valuable. About half of the researchers noted high value for the GDN monthly newsletters. Of the registered users, 67 percent found value of announcements for competitions and conferences of very high value, which implies GDN should continue to post announcements for conferences and competitions and disseminate the newsletters.

specific research themes across regions. GDN might consider exploring the value of increasing resources in generating documentation of such documents to increase its outreach and demand for GDN services and GDN-sponsored research.

Suggestions that should be considered in the GDN evaluation is whether the portal's interface is sufficiently user friendly and consistent with other sources in addition to providing access to national datasets.

2.3.4 Policy Impact

There is modest evidence that GDN's efforts thus far have resulted in increased capacity for policy relevant research as well as more modest evidence of outreach to policy makers, but not always supported by the GDN-funded processes. GDN activities, training, and mentoring generally generate more academically than policy focused outputs, with two of the GRPs representing exceptions. This is a weakness of GDN in this area and reflects in the higher output for academic-style research than policy briefs or more locally policy relevant outputs despite interest in this area among survey respondents, as discussed in the section on overall success. It also reflects the potential conflict perceived by traditionally trained economists between the objective of high policy relevance and impact and the objective to foster highest quality research. There is only pointillistic evidence of that capacity leading to research that helped inform policy making – a first step toward potentially affecting policy. There is no compelling evidence that GDN-produced research and efforts themselves have produced a “policy impact,” which the evaluation considers an inappropriate indicator for an apolitical organization working at the global or cross-regional level. The evaluation focuses on capacity for policy relevant research and potential for informing policy making and linking researchers and policy makers as the most appropriate measures. The GDN-funded research that most likely will result in policy-related effects – training resulting from the BRP research and the Health project – are too nascent to count toward this goal for the current evaluation but promise to help GDN in this area for its next evaluation. GDN as a method for distributing policy-relevant summaries is likely to represent a key area where GDN could augment its output in terms of policy relevance.

The political science and political economy literature as well as GDN's own Bridging Resource and Policy Synthesis Report (Livny, Mehendale, and Vanags, 2006) suggest that policy impact is very challenging to generate systematically through apolitical independent research processes and that independent research ideas and policy research institutes affect policy only when several factors coincide. Major policy shifts may require special windows of opportunity, although there may be more room for research-based efforts to have influence on more incremental, less politicized changes.⁵⁰ In fact, policy relevance and policy impact may be considered as part of a continuum rather than entirely distinct categories, and measurement problems in gauging policy influence abound.

The evaluation does not attempt systematically to measure policy impact from these identified individuals or incidents given the inherent difficulty in doing so, as expressed in the literature. Most RNP coordinators, and many other stakeholders as well, claim developing policy-relevant capacity either is or should be GDN's focus anyway for its activities, not policy impact. Respondents noted little evidence that policy makers frequently “read research books” generally, let alone those generated by GDN's processes. As a practical matter, GDN would not be able to track the use of their products to policy makers.

⁵⁰ For the early literature, see, for instance Kingdon (1984); Baumgartner and Jones (1993), and Downs (1972). For a discussion from the point of view of a funder of such research efforts, see also Carden *et al.* (2002).

The grantee survey kept these constraints in mind in its design, asking RNP coordinators for examples of policy influence or impact, following up from the prior evaluation. This evaluation also includes surveys asking grantees about contacts made and the extent to which capacity is built to highlight policy relevant points in research.

Given the relatively young age of GDN's capacity building activities, one would expect that noticeable, if not measurable, effects should begin occurring only recently. Indeed, only two RNP coordinators interviewed provided examples of GDN-funded efforts creating capacity that has been directly drawn upon later for national-level policy development.⁵¹ LACEA indicated indirect evidence of policy influence through training researchers who move in and out of political positions and thus make more informed decisions and through funding a policy-accessible journal that likely has been read by policy makers. AERC provides an interesting potential model for other RNPs to consider in that they have connections with national policy research institutes with direct links to government representatives. This has allowed AERC periodically to hold dissemination workshops to disseminate RRC results to policy makers and to encourage them to consider commissioning funded researchers to fill policy gaps. The evaluation would categorize this generally as influencing the policy dialogue, although there could be unknown examples of policy influence. The evaluation found limited examples of grantee survey respondents in other regions indicating having been contacted by donors or policy makers regarding the work funded by GDN.

As reported earlier, grantees from CIS report generating twice as many policy briefs as other regions on average due to the concerted efforts of the former coordinator for EERC⁵², and those from Sub-Saharan Africa and, to a lesser extent, MENA ($p \leq 0.10$) evaluate the grant process more highly on average in terms of increasing contacts with policy makers, although these do not constitute evidence of policy effects but its latent potential. The finding for Sub-Saharan Africa is likely due to the close contacts between AERC and applied governmental organizations.

Among other policy impact related activities included in the survey, only around a third of RRC and GRP respondents indicated they wrote articles for the popular media (such as newspaper or magazine articles) and about a quarter indicated they made presentations to the media or press.⁵³ Grantees responded that they distributed the report, paper or policy brief to potential policy makers and discussed ideas related to the funded research with slightly more than one potential policy makers on average, with grantees from the CIS indicating a greater probability of having done so than others. GRP grantees more frequently reported distributing their research to potential policymakers than RRC grantees and also more frequently reported conducting policy briefings or round table discussion than RRC grantees.⁵⁴ They also more frequently reported receiving help from the grant making process in doing so. A third of respondents indicated that any policy-oriented outreach was done entirely on their own, with another half indicating the RNPs helped to a small to moderate extent with no differences found across regions. The evaluation noted that policy outreach or involvement of policy makers understandably is not really a part of most RNP efforts and does not necessarily fit in easily with their more

⁵¹ Specifically, the claim was that the researchers funded by the RRC process developed specific expertise on an issue through their papers that later were the cause for them to be identified for involvement in providing research or expertise support.

⁵² Although interestingly these grantees also provided lower value ratings than those from other regions regarding contribution of the grant process to making policy contacts indicating that policy briefs themselves may not bring researchers into contact with policymakers, assuming this is the intention.

⁵³ These methods of publication are, in fact, often discouraged generally by the traditional academic process, although the evaluation did not assess the extent to which that is true among GDN's processes.

⁵⁴ And at the $p \leq 0.10$ level, they also reported more frequently conducting press conferences or other media outreach.

academically-focused capacitation efforts or contact base.⁵⁵ This suggests separate efforts must be planned or organized if policy interaction or outreach is desired through RNPs or that this outreach should be thought of entirely separately, such as through GDNNet.

Both GRP and RRC grantees on average rated the effect of the grant process lowest in terms of increasing contact with policy makers (somewhat less than medium value added) followed by ability to introduce recommendations into the policy debate (medium value added) and highest in terms of ability to extract policy recommendations (between medium and high). Although the latter is the area of greatest success, it also represents an area where GDN and the RNPs could further increase policy relevance of their activities, which could build capacity for policy impact in the future. The various types of AMC grantees rated the effect of their awards more highly than research grantees in terms of contribution to contact with policy makers.

GDN also uses GDNNet as a tool to increase policy communication and influence of southern researchers and their GDN-funded or independent research papers. The evaluation was unable to find an appropriate measure of its doing so and instead notes its latent potential in this direction. The separate evaluation of GDNNet should consider exploring the value of developing such measures. Given the discussion on policy windows, it is unlikely to be fruitful, but GDN could explore how think tanks attempt to demonstrate influence to the extent that they attempt to do so through evaluations. GDNNet is likely to serve as a key, cost-effective method to improve the extent to which GDN-funded activities inform policy dialogues and a way to unify this effect across RNPs with varying, but generally low, interest in doing so.

GDN is not a think tank and is supposed to be apolitical operating considerably above the national or local levels at which policies are affected. Further, affecting policy is an unpredictable, problematic issue to measure even for organizations dedicated to doing so. Even though donors may be interested theoretically in “policy impacts,” they are not easily fooled into believing an organization such as GDN affects policy and can become confused if this is listed as a goal. Thus, the evaluation considers GDN’s potential objective of “policy impact” both an unrealistic and inappropriate target that the board should clarify is not a target objective. GDN can claim and should claim, however, to be building a base of policy relevant literature and future researcher leaders who can and are informing better policy at national and sub-national levels.

2.3.5 Effect on Careers

This section of the evaluation highlights the intermediate outcome related to the effect of the grant on career advancement of the research grantees. The changes in career can be attributed to exogenous factors besides a direct effect of the capacity built. As a result, the evaluation team collected both qualitative data and quantitative data from grantees on the specific effect of career advancement and professional visibility after the completion of the GDN funded grant. The evaluation team had limited informal discussions and held focus group discussions or interviews with grantees across only some regional network partners (CERGE, EERC and SANEI) on a site-by-site basis and the GRP grantees from the Health project concerning the changes in career noticed by the grantees.

Qualitative findings on the perceptions on the impact of the grant on researchers in relation to their career was collected from interviews held with RNP and GRP coordinators and interviews or focus groups held with mentors/resource persons (academic or policy experts). Most of the RNP coordinators and mentors were relatively unaware of the impact of the grants on the careers of the researchers due to the lack of data. Anecdotal evidence from two regions suggest that the career of the grantee researchers positively benefit from the grant either due to existing publications generated due to the funded research (see earlier discussion) or increased

⁵⁵ Several stakeholders interviewed who are involved with the processes note in fact that many of the papers involved are not of the quality one would want to inform policy makers.

capacity for generating future publications ultimately leading to higher career mobility (promotions and new job offers).

The quantitative indicators that assess the effect of grants on the career of participating researchers include perception on career advancement and professional visibility as an outcome of the grant process and building contacts with other researchers. Forty percent of the 196 GRP and RRC grantee respondents rated the contribution of the grant on their careers generally as high or very high and 31 percent rated it as medium with 26 percent rating it low or very low. No difference was found by organizational type. RRC grantees attributed a higher rating of the grant's effect on their career even controlling for the lower average age of RRC grantees. Researchers from Sub-Saharan Africa rated the effect of the grant on their career relatively higher than researchers from other regions. This finding appears to corroborate with other positive responses from researchers from this region and may reflect a combination of the intensity of mentorship provided by the RNPs in Africa and the perceived lack of capacity building activities in the African region otherwise. Surprisingly, the AMC respondents rated the value of the awards and medals as higher than GRP or RRC grantees in terms of increased career advancement.

According to survey respondents, higher visibility did not appear to be the primary reason for perception of career advancement, although it may itself represent the potential for further advancement. Overall, only 22 percent of respondents rated increased professional visibility due to the grant as high or very high, with another 31 percent rating it medium. Interestingly, the RRC grantees attributed a slightly higher rating than GRP grantees (at the $p \leq 0.10$ level) for increased professional visibility as a result of their grants. Research grantees from Sub-Saharan Africa gave a higher rating than grantees from other regions to the contribution of the grant towards a positive impact on their professional visibility, while those from Central Europe rated it lower than others along this dimension. AMC grantees rated the value of increased professional visibility higher than those of the RRC and GRP grantees.

The evaluation also considered the value grantees find from the grant in terms of contacts with other researchers as an indirect, long-term potential indicator of its career assistance. Thirty-nine percent of grantees rated the value of the grant high or very high in terms of contacts with other researchers and another 35 percent medium. Again, grantees from Sub-Saharan Africa region attributed higher value to the grant in opportunities for interacting with other researchers, while those from East Asia rated it lower. AMC respondents rated more highly the value of the award or medal in making contacts with researchers than other respondents. This finding might be a function of the interaction and networking opportunities available for the AMC respondents and Sub-Saharan African researchers at the GDN annual conferences.

2.4 Reach and Multidisciplinarity

GDN is making concerted efforts for its more centrally run activities to have broad representation across regions, with women, younger researchers, and increasingly to target people from a variety of disciplines in addition to economics. GDN's record with delegated activities – the RRCs – in broadening beyond targeting only economists is more mixed with some regions broadening considerably and others not at all. The issue of interdisciplinarity or multidisciplinarity remains from the last evaluation one of the greater areas of disagreement if not tension within the organization and one on which the organization still has not come to a stable point. An area to consider for greater emphasis in coming years is to track not only number of countries but to develop measures of higher needs countries or even areas within countries within regions.

To explore the issue of GDN's reach, the evaluation examines first the extent to which inclusiveness has been systematically taken into account in the process of selecting grant recipients, topics for research and conference agendas. Included in this examination are complicated perceptions of stakeholders towards GDN regarding this issue. It then examines the extent to which GDN activities have reached all regions; younger, less established

researchers; women; and traditionally under-represented (non-economic) disciplines. The term inclusiveness has not been formally defined by GDN in the evaluation TOR, which may be explained by its different potential meanings by region and by GDN's lack of resolution of its goal regarding multidisciplinary.⁵⁶ The evaluation thus operationalizes the definition moderately broadly to include younger researchers, gender, under-represented countries or areas within a countries, and less-represented (non-economic) disciplines. Regarding criteria for judging multidisciplinary, it posits two potential thresholds: 1) primarily economists with many other disciplines also involved as relevant and 2) no one or two disciplines representing hegemonic dominance.

The issue of reach generally, and multidisciplinary⁵⁷ specifically, continues from the prior evaluation to be an issue of between mild and considerable tension, as expressed in interviews and survey results of current and recent RNPs and board members. It does appear that the tension over disciplinary focus has decreased somewhat in centrality.⁵⁸ Nonetheless, GDN's efforts towards multidisciplinary remains an active internal debate.

The reach of GDN's funding is considered first through its indirect chain through RNP's RRCs and then through GDN's own actions and activities. The evaluation found that most RNPs take some aspect of inclusiveness, but not necessarily multidisciplinary, into account. Some specifically exclude applicants from cities or institutions considered to be well represented with sufficient capacity. Some use disciplinary inclusiveness as one criterion assuming that applicants meet minimum thresholds and choosing the best applicants among them. Most state they favor working with and building capacity of younger, impressionable scholars, even if they do not have explicit age cut offs. Some made it clear that, beyond favoring younger researchers, inclusiveness is not and would not be a criterion used except under duress.

Stakeholders uniformly indicate that GDN has been actively attempting to cajole RNPs to broaden disciplinary focus, with mixed success in terms of results and effects on its relationship with RNPs, as discussed later. RNPs whose missions traditionally have been focused on economics are considerably less likely to feel compelled to broaden their mandate and are less likely to be successful in doing so due to selection of economics-focused themes or problems with outreach. The average proportion of RRC grantees outside of economics between 2003 and 2006 ranges from 67 to 0 percent with three year averages for FY 2004-2006 for RNPs ranging from 0 percent to 53 percent and 38 percent overall. AERC (and, at least in terms of annual meeting attendees, LACEA) are at the bottom followed by SANEI. (See table O-1 in Annex O.) If one assumes as given that RNPs founded with an economics focus should and will select based on use of methodologies favored by economics, this trends suggest strong progress was made along this indicator with EADN, relative parity exists between economics and non economic disciplines in OCEANA, EADN, CERGE, and ERF⁵⁹, and insufficient progress with SANEI, AERC, and LACEA.⁶⁰ The evaluation is told GDN recently has undertaken an unofficial policy of encouraging the selection of themes in calls for papers that are more naturally interdisciplinary in subject to help encourage more people outside of economics to apply, but evidence is not available yet to support this

⁵⁶ GDN's statement on governance defines the principle of "plurality" as "encompassing a *diversity* of disciplines and paradigms" including heterodox traditions within economics. [Refer to the link for further information](#)

⁵⁷ Multidisciplinary is operationalized in this analysis as attracting and including researchers from multiple disciplines or examining topics from different disciplinary perspectives and not necessarily interdisciplinary projects.

⁵⁸ This may be due to board member rotation, the results of concerted efforts by the secretariat to meet goals and fulfill plans developed in conjunction with the board to broaden disciplinary representation, and greater "tolerance of ambiguity" in disciplinary focus over time as a status quo.

⁵⁹ Especially given the low proportion of non economists submitting to ERF in 2005.

⁶⁰ The data the evaluation received from GDN on proportion of applicants outside of economics is partially incorrect but suggests that those RNPs selecting non-economists generally confirm the impression of the evaluation team of being between neutral and positively inclined toward multidisciplinary in their selection processes.

expectation. Additional outreach beyond just the economics field also would be necessary to foster appropriate change.

Since the previous evaluation, GDN itself has continued to move toward activities that encourage greater inclusiveness in terms of favoring inclusiveness of women and of multiple disciplinary trainings having representation although with economic methodologies and participants emphasized. One area of progress and the key way it signals doing so is through choosing themes for its conferences and GRPs that are less focused on solely economic issues. The results of doing so are discussed later in this subsection. GDN does not apply a specific age limit to its GRPs. The use in GRPs of teams that often include more senior researchers managing research teams, leads to a mix of age ranges involved in some way with the GRPs. GDN only specifies an age limit (of 45 years) for the Awards and Medals competitions. Before the last evaluation GDN also began to roll out regional windows that provide resources and opportunities broadly across the social sciences, maintaining control of the website itself to maintain breadth of approach when regional partners are unwilling to do so themselves.

Although it appears GDN signals somewhat greater openness than during the period of the previous evaluation through staffing and board membership outside of economics, the signal still suggests hegemony of economics and a preference for traditional neo-liberal viewpoints. One periodic complaint among some stakeholders is that GDN has not made significant progress in its annual convention plenary sessions in promoting viewpoints outside of the “Washington orthodoxy.” This issue was not analyzed broadly or systematically, but non-random observation of research papers suggest that this complaint would not register as strongly, although perhaps still to some extent, for the funded research grants. GDN’s board membership and staffing also reflect a bias toward economics with 11 of 16 members trained as economists. The organization also continues to signal an emphasis on economics by choosing economists as president and maintaining the title of the second in command as “Chief Economist.” While this represents the traditional approach for the World Bank and does suggest disciplinary expertise, it runs counter to the generalist titles among other development-based NGOs and organizations. It also clarifies which discipline is on top. GDN’s recent addition of a North American partner focused on economics is helpful in several ways, but also adds clarity to GDN’s disciplinary focus. GDN signals political science as its perhaps second favored discipline since the two other top (current or recently departed) staff with a title specifically representing a discipline (“senior political scientist”) come from political science.

In interviews with key stakeholders as well as stakeholder survey results, many note both a change toward more inclusive themes and greater inclusiveness in terms of disciplines involved.⁶¹ The inclusiveness is primarily in the direction of political science or political economy, as three board members come from political science or have some political science training. Further, the Impact project is widely perceived as a “political economy” topic. Several interviewees who indicated that the more purely economics-centric RNPs are becoming more intellectually broad always point to increasing interest in political economy research.

Nonetheless, non-economist stakeholders generally indicate that GDN’s efforts in the direction of disciplinary reach have been insufficient, while some economists state that GDN should move back towards its comparative advantage focusing on economics (albeit using a broader definition of what constitutes economics than traditional in the past). Several stakeholders, including key donors, interviewed agreed with one of the key conclusions from a “Friends of GDN” retreat (attended almost exclusively by economists predominantly from

⁶¹ Results must be understood given that perceptions are affected by the breadth of understanding of GDN’s various activities generally, as well as lags over time in between actions taken and effects seen. It also seems to include a lot of focus from some on the RRCs, reported earlier.

the World Bank⁶²) that GDN should consider interdisciplinarity as an *instrument* or *tactic* rather than a goal itself and not pursue multidisciplinary as an objective.⁶³ In contrast, some current or recent former board members feel that promoting inclusiveness generally, and disciplinary breadth in particular, is an important part of GDN's *mission* and thus insufficiently fulfilled.

In the stakeholder survey, respondents rated GDN's emphasis in practice on "core applied economic research" about half way between "mostly" and "about half-way" (average 2.4) and slightly above that of "any social science-focused development research" (2.8) and "multidisciplinarity" (2.7) with relatively high variation (a standard deviation – a measure of variation around this average – of about 1.0). (See Annex P for further details on stakeholder responses regarding GDN's objectives in practice and ideally.) Among all the items in the stakeholder survey about GDN's activities or emphases in practice, the item on extent to which GDN in practice emphasizes core applied economic research received by far the highest percentage of respondents indicating any item's level should be decreased – 20 percent of all respondents who marked a preference for change for at least one category.⁶⁴ However, even more respondents – 29 percent – indicated it should be increased and the remaining 51 percent did not indicate a need for change. The evaluation even encountered two people interviewed at once who disagreed with each other over this issue.

GDN's trends in terms of both attracting and awarding grants through the GRPs and AMCs to people outside of economics suggests progress. GDN's 2006 document on multidisciplinary research (see Annex Q) indicates progress in selecting GRP and AMC themes that appeal to non-economists and in more non-economists applying to and successfully being selected. The previous internal evaluation notes (pp13-14) efforts made in 2003 to advertise competitions outside economics networks. The themes of the two most recent GRPs indeed suggest the potential to attract and include more than just economists and interdisciplinary teaming for the health GRP, although it has been noted by several stakeholders that the health GRP still emphasizes the need for having economists or similarly trained members on each team. GDN also conducted a workshop at their 2007 conference on ways of organizing future Global Research Projects to increase development effects. The workshop findings (summarized in the "Multidisciplinary Research" document) suggested broadening the methodological approaches favored in GRPs away from emphasis on regression analysis and purely quantitative techniques. In theory, doing so could increase both GDN's disciplinary breadth in numbers as well as the perception it is moving beyond the "Washington orthodoxy." It is indicated to the evaluation that follow-on activities for the workshop are planned including greater methodological breadth for the new GRP on Migration, "Development on Move." Future evaluations will have to judge the success with which GDN has responded to these recommendations.

GDN's annual conferences have been relatively steady in recent years in terms of about a 50 to 60 percent rate of participation of economists. If the concept is for GDN's conferences to be primarily economists with many other disciplines also involved, GDN has met that definition. It has not met the definition of a conference in which no one discipline appears clearly dominant. A model GDN may wish to emulate if that is the case would be the Association of Public Policy Analysis and Management, whose Fall 2007 conference had the following

⁶² GDN has responded to this parenthetical phrase by indicating that "the list of Friends included non-World Bank staff, other academics, policy makers etc. and a mix of disciplines." The evaluation is told that some who could not attend sent in written comments and suggestions and will stay involved in some way, but the evaluation team can only react to the evidence available at the time.

⁶³ A key rationale provided is that it is difficult for an organization as small as GDN to have an impact across regions even within one discipline, and that it is considerably more difficult to have an impact across several disciplines, particularly when they are considered further from the model of quality GDN implicitly promotes.

⁶⁴ That is, the denominator excludes the few who may have rushed through this part of the survey without answering any preferences for increasing or decreasing current emphases and those who did not respond to the questions at all.

disciplinary training breakdown: Economics – 26 percent, Public Policy – 25 percent, Political Science – 9 percent, Public Administration – 8 percent, Sociology – 7 percent, Psychology and Social Work – each 3.5 percent, Urban Studies, Statistics, Demography, Law – between 1 and 2 percent.⁶⁵

GDN's document "Gender Representation in GDN Activities" (See Annex R) indicates both women and men are involved in activities sponsored directly by GDN, with potentially increasing numbers of women across GRPs, AMCs, and conferences. The evaluation has no baseline information on the representation of women as researchers in the Global South to use as a benchmark for comparison aside from potentially using figures from its GDN database. These are flawed in that they are voluntary data provided by only a fifth of registered users, and likelihood to respond might be negatively correlated with whether one is female. This is especially true given the low proportion (12 percent) claiming to be female in GDN. The gender data found in the broad constituency survey of GDN subscribers is another potential source, which indicates that the proportion of female researchers using GDN is closer to 25 percent, with considerable regional variation from 15 to 18 percent (Sub-Saharan Africa and Southeast Asia) to 43 to 48 percent (East Asia and Central and Eastern Europe). The percentage of RRC grantee respondents also varied considerably by region from 17 percent (AERC) to 41 to 43 percent (CERGE and EADN), although the proportions of grantee respondents generally were roughly equal to or above those found generally in the broad constituency survey by region with the exception of female GRP grantees from the LAC region. (See the table in Annex O, Economics Disciplinary Training and Gender Figures by Region.) This data suggest that the gender ratio in GDN-funded activities is relatively neutral in gender bias and thus inclusive of both genders, and provide only one instance of evidence that selection and review processes adversely affect either gender disproportionately. The evaluation did not undertake measures to assess the extent to which these figures are representative of GDN or the gender balance in the field more broadly in the developing and transition world and thus whether GDN should consider ways to increase female representation, although the figures for Sub-Saharan Africa and South Asia in the table referenced above should be studied more carefully.

The evaluation also briefly examines the gender, disciplinary, and regional representation of data for GDN's activities gathered through its surveys recognizing caution must be taken in inference from them due to unknown biases in response by gender, probable over representation among less established researchers, and probable under representation of researchers from Africa and possibly parts of Asia due to potential Internet connectivity limitations. Results generally show diversity by age range when awarded with strong representation of younger researchers. It also shows strong diversity by region despite concerns about under representation of researchers from Africa and possibly parts of Asia. (See Annex S for tables.)

GDN data suggest broad representation across regions by activities as well as a large number of countries represented in one way or another. GDN has not requested nor has a methodology for examining representational equity within regions, such as the Caribbean in LAC or the "stans" or Caucuses in CIS, to name a few. A few stakeholders note the difficulty in GDN's approach for building capacity in such areas that lack strong training baselines while still choosing and promoting as high quality research as possible. It appears that at least some RNPs spoken with try to weight more heavily people from such underserved areas in their regions and thus may be doing as reasonable a job as feasible in serving these areas given these sometimes conflicting objectives. Nonetheless, GDN should consider more refined measures to check its success in this area.

2.5 Cost Efficiency and Cost Effectiveness

The evaluation is tasked with undertaking cost effectiveness analysis to help assess the extent to which activities run by GDN are run cost effectively. The evaluation is hampered by the lack of benchmarks from similar types

⁶⁵ Email correspondence with Erik Devereux, Executive Director, APPAM, 4 December, 2007. No comparable data are available for APPAM's earlier years.

of research organizations to serve as comparisons. Further, the activities involve different outputs and outcomes, making even intra-organizational comparisons difficult, especially with GDNNet. Some activities involve multiple outputs, including some that are not readily quantified, that themselves also are not comparable and appear that they should not be added together. The cost effectiveness analysis itself does not formally include relative measures of different value within and across activities. Because variation in survey responses about the value of activities was not great, including such measures would result in largely similar findings that would be somewhat more favorable to RRCs generally. Lacking information on relative value of outputs, the evaluation considers these measures more measures of cost efficiency than cost effectiveness.

The analysis of RRCs finds that there is considerable regional variation in cost effectiveness in this very basic cost efficiency sense of the term. The medals competition appears about as cost effective as RRCs to the extent that it provides similar, less tangible benefits, which is not clear. The MIR and Impact GRPs' cost effectiveness were in the general range of that found for RRCs only if considering total number of team members, which is likely to overestimate their cost effectiveness by assuming benefits are distributed roughly equally across all team members. In theory, the greater and more consistent mentoring through the GRP process should have resulted in greater value added, if not superior research outputs in the end depending on the base level of researcher capacity. The BRP and Health GRPs are the least cost efficient in the strict sense, especially if one also considers academic-style contribution to a field of knowledge. These two GRPs are intended, however, to provide different outputs – the BRP should contribute to GDN-funded training for research institutes and the Health project primarily to donor and policymaker understanding of replicability and scalability of health MDG-related interventions (as well as to applicability of some methods to such types of evaluations – increasing costs further). The subsequent evaluation will have to determine the ultimate cost effectiveness of these two projects to see if they result in these expected outcomes.

The evaluation indicates that the least cost efficiency part of the conference are the business meetings, which it assumes GDN would need to conduct anyway and provide some side benefits in non-measurable conference outputs. The question then would be whether its remaining \$1,200 to \$2,400 cost per measurable “output,” combined with the conference’s non-measurable outputs, is justified were the Bank or other key donors to cut funding severely. GDN also should compare these cost efficiency figures with those of RRCs, which are rated by stakeholders somewhat higher in value for target recipients in deciding whether to keep the conference at its same size or temporal frequency. RRCs are on average an order of magnitude less cost efficient, strictly speaking, and do not allow cross-regional networking and other public goods. They are perceived, however, as providing vastly more direct capacity building over a longer terms and sometimes considerable intra-regional networking. An alternative is to try to generate more capacity building through GDN’s annual conference, which might still provide much, but not all, of its current non-measurable benefits but would require time and resources to do so. The measure of cost efficiency based purely on the number of papers generated (about \$10,500) suggests low cost efficiency relative to RRCs given that no real capacity building is associated with it.

2.5.1 General Questions

In considered cost effectiveness, the evaluation first examined three basic questions:

- whether administrative costs seem reasonable relative to programmatic costs,
- whether there are obvious cases of inefficiency or wasted resources, and
- whether benefit streams are predictable and delivered on-time.

The evaluation’s limited analyses on these subjects suggest generally reasonable administrative costs, and probably less waste in conference and GRP management than in the past. GDN’s delivery of services generally is predictable, with a few notable exceptions with network partners and slow or delayed roll-out of GDNNet services largely outside of its control. Due to space considerations, these questions are handled in full in Annex T with specific examples to the contrary of these generalizations noted.

The section on activity analysis focuses primarily on measuring average unit cost per output across all the core activities run by GDN. The average cost for each activity is calculated by dividing the total cost incurred, segregated by output type if required, with the total number of each primary measurable output achieved. The time period taken for the analysis is fiscal years 2005-2007 for all the activities except for GRPs, because they are multiple-year projects and not activities that have annualized outputs.

2.5.2 Annual Conference

As discussed in annex L and N, GDN's annual conference includes a host of events with multiple goals and outcomes for all of GDN's constituencies. The cost effectiveness analysis for the conference seeks to measure and compare the total average cost (three years) of different programs that are run in the conference, in order to assess which program has been economically most worthwhile or most cost effective in achieving GDN's objective. More information on how figures were calculated is located in annexes U and V. GDN's audited accounts show decreases in percentage of GDN's total budget expended on the annual conference from FY 2004 to FY 2007 and a decrease in costs from FY 2004 (see Annex W). Table W-1 in the annex demonstrates the average cost over the 2005-07 period computed separately for different events and programs that are run in the conference. These include per capita costs for net conference costs, per paper costs, and per capita costs for side meetings. There is no objective way to determine whether the value judgements discussed previously that stakeholders (just above "moderately valuable") and conference goers (between moderate value added to high value added) gave to the conference are sufficiently high to justify the resources absorbed by the conference. These valuations should be kept in mind, however, when reviewing the cost efficiency numbers to consider cost effectiveness of the expenditures.

The method of calculation that makes the conference seem least efficient is the one that assumes the main or only value of the conference is paper output. The average per capita cost from FY 05-07 for the net costs of the conference is calculated by excluding the cost for board, RNP coordinator, and donor side meetings as well as post-conference GRP workshops, which are added into GRP costs, as discussed further in annex U. The per paper average cost is about \$10,500. Given the low capacity building involved for most papers and the variable quality discussed earlier, this would represent a cost inefficient process were this the most appropriate measure.

General networking and knowledge "cross-fertilization" across regions is another outcome of the conference, arguably best represented by total per capita costs. The average per capita cost for the conference including all activities is \$ 2,315, and the per capita net conference cost is \$1,256.

Given that some stakeholders indicated the primary value of the conference is networking and knowledge cross-fertilization for as well as, to a lesser extent, building capacity of more junior researchers, the evaluation calculated net conference costs of building capacity and networking for researchers below age 40. This is the target average age GDN is aiming for in its 2008 conference and only five years less than its maximum age for the Awards and Medals competitions. This unit cost of building capacity and networking among researchers under 40 is \$2,439. Extending the definition of junior researchers to people below 50 results in unit costs of \$1,557.

As discussed previously, these net conference costs exclude the cost of side business meetings conducted at the conference. The evaluation makes a rough estimate of these costs separately under the assumption that GDN otherwise would still hold a semi-annual board meeting and annual RNP head and donor council meeting. The

estimated average per capita cost for holding these side meeting is about \$4,000.66 Other non-measurable outputs that result from holding the side meetings in concert with the conference include building reputational effects, increasing probability of increased donations, increasing the buy in and understanding of key stakeholders, and the public goods of allowing others to network with donors and key stakeholders and hold pre- or post-conference workshops. The evaluation cannot provide per unit costs for any of these, but it notes that the other cost calculations implicitly share the costs to provide them and should be so noted as caveats.

The average per capita cost for these business meetings is considerably higher than the cost of networking and building research capacity and inflate the total costs (and per capita unit costs) of the conference, as they are not a cost efficient activity. They cannot, however, be ignored as a cost even if they are segregated in this analysis from other conference outputs. Thus, any net conference cost figures quoted (\$1,256, 1,557, or \$2,439) must also be accompanied by the per capita business meeting costs to understand total conference costs appropriately.

2.5.3 GRPs

Table V-2 in annex V demonstrates the average cost computed for each GRP that overlaps significantly the 2004-2007 period focusing on their total cost rather than an annual cost. Costs include direct and indirect costs and do not account for in-kind assistance of resource people. These represent an increasing proportion of GDN's budget, increasing to 27 percent of GDN's total budget by 2007. The per unit figure is computed two ways, by number of projects and number of total team members (primary and co-authors). The former assumes the primary investigator is the primary beneficiary of any capacity built and that the project itself represents one intellectual piece of knowledge added to the field. The latter assumes spillover effects to co-authors. The evaluation assumes these figures represent upper and lower bounds bracketing the "true" cost efficiency in that the former figure underestimates somewhat the total benefit of the project, while the latter probably overestimates benefits by assuming that all the team members benefited equally from the project. More importantly, neither figure adjusts for quality or intellectual contribution within its field of the resulting outputs across GRPs, which we would encourage be added in future analyses through an independent review process improving upon that taken in the proposal to paper review.

The analysis demonstrates that MIR represents the most cost efficient GRP embarked upon by GDN during this time period using either calculation method but particularly if one calculates using total number of team members (\$13,533). The total average cost per project (\$54,211) is only marginally less than that for the Impact project. The MIR project was not a large undertaking, has funded just six projects and does not appear widely publicized by GDN.

The Impact project has funded the highest number of research studies in different countries and the highest number of total team researchers. The total average cost per research team for the impact project is \$23,978 and \$57,830 by number of grants. The proposal to paper analysis suggests at least marginal value added to the team's work (although anecdotal sources suggest considerably more) and overall good knowledge added to the field from the Southern perspective.

The evaluation also notes that the size of the grants for Impact and MIR appear comparable to smaller non-mentored grants available to Northern researchers through sources such as the National Science Foundation. Were the point of GDN and the funders merely to generate knowledge from any source or perspective and not to

⁶⁶ The evaluation suspects this represents a slight underestimate assuming the costs of sending all board members to the US is lower on average than sending them to other locations, but it provides its assumptions in the annex and related spreadsheet available on request.

build capacity in developing and transitional countries, the projects probably would appear somewhat less efficient than these other options given that the output quality is unlikely not be on par on average. However, the whole point of these projects is capacity building, cross-regional knowledge exchange, and knowledge creation from the Southern perspective, which are not easily quantified so that the cost efficiency comparison should reverse given these additional “outputs.” Given these additional outputs, other potentially appropriate comparisons might include post doctorate positions combined with large workshops although noting that Northern post doctorates assume most capacity already has been built unless researchers are changing fields. The evaluation recommends GDN attempt to see if comparable figures for comparisons might be available.

The relatively low costs for MIR suggest GDN should consider whether smaller projects can represent cost effective alternatives to the larger projects when funding may be more limited or if GDN would like to attend to a greater number of issues of import, since it is not clear from a very broad perspective that there are considerable economies of scale.

The per capita average cost for the BRP and Health projects are considerably higher in comparison with the MIR and Impact projects on a per project basis (by a factor of 0.6 to 1.5), and more so comparing by total team members, where total per unit costs are 2.2 to 6 times as high as the other projects. A few stakeholders questioned the intellectual value added to the field of the smaller BRP project papers if not the overall subject matter, although none had read the synthesis paper.⁶⁷ Unless the intellectual contributions were great from this project, on this basic measure, the BRP project appears inefficient.

The evaluation team’s review, however, suggests that one of the primary benefits from the BRP project is to decide on and provide learning experiences to lead into follow up trainings. To the extent that GDN is successful in providing trainings that help bridge the gap between research and policy institutes in the South, this project would appear more cost efficient. The evaluation thus performed a basic thought experiment, where it increased the number of output beneficiaries depending on the number of trainings held assuming that each training has a “value” in terms of both direct and indirect benefits set arbitrarily at equivalent to 2.5 regular grants because of the potential spill-over effects on policy. If GDN performs only the one planned training, the resulting per “grant” cost reduces from about \$92,000 to about \$82,000. If GDN is able to conduct a series of trainings successfully, the “cost” per “grant” cost reduces to about \$57,000 with five trainings or about \$41,000 with ten trainings, comparable to or slightly more efficient than other projects. It is up to the next evaluation to decide whether GDN ultimately has successfully made this relatively inefficient project more efficient resulting in research policy linkages.

Like the BRP project, the Health project undertaken by GDN is also somewhat different from other GRPs and represents the most expensive GRP. The average cost for the Health project is \$134,908 per project and \$95,228 per research team member. The Health project is designed to generate capacity in evaluations as well as to inform national policymakers and international donors what types of interventions are most likely to be scalable and replicable and to inform researchers on accuracy of quasi-experimental designs in conducting evaluations. To the extent to which it is successful, additional outcomes should be added to the equation, reducing the per unit costs of the project. Since the Health project is ongoing, the evaluation does not have adequate information to evaluate the impact of the project and to assess whether the project has been a least or most cost effective undertaking.

⁶⁷ The evaluation found the synthesis paper clear and thorough research with a strong understanding of the inherent problems in the subject especially as related to developing and transition countries and a good analysis of potential next steps. The primary author, however, no longer is associated with GDN, diminishing somewhat the potential value for follow up trainings.

2.5.4 *RRCs*

The cost effectiveness analysis for the Regional Research Competitions measures and compares across regions total average cost per unit of output. GDN's audited accounts show an overall decrease in the total funds provided for RRC, from 2004 to 2006 a total average cost spent on RRC's of about 23 percent of GDN's total funds over the period. The allocation of funds is not uniformly disbursed across RNPs, with the formula having been revisited and revised slightly in 2007. The calculations are made similarly to those performed for GRPs both by grant and by total number of team members to provide upper and lower bounds of benefits per unit output.⁶⁸ The evaluation's observations suggest less consistency in mentoring and opportunities for networking across RRCs relative to some GRPs with some RRCs arguably providing as strong or stronger mentoring and/or networking and others less on one or both dimensions. This implies different value added, not captured in this analysis.

The average cost for holding RRCs across regions is lower than for GRPs and range from about \$10,000 to \$29,000 per project depending on region. CERGE-EI and EERC provide the smallest average grants, while EADN and ERF the largest. Without accounting for differences in value added, initial starting points, or quality of papers produced, CERGE-EI, ODN, and EERC appear more efficient in this simple measure. The cost effectiveness measures using total number of team members result in lower per unit costs, bringing all such costs between or below those for the MIR and Impact projects. The evaluation did not conduct measurement of value added or research outputs across all RNPs and GRPs. In theory, however, MIR and Impact provided more and more consistent mentoring and value added than some of the RNPs, particularly those that provide little mentoring, such as SANEI. Given the efficient and consistent mentoring process for recipients needing it in some regions, the RRCs appear efficient

2.5.5 *AMC and MIDP*

The evaluation performs similar calculations with the AMC and MIDP awards as the others, but it focuses on the awards and the medals competitions, as they are more comparable to the other activities. The analysis assumes fungibility of resources even though all of these are funded through an individual earmarked grant. Additional benefits include networking possibilities through attending the annual conference, but this also implies slightly higher travel costs. Analysis by total number of grants suggests grant sizes for the awards competition on par with GRPs and those for the medals competition comparable to RRC grants. The per team member figures for the awards competition is closer to the upper end for RRC grants. Unless the reputational benefits to grantees from the award is great, the awards grants do not appear a cost effective alternative to GDN's other activities. Since the cost effectiveness measure for the medals competition is similar to that for RRC grants, it would be as cost effectiveness as RRCs if it provides similar other benefits. The evaluation has not attempted to assess level of mentoring and networking provided through the medals grants, but it assumes them to be less than for RRCs on average and so somewhat less efficient than some RRCs although more so than others.

2.5.6 *GNet*

The evaluation's cost effectiveness analysis for GNet is intended as a framework for the independent GNet evaluation to expand upon given limitations in data availability to the evaluation and time to perform a more

⁶⁸ There are a number of qualifications on how the calculations had to be performed given that one RNP gave grants once every two years and that disbursements did not match allocations. The evaluation primarily uses total allocations given their availability and association generally with number of recipients awarded. See Annex U for details.

thorough analysis. The quantitative analysis covers only annual marginal additions to the knowledgebase, which represents only about a tenth of GDN's overall direct costs over FYs 2005 to 2006. The activities are not easily comparable with other GDN activities and thus the GDN evaluation should consider generating comparative figures for other organizations with similar activities, where possible.

The cost effectiveness analysis suggests an average of about \$400 for each researcher profile added, \$34 for each working paper added, and \$118 for each organizational profile added (see Table V-6 in Annex V). The evaluation is aware that adding and maintaining researcher profiles over time is a surprisingly time intensive activity given the lower Internet savvy of new Southern researchers relative to their average Northern colleagues with longer experience. Although the evaluation has not analyzed that figure closely, it appears that the costs are quite reasonable given these large time costs. GDN clearly should brainstorm about ways to reduce these costs, although it offers no suggestions and does not expect it to be able to reduce them greatly.

2.6 Overall Success

2.6.1 Overall Success Aside from Cost

In terms of impact, the evaluation also examined the overall success of GDN activities as well as the relative marginal benefit across activities. This is analyzed using professional judgment, asking key stakeholders through interviews and a survey about the success of GDN's key activities and reprioritization desirable, and asking grantees their relative preferences regarding funding of activities.

Key stakeholder groups most likely to have a sufficiently broad perspective of GDN to provide cross-activity value judgments were asked the value of GDN-sponsored activities on average for targeted participants. For the annual conference, the survey further disaggregated targeted participants into three groups: researchers; board members, RNPs, and donors; and GDN itself. Opinions varied widely.

Each stakeholder group separately rated all activities at least at the "moderately valuable" level, admittedly not a very high criterion. On average⁶⁹ though – even excluding GDN's staff and interns, who consistently provided higher average ratings than other stakeholders – the stakeholder groups provide an average rating across each activity no less than between "moderately valuable" and "between very and moderately valuable" (at 2.3, this is closer to the latter). The overall average across all activities is similar (at 2.2). Despite this apparent uniformity, a moderate degree of disagreement remains between and within groups, with moderately large standard deviations across all non-staff respondents for each activity of between 0.9 and 1.2. These represent good perceptions of value and yet with room for improvement.

The evaluation next considered grantee overall satisfaction with GDN's portfolio of activities. Three quarters of grantees rated their satisfaction with GDN's overall portfolio between medium and high with another 20 percent at very high with no statistical difference among groups.⁷⁰ Their average ratings, ascribing zero to four to each response, is above medium and three quarters of the way toward high. Respondents who have attended GDN's conferences rated their overall satisfaction slightly higher than non-attendees. Among all respondents including non-grantees, survey respondents from Sub-Saharan Africa and South Asia rated overall satisfaction higher than other groups, while those from LAC and MENA rated their satisfaction lower.⁷¹ Respondents who only know GDN through GDN's website rated their overall satisfaction slightly lower than others with an average right between medium and high. While a difference is expected, it is smaller than one might expect given the much higher inputs provided to grantees and the greater efforts toward capacity building. While the level of satisfaction

⁶⁹ Assuming a cardinal score of 1 to 5 with 1 representing "very valuable."

⁷⁰ No differences were found by age, position seniority, years doing research, or discipline.

⁷¹ Using the Wilcoxon signed-rank test. Overall satisfaction among CIS respondents was lower than average at $p \leq 0.10$.

overall is good, as a maturing organization, GDN should be working toward a more homogeneous product and thus, hopefully, a more highly positive response from those it is most focused on serving.

It would be unrealistic to expect uniformly high scores given the modest type and scale of activities GDN is undertaking relative to the long-term capacity building needs and knowledge building needs at which GDN's mission aims. Therefore, the evaluation interprets these fairly consistent scores (somewhat above the moderate/medium value level but not quite at high value level) across key stakeholders and target beneficiaries as indicating both positive value added from GDN's funded activities as a whole but with room remaining for improvement and greater success.

2.6.2 Relative Valuations

The evaluation next considers relative value across GDN's activities. Economic theory suggests that GDN's portfolio would be optimized if the marginal net benefits were equal across all activities. In a strict sense, it is infeasible or impossible to collect comparable data on costs and benefits and that one can compare or unitize in some way benefit gained across different potential beneficiaries. This is further complicated by the fact that some activities are intended to achieve more than one objective, especially the conference and GDNet. This analysis is not intended to suggest complete fungibility of resources, although respondents may, to some extent, assume so in their valuations. It also does not suggest GDN should not pursue a portfolio approach, as the evaluation supports earlier. However, it is important conceptually to examine relative values of the current parts of the portfolio for those resources that are fungible and to consider what might be added or changed in emphasis.

The responses of key stakeholders regarding the value of GDN's activities to targeted groups is one potential measure of relative valuations. Caveats include that many interviewees indicated lack of familiarity with GDNet given that they generally are not its target audience and that some respondents may have focused on valuations only as manifested in their regions. Stakeholders also were asked more directly an open ended question in the survey about changes or reprioritization desirable among GDN's major activities, and many of those interviewed were also asked to comment on this subject. Some key stakeholders, through interviews or an open ended survey question, indicate strongly their belief that the RRCs represent the highest marginal net benefit for junior researchers. It also was the highest or second highest ranked activity and target group across all four stakeholder groups surveyed.

There was no clear rank ordering in terms of value of activities of most other activities across all stakeholder groups, which were grouped tightly together. Somewhat surprisingly, the activity/target group combination that had the next highest average score and ranking was the annual conference, although the difference in ratings between it and even the lowest rated activities was only statistically significant at the $p \leq 0.10$ level.⁷² Even some of those most critical of the annual conference in interviews rated it "moderately valuable." The evaluation believes that the survey suffers less from "grade inflation" such that "moderately valuable" should be interpreted as a vote of at best weak support for an activity and more that those most critical think the annual conference is indeed "moderately valuable" but not worth the per unit cost, which the question did not address specifically.

Grantees were asked a more specific question related to relative valuations of GDN activities. The survey asks them if GDN had additional money to assign to any activity to benefit developing or transitional world researchers such as themselves, where would they advise GDN to direct that money. Item ordering was rotated across respondents. Up to three choices were allowed followed by the option to rank them (see annex X, questions 29-31). This question allows from the researcher perspective marginal benefit comparisons across

⁷² Even though small differences were found by stakeholder group for the annual conference by target group, the average across all groups was essentially the same and so are not discussed separately in detail.

activities to some extent by including the annual conference, RRC-style and GRP-style funding, separate policy or methodological trainings, increasing journal access, and augmenting dataset coverage. Based on suggestions from key stakeholders, the survey also included the alternative of using additional money to send researchers instead to other conferences.

The highest number of responses when three options were allowed were for both training researchers and institutes on how to influence policy and policy debates and funding workshop trainings on technical issues more effectively. The latter response is consistent with GDN's individual training needs survey. The next highest responses were funding additional research grants within their region and increasing access to journal articles or working paper series for those without access. When researchers were asked to choose their top priority, training on how to influence policy was highest followed by workshop trainings on technical issues and funding additional research grants within their region.⁷³

The next tier down of priorities among respondents were additional GRP type competitions and funding additional researchers to go to the annual conference.⁷⁴ In contrast, the annual conference was relatively low on the list in the top priority listing. The final tier was increasing access to national datasets and sending researchers to other academic conferences. This generic alternative to the GDN annual conference did not fare high, although it may have been hampered by not being tied to their own discipline. The evaluation suspects that increasing access to national datasets, mentioned by some key stakeholders, would create considerable value to a smaller number of respondents over the long-term, while most respondents focused on shorter-term and more immediate payoffs.

These stakeholder responses, combined with the cost effectiveness data and interview responses suggest the following observations. More direct trainings, especially if paired somehow with its other capacity building activities, is expected eventually to provide a high return if costs can be kept reasonable. The suggestion to pair it with other activities is a way to do just that in the same way that the annual conference could be seen as a more reasonable expense in the cost effectiveness analysis due to pairing it with the "side meetings" GDN requires anyway. Some stakeholders have suggested pairing these activities but not limiting the training just to grantees and not limiting attendance at the workshops to grantees either to expand the number of people who can benefit from watching the process and apply in the future.⁷⁵

⁷³ The same four discussed here appear at the top of the second priority list.

⁷⁴ This finding is consistent even among those who attended the annual conference.

⁷⁵ Conceivably, the same could be true with GRP workshops to create collateral benefits to people not granted but interested in the subject and process in locations where they are held.

3.0 FINDINGS: OVERALL PROCESS AND PERFORMANCE

3.1 Financial Sustainability and Risk Diversification

GDN has struggled toward increasing its financial sustainability and diversification. GDN failed to meet its over-ambitious financial goals of its 2005-07 strategic plan, but it has made significant, if qualified progress in diversifying its donor base for FY 2007 and 2008. Although GDN's viability looks strong and potentially promising to at least maintain scale in the near term, as with most grant-funded organizations, GDN's long-term financial sustainability is uncertain and will remain so unless and until the donor base of sizeable donors is widened and funding becomes less volatile. GDN will have to remain flexible in scaling activities up or down annually response to unavoidable fluctuations in overall funding. To a large extent, however, GDN appears able to do so.

The first step in considering GDN's overall performance and processes is to examine how well diversified and sustainable is GDN's financial position. Since there is no universally accepted definition for these terms in reference to organizations, the evaluation considers three definitions of risk diversity and sustainable funding: 1) How well has GDN been able to mobilize funds over past three years to implement its primary planned activities? 2) Has GDN met its own financial targets? 3) Is there a sustainable portfolio of donor support with sufficiently more diversification and smaller variation?

3.1.1 How well has GDN been able to mobilize funds over past three years to implement its primary planned activities?

The first definition – how well GDN has been able to mobilize funds to implement its activities – is stronger and more demanding if one focuses on funding all desired activities to the absorptive level of potential beneficiaries. A weaker, more reasonable, measure focuses on the ability to fund primary activities to a level deemed satisfactory to key stakeholders. GDN certainly has not been able to implement all activities desired to the high expense levels that beneficiaries could absorb, especially if one considers needs outside just economics. Stakeholders and GDN itself have provided the evaluation team with many ideas that arguably could fit comfortably within GDN's mission – a few of which would require enormous resources were they taken to full scale across regions. In terms of the weaker definition, GDN's revenue dropped each year from fiscal year 2004 to 2006. It dealt with that problem, as suggested in the strategic plan, through cutting back the scale of some activities during that time period but continuing to fund them to a lesser extent through shifting funds and eventually looking for additional funding sources. Although stakeholders may indicate preferences for changing in funding priorities or additional spending by GDN, they do indicate that GDN is meeting all of its current set of planned activities at at least a minimally acceptable level.⁷⁶ GDN thus succeeds in the weaker sense of this definition of financial sustainability but does not in the stronger sense of the definition.

3.1.2 Has GDN met its own financial targets?

From GDN's 2005-07 strategic plan, its primary financial targets, as with many organizations, aims ambitiously for robust growth of preferably 50 percent or if not 30 percent of the overall budget from FY 2003 to FY 2007. The budget for FY2003 was GDN's largest at \$10 million in contributions, and the assumption of evenly continuing levels of World Bank funding in the plan was errant. Thus, GDN did not come close to meeting either ambitious target. Its best year was FY 2007 at \$9.3 million after three years of declining revenues.

Because World Bank funding has declined by more than half over the stated period (from \$7.9 to \$3.7 million), the evaluation also examined GDN's non-Bank funding. Due primarily to the Gates Foundation funding, this

⁷⁶ This section ignores the distributional issue of whether resources would be better spent on other activities.

part of GDN's moved up from a slight increase between FY 2003 (\$2.2 million) and FY 2006 (\$2.5 million) to doubling in FY 2007 (\$5.6 million). The FY 2008 non-Bank budget is projected to be somewhat lower as of the report date at \$4.3 million, but could increase since many donors pledge funds after GDN's fiscal year begins. Until FY 2007, this non-Bank funding increase was more than offset, however, by steady decreases in funding from the World Bank and the dramatic reductions in funding from a dependable, large bilateral donor (Italy) that dropped GDN from its portfolio after FY 2005 due to internal politics according to stakeholders. Further, to a large extent, the recent increase in non-Bank funding largely is through earmarked funding, constraining GDN's ability to shift resources among activities. This constraint matches an unfortunately common trend among donors to provide only earmarked money and means GDN cannot move all funding fungibly to meet highest perceived needs or desires of stakeholders.

Some key stakeholders have stated that GDN's mission is not one likely to connect with a highly diverse and rapidly growing donor base. This is corroborated by GDN's lack of success despite concerted efforts in the corporate sector – a suggestion highlighted in the previous evaluation. The implication is that a target for strong growth is unrealistic in the face of a continuing slight downward trajectory of Bank funding, and maintaining or slightly increasing the budget over time is more realistic and could itself be considered a success given its niche mission and difficulty in selling a mission based on research given current funding priorities.

The evaluation also examined briefly the subgoal in the strategic plan of reducing dependence on World Bank funding “from more than 65 percent [in actuality, 78 percent] to 40–45 percent” of its total budget. It must be noted that stakeholders indicate that the desired target negotiated within the Bank was, in fact, to be under a third and the general rule for funding for the Bank's Development Grant Fund is that its funding should constitute 15 percent or less of total funding. Here, GDN has a partial, qualified success with the Bank's share of funding dropping to 40 percent in FY 2007.⁷⁷

One qualification is that the strategic plan unrealistically assumed constant levels of funding from the Bank, but part of the way that GDN has “met” its goal of reduced dependence on World Bank funding was through the Bank reducing contributions, which was not the planned method to reducing dependence. Had Bank funding instead remained constant, its share of contributions to GDN would have dropped from 78 percent only to 59 percent. Another qualification is that GDN has not yet reached the real target of 33 percent, although GDN's calculations indicate that if all additional pledged and active prospective donations are realized during the year, the goal would be nearly in reach. The evaluation supports this contention but also notes that, in rare cases, funding also can decrease. A final note of concern is that if GDN's support within the Bank wanes, the demand may be for funding levels to reduce to no more than a typical 15 percent of the budget or less, given the constant competition for Development Grant Fund money.

On the other hand, although the TOR does not call for comparisons with other organizations, discussions with World Bank staff suggest GDN is in not at the bottom of the pack of Bank spin-offs in terms of creating an independent portfolio. The evaluation notes that stakeholders at the Bank, while expressing concern about GDN's progress, also indicate that it has performed better than some other Bank creations that also have had difficulty generating wide ranging support to reduce their dependence on Bank funding.⁷⁸

⁷⁷ The measure is slightly higher at 45 percent for all confirmed funding at the beginning of FY 2008, although the final figures could change most likely downward.

⁷⁸ The evaluation was unable by the report deadline to gather sufficient information to make appropriate comparisons with organizations such as the Development Gateway Foundation, Africa Virtual University, Consultative Group to Assist the Poor, or other organizations.

3.1.3 Is there a sustainable portfolio of donor support with greater diversification and smaller variation?

The third definition considered is whether there is a “sustainable” portfolio of donor support with sufficiently greater diversification and smaller variation than in the past. The first part of this definition (whether there is a sustainable portfolio of donor support) is subjective and could be described in the affirmative and negative given the evidence, and indeed has been by various stakeholders. There has been greater donor diversification, even if not fully meeting stakeholder desires. The answer regarding lower variation is mixed in that financial fluctuations in GDN’s donor database have decreased if one includes the World Bank’s lower recent funding levels, but it has increased in terms of fundraising among donors other than the Bank. The latter includes both “good” variation in terms of large recent multiple year donations and “bad” variation in the loss of a large consistent donor during the period in question. To the extent that GDN is able to get more large, volatile funding for specific activities, the larger budget would help offset some of the fluctuation from lack of dependence on a few large donors.

The stakeholder survey asked respondents a question speaking to the first two parts of this definition. It asked to what extent GDN’s financial position is stable with a sufficiently diversified donor base sufficient to implement planned activities and goals. This is one of the three areas in the survey, where stakeholders rated GDN lowest. Even including staff, the average rating⁷⁹ across GDN stakeholder groups is about half way between “reasonable” and “good.” Stakeholders indicate they feel more comfortable than in previous years that GDN’s portfolio is appearing more sustainable than many feared heading into FY 2006.⁸⁰

One measure of greater diversity would be diversification by donor type.⁸¹ In response to funding concerns, including that highlighted in the previous external evaluation, GDN placed concerted efforts in increasing funding from bilateral donors and corporate donors. The efforts with bilateral donors appear to be mixed but largely succeeding, although not at as high levels as preferable, while efforts with corporate donors were largely unsuccessful. GDN generally seems to be making modest progress in its relationships with donors recently. Nonetheless, during this three year time period, a political shift at its largest bilateral supporter in FY 2005 (Italy) decreased its support from over \$1.1 million (earning it a seat on GDN’s board) ultimately to zero. No substitute very large bilateral donor has emerged since then, although it appears GDN is making inroads and increasing support from some key medium-sized donors, particularly Netherlands, DFID, and AUSAID, which have provided multiple year funding.

During the most recent three fiscal years, GDN formed a corporate council but only succeeded in raising significant money from one already supportive corporation and, ultimately, one corporate member’s foundation. The latter was sizeable though – representing \$5.9 million from the Gates Foundation and represents an advisable target for future fundraising. Nonetheless, the Evaluation team’s understanding of GDN’s experience looking for funding directly from the corporate sector as well as the team’s own understanding of fundraising with private sector donors suggests that GDN was correct in not continuing to pursue corporate funding as a main emphasis.

⁷⁹ Converting responses into cardinal measures from 1 to 7.

⁸⁰ Although a question directly comparable to that from the previous evaluation was not asked, two-thirds of the board members solicited in the prior evaluation indicated they were not satisfied with GDN’s efforts at raising funding from new sources. Only one board and one RNP respondent to the 2007 survey rated GDN’s financial base below “reasonable” suggesting some progress on this measure.

⁸¹ The evaluation does not consider diversification of income type by income such as user fees, as such income seems to violate its objectives.

A measure on which GDN has seen modest success is in basic measures of growing the donor base. In the two most recent fiscal years, the average total number of donors of any size and the number contributing at least \$100,000 both increased somewhat from the FY 2002 to 2004 period and increased dramatically over the problematic FY 2005. (See Table 3.1-1, below.) For example, the number of donors contributing at least \$100,000 averaged 7.3 from FY 2002 to 2004, fell to 5 in FY 2005, and then rose to an average of 11 in FY 2007-2007. The average number contributing at least \$250,000 decreased between periods, as GDN recovered from a drop-off after the highly successful FY 2004. The total numbers appear somewhat lower for FY 2008 in part because not all donors may have committed to funding in a way that maps with GDN's fiscal year.

Table 3.1-1: Number of GDN Donors by Year and Averages across Three-year Periods

	FY'02 (01-02)	FY'03 (02-03)	FY'04 (03-04)	FY'05 (04-05)	FY'06 (05-06)	FY'07 (06-07)	FY'08 (07-08)*	Ave. FY02-04	Ave. FY05-07	% Change
# of Donors	14	11	17	11	17	21	10	14.0	16.3	17%
# of Donors, > \$100,000	7	5	8	5	9	12	6	6.7	8.7	30%
# of Donors, > \$250,000	3	4	7	3	4	5	5	4.7	4.0	-14%

*Note: Committed as of September 2007.

On a more troubling note for RRCs and the annual conference, the amount of unrestricted funding has been decreasing steadily for GDN over time – partially a symptom of a broad increase in donor preferences for earmarking funds. This suggests that GDN's activity mix is constrained by preferences of donor as well as targeted constituents. GDN has indicated it is looking for additional sources of unrestricted income and for funding of selected new activities that also parallel constituent preference.

To try to provide an objective measure to prior observations about volatility that can be used in future evaluations, this study examined two measures of funding volatility. The first is variation across all active donors taken together for each given year.⁸² Smaller values for this measure favor greater homogeneity in funding levels across donors (implicitly valuing more highly lower dependence on a few large donors). Using four different calculation methods, the evaluation found the same general result – that the average year-over-year standard deviation has decreased from the FY 02-04 to FY 05-07 time periods considerably by between 30 and 38 percent. (See Table Y-2 in Annex Y.) This is largely due to decreased funding levels by the World Bank, and if the Bank is excluded, this measure of variance among all other donors actually has increased due to the high difference between the top “other” donors 2004 (Italy) and 2006 (Gates Foundation) versus all others. This suggests increased variation in funding levels across donors besides the Bank, but decreased variation across donors including the Bank – on net a positive in the short-run as long as GDN can find substitutes for its current big non-Bank donor in the future.

The second measure is variation (standard deviation) for giving by each individual active, moderate sized donor across three year periods.⁸³ Smaller values for this measure favor greater consistency in funding by each donor across time.⁸⁴ The average standard deviation measured this way similarly has decreased by 38 percent. Again, this is largely due to decreased funding levels by the World Bank, and if the Bank is excluded, this measure of

⁸² This is measured as standard deviation (a measure of the spread of data about the mean) across all donors, either excluding all non-donating years or including one year after funding stopped and any non-donating years between any donating years.

⁸³ Including up to one year after funding ceased and any years without donations between donating years.

⁸⁴ Drawbacks include that it provides inaccurate measures if a donor provides funding designated for two years in a single year.

variance remains about even between the two time periods, suggesting no change in funding volatility across donors besides the Bank. GDN management and other stakeholders indicate it is unlikely they could locate donors willing to establish a sufficient endowment that would reduce this measure.

The evaluation did not find evidence of continuing reluctance among donors to contribute to GDN due to the “World Bank whiff” smelled so strongly in the prior evaluation (no responding donor indicated dissatisfaction in GDN’s independence from the Bank) nor direct evidence of reluctance to pay overhead rates to the secretariat (although noting again the strong bias in having time only to interview generally strong supporters). It did note a potential weakening of ties between GDN and its primary funder – a function of its moves from the Bank, personnel changes in both organizations over time, and decreasing inter-dependence in part to meet stakeholder demands for independence. While this may have increased GDN’s legitimacy among its constituency, there is potential that it ultimately may endanger its connection to its most consistent and important funding source due to continuing demands within the Development Grant Fund to decrease funding for existing donees, especially long-term ones such as GDN. Given the current trajectory of this relationship, it makes even more crucial GDN’s ability to succeed in capitalizing on the perception of its independence from the Bank,

3.2 Governance, Organization and Interaction with Regional Partners and Stakeholders

The evaluation team has attempted to assess the governance and management of GDN by soliciting qualitative and quantitative responses from key stakeholders that include the board members, GRP coordinators, RNP coordinators, donors and GDN staff in gauging the organizational traditions that effectively support the functioning of GDN and the meeting of its strategic objectives. The first section reflects the aspects of governance of GDN and the ways in which objectives are fulfilled. The second section of discusses the management processes undertaken by GDN that supports GDN to implement its objectives efficiently and effectively.

3.2.1 Organizational Structure

The analysis of organizational structure indicates there is some disconnect between GDN management’s description of GDN being “an efficient administrative mechanism to conduct research” and a quasi-think tank model assumed by some stakeholders, who suggest greater research capacity in the secretariat. The evaluation does not advocate the latter view but does suggest it as an option worth exploring as a way to increase the extent to which highest quality work is facilitated as an objective. It also notes the value of adding staffing with fundraising expertise, although the cost-benefit ratio for an expensive position such as that is unclear for a research organization and could be considered as a provisional appointment for 18 to 24 months based on success.

As part of its governing structure, GDN has a board that provides strategic direction and management oversight. GDN has three offices, with the Secretariat based in New Delhi, India, the GDNNet office in Cairo in Egypt and a smaller holding office in Washington D.C. Management responsibilities and roles are well defined in job descriptions, and top staff coordinate regularly across the three offices. The president has been involved in each of GDN’s activities that include financial, programmatic and administrative functions.

GDN has a small program staff in Delhi that handles the research activities, finance, and communications with stakeholders. The staff primarily consists of administrative and logistical staff as well as interns handling arrangements most notably for the annual conference, which absorbs much of the third of the year leading up to

it. The finance staff attend to financial management by conducting reviews, budgets and overseeing regulatory compliance issues. The small technical staff in Cairo are managed by the GDN Director.

Some stakeholders indicate staff research capacity is too limited, which is reflected to some extent in the stakeholder survey responses with board and RNP coordinators on average rating GDN in this area as just below “good.” GDN’s former president accurately indicated that GDN was, at its essence, an “efficient administrative mechanism to conduct research” globally and across regions rather than a research organization itself. Thus, there is some disconnect between the image and goal of GDN’s administrative organization and stakeholder perceptions. As an organization and network that promotes and facilitates research in developing countries, several stakeholders believe GDN should consider hiring additional senior research staff to be engaged primarily in research activities without having to handle administrative functions for the conference or else providing greater intellectual oversight for the conference and GRPs. The concept is that doing so would allow GDN to move toward its objective of facilitating or producing highest quality research. The evaluation does not necessarily advocate such a quasi-think tank model, although it does suggest it as an option worth exploring under the new president’s strategic review.

As an organization that ideally should diversify its donor base and size, GDN should consider carefully hiring a next candidate hire with fundraising expertise. One model is to hire a fundraiser as a fulltime staff member or consultant based in the US or Europe dedicated to the process of identifying and establishing contact and maintaining communication with donors. Such a position, however, might be weaker handling programmatic interactions, would have to have relationships with bilateral and multilateral donors which most do not, and tends to be expensive, paying for itself over time. There is no simple answer regarding the cost-benefit ratio for such a position, but many organizations make such appointments provisional for 18 to 24 months based on success.

3.2.2 Board Composition

The analysis of board composition suggests significantly lower division in the board than previously. Key issues appear to be the need for greater transparency in the process of choosing the board members and processes to guarantee some members have experience in managing organizations or boards. An additional donor representative could be considered to help groundtruth GDN’s ideas from the perspective of funding. Despite board member interest in working more on strategic issues, the board is unlikely to be able to do so meeting only twice a year unless meetings are extended in number of days, one or two virtual meetings are held during the year, and more committees are formed to handle the business issues. The evaluation also has not been made aware of a general conflict of interest policy for board members, in line with best practice.

GDN has a fourteen member board with representation from three disciplines and across all regions in which GDN operates. The composition of the board consists of primarily economists with representation from network partners from the South and non-recipient partners in the North and two social scientists and a political scientist. Stakeholders view the board as independent, democratic and legitimately representing the world’s research community. Unless special exception is made, members rotate after two two-year terms, increasing democracy but reducing institutional memory. The board convenes twice a year with one of the meetings occurring at the GDN annual conference. GDN has moved to address some concerns raised by the previous independent evaluation in 2004 related to lack of representation of donors, although some concerns remain true for the current evaluation period including limited membership with management experience.

The evaluation encountered general agreement among most board members on satisfaction of the representation on the GDN board since they are mostly researchers and want mostly researchers on the board. However,

opinions on the recruitment of board members were mentioned to be opaque by some stakeholders who suggested a need for greater transparency in the process of choosing the board members. Regarding representation of donors, some preferred only advisory representation of donors and not a fixed proportion on the board, while others with greater fundraising experience indicate that donor representation implies improved prospects for mobilizing funding and motivation and perceived need for change was low in this regard; and the evaluation team's analysis concurs. An additional donor representative would help groundtruth GDN's ideas from the perspective of funding, although most board members prefer adding more researchers. The key membership issue appears to be that recruitment processes do not appear to guarantee any members have experience in managing organizations or boards, which is expected to become a problem with the upcoming retirement of one of the board members. In the evaluation team's experience, this is a key problem that weakens board governance, oversight, and direction setting, making GDN a primarily management-run operation.

Board members' key concern with their support to the organization indeed is that the board focuses too much time on management issues rather than on strategic direction. Divisive issues during the previous three years included GDN's legal status, GDN's move to Delhi, increasing Southern representation on board, and two objectives: 1) the tension between ensuring capacity building and producing highest quality research output and 2) disciplinary approach. The most divisive issues had been settled, however, before many current members arrived, leading to a less divided current board. Newer board members reflected the need to expand the notion of development to political science and sociology. There were some, but limited concerns with the strong World Bank ties among a few board members and President, which could raise suspicions among potential donors, although the evaluation encountered no such concern among the very limited set of interviews with donors it contacted.

Since the board convenes only twice a year, even with email, the intensity of the board's engagement and its oversight capacity is limited. Further, the high name recognition of some members also means that they have limited time to attend to GDN and have competing fundraising objectives that limit their promotion of fundraising for GDN. Most board members believe that meeting twice in person with emails in between is adequate given the busy schedules of board members but that the board needs to organize itself to spend more time providing inputs on strategic issues rather than attending to management issues. The board has only two standing committees – the audit and program committees – plus an executive committee all created in response to the last evaluation in lieu of the concept of creating a second business board, as recommended in the previous evaluation but not supported by the board. There was some concern that board members become involved in micromanagement in response to divisive issues, and that it might better be able now to focus on strategic issues. In the evaluation team's experience, it is difficult to understand how the board will be able to move in such a direction meeting only twice a year unless meetings are extended in number of days, one or two virtual meetings are held during the year, and more either committees are formed to handle the business issues or the committees become more active by being required to meet more frequently. Each option should be considered by the board chair and new president.

The evaluation has not been made aware of a conflict of interest policy for board members, in line with best practice. If not already adopted, GDN should develop and adopt one. Some organizations are moving to having board members sign a conflict of interest policy annually, and GDN should consider doing the same.

3.2.3 *Interaction with RNPs*

The evaluation next considers the extent to which GDN's interactions with its RNPs is effective. The evaluation does not find a clear reason to change models dramatically, although GDN should clarify the extent to which it is a network of networks – that is, more of a partnership or more of a grantor/grantee relationship, increasing contacts if more of a partnership approach is chosen. The evaluation supports GDN's new efforts toward creating dedicated work plans with joint fundraising with RNPs, which could prove financially rewarding and decrease perceptions of potential competition for fund raising and increase funds that can be leveraged regionally. GDN should consider a more detailed, but still streamlined, reporting system to better monitor the grant fund distribution to categories of researchers and overhead costs, possibly better matching systems used by other donors. GDN also should be careful not to reduce overhead costs too much to partners despite shrinking budgets.

In the stakeholder survey, the board members, RNP and GRP coordinators, donors and GDN staff rated the clarity of roles and responsibilities among GDN's functional units and partnerships as between good and reasonable. This suggests an area where further clarification and improvement is possible. Working across regions, GDN cannot afford to have a uniform mission but a broader mission that encompasses flexibility for each partnering constituency. One aspect of these interaction is what is the appropriate model for interaction with its partners. The evaluation notes four such models:

1. One essentially permanent partner per region unless there is lack of agreement over key goals or another organization would better be able to meet goals.
2. One relatively permanent partner with other associate partnerships with other types of policy-relevant organizations to create network synergies found with its current African partner.
3. More than one relatively permanent partner to achieve multidisciplinary, with both competing for funds and contributing suggestions to GDN as appropriate.
4. Keep RNP competitive – literally competed every few years.

GDN funding for RNPs is relatively modest and is not expected to increase as non-earmarked funds decline, making the third model less compelling. Only very few stakeholders suggested the fourth model, which the evaluation believes is impractical given the institutional memory issues required to run annual RRCs and regional windows efficiently and effectively. It also does not solve the issue of the appropriate criteria for selection and presumably might result in the same RNPs with lower levels of partnership commitment. Some stakeholders find the second model compelling, but GDN and its RNPs would have to figure out whether such a model is feasible and how to engage and encourage such partners as well as whether each RNP really has interest in such working relationships. Doing so would augment GDN's ability to claim increased policy dialogue though.

Even if GDN sticks with its current institutional arrangement (the first model), a related question is the extent to which GDN should be a network of networks – that is, more of a partnership or more of a grantor/grantee relationship. For instance, until recently, GDN has had no agreements with its network partners regarding joint resource mobilization efforts. GDN has just signed an agreement with one RNP toward a dedicated work plan with joint fundraising, which could prove financially rewarding and decrease perceptions of potential competition for fund raising and increase funds it can leverage at the regional level although the division of responsibilities for fund raising may not be spelled out completely. Supporting participation in each others' activities comprises a positive direction. There appears to be considerable interest among partners to extend this pilot idea, and the evaluation supports GDN's continued efforts to do so.

GDN maintains direct links with the RNPs through dedicated staff members for reporting and maintaining communication with the RNPs. Nonetheless, most RNPs believe that GDN requires more frequent and interactive contact with the RNPs for a good partnership. According to RNPs, interaction is limited to emails

and infrequent telephone calls and one rushed meeting with the secretariat staff and board members, and little is done to encourage or foster cross-RNP interactions. Further, some RNP coordinators mentioned specifically concerns that access to the new president has been more indirect requiring more scheduling through an assistant, although the evaluation team has not followed up to see whether this may just be a process of adjustment in part involving multiple time zones until the new president arrived in India.

A few RNPs and stakeholders are concerned about the lack of monitoring by GDN on the utilization of grant funds and suggest a more detailed reporting system to better monitor the grant fund distribution to categories of researchers and overhead costs. The evaluation notes that the system should not be overly complicated and hopefully could better match that used by other donors – a complaint that arose, albeit infrequently – in both the previous and current evaluation.

Over the last two years, in order to strengthen transparency of outcomes and financial accounting procedures, GDN has incorporated the following changes in the RNP contracts after the evaluation in 2005. The clauses included acknowledgement of GDN as a source of funds by RNPs and the grantees, the addition of the categories of gender and age of the grantees in the reporting format, and a requirement of RNPs to maintain accounts a year after reporting to GDN.

Most RNPs agreed that the proportion of grants for RNPs to support the networks themselves rather than deliver products has shrunk noticeably. This decreases motivation for RNPs to remain as networked, dynamic partners and ultimately could imperil relationships if RNP boards perceive the networks as subsidizing GDN.

Suggestions were highlighted as to making better use of grant funds through inter-regional cooperation; however, the stakeholders did not provide a clear mechanism for doing so, and the evaluation does not have evidence regarding the effectiveness of such an engagement. The evaluation team also suspects the secretarial lacks time to implement such changes as well as plans to augment individual training given the current annual conference schedule.

Some RNPs think that the allocation of RNP funds has been reasonably fair across RNPs while others expressed a need for allocation of funds to RNPs to be held on a competitive basis or a needy basis. GDN's board, however, has spent considerable time in analyzing and revising its formula to account for these competing objectives. Such democratic revisions mean that neither objective is fully met, but the evaluation does not foresee a dire reason to revisit this policy in the immediate future rather than focus on other priorities.⁸⁵

3.2.4 Interaction with Target Groups and Other Organizations and Communications

The evaluation team collected data on the ways in which GDN interacts with its target groups and establishes policies for inclusion of stakeholders in programmatic activities. The extent to which GDN seeks the participation of targeted beneficiaries in planning activities was the lowest rated area in the stakeholder survey among board members, although the evaluation did not uncover many specific complaints in this area. As with the prior evaluation, this evaluation does not see a workable alternative model for administering GRPs given their size and subject matters, and it believes GDN does not overly centralize intellectual leadership currently. Stakeholder survey results suggest GDN might more actively participate in other existing networks within its field and build further strategic alliances over time. The evaluation team also notes branding and self-marketing as an area for additional review and improvement. GDN has responded to the evaluation team noting their

⁸⁵ An additional suggestion from a couple of stakeholders included using grant funds for inter-regional cooperation, however the stakeholders did not provide a clear mechanism for doing so, and the evaluation does not have evidence regarding the effectiveness of such an engagement. The evaluation team also suspects the secretarial lacks time to implement such changes as well as plans to augment individual training given the current annual conference schedule.

awareness of the branding issues given that branding is discussed in background notes for the GDN strategic retreat in Washington, DC in 2007.

GDN's interactions with its target RRC groups is mediated through the RNPs, mentors or the major communication activities like GDNNet, the annual conference, and regular electronic and print newsletter. GDN also periodically conducts focus groups of RRC grantees through programmatic staff, who attend RRC workshops, and it has more direct contacts with GRP and AMC grantees and conference participants, since those activities are more centralized. It also appears to be setting a process for triennial stakeholder surveys. These are all appropriate ways to solicit input and feedback.

GDN has contacted the researcher base through surveys. This does not mean, however, that there is a continuous long-term two-way interaction process with research grantees, especially if grantees do not sign up for electronic newsletters. Discussions with GRP and RRC grantees suggest that RRCs should continue to remain in touch with former grantees after the completion of the grant with more of them taking roles of discussants, moderators, and mentors to younger researchers in workshops and conferences, as done in some regions to further augment the capacity building process.

In the stakeholder survey, RNP, GRP, and regional window coordinators rated the extent to which GDN seeks the participation of targeted beneficiaries in planning activities and enjoying their trust on average a bit higher than good, while board members rated the extent at just below reasonable. This is, in fact, the lowest rated area in the stakeholder survey among board members. Although the evaluation did not uncover many specific complaints in this area through interviews or the survey,⁸⁶ GDN may wish to follow up with board members on their concerns in this area.

For GRP projects, GDN enters into a contractual agreement mostly with individuals or otherwise with institutes for provision of technical project management support that clearly outlines delivery schedules and responsibilities for performance. Some stakeholders believe that the strategic objective to engage and build capacity of researchers in the Global South is weakened by centralization of intellectual leadership by GDN for GRPs and to some extent the annual conference making researchers merely static research assistants to western academics. The concerns in decentralization, however, are increased variation in quality of the research conducted and larger time and financial costs to do so than is reasonable for the smaller GRPs following its first GRP. As with the prior evaluation, this evaluation does not see a workable alternative for the current size GRPs, and it believes GDN does not overly centralize intellectual leadership currently.

In the stakeholder survey, the board and RNP coordinators on average rated the extent to which GDN actively participates in existing networks within its field and builds strategic alliances between good and reasonable – one of their lowest ratings across the dimensions analyzed. Specific criticisms are that GDN should consider membership in additional consortiums and participation in conferences aside from its own and RNPs, where there might be opportunities to promote the network. Increased effort in this area would require time from more senior-level staff and should tie in to its branding, communications, and strategic plans. Working with additional regional networks at a minimum to promote its grant competitions is also worthy of consideration. GDN's recent addition of a North American partner is at least mildly helpful regarding participation in additional networks.

⁸⁶ To the extent that the team received specific critiques, one was that there is scope for improvement in this area with respect to the Sub-Saharan and Pacific regions.

GDN has not asked the evaluation to consider its success in branding and self-marketing. The evaluation team noted this as an area that GDN should consider reviewing in greater depth separately. In response, GDN notes that this is a central area in its own concurrent strategic review and was discussed to some extent during the 2007 GDN Friends Retreat in Washington. As GDN appears aware, although GDN has a communication officer who coordinates the newsletters and website activities of GDN's portfolio, it does not have a coherent motto or branding strategy.⁸⁷ GDN noted to the evaluation team that not all grant recipients know their grants come from GDN, although the organization has recently changed its contracts with RNPs in 2007 to require grant recipients to indicate GDN as a sponsor. Nonetheless, GDN's name had not been required on flyers, posters, email solicitations or RNP workshop and conference materials and there is no standardized format for branding and acknowledgement of GDN. In response, GDN notes that most, but not all, RNPs now use GDN's logo for their workshops and other allied activities. These reasons represent some of the reasons stakeholder indicated reputation and communications as one of the weakest areas in the stakeholder survey with almost half of respondents rating GDN as only "reasonable" or lower with an overall average only about "reasonable." The evaluation team applauds GDN's understanding that this is an area for further development.

3.3 Management Processes

3.3.1 Organizational Policies – Transparency/Disclosure and Democracy

The stakeholder survey asked about two dimensions of application of organizational policies: 1) transparency and disclosure and 2) democracy. The evaluation found that a desire among RNP heads and board members for a more participatory approach in selection of GRP and annual conference themes, especially to the extent that GDN decides it is a network of networks rather than a network overseeing grantee networks. GDN recently has shown increased openness to donor inputs, although it is unclear that donors consider the input is meaningful. In terms of GDN's financial policies, GDN conducts annual audits and conducted in the summer of 2007 a review by the Bank's Development Grant Fund unit of its financial policies. Although the evaluation did not receive from GDN the DGF group's report, GDN appears to be responding to the recommendations by incorporating the suggestions into mainstream financial policies. The DGF's review suggests GDN has reasonable to strong systems given its age and size.

Board members on average rated GDN just below good on transparency and disclosure, while RNP coordinators rated it just above good, and staff and GRP and windows coordinators rated GDN slightly higher. Responses regarding democratic application of rules were slightly lower. Board members on average rated GDN half-way between reasonable and good on transparency and disclosure, and RNP members rated it good. These responses suggest some room for improvement, especially on being more participatory and democratic. Discussions with RNP coordinators and board members highlighted the need for a participatory approach in selection of GRP and annual conference themes, which are seen as somewhat transparent but not very participatory, again to the extent that GDN decides it is a network of networks rather than a network overseeing grantee networks. The RNP coordinators recommend that GDN should invite them to select themes or select a theme from a pre-selected group of themes. It has also been suggested to the evaluation team that if GDN can figure out a way to better incorporate RNP input in selecting GRP grantees, it might be able both to increase perception of democracy as well as to better standardize researcher quality in the GRPs assuming they have or can make sufficient connections across relevant disciplines. In terms of GDN's openness to donor inputs, the evaluation notes the recent formation of the Donor Advisory Council as an important step, although it is not clear that donors feel the input is meaningful and paid attention to given their limited role on the board.

⁸⁷ The evaluation did not analyze specifically whether its current communications plan matches that of other similarly situated organizations, but it suspects GDN could improve in this area.

In terms of GDN's financial policies, GDN has an accounting manual that incorporates the financial management policies and a financial controller. GDN conducts annual audits and conducted in the summer of 2007 a review by the Bank's Development Grant Fund unit of its financial policies. The evaluation did not receive from GDN the DGF group's report. According to GDN's response, the DGF group made four recommendations. The two non-technical recommendations include 1) conducting external verification of grants and 2) procuring a management recommendation letter from RNPs such that if any recommendation has been made by an auditor to RNPs, the implementation of the recommendation can be checked by GDN. GDN has responded to these recommendations by incorporating the suggestions into mainstream financial policies. GDN does not appear currently to be spending ahead on expected but unrecognized revenue, as pointed out as a concern in the previous external evaluation, although that is always a potential concern and source of caution with an organization like GDN with predictable and set schedules of deliverables (one of the evaluation criteria) but with not always predictable donor schedules. The DGF's review suggests GDN has reasonable to strong systems given its age and size. The evaluation did not duplicate the DGF's review and merely recommends that GDN continue to improve its systems as recommended in the review (see Annex Z for DGF review).

3.3.2 Management and Feedback Processes

Stakeholder survey results and interviews suggest that stakeholders feel GDN has some feedback learning processes built into the system but that they could be expanded. At each board meeting, GDN staff presents updated information on activities to the board members in the form of a progress report. References to targets on inclusion of categories of conference participants have been reflected in the databases on statistics on conference participants for the last two years. GDNet's quarterly report as well mentions status report against targets in formulated plans, and the statistics for GDNet and Knowledgebase users are circulated by the Cairo office on quarterly basis. GDN's response notes that the conference C-Tool database has the profile information of conference attendees starting from 2003 available at any time. In the stakeholder survey, board members and regional window and GRP coordinators on average provide GDN a score a bit below "good" on the dimension of evaluation and feedback.

However, RNP coordinators on average provide among their lower scores to GDN on this dimension, providing a score of just above reasonable. In terms of feedback with RNPs, meetings held of board members and RNP coordinators are generally considered too large and brief given the conference's many distractions to provide effective feedback and two-way and horizontal interaction and thus makes communication more top-down. GDN's staff have a presence and participate in RRC workshops, although it appears that the interaction has been modest in most cases and might be increased somewhat to foster dialogue and effective feedback. There also is annual reporting from RNPs on funds spent, but this represents monitoring rather than a feedback mechanism. GDN previously conducted a needs assessment questionnaire with RNPs presented at the meeting in Dakar, Senegal 2005⁸⁸ as a one-time feedback mechanism. GDN should look for ways to build and improve relationships with RNPs through establishing more regular, horizontal interactions beyond the annual conference meeting. GDN should consider experimenting with ideas such as at least biannual meetings just of RNP coordinators and key secretariat staff and selected board members potentially paired with or following the board meeting. Such a meeting would facilitate both increased feedback as well as greater inter-RNP learning.

Stakeholders indicate some need for additional attention to administrative routines with both board members and RNP coordinators giving GDN on average a rating about halfway between "good" and "reasonable." Other stakeholder groups give GDN a rating just above "good." GDN was recognized by stakeholders as having done a good job in adjusting World Bank systems designed for much larger organizations to provide administrative systems sufficient for its purposes. A few RNP coordinators in particular have called for some improvement in

⁸⁸ Refer to Annex AA on Dakar Questionnaire for RNPs.

forms, although the only specific complaints encountered were the bureaucratic approach to its implementation and greater harmonization with other donor forms.

Management and feedback to and from staff also is important to consider in terms of regular management. The roles of staff in both the Delhi and Cairo offices are clearly defined in job descriptions. GDN conducts performance reviews of operational staff on a semi-annual basis. The evaluation has not reviewed this process. Although the evaluation team believed it is an older style top-down review rather than a more modern, albeit time consuming, 360 degree review, GDN has responded that it a 360 degree performance appraisal was done in spring 2007. The staff evaluations precede the drafting of annual work programs for each staff member. The evaluation team did not examine closely the procedure but did not encounter specific complaints about it. The evaluation also has identified a potential concern that raises should reflect changes in responsibilities over time rather than just performance over previous years.

3.3.3 *Secretariat and Daily Management*

Since its move to Delhi in 2005, GDN has a staff manual outlining policies for proper office behaviour, hiring, travel policies, and administrative guidelines. While GDN has undertaken advisable external reviews of its financial and accounting systems, the evaluation's understanding is that it has yet to do so in terms of its human resources practices and policies. There are many areas that such an external review could and should cover including annual review process and all human resource policies.

Several areas the evaluation notes include making sure GDN has clear and implemented policies on selection and supervision of family or friends, establishing a policy on sexual harassment, formalizing how staff would handle problems with supervisors or top management, and the consistent application of travel rules. The evaluation team suggests the inclusion of three additional policies missing in the current staff rules of GDN. These include clauses related to presence of a clause on sexual harassment, a clause on panel recruitment of staff and consultants to be hired by GDN, and a requirement to excuse oneself from selection processes involving friends or relatives. The clause on sexual harassment should be designed to ensure a safe working environment for all staff members and set appropriate methods for redressal of grievances in case of an instances of sexual harassment. Appropriate training of some type in the meaning of such policies is also standard practice. GDN also should identify a board member or other person outside of management as an ombudsperson to handle serious or personal concerns with management, should they arise. In addition, the team recommends that staff should not be allowed to serve as supervisors of relatives or people with whom they previously are friends.

Further, to make cost effective use of resources, GDN should review carefully the extent to which exceptions are made by the president to the economy class travel policy and the extent to which stakeholders also are given exemptions. Although only the president officially always travels business class, this is not the case in practice. Inconsistency creates a sense of unfairness across staff and stakeholders, but widespread exceptions drive up costs for an organization that should not be spending high rates on travel instead of on programmatic needs. In its review, GDN should consider as comparable organizations for consistency other similar NGOs or smaller international organizations, which generally do not fly many people business class, and not the World Bank which is considered more profligate in the development community. The evaluation also has not received from GDN an indication that salaries are based on local market rates, and GDN should consider a market survey of similar organizations in India and internationally to make sure at least local salaries are in line with comparable organizations, as it suspects that international hire salaries are not. The evaluation believes a proper human resources review could provide many additional improvements.

The previous president established working groups consisting of different staff members for handling certain recurring activities or long-term issues, and these groups met between weekly and monthly to chart out plans

and implement them depending on needs. Although GDN is a relatively small organization, such management processes are appropriate systems for managing an organization's recurrent needs.

On the technical side, the evaluation team observed that GDN requires an improved and more regularly updated internal database at GDN and better maintenance and consistency of information across staff. The evaluation sometimes had to contact several different staff to gather similar information across time, and so it recommends GDN update the database and inform staff of updated figures more regularly. It was noted that GDN has a legal status in the US while it operates in Delhi, India. In circumstances of legal disputes, the costs and implications of settling outside of India would be relatively costly for GDN since all of its employees are employed through GDN India and not GDN Inc, USA. Now that GDN has acquired international organization status, it should conduct a legal review of its options.

Additionally, senior staff is best used for interfacing with donors. The management of GDN is conducted through regular working group meetings wherein staff are divided into groups to implement activities. It was noted frequently by staff and stakeholders that GDN staff is overburdened during the period prior to the annual conference, which can result in the neglect of other activities on occasion and the potential for crises if any staff have personal issues arise. GDN's use of interns and temporary staff is helpful and adds cost efficiency but decreases somewhat consistency across time. Additional local staff would be helpful if affordable and the conference remains every 12 months rather than another time period such as every 18 months.

4.0 CONCLUSIONS AND RECOMMENDATIONS

4.1 Objectives, Activities, and Effects

4.1.1 Objectives and Strategies

The evaluation is asked to assess a large number of criteria that are designed ultimately to reflect on whether GDN is achieving its stated objectives.⁸⁹ These include effects of GDN programming and its processes. Most information suggests that GDN is moving in the right direction but with considerable additional work to convince stakeholders more fully it is doing so efficiently, effectively, consistently, and democratically across its different regions and activities. GDN is perceived as providing mostly unique and theoretically valuable services that arguably would generate greater impacts were more funds available. It is also seen as a potential vehicle for providing additional public goods for its targeted beneficiaries.

An organization's objectives define how one determines how well it is reaching goals, and the evaluation's findings suggest less contention within GDN about objectives over time but that considerable disagreements and lack of clarity remain due to the large number of objectives it claims and the innate conflicts among some of them. These disagreements are generally considered to be somewhere between somewhat negative to highly problematic, as characterized in the previous evaluation and by those with experience with GDN in the early part of the time period considered.

Many, potentially competing, objectives allows greater leeway in programming by GDN's management and ability to create new or innovative programming. However, it also can create great differences in perceptions or even conflicts about the programming and what the proper direction is for the organization. For instance, two different key stakeholders from the same region interviewed simultaneously characterized GDN's Health GRP as alternatively outside of GDN's mission or perfectly in line with its mission. The clearest area where objectives continue to cause problems is in the question of multidisciplinary, which does not appear settled despite modest management and board efforts to reach a middle ground.⁹⁰ The evaluation does not foresee an easy solution for GDN aside from either stakeholders determining objectives implicitly through its choices or having the board revisit key objectives. Other disagreements, such as policy versus academic focus are less pronounced but might increase to the extent that GDN management and board try to pursue to further the objective of informing policy dialogue or having a policy impact. The primary recommendation relevant to GDN's objectives is:

- Clarify further whether GDN regards multi- or poli-disciplinary as an objective or an instrument. Also clarify the extent to which GDN generally plans on focusing on building capacity primarily in economics with other disciplines serving mostly as an instrument to answer questions of interest or whether GDN really is trying aiming for building capacity broadly across the social sciences.

⁸⁹ Technically, the evaluation is assigned the question "Has GDN achieved its stated objectives?" a far more ambitious and unrealistic question. Based on its research and findings, the evaluation operationalizes GDN's agreed upon objectives in practice as "building methodologically sophisticated policy-relevant social science research capacity and facilitating exchange of knowledge across the Global South."

⁹⁰ In response, GDN has noted that "the board paper on multidisciplinary makes clear that the choice of discipline(s) should be driven by the research question. This implies that multidisciplinary is an instrument rather than an objective..." and thus that the recommendation is therefore addressed. Unfortunately, the evaluation team has found that many stakeholders do not agree that the issue is so simply solved by the presence of 2004 board paper. The issue even arose during the recent 2007 Friends of GDN Retreat as a "key take-away message" noted by GDN's president.

4.1.2 *Relevance*

Given these uncertainties and disagreements regarding objectives, target beneficiaries, and dominant needs, GDN's activities generally can be considered relevant to either one stated or implicit objective or another, one set of beneficiaries or another. Thus, all of GDN's current activities as well as planned activities could be justified in terms of their relevance to at least some extent. GDN's general contention that a mix, or "portfolio," of activities best meets its objectives has merit based on the concept not only that the needs vary across and within regions but also across these different target beneficiaries and across funding constraints by donor that change over time. It also allows GDN to provide complementary services, such as much of the portfolio underneath GDNet, that can increase and prolong GDN's effect over the career of individuals, if not institutions. This approach also permits GDN to diversify itself through different donor channels so that it can meet its commitment to the World Bank in reducing dependence to the extent that board and management agree to the new directions fitting within the mission, if not maximizing cost effectiveness and effect of resources.

4.1.3 *Outcomes, Effects and Impacts*

4.1.3.1 *RRCs and GRPs*

The evaluation's review of outcomes and impacts of GDN's capacity building activities suggest increases in knowledge created and capacity built for many individuals funded across most regions. The evaluation found evidence that inconsistency in levels of support and mentoring across activities affect perceived value added, with particular concerns in the Asian regions. GDN has accurately assessed at least one of the top priorities of stakeholders that could prove a cost effective way to broaden or deepen the effects, if not impact, of its portfolio through additional individual-level training. Little evidence was found from pilot test research of capacity built from the GRP and RRC processes as measured in improvements from proposals to research paper. Limitations of proposals may explain the predominantly non-findings, especially given the generally supportive responses from grantees, key stakeholders, and resource person, including those in the one included in the pilot test. The evaluation also did not find positive effects from GDN's RRC and GRP capacity building through a non-experimental comparison using matched respondents from the broad constituency survey.⁹¹

Many recommendations were cited from stakeholders and from the evaluation team's observations that could affect outcomes and impacts from the RRC and GRP capacity building activities. Some key ideas GDN should consider include:

- Giving greater consideration from the outset of future GRPs toward strategies for disseminating research findings to the broader policy and development community.
- Consider facilitating the coordination of common themes sponsored across RNPs to increase the potential for publishability of outputs together in journals.
- RNPs may wish to sponsor one or two specific sub-themes each grant period to increase networking and interaction of grant recipients within the region.
- Develop better delivery vehicles for clarifying the policy relevance of papers from RRCs, if not GRPs. These could include asking RNPs to require policy briefs as outputs or providing on-line training modules in generating policy briefs combined with hour or two trainings/briefings on how to do so at workshops. Another vehicle would be to create policy briefs centrally (although potentially outsourcing them) as done by some other organizations to be posted or distributed by or through GDNet.

⁹¹ Results all were either negative or statistically insignificant across seven outcome measures and using six different matching methods.

- Follow through with the findings of its “Workshop on Methodology for Comparative Analysis” to be more proactive in promoting a diversity of methodologies in the choice of intended research design through its GRPs and other activities.
- If the primary goal of GRP workshops really is to focus on the capacity building of attendees rather than primarily to serve its own conference’s other goals, GDN should consider holding some of them before or after a disciplinary conference chosen appropriately for the theme.
- GDN already appears to be laying the ground work for establishing some type of individual-level training program. While methodology and technical issues appear the clear priority, GDN also should consider implementation of a small technical assistance and training program aimed at the use of electronic help desks and web-related technologies including GDNNet that could be set up as modules that accompany RRC workshops. Stakeholders also indicate interest in training for individuals, in addition to GDN’s plan to focus on institutions, on how to increase the impact from policy relevant findings. GDN should explore ways to use GDNNet as a tool for conducting relevant workshops in order to cost effectively widen its reach and broaden its portfolio’s effects.

4.1.3.2 Annual Conference

The most contentious activity among stakeholders interviewed remains the annual conference, although there appears to be no groundswell for any change in approach or frequency and sufficient support for its continuance roughly as is with a little more democracy in program decision making. The evaluation finds evidence that at least some post-conference inter-regional networking does occur, although efforts to encourage cross-regional networking during the conference could be improved. The evaluation indicates that the least cost effective part of the conference are the business meetings, which it assumes GDN would need to conduct anyway and provide some side benefits in non-measurable conference outputs. The question then would be whether its remaining \$1,300 to \$2,400 net cost per measurable networking “output,” combined with the conference’s non-measurable outputs, is justified were the Bank or other key donors to cut funding severely. Measuring conference outputs by networking rather than by papers, RRCs appear on average an order of magnitude less cost efficient, strictly speaking, and do not allow cross-regional networking and other public goods. RRCs are universally perceived as providing more direct capacity building, however. GDN could try to generate more capacity building through GDN’s annual conference, which might still provide much, but not all, of its current non-measurable benefits but would require time and resources to do so.

GDN stakeholders and experience from other conferences provide a host of recommendations; key considerations include:

- GDN may wish to experiment a few years with scheduling to see if it finds alternatives helpful. One of the primary alternatives proposed is an 18 month schedule that would consume fewer resources, allow secretariat and board to focus on other issues as well, potentially increase uniformity of paper quality within sessions as well as generally, and allow scheduling across regions to vary by appropriate season as well.
- GDN also could experiment one year pairing itself with a relevant academic conference and surveying respondents’ similar questions that year and the prior year to be able to make comparisons about the added value of the whole conference experience for intended beneficiaries with an academic conference versus a conference such as ABCDE.
- Consider consistent specific track themes across years, as found at most other regular conferences to attract more similar papers over time, increasing networking among clearer sets of audiences.
- Consider investing resources in publishing and distributing proceedings and papers in a post-conference CD (if possible, pre-conference) and on-line to extend the long-term value of the conference.
- If GDN desires to pursue both capacity building and highest-quality knowledge building objectives for the conference, GDN should experiment with purposely focusing on increasing capacity building in at

least selected sessions in a more meaningful and effective way than traditional presentations and discussants while taking measures to ensure more consistently higher quality in other sessions.

4.1.3.3 GDNNet and Upcoming GDNNet Evaluation

GDN's GDNNet represents promising tool for broadening its portfolio's effects in ways that also expand its donor pool. GDN's current plans to extend either the time period of the breadth of its mentoring through GDNNet represents a wise pilot test. Over time, GDN should explore further ways to use GDNNet as a tool for conducting relevant workshops in order to cost effectively widen its reach. The evaluation concurs with the recommendation of Dave Witzel at the September 2007 Friends of GDN meeting in Washington, DC that the organization should consider some "experiments" that fit within its mission and could increase networking with other organizations and increase name recognition – two areas rated lowest by stakeholders. Recommendations for GDNNet include:

- Consider ways to more make GDNNet a portal for Southern researchers to leading working paper and database collections worldwide.
- Change grantee contracts so that GDNNet serves as a central location for datasets funded by GDN process so that it can confirm its policy of open access data is followed and to reduce search costs for researchers.
- Continue to pursue ways to extend the reach of journal availability through GDNNet or otherwise, considering the use of key stakeholders and board members to serve as liaisons.
- Assess the extent to which more senior researchers might be aided by serving as a focal point for changing access policies of national datasets by setting up an international model program of data access from national institutions in the Global South, which currently often severely limit access.

The evaluation has made preliminary steps that should assist a more focused evaluation of GDNNet not just as a component of GDN but as an effective interface with its potential beneficiaries. Recommendations specifically for the GDNNet evaluation include:

- Assessing the extent to which the portal's interface is sufficiently user friendly not just for targeted beneficiaries' web skill levels but also their technological limitations. GDN has responded their awareness of this issue given that it is discussed in background notes for the GDN strategic retreat in Washington, DC in 2007.
- Assessing whether GDNNet has been too "top-down" or too focused on individuals in its approach to organizing interactions of beneficiaries and insufficiently attending to "bottom-up" social networking possibilities.
- Attempting to develop good baseline comparative measures of GDNNet across its current and planned uses. To the extent possible, comparisons should be made relative to websites with similar missions and, as a portal for accessing research to which it links, to direct access figures from the developing/transition world to the data available through these sites. Appropriate comparisons are important, as access figures do not uniformly rise for research websites and to benchmark where GDNNet's components' highest value is found.
- Identifying and clarifying GDNNet's various intervention type; for instance, sharing knowledge, providing tools, providing training, reputation building, linking researchers, filtering research, translating research, creating efficient means of researcher interactions, etc.
- Determining most appropriate metrics by type of intervention and interpreting them for a more general audience. Measures focusing on relevance to developing and transition countries should be kept in mind.

- Check whether the names of respondents suggest that there is a gender imbalance in registered users from baseline expectations in the South and, if so, whether there is a systemic reason that could be addressed.
- Extend the evaluation's cost effectiveness analysis for remaining GDN activities using comparable figures for other organizations with similar activities, where possible.

4.1.3.4 Institutional Support

Although GDN is hoping to embark on a pilot test of assistance to a new research institute as an extension of its BRP project, it currently does little systematically in institution building outside of having set up and assisting its own RNPs. The GDN model is set up well for capacity building of academic-style individual researchers, but not for institutional capacity building for which it would have to find efficient vehicles to deliver training, such as public research institutes in western countries, perhaps paired with one from a developing or transition country. Were GDN certain this is an area for additional growth, it might wish to consider a person with such experience as a replacement for its previous first senior political science position. Two recommendations related to institutional work follow:

- Consider facilitating research partnerships between Northern and Southern researchers and institutes through its current ongoing activities – especially GRPs. In this way, GDN can learn lessons from its own partner, the India Development Foundation, in creating collateral benefits from its resource people, especially when they share similar research interests with other developing or transition country institutes. This could help better fulfill the organization's objective to generate high quality research to accompany capacity building. GDN would have to make sure the structure did not put Southern researchers only in a primarily “research assistant automaton” role, as concerned some stakeholders.
- In the long run, consider raising funds in coordination with RNPs to work with RNPs to organize regional or national conferences among policy makers, researchers and possibly civil society organizations to discuss key policy issues of interest to increase dialogue of these institutions with researchers.

4.1.3.5 Policy Impact

The evaluation provides an argument in the section on “policy impact” regarding why it considers policy impact both an unrealistic and inappropriate potential objective that the board should clarify is not a goal. The evaluation finds evidence though that GDN can claim and should claim to be building a base of policy relevant literature and future researcher leaders who can and are informing better policy at national and sub-national levels. Nonetheless, strong policy relevance of its training and output remain a weakness of GDN's and reflects in the higher output for academic-style research than policy briefs or more locally policy relevant outputs despite interest in this area among survey respondents, as discussed in the section on overall success. It also reflects the potential conflict perceived by traditionally trained economists and some other social scientists between the objective of high policy relevance and policy impact and the objective for high quality research. The GDN-funded research that most likely will result in policy-related effects – training resulting from the BRP research and the Health project – are too nascent to count toward this goal for the current evaluation but promise to help GDN in this area for its next evaluation. GDN needs to do a better job, however, in trying to capture these examples systematically. GDN as a method for distributing policy-relevant summaries is likely to represent a key area where GDN could augment its output in terms of policy relevance.

4.1.3.6 Reach

GDN has been making concerted efforts for its more centrally run activities to have broad representation across regions, with women, younger researchers, and increasingly to target people from a variety of disciplines in

addition to economics. In this respect, GDN's record and efforts are strong. GDN's record in broadening beyond targeting only economists is more mixed with its delegated activities, the RRCs, varying greatly by region. This variation represents a key argument in addition to efficiency and less economics focus for keeping GRP management centralized, if not controlled directly by the secretariat itself. The evaluation also points out an area to consider for greater emphasis in the future is to track not only the number of countries but to develop measures of higher needs countries or even areas within countries within regions. GDN should continue its recent unofficial policy of encouraging the selection of themes in calls for papers that are more naturally interdisciplinary in subject, although this will have to be supported as well with outreach beyond just the economics field.

4.2 Governance, Organization and Processes

4.2.1 GDN's Responses to Prior Evaluation, Including Governance

Donors have indicated interest in the evaluation as a way to assess how GDN has responded to governance and process issues and recommendations from the previous evaluation. GDN has responded to several, although not all, issues and recommendations discussed in the previous evaluation, with mixed success.

1. The GDN board created audit and executive committees in lieu of having two boards – as recommended in the previous evaluation, but which was not supported by stakeholders.
2. GDN staff have made concerted efforts to raise money from corporate sources, although mostly unsuccessfully.
3. GDN has reduced somewhat their conference expenditures sufficiently to satisfy some, but not its most ardent, critics.
4. GDN did not convene a grand stakeholder and donor meeting but appears to be listening somewhat more closely to donors having formed a Donor Advisory Council that meets annually.
5. Like most organizations, GDN has been unable to find a large donor to endow the organization to stabilize funding, underwriting indirect costs.
6. Current levels of funding are no longer uniformly “woefully insufficient” for GRPs although remain considerably less than desirable especially for RRCs, and GDN has shown that its activities are scalable during fiscal crises.

These activities indicate responsiveness, but further work remains and some innate difficulties and contradictions may not be easily overcome. For example, GDN's board has slightly improved its governance process with the addition of the audit and executive committees, but the evaluation is informed that the executive committee functions more as a stop gap between meetings than as a committee that handles important business matters regularly. Further, there are insufficient structures to guarantee high levels of institutional memory or board membership requirements to guarantee improved board management capacity relative to the last evaluation. Combined with infrequent meetings of very busy members, this relatively weak board governance structure allows management and staff considerable latitude. This has the advantage of allowing management entrepreneurial opportunities but less ability of the board to guarantee that the organization is meeting stakeholder needs in line with current best practice.

Changes regarding governance GDN should consider include:

- Recruit at least one member with a strong financial background to lead more proactive financial oversight by the board; For example, the board of the International Food Policy Research Institute's terms of reference require that four members possess such a background.
- Require a certain number of members overall, through specific positions, to require experience in managing research institutions or NGOs.

- Clarify for prospective incoming members appropriate time expectations for board membership such as through a frequently asked questions document and a nonbinding agreement about board member responsibilities.
- Providing new board members a strong orientation to the organization within their first month on the board.
- Work with nominating organizations to make recruitment of board members more transparent.
- Consider one additional donor representative on the board, which would help further groundtruth GDN's ideas from the perspective of funding.
- Set the conditions for greater board activity in general and on strategic issues in particular through extended in number of days of in-person board meetings, add at least one virtual meeting during the year even if not all members can attend, and form additional permanent committees to handle issues outside of the main meeting.
- Make sure the board has a proper conflict of interest policy, in line with best practice, if not already adopted.
- Although unlikely, GDN should explore whether organizations nominating members could in some way pay for some time of the researcher members to allow members to be able to provide additional effort towards their GDN responsibilities rather than seeming like an additional burden on top of already too busy schedules.

4.2.2 Relationships with RNPs

GDN's relationships with its RNPs is somewhat unclear in that it sometimes claims to be a network of research networks, but the relationship is often one-sided with limited two-way feedback and differing agendas. The latter tension between establishing relationships with economics shops and promoting a poli-disciplinary outreach approach creates special problems, and declining unearmarked funding leads to declining reasons for partners to feel and act as partners. The evaluation recommends a few clarifications in GDN's relationship with its RNPs:

- Consider clarifying extent to which GDN is a network of networks versus a centralized organization with networks as grantees, rather than partners. In doing so, clarify what the financial and programmatic implications should be of the two choices in terms of financial support besides funding RRCs.
- Provide greater institutional support for RNPs that rotate administrative centers to increase continuity during transitions. Look for partnerships and other ways to generate greater impact in the Asia region in building research capacity in order to increase the probability that donors interested in Asia will continue funding.
- Improve learning across regional networks through an additional, dedicated meeting outside of the annual conference, where there are fewer distractions and coordinators can discuss and strategize better about ways to improve service and potentially work together across regions.
- Although GDN has proven scalability of its RRC activities under financial stress, it must be careful of cutting overhead costs from RRCs too much for partners to weaken the partnership and leverage GDN has with them.
- Look for ways to augment input in decision making process from RNPs on GRP themes or candidate selection and recommendations even if keeping management centralized for smaller GRP themes.
- Consider additional monitoring on how grant funds are used by requiring more detailed reporting to monitor the grant fund distribution to categories of researchers and overhead costs. Consider adopting the common reporting framework used by many other donors.
- Widen its new pilot idea of working with RNPs in designing a coordinated fundraising plan toward a dedicated work plan with joint fundraising in order to reduce perceptions of potential competition and increase funds it can leverage at the regional level.

- Look for ways to build and improve relationships with RNPs through establishing more regular, horizontal interactions beyond the annual conference meeting. Ideas suggested include considering experimenting with ideas such as at least biannual meetings just of RNP coordinators and key secretariat staff and selected board members potentially paired with or following the board meeting. Such a meeting would facilitate both increased feedback as well as greater inter-RNP learning.

4.2.3 Organizational and Management Processes

Despite these suggestions for management changes, stakeholders do not indicate concern about GDN's previous management, which they strongly praised for its ability to bring the organization through its most challenging period of difficult decision making and GDN's sometimes competing objectives. GDN's management has contended with a situation that metaphorically might be compared to trying to solve a large series of equations with disagreement over constants, number of variables, and perhaps arithmetic expressions, resulting in widely varying solutions by person. Despite this untenable challenge, stakeholders on average provide GDN "good", if not "excellent" ratings along most dimensions of its work – itself a testament to strong leadership given the challenging issues GDN contended with over recent years.

Nonetheless, the evaluation has uncovered a few areas where improvements in organizational and management processes that GDN should be considered. These include:

- Increasing transparency and widening the democratic processes in decision making processes where possible especially regarding GRP themes. Among several alternative models that could increase stakeholder ownership includes having GDN serve as an intermediary between stakeholders and donors, fielding proposals for global or multi-regional GRP themes among stakeholders and shopping them with donors.
- Given GDN's more hierarchical, bureaucratic heritage from the World Bank, being attentive to remaining as non-hierarchical as possible given its small size and need to have partners feel ownership and involvement to remain engaged as partners.
- Engage individuals from all levels of the organization to participate more fully in planning processes. For instance, a planning and teambuilding retreat for all secretariat members would make GDN more akin to following best practice of smaller development NGOs than a World Bank-type highly hierarchical organization. The trade-off for the lower efficiency and slower processes should be greater buy-in, ownership, and lower staff turnover. On the other hand, this
- GDN should hire a consultant to work with the organization to develop a coherent branding strategy and advise it regarding a communications plan. Create a standardized format and set of requirements for branding and acknowledgement of GDN.
- Consider hiring a staff member with deep research experience if GDN wishes to focus more on providing centralized intellectual direction to new activities and additional intellectual oversight for conference activities and a focus on the objective of facilitating or producing highest quality research. The evaluation does not necessarily advocate a mini-think tank model, although it does suggest it as a suggestion worth exploring under the new president's strategic review.
- Alternatively, or perhaps additionally, consider hiring a fundraiser as a fulltime staff member or consultant based in the US or Europe dedicated to the process of identifying and establishing contact and maintaining communication with donors. Such a position, however, might be weaker handling programmatic interactions, would have to have relationships with bilateral and multilateral donors which most do not, and tends to be expensive, paying for itself over time. There is no simple answer regarding the cost-benefit ratio for such a position, but many organizations make such appointments provisional for 18 to 24 months conditional on success.
- Consider ways to create additional strategic partnerships and linkages with additional consortia and participation in other conferences.

4.2.4 *Secretariat and Daily Management*

In terms of daily management of the secretariat and general administration, GDN has been steadily establishing appropriate administrative processes. GDN itself has been undertaking external financial and process audits superior to that possible through the evaluation process, and GDN has been responding to recommendations. Areas for attention in alignment with organizational best practice, especially given the small secretariat size, include:

- Assuring appropriate policies are in place on issues such as a sexual harassment policy.
- Identifying a board member or outside party as an ombudsperson for staff concerns with direct supervisors or management that should not be addressed directly.
- Undertaking a more complete external review of human resource policies similar to its financial process review from the World Bank.
- Establish clear and implemented policies on selection and supervision of family or friends
- Making sure policies are implemented as consistently as possible.
- Undertake a market survey of similar organizations in India and internationally to check on comparability of local and international salary rates.
- Review application of travel rules for consistency and cost effectiveness in line with other similar NGOs or smaller international organizations and not the World Bank.
- Conduct a legal review of its situation and options given its newly acquired international organization status in reference to liabilities, taxes, etc.

4.2.5 *Recommendations for Future Evaluations and Other Technical Issues*

Finally, this evaluation provides several recommendations for the next evaluation and some minor technical recommendations in annex AB.

REFERENCES

- Baumgartner, F.R. and B.D. Jones, (1993). *Agendas and Instability in American Politics*. Chicago: University of Chicago Press.
- Carden, F., Deby, D., Earl, S., Gillespie, B., Neilson, S. & Smutylo, T. (2002). *IDRC Supported Research in the Public Policy Process: a Strategic Evaluation of the Influence of Research on Public Policy*. IDRC Evaluation Unit.
- Downs, A. (1972). "Up and Down with Ecology: The 'Issue Attention Cycle,'" *The Public Interest*: 28.
- Heckman, J.J., Ichimura, H. and Todd, P.E. (1998), "Matching as an Econometric Evaluation Estimator", *Review of Economic Studies*, 65, 261-294.
- Kingdon, J. (1984). *Agendas, Alternatives and Public Policies*. New York: Little Brown & Company
- Livny, E., Mehendale, A., and Vanags, A. (2006). "Bridging the Research Policy Gaps in Developing and Transition Countries: Analytical Lessons and Proposals for Action: A synthesis of findings from the Global Development Network's Bridging Research and Policy Project." GDN.
- Rosenbaum, P.R. and Rubin, D.B. (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika*, 70, 1, 41-55.