

Do Faith-Based NGOs Represent a Replicable Example For the Delivery of Public Services? An Application to Health Care Delivery in Developing Countries

Kenneth L. Leonard, University of Maryland

Abstract: *In this paper we examine the evidence that faith-based NGOs can provide a working example of a service delivery organization in a developing country context. Though it is well known that such organizations can provide high quality care, and in particular can serve as highly cooperative collaborators for research, it has not been generally shown that the lessons learned from such organizations are replicable. We show that for health care, and in Tanzania, the faith-based organizations succeed because they motivate the doctors under their employ to work hard. Though there is evidence that some doctors are extraordinarily and intrinsically motivated, these doctors are as likely to work in the public service as they are in faith-based NGOs.*

Nongovernmental organizations (NGOs) play an important role in many developing countries and provide a large proportion of public services such as health care and education (Berman et al., 1995; Berman, Nwuke, Rannan-Eliya, & Mwanza, 1995; Gilson et al., 1997; McPake, 1997).¹ There is considerable evidence that many NGOs provide significantly better services than do their public counterparts and that the average NGO facility (school or clinic) is superior to the average public facility (Gilson et al., 1997; Leonard & Masatu, 2007, 2008b; Leonard, Masatu, & Vialou, 2007; McPake, 1997; Mliga, 2000; Reinikka & Svenson, 2004). Their important contribution to human capital accumulation demands attention by itself, but much recent research on NGOs has focused on their flexibility and therefore their use as laboratories for experimentation (Dean, Schaffner, & Smith, 2005; Miguel & Kremer, 2004). In what is rapidly becoming a gold standard for research, some NGOs allow the researcher to control the implementation and timing of proposed reforms to create a randomized controlled experiment. Such research allows for rigorous investigation of potential reforms and improvements.

Author's Note: *The data collection on which this paper is based was funded by NSF Grant 00-95235 and The World Bank, and was completed with the collaboration of M. Masatu (M.D.) and the assistance of R. Darabe, M. Kyande, S. Masanja, H. M. Mvungi, and J. Msolla. We extend our appreciation to the Commission for Science and Technology (COSTECH) for granting permission to perform this research. Author's address: 2200 Symons Hall, University of Maryland, College Park, MD 20742. email: kleonard@arec.umd.edu.*

For NGOs to represent useful laboratories for experimentation, it must be true either that NGOs are not fundamentally different from other service providers (that lessons are replicable) or that NGOs can successfully deliver more services (that NGOs are scalable). However, many if not most NGOs in developing countries focused on health care and education are faith-based organizations. The fact that these organizations and the motives of their leadership are explicitly religious may invalidate either or both of the above assumptions. The results of controlled experiments may not be replicable if proper implementation relies on the assistance of NGO staff who are fundamentally different from staff in all other organizations and NGOs themselves may not be scalable if they rely on small teams of close knit individuals. Thus, it is important to ask whether the superior services provided by NGOs and the gains seen in policy experiments with NGOs result from flexible organizational structures that can be imitated, or from organizations composed of self-motivated or altruistic “volunteers”—something difficult to imitate.

In this paper, we examine the evidence for the replicability of NGOs using data on doctors practicing medicine in the Arusha region of Tanzania. These doctors work for various organizations, including public and private facilities and five faith-based NGOs (the Lutheran, Roman Catholic, Seventh Day Adventist, and Church of Gospel International churches, and Ithna Asheri Mosque.²) The study measured the ability (the capacity for quality) and the practice quality (the actual quality delivered) of doctors—in particular, the gap between ability and quality. We show that doctors who work in NGOs are not different from doctors who work in the public service; the abilities of doctors in NGO facilities are the same as the abilities of doctors who work in the public sector. However, the practice quality of doctors in NGOs is significantly higher than that of doctors in the public service. In addition, there is evidence that some doctors are different from the average doctors in that they are motivated by altruistic or professional values. However, in Tanzania at least, these doctors do not work exclusively in faith-based NGOs; they are almost as likely to be found in the public service. Notwithstanding the presence of these altruistic doctors, the average doctor in either NGOs or the public service does not behave in an altruistic fashion. For these doctors, NGOs succeed because the organization is designed to encourage and reward effort, whereas the public service does not. Overall, these results suggest that the success of NGOs can be replicated because it does not depend on staff who are fundamentally different from staff in the public service.

In addition, we review the potential for methodological innovation in the study of the motivation of agents who work for NGOs, the private,

or public sector. In particular, we found that our study of the behavior of doctors had a significant impact on their behavior; when doctors know they are being observed they change their behavior. This impact is particularly striking in our study of doctors because the researchers were also doctors. In a setting such as medicine, where professional and ethical standards are part of all training, this peer-scrutiny has a particularly strong effect on behavior. Because this effect is temporary, we were able to observe the behavior of doctors as they reacted to this peer-scrutiny as well as when they were no longer reacting to the presence of the research team. The differences between the behaviors of doctors under normal circumstances and their behaviors when they face peer-scrutiny allow us to speculate on the ulterior motives of doctors. Professionalization in medicine is designed to encourage doctors to treat patients as they themselves would like to be treated: to hold the interests of the patients paramount. This is exactly the same behavior that an altruistic or intrinsically motivated doctor would exhibit. Thus, when our research team arrives, doctors behave as they would if they cared only about the interests of patients. This quality, therefore, reflects the natural capacity or competence of a doctor as well as the fact that they are exerting significant effort. The gap between this behavior and normal behavior is therefore a measure of just how motivated the doctor really is. Thus, we will argue that it may be impossible to research NGOs without impacting their behavior, but that if the researcher can measure behavior (or outcomes) with and without peer-scrutiny, the differences between these two behaviors (or outcomes) can reveal important details of an organization's effectiveness and motivation.

1 Empirical Setting and Methodology

The data used in this paper were collected over a period of two years from October of 2001 through March of 2003. Thirty-nine health facilities in the rural and urban areas of Arusha region were visited twice over this two-year period.³

1.1 Measures of Quality

The research team used direct clinician observation (DCO) to measure the actual performance of doctors with their regular patients. DCO measures compliance with Tanzanian protocol and is designed to be sensitive to the limited resources available in the facilities we survey. Every doctor visited was trained in protocol and had the resources at his or her disposal to follow it. Protocol requires history-taking (such as asking the patient the duration of the illness or whether diarrhea is accompanied by vomiting)

and physical examination (such as taking the patient's temperature or auscultating the chest). With the DCO instrument, a doctor on the research team sits in on the examined doctor's consultations. For each consultation, the observer fills a protocol checklist designed to match patients presenting with fever, cough, or diarrhea. For other conditions, there is a more general history-taking protocol and one physical examination protocol item.⁴

In addition, each of these doctors was also evaluated using vignettes, which are case-study patients presented by an actor. Vignettes have gained increasing popularity as a tool for quality evaluation both in developing and developed countries (Das & Hammer, 2005, 2007; de Geyndt, 1995; Epstein et al., 2001; Kalf & Spruijt-Metz, 1996; Koedoot et al., 2002; McLeod et al., 1997; Murata, McGlynn, Siu, & Brooks, 1992; Murata et al., 1994; O'Flaherty, Lerum, Martin, & Grassi, 2002; Peabody, Rahman, Fox, & Gertler, 1994; Peabody, Gertler, & Leibowitz, 1998; Peabody, Luck, Glassman, Dresselhaus, & Lee, 2000; Tiemeier et al., 2002). There are many possible ways of implementing a vignette; we use the unblind case study with an actor. There are two researchers present: a "patient" and an examiner. The examiner, after introductions, never speaks; he only observes. The "patient" presents herself as a patient would, entering the room from outside and leaving after the consultation. She describes her symptoms and answers questions as a patient would. It is explained to the clinician that he must do physical examination by posing questions. The patient then answers the question verbally. For instance, if the clinician says "I would take the patient's temperature," the "patient" would say "the temperature is 38.5." The examiner then fills a checklist of the expected inputs, including expected history-taking questions, physical examination items, and health education points. Each clinician was tested in their ability for six typical cases: malaria, pelvic inflammatory disease (PID), diarrhea, pneumonia, flu, and worm infestation. For the purposes of evaluating the differences between ability and practice quality, our work focused on the malaria, diarrhea, and pneumonia vignettes because they correspond well to categories in the DCO instrument. 80 doctors were observed directly and evaluated with vignettes representing a total of 928 consultations.

Additional data were collected in urban Arusha in 2005, using the retrospective consultation review (RCR) instrument (see Leonard & Masatu, 2006). This instrument uses the same checklist as the DCO instrument and is filled by interviewing patients who have just left the consultation. The RCR questionnaires were administered to 320 patients at 11 facilities in urban Arusha in Tanzania. 211 of these patients visited one of the 12 clinicians directly observed by the team and the remainder

visited clinicians at the same facilities who were never observed. On average, we have data on 6 consultations before the team arrived and 11 after we arrived. For consultations that were also observed by the research team, Leonard and Masatu (2006) show that the results from these two instruments are well correlated.

1.2 Doctors and Organizations

The doctors in our sample include nurses of various specializations, clinical assistants, clinical officers, assistant medical officers (AMOs), and medical officers (MOs). Clinical assistants have an elementary school education and three years of medical training. Clinical officers traditionally have O level education and two years of medical training. AMOs are clinical officers with two additional years of training. MOs have both an A level education and five years of university-level medical training. Nurses are not supposed to diagnose, but in the rural areas they are frequently the only health personnel present and they do diagnose patients in these circumstances. Most doctors in the sample, as in Tanzania, work in the public service in government-run health facilities. In addition, there are seven other organizations delivering care in the area: one parastatal hospital (owned by the government but operated as an independent entity), one private facility, and five faith-based NGO organizations operated by the Lutheran, Roman Catholic, Seventh Day Adventist, and Church of Gospel International (COGI) churches⁵, and the Ithna Asheri Mosque.

1.3 Measures of Decentralization of Authority

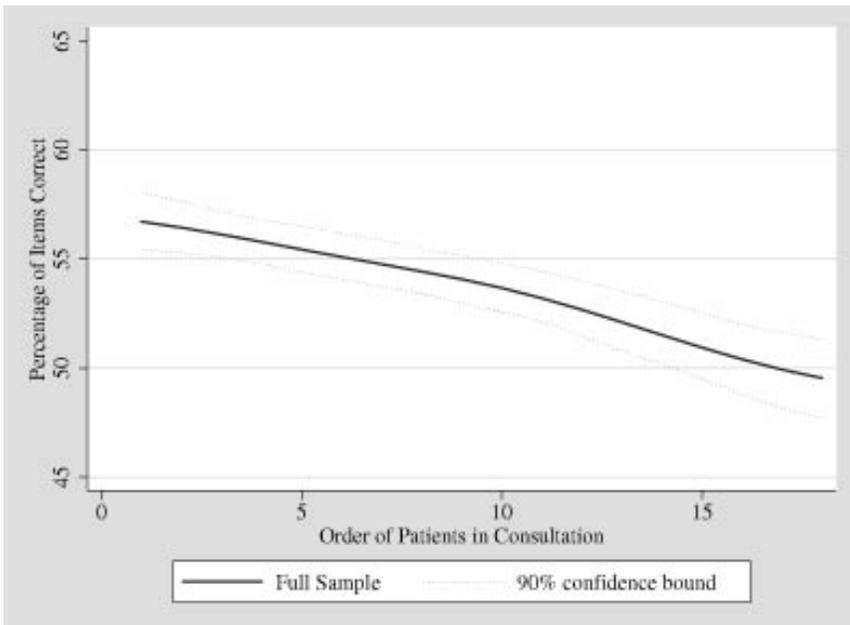
Leonard et al. (2007) introduce the degree of decentralization in decision making authority as a proxy for motivation. The index of decentralization is derived from a series of qualitative variables collected by Mliga (2000) during a study of the management structure of health organizations in Tanzania, including the facilities analyzed in this study. The variables used to create an index of decentralization include: a dummy variable indicating whether the chief of post can hire and fire personnel; the level at which salaries are set (national / regional / local); the degree to which the chief of post can (or must) use local funds to pay salaries and buy medicines; and the level at which choices about staffing are made (national/regional/local). The index is derived from factor analysis of these four variables (Leonard et al., 2007). The index varies by organization, as well as across facilities within organizations, but does not vary within facility. Examining facilities by degree of decentralization puts NGO facilities on a scale between private facilities and public facilities, solely

in accordance with the organizational structure of the facility. To the extent that this index helps to explain the quality of care provided across public, NGO, and private facilities, it suggests that management, not extraordinary staff, explains the differential quality of faith-based NGOs.

1.4 The Hawthorne Effect

One of the more striking features of the data is the steady decline in the quality of care provided by most of the doctors over the period that the research team observed their consultations. Figure 1 shows the average impact of this dropoff. Quality, as measured by percentage of items required by protocol that are actually implemented, falls by about 5 percentage points (10%) over 10 to 15 consultations.

Figure 1: The Hawthorne Effect on Quality



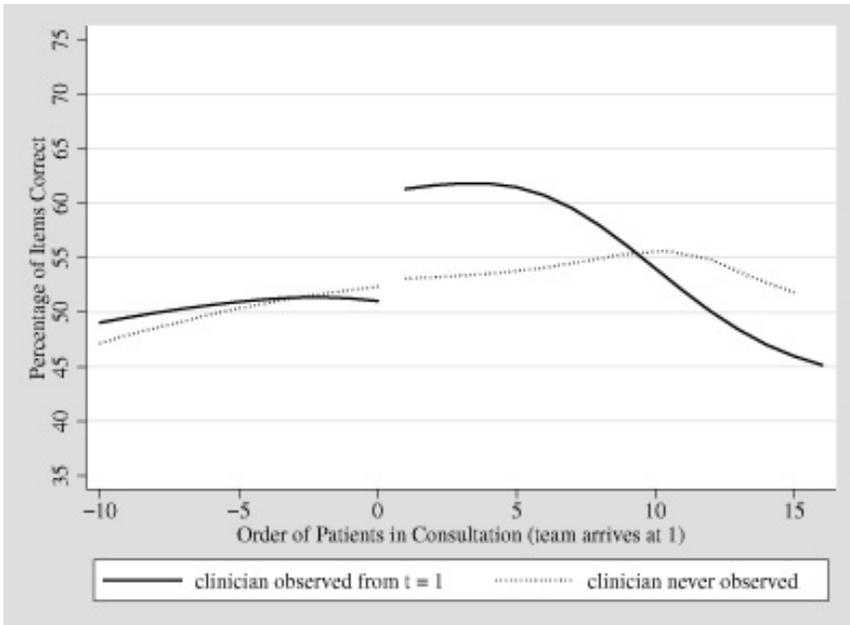
The figure shows smoothed average percentage of items required by protocol as measured by the research team doctor present in the consultation. This quality measure is graphed against the order of consultation, from the moment the research team arrives. The solid line shows smoother average percentage of required items that were provided and the dashed lines show the upper and lower bounds of the 90% confidence intervals (derived from a bootstrap with 500 samples.)

This dropoff is due to a temporary Hawthorne effect⁶: quality is falling from an abnormally high level caused by the arrival of the research team towards the normal level of quality. Leonard and Masatu (2006) document the full pattern of the Hawthorne effect with a small sample of doctors practicing in the Arusha region in Tanzania. They measured the quality of a consultation using a patient exit interview and showed that this instrument is a good approximation of the data on quality collected by doctors on the research team. Because they used a patient exit survey, they could collect data for three types of patients: patients who had consultations before the team arrived at a facility, patients consulted after the team arrived whose consultations were observed by the research team, and patients consulted after the research team arrived whose consultations were not observed by the research team. Patients in this third group were seen by doctors who were not evaluated by the research team, but who practice at facilities where other doctors were evaluated. Leonard and Masatu (2006) validate the Hawthorne effect by showing that quality increases significantly when the team arrives. Figure 2 shows the pattern of quality as estimated from patient responses for observed and unobserved doctors. For doctors who were observed, figure 2 shows a significant jump in quality when the team arrived. However, for doctors who were never observed, there is no significant change in quality. Figure 2 also shows that the Hawthorne effect is temporary; quality rapidly returns to levels similar to those found in the absence of the research team.⁷

The temporary duration of the Hawthorne effect allows us to observe differences in the behavior of most doctors. In the smaller sample for which data was collected before the arrival of the team, the size of the Hawthorne effect can be seen in the gap between behavior before arrival and behavior right after the team arrives. For the larger sample, for which there is no data before the team arrives, the size of the Hawthorne effect can be seen in the fall in quality as the Hawthorne effect wears off. Leonard and Masatu (2008) show that this gap is similar to the gap between ability as measured by vignettes and practice quality. Thus, the Hawthorne effect encourages doctors to display their best effort; doctors who routinely practice at levels below their best display a much larger gap, and doctors who routinely practice at levels that are close to their best display a smaller or non-existent gap.

1.5 Heterogeneity and Altruism

In our studies of doctor behavior, we found that behavior for diagnostic inputs (such as history-taking and physical examination) is systematically different than behavior for health education and communication. The

Figure 2: The Full Pattern of the Hawthorne Effect

The figure shows smoothed average percentage of items required by protocol as measured from patient exit interviews performed immediately after the consultation. The dashed line shows percentage provided for patients seen immediately before and after the research team arrives at a facility who visited a doctor who was never directly evaluated by the research team. The solid line shows the percentage provided for doctors who were observed by the research team starting at $t=1$.

latter involves explaining treatments and diagnoses to patients, giving them information about the typical causes of the illness, and behavior modifications that can reduce the chance of future illness. Health education is an important part of all consultations and all doctors are trained in its use. However, unlike diagnostic effort, it appears that organizations do not exert effort encouraging health education, even when they do exert effort to encourage diagnostic quality. Despite the lack of organizational encouragement, some doctors do provide high levels of health education, suggesting that they are different than other doctors.

Figure 3 shows three clinicians who represent three types of behavior that we observe in the data. Clinician 33 (leftmost) exhibits high and constant physical examination, but rapidly falling health education. He provides health education only for the first few observations. Clinician 32 provides

low and falling levels of both physical examination and health education. Clinician 81, on the other hand, provides high and constant levels of health education as well as physical examination. There are no doctors who provide high health education but low physical examination, suggesting that doctors who care to provide health education will automatically be driven to provide physical examination, though the inverse is not true. Note that, if we examine only diagnostic quality (physical examination), there are only two types of doctors (clinicians number 81 and number 33 are very similar). Thus, looking at health education reveals a second dimension of provider behavior.

If there are altruistic doctors in the study—doctors who always act in the interest of their patients—they will provide health education to their patients. Thus, doctors who provide high and constant levels of health education (such as clinician 81), could be altruistic, but doctors who provide low and/or falling levels of health education, are clearly not altruistic (such as clinicians 33 and 32). In this way, we use patterns of health education to categorize doctors as “potentially altruistic” and “not altruistic.” Importantly, describing doctors as potentially altruistic does not mean that they are altruistic; there are many reasons why doctors provide high levels of health education that may have nothing to do with altruism.

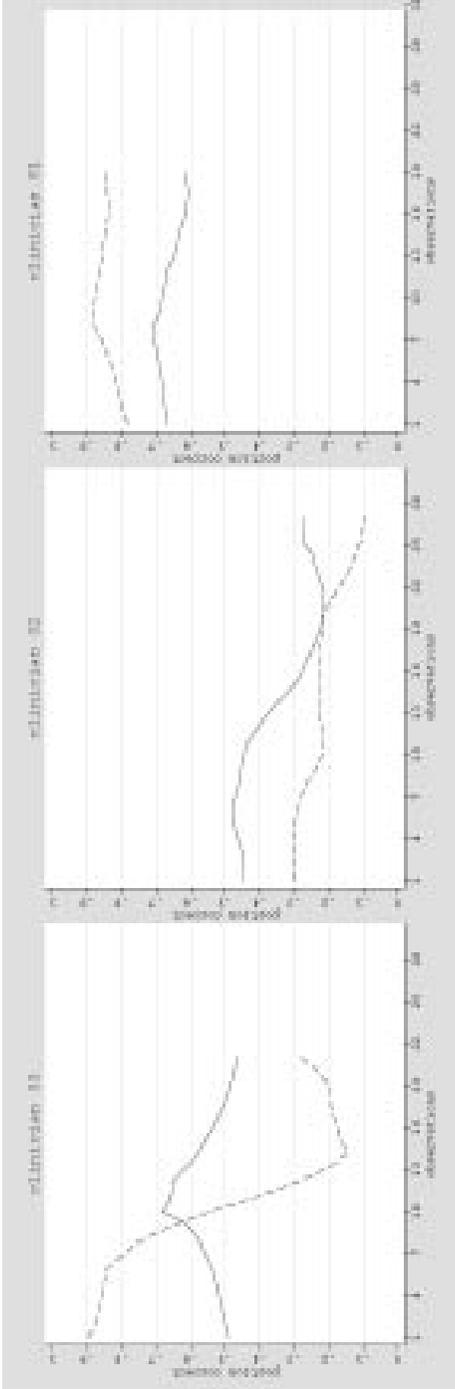
Leonard and Masatu (2007) categorize doctors following the intuition advanced in figure 3, using objective methods to make the assignment. The provision of health education is predicted in a probit model with random effects at the doctor level, a doctor specific intercept and slope (with order of consultation), and patient characteristics. All doctors with a statistically negative slope (a fall in quality over observations) were assigned to “not altruistic.” Doctors who exhibited both a positive (or flat) slope and an intercept that was above the median were assigned to potentially altruistic. Approximately 25% of the sample is defined as potentially altruistic in this manner. Assignments are made without reference to ability, cadre, organization, or provision of either physical examination or history taking.

2 Empirical Findings

2.1 Distribution of Ability

Despite the fact that all doctors practicing medicine in Tanzania are trained to be able to diagnose and treat the common medical conditions reported at health facilities, competence in the treatment of these conditions

Figure 3: Three Examples of Clinician Behavior on Physical Examination and Health Education



The dashed line is health education and the solid line is physical examination. The smoothed lines shown represent the change in the number of physical examination and health education items answered correctly as a percentage of possible items. The lines are derived from scores controlling for patient characteristics, using a local average regression with an Epanechnikov kernel and a bin width of 6 observations.

is disturbingly low. NGOs do not escape criticism on this front. In fact, since NGOs are more likely to hire doctors with less training, and training is positively associated with competence, many NGO facilities display low levels of competence. Leonard and Masatu (2007) show that the biggest gap in competence in Tanzania is between rural and urban facilities, with urban facilities generally displaying much higher levels of competence. In addition, the gap between rural and urban competence is almost exactly the same for NGO and public facilities. The fact that competence is lower in rural facilities most likely reflects the fact that it is difficult to get doctors with more training to locate in rural areas. The evidence suggests that both NGOs and the public sector face the same difficulties. In Tanzania, doctors in the public and NGO service are trained in the same medical schools, and our data suggests that, at least as doctors, they are similar.

2.2 Distribution of Practice Quality

On the other hand, the distribution of practice quality is different from the distribution of competence (or ability). Leonard and Masatu (2005) show that the gap between ability and practice quality is relatively large for doctors as a whole: the average doctor does significantly less for his regular patients than he does for the case study patient, even when these two patients are very similar. However, the size of this ability-practice quality gap varies across locations and organizations. The gap between rural and urban quality remains for practice quality, but it is significantly smaller for NGOs than it is for the public sector (Leonard and Masatu, 2007). In general, doctors who work in NGOs provide higher quality care than doctors in public sectors, even after controlling for ability (Leonard et al., 2007).

2.3 Making the Case for Practice Quality

The current enthusiasm for the use of vignettes demonstrates the increasing popularity of measures of competence in health care, particularly in developing countries. Das, Hammer, and Leonard (2008) discuss the recent scholarship on quality in developing countries and suggest that the current state of knowledge is so low that important gains can be made simply by understanding the distribution of competence. This is particularly important to the study of the emerging private market, because these doctors choose where to locate their practices and there are important associations between competence and location that we do not yet understand. However, in Tanzania, both the public sector and the NGO sector are engaged in roughly the same activities: providing reasonable quality care to poor and rural populations. Moreover, they appear to make

the same choices with respect to the competence of their providers. This suggests that practice quality, rather than competence, is more interesting to the investigation of health care quality in NGOs in Tanzania.

Additional evidence of the importance of practice quality comes from examining the relative importance of policies to increase training (raising competence and therefore practice quality) and policies to increase motivation (raising practice quality without raising competence). Five years of additional training increases the competence of the average doctor by one standard deviation, and increases the probability that he or she would correctly diagnose the illness of a case study patient by 5 percentage points. However, practice quality (what they would actually do) increases by much less and the probability of correctly diagnosing an actual illness increases by only 1 percentage point (Das, Hammer, & Leonard, 2008). Policies that directly address practice quality (rather than competence) are more likely to have a real impact on health outcomes.

2.4 The Role of Decentralization in Practice Quality

Leonard et al. (2007) examine the difference between the ability of doctors, as measured by vignettes, and the practice quality as measured by direct observation. They find that the location of decision making authority has important implications for the difference between these two measures of quality. In particular, doctors who work in facilities with decentralized decision making authority practice at higher levels of quality than doctors who work in facilities with centralized decision making authority. Leonard and Masatu (2008b) examine the same patterns, but using the change in quality from the Hawthorne effect to measure the difference between ability and practice quality. They find that, as with the difference measured by vignettes, doctors who work in facilities with decentralized authority exhibit a much smaller gap in quality over the course of observations.

2.5 Heterogeneous Doctor Types

Leonard and Masatu (2008a) examine the gap between ability and practice quality, controlling for two types of doctors: those who are potentially altruistic and those who are not altruistic. Not surprisingly, potentially altruistic doctors are superior to other doctors both in ability and in practice quality. Importantly, for these potentially altruistic doctors, there are no differences in either ability or practice quality across facilities as measured by the index of decentralization. On the other hand, the gap between ability and practice quality varies significantly across organizations for doctors who are not altruistic. It is not the case that potentially altruistic doctors work in different types of organizations from other doctors.

21% of doctors in the public sector are potentially altruistic compared to 30% of doctors in NGOs.

3 Discussion and Conclusion

3.1 Policy Implications

In this setting, the success of NGOs appears to be replicable, at least in the sense that NGOs manage to encourage normal doctors to provide high quality care. Thus, the evidence suggests that the same doctors who are currently working in the public sector would provide higher quality care if they worked in NGOs. Whether the public sector could achieve the same results as NGO organizations is still open to debate. The structure of public service makes systematic reforms unlikely, but there are intermediate solutions to improve quality. First, the public service can contract out many services to NGOs (Gilson et al., 1997; Leonard, 2002) effectively using NGO management capacity to improve the quality of public sector doctors. Gilson et al. (1997) suggest that in those settings where such policies have been tried, the quality of public sector doctors seconded to NGO managed facilities has improved. Second, the public sector can provide funds directly to NGOs and expand the current coverage of NGOs. Since this would require hiring additional doctors, it is important that we were able to establish that NGOs know how to induce high quality care for normal doctors, and are not reliant on a small group of exceptionality-motivated doctors. Third, NGOs can expand their coverage by franchising their name and therefore management structure to private doctors. In Tanzania, where such franchising has been legally encouraged, the only organizations that have embraced the practice are those that do not have any services of their own to offer. Thus, there appears to be some reluctance on the part of NGOs with a traditional public good provision agenda to expand in this direction, though improved legal structures might address some of these concerns.

The degree to which decision making authority is decentralized is an important aspect of management in NGOs. However, this is not the only difference between NGOs and the public service, and considerably more research is necessary before suggesting that decentralization would improve the quality of care in the public sector in Tanzania. Indeed decentralization has often had disastrous results (Bossert, 1998).

3.2 Methodological Notes on the Study of Faith-Based NGOs

One of the more pleasant aspects of studying NGOs is that the researcher often shares some fundamental values and/or interests with the officers

and staff of the NGO. This is particularly the case in faith-based NGOs as studied by researchers with similar motivation. However, this common interest is likely to affect the research, and if the researcher is not careful, it could lead to falsely optimistic results. Naturally, all researchers must be careful of making subjective evaluations of organizations when there is such a strong disposition to find positive impacts. However, the danger of over-optimism extends even to quantitative and objective analysis because the mere presence of the researcher can change the behavior of the subject.

The Hawthorne effect is stronger if a doctor is being observed by another doctor than if the doctors are being observed by “lay” researchers (van den Hombergh, Grol, van den Hoogen, & van den Bosch, 1999). Thus, shared motivation, training, or ethics can create a demand for quality during the study. The temporary duration of the Hawthorne effect in our studies is probably a result of the fact that researchers were trained not to give feedback, which may have slowly changed the way the subjects understood and responded to the research. Thus, we believe that the presence of an outside researcher who is explicitly interested in the role of faith-based organizations will induce a quality response, even if the researcher is careful. The only way to reduce the possibility of the Hawthorne effect is to be completely neutral and passive. However, most researchers (this author included) find that the interaction with the subjects is at least as informative as the objective research design and should be reluctant to forgo this experience. The objectives of neutrality and deeper understanding are at cross purposes.

In our work, we did not eliminate the Hawthorne effect, but instead found two ways to integrate it into our study, and in the end revealed more about behavior than if we had eliminated it. We integrated the Hawthorne effect by examining doctors for long enough that they grew used to increased scrutiny, and by collecting data on the behavior of doctors before we arrived at the facility. Thus, we were able to observe both “normal” and peer-scrutiny induced quality. Peer-scrutiny induced behavior is a good measure of the capacity of agents to perform tasks because it reveals the difference in quality across subjects who are providing maximum effort. The gap between normal and peer-scrutiny behavior, in turn, reveals the importance of organizational incentives as well as intrinsic motivation or altruism. For those activities monitored and regulated by the institution, a smaller gap between normal and scrutiny-induced behavior demonstrates superior organizations. For those activities that are not monitored or directly rewarded (but that remain important), a smaller gap demonstrates the presence of intrinsic motivation.

Endnotes

- 1 The average NGO probably contribute little to the direct provision of public services. Barr, Fafchamps, and Owens (2005) show that, in Uganda, most NGOs are small, underfunded, poorly organized, and focused on topics such as advocacy. However, some NGOs are considerably better funded and ambitious. In this paper, we focus on NGO presence in important services and thus focus on this latter type of NGO.
- 2 Ithna Asheri is a Shia branch of Islam, and the largest school of Shia thought.
- 3 Two facilities visited in the first round were closed by the time we visited in the second round.
- 4 For further details of the study, see Leonard and Masatu (2005) and Leonard et al. (2007).
- 5 The COGI facilities are actually private facilities that have franchised the church's name, allowing them to practice care under a different tax status. Kanji, Kilima, and Munishi (1992) document this common pattern and suggest that facilities that are supervised by an NGO that does not operate any independent health facilities are subject to no medical supervision or scrutiny from the franchise organization and therefore can be considered private facilities. Therefore, the COGI facilities are better characterized as private facilities, not NGO facilities as we mean them in this paper.
- 6 The Hawthorne effect refers to a situation in which an individual's behavior changes when they realize they are being observed. It is characterized by a positive but temporary change in some measurable behavior in a situation in which there was no deliberate attempt to affect behavior (Benson, 2000; Mayo, 1933).
- 7 Leonard and Masatu (2006) use regression analysis to verify the significance of the change in quality before and after observation, the gradual fall in quality as time passes for observed doctors, the unchanging quality before the team arrives, and the unchanging quality for doctors who are never observed.

References

- Barr, A., Fafchamps, M., & Owens, T.** (2005). The governance of non-governmental organizations in Uganda. *World Development*, 33, 657–679.
- Bennett, S., McPake, B., & Mills, A. (Eds.).** (1997). *Private health providers in developing countries: Serving the public interest?* London: Zed Books.

- Benson, P. G.** (2000). The Hawthorne effect. In W. E. Craighead & C. B. Nemeroff (Eds.), *The Corsini encyclopedia of psychology and behavioral science*. New York: Wiley.
- Berman, P., Nwuke, K., Hanson, K., Kariuki, M., Mbugua, K., Ngugi, J., et al.** (1995, April). *Kenya: Non-governmental health care provision* (Technical Report, Department of Population and International Health, Harvard School of Public Health). Cambridge: Harvard University.
- Berman, P., Nwuke, K., Rannan-Eliya, R., & Mwanza, A.** (1995, January). *Zambia: Non-governmental health care provision* (Technical Report, Department of Population and International Health, Harvard School of Public Health). Cambridge: Harvard University.
- Bossert, T.** (1998). Analyzing the decentralization of health systems in developing countries: Decision space, innovation, and performance. *Social Science and Medicine*, 4 (10), 1513–1527.
- Das, J., & Hammer, J.** (2005). Which doctor? Combining vignettes and item-response to measure doctor quality. *Journal of Development Economics*, 78, 348–383.
- Das, J., & Hammer, J.** (2007). Money for nothing: The dire straits of medical practice in Delhi, India. *Journal of Development Economics*, 83(1), 1–36.
- Das, J., Hammer, J. S., & Leonard, K. L.** (2008). The quality of medical advice in low-income countries. *Journal of Economic Perspectives*, 22(2), 93–114.
- Dean, J. M., Schaffner, J., & Smith, S. L. S. (Eds.)**. (2005). *Attacking poverty in the developing world: Christian practitioners and academics in collaboration*. London: World Vision and Authentic Media.
- Epstein, S. A., Gonzales, J. J., Weinfurt, K., Boekeloo, B., Yuan, N., & Chase, G.** (2001). Are psychiatrists' characteristics related to how they care for depression in the medically ill? Results from a national case-vignette survey. *Psychosomatics*, 42(6), 482–489.
- de Geyndt, W.** (1995). *Managing the quality of health care in developing countries* (World Bank Technical Paper 258). Washington, D.C.: The World Bank.
- Gilson, L., Adusei, J., Arhin, O., Hongoro, C., Mujinja, P., & Sagoe, K.** (1997). Should African governments contract out clinical health services to church providers? In S. Bennett, B. McPake, & A. Mills, (Eds.), *Private health providers in developing countries: Serving the public interest?* London: Zed Books.
- van den Hombergh, P., Grol, R., van den Hoogen, H. J. M., & van den Bosch, W. J. H. M.** (1999). Practice visits as a tool in quality improvement: Mutual visits and feedback by peers compared with visits and feedback by non-physician observers. *Quality in Health Care*, 8, 161–166.

- Kalf, A. J. H., & Spruijt-Metz, D.** (1996). Variation in diagnoses: Influence of specialists' training on selecting and ranking relevant information in geriatric case vignettes. *Social Science and Medicine*, 42(5), 705–712.
- Kanji, N., Kilima, P. M., & Munishi, P. M.** (1992). *Quality of primary curative care in Dar-Es-Salaam* (Working Paper, Urban Health Program). London: School of Hygiene and Tropical Medicine.
- Koedoot, C. G., de Haes, J. C. J. M., Heisterkamp, S. H., Bakker, P. J. M., de Graeff, A., & de Haan, R. J.** (2002). Palliative chemotherapy or watchful waiting? A vignettes study among oncologists. *Journal of Clinical Oncology*, 20(17), 3658–3664.
- Leonard, K. L.** (2002). When states and markets fail: Asymmetric information and the role of NGOs in African health care. *International Review of Law and Economics*, 22(1), 61–80.
- Leonard, K. L., & Masatu, M. C.** (2005). The use of direct clinician observation and vignettes for health services quality evaluation in developing countries. *Social Science and Medicine*, 61(9), 1944–1951.
- Leonard, K. L., & Masatu, M. C.** (2006). Outpatient process quality evaluation and the Hawthorne effect. *Social Science and Medicine*, 63(9), 2330–2340.
- Leonard, K. L., & Masatu, M. C.** (2007). Variation in the quality of care accessible to rural communities in Tanzania. *Health Affairs*, 26(3), w380–w392.
- Leonard, K. L., & Masatu, M. C.** (2008a). *Potentially altruistic doctors and the role of incentives in health care* (University of Maryland Working Paper 08–10). College Park, MD: University of Maryland.
- Leonard, K. L., & Masatu, M. C.** (2008b). *Using the Hawthorne effect to examine the gap between a doctor's best possible practice and actual performance* (University of Maryland Working Paper 08–04). College Park, MD: University of Maryland.
- Leonard, K. L., Masatu, M. C., & Vialou, A.** (2007). Getting doctors to do their best: The roles of ability and motivation in health care. *Journal of Human Resources*, 42(3), 682–700.
- Mayo, E.** (1933). *The human problems of an industrial civilization*. New York: MacMillan.
- McLeod, P. J., Tamblyn, R. M., Gayton, D., Grad, R., Snell, L., Berkson, L., et al.** (1997). Use of standardized patients to assess between-physician variations in resource utilization. *Journal of the American Medical Association*, 278, 1164–1168.
- McPake, B.** (1997). The role of the private sector in health service provision. In S. Bennett, B. McPake, & A. Mills (Eds.), *Private health providers in developing countries: Serving the public interest?* London: Zed Books.

- Miguel, E., & Kremer, M.** (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1), 159–217.
- Mliga, G. R.** (2000). Decentralization and the quality of health care. In D. K. Leonard (Ed.), *Africa's changing markets for human and animal health services*. London: Macmillan. Also available at <http://repositories.cdlib.org/uciaspubs/editedvolumes/5/>.
- Murata, P. J., McGlynn, E. A., Siu, A. L., & Brook, R. H.** (1994). *Prenatal care: A literature review and quality assessment criteria* (Research Report, Rand Corp). Santa Monica, CA: Rand Corporation.
- Murata, P. J., McGlynn, E. A., Siu, A. L., Raube, K., Sherwood, M. J., Roth, C. P., et al.** (1994). Quality measures for prenatal care. *Archives of Family Medicine*, 3(1), 41–49.
- O'Flaherty, M., Lerum, K., Martin, P., & Grassi, D.** (2002). Low agreement for assessing the risk of postoperative deep venous thrombosis when deciding prophylaxis strategies: a study using clinical vignettes. *BMC Health Services Research*, 2(16), 1–3.
- Peabody, J. W., Gertler, P. J., & Leibowitz, A.** (1998). The effects of structure and process of medical care on birth outcomes in Jamaica. *Health Policy*, 43(1), 1–13.
- Peabody, J. W., Luck, J., Glassman, P. A., Dresselhaus, T. R., & Lee, M.** (2000). Comparison of vignettes, standardized patients, and chart abstraction: A prospective validation study of three methods for measuring quality. *Journal of the American Medical Association*, 283, 1715–1722.
- Peabody, J. W., Rahman, O., Fox, K., & Gertler, P.** (1994). Quality of care in public and private primary health care facilities: Structural comparisons in Jamaica. *Bulletin of the Pan American Health Organization*, 28, 122–141.
- Reinikka, R., & Svenson, J.** (2004). *Working for God?* (Discussion Paper 4214, Center for Economic Policy Research). Washington, D.C.: Center for Economic Policy Research.
- Tiemeier, H., de Vries, W. J., van het Loo, M., Kahan, J. P., Klazinga, N., Grol, R., et al.** (2002). Guideline adherence rates and interprofessional variation in a vignette study of depression. *Quality & Safety in Health Care*, 11(3), 214–218. ■

Research Partnerships Between Faith-Based NGOs and Academic Researchers: An Example from Food Security and HIV and AIDS Research in Delhi, India

Paul E. McNamara, University of Illinois at Urbana-Champaign; Joel Cuffey, International Food Research Institute, Gulu, Uganda; Anil Cherian, Emmanuel Hospital Association, New Delhi, India; Saira Paulose, Emmanuel Hospital Association, New Delhi, India

Abstract: *A divide commonly exists between Christian development practitioners and academic researchers, making it difficult to collaborate and engage in potentially illuminating research projects. While solid reasons exist to explain the difficulties in linking academics and NGOs, partnerships offer the potential benefit of the generation of knowledge that can both inform development practice as well as change modes of thinking in scholarly communities. Further, we argue that such partnerships offer the possibility of a way of doing research that is distinctly Christian and thus can be a witness to peers in the development community as well as in academia. This paper presents a theological basis for a research partnership, details some of the challenges faced in building such a partnership, and describes an example of one such partnership between staff of the Emmanuel Hospital Association in India and researchers from the University of Illinois at Urbana-Champaign.*

Among Christians interested in engaging the world, practitioners and academic researchers often segregate themselves into camps that have little to do with each other. University-based researchers can explore the “bigger picture” while failing to address the needs of practitioners. In turn practitioners, feeling that research has little use for them, can ignore potentially illuminating scholarship. Partnerships between academic researchers and Christian nongovernmental organizations (NGOs) provide a means to bridge this divide. Such partnerships offer the possibility of the generation of knowledge that can both inform development practice as well as change modes of thinking in scholarly communities. Furthermore, such partnerships offer the possibility of a way of doing research that is distinctly Christian and thus can be a witness to peers in the development community as well as in academia. Despite the potential of research partnerships between academic researchers and NGO staff, relatively few examples exist that have been documented and reported in the academic literature. This paper serves to describe and report on an example of one such partnership.

This paper illustrates such a partnership in health economics. First we will address the theological underpinnings for such a partnership. Then