

# Providers' Compliance with the Balanced Counseling Strategy in Guatemala

Federico R. León, Carlos Brambila, Marisela de la Cruz,  
Julio García Colindres, Carlos Morales, and Benedicto Vásquez

*The balanced counseling strategy developed in Peru improved family planning care and clients' knowledge of their contraceptive method choice, but few providers adopted it. To expand its use, an algorithm was introduced and training, job aids, and reinforcement were supplied to Ministry of Health providers, most of whom were paraprofessionals, from two areas (40 clinics) in Guatemala. Mystery clients made pretest and post-test visits to these clinics and to providers from a nonequivalent control group (40 clinics). The results showed that the strategy was used in 85 percent of the controlled consultations at the experimental clinics. Use of the strategy improved quality of care regardless of the provider's performance at baseline and regardless of ethnic or regional differences. Counseling session length increased by nine minutes, but real-client load did not change. Guatemalan clients can be expected to benefit from the strategy. The increased session length has not yet caused problems, but it may pose policy dilemmas in the future. (STUDIES IN FAMILY PLANNING 2005; 36[2]: 117–126)*

A recent literature review classified the interventions designed to improve the quality of family planning care into three types: systemwide interventions, interventions to expand contraceptive choice, and interventions relying on specific tools (RamaRao and Mohanam 2003). This article describes an effort to refine and assess a novel intervention of the third type. This new intervention is known as the balanced counseling strategy.

The strategy was developed to meet the challenge revealed by the findings of a study of client-provider interactions at Peru Ministry of Health (MOH) clinics in

which the questions asked by providers were medical in nature and ignored the client's reproductive intentions and whether her partner cooperated with her in regard to family planning. The providers, using a flipchart, behaved as though they had to describe to the client the attributes of all the methods offered by the program, regardless of the client's particular needs (León et al. 2001). The client was expected to make a contraceptive choice by evaluating the full set of attributes, an expectation that implied an underlying linear decisionmaking model that placed exaggerated demands on her information-processing capabilities. People do not make such complex calculations, but rather make choices by means of sequential decisionmaking (Simon 1955 and 1978). Because most of the session time was spent describing all method options, the client obtained as little information about the method she chose as about methods that were irrelevant to her (León 1999).

To address this problem, the balanced counseling strategy reorganizes the family planning consultation with the aid of two techniques that, by means of sequential decisionmaking, help to simplify the client's task of choosing a method (León 2002). First, the consultation is reorganized around a process of elimination. At the outset, needs assessment works as a process for discarding those methods that the client and the provider identify as inappropriate in her case (for example, sterilization, if the client wishes to have children later). In the

---

*At the time this study was conducted, Federico R. León and Carlos Brambila were Program Associates, and Marisela de la Cruz was Project Coordinator, Frontiers in Reproductive Health Program, Population Council. Julio García Colindres was Director, Carlos Morales was Family Planning Specialist, and Benedicto Vásquez was Chief of Family Planning, Reproductive Health Division of the Guatemalan Ministry of Public Health and Social Welfare. Correspondence should be directed to Federico R. León, Avenida Javier Prado 7427, Lima 3, Peru. E-mail: frleon@amauta.rcp.net.pe. Requests for the job aids and administration manual of the strategy can be sent to the Frontiers in Reproductive Health program, Population Council, 4301 Connecticut Avenue, NW, Suite 280, Washington, DC 20008. E-mail: English: Iraney@popcouncil.org; Spanish: tmartin@popcouncil.org.mx.*

choice phase, the provider describes only the relevant methods and conveys information that is essential for a preliminary choice. Then the provider and client focus on the chosen method. If this method is contraindicated or if the client rejects it once she learns more about it, she can return to the choice phase.

Second, the balanced counseling strategy incorporates visual aids that assist client and counselor alike. Because workers require challenging tasks (Locke and Latham 1990) that they feel able to accomplish in order to be motivated (Vroom 1964), León (2002) developed two sets of job aids to enhance providers' counseling skills and their feeling of self-efficacy. One set includes 11 hand-sized cards, one for each method offered by the family planning program, which the provider displays on the table. When a method is identified as inappropriate for the client, the provider asks the client's consent to discard that card, telling her why the method would not be suitable. The provider reads the remaining cards aloud (or gives the client the cards to read), detailing four fundamental attributes per method, and asks her to make a choice. The second set of aids consists of pamphlets for the client, one for each method, which incorporate the method's contraindications, instructions for use, side effects, alarm signs, and follow-up indications as described in the Norms of the National Family Planning Program (Peru MOH 1999). Once the client makes a choice, the provider uses the corresponding pamphlet to evaluate contraindications and give her detailed information on the method she has chosen. The provider explains to the client that she can take the pamphlet home to use as a technical guide.

A field test at Peru MOH clinics showed that providers trained in the strategy improved the quality of the care they offered when they used the method cards and pamphlets in their interactions with clients. In clinic exit interviews, clients who chose hormonal methods or the IUD (the choices of the vast majority of clients) increased their knowledge of the method they had chosen when the provider used the job aids. However, the intervention was limited in its success by a number of shortcomings in the implementation of the strategy. For example, because only 37 percent of the trained providers complied with the requirement that they use these aids, only a minority of clients benefited from the intervention (León 2002). Additionally, the trainer-trainee ratio was poor, the method pamphlets for clients were introduced late during the training, and reinforcement by family planning supervisors to monitor on-the-job service provision did not materialize. Therefore, an accurate assessment of the full benefits of the strategy could not be made from the field test in Peru.

The strategy generated interest at the Guatemalan Ministry of Public Health and Social Welfare (MPHSW). Concern was raised, however, by the noncompliance problem observed in Peru, which had the potential to be greater in Guatemala because the overwhelming majority of MPHSW providers (about 70 percent) are auxiliary nurses with limited education. The behavioral guidelines for providers for implementation of the strategy were few and could cause confusion in this population. Because strengthened training alone might not be enough to enhance providers' confidence and ensure use of the new counseling model, the solution developed consisted of a service algorithm that offered step-by-step instructions (see Appendix Figure A1).

This study tested whether the balanced counseling strategy, when introduced by means of intensive training and reinforcement, can yield a satisfactory level of compliance among providers and a substantial enhancement of the quality of care. Possible unintended effects of the new paradigm, such as increased counseling session length and decreased number of clients counseled, were also targeted for evaluation.

## Methods

A nonequivalent control-group quasi-experiment with pretest and post-test components (Cook and Campbell 1979) was replicated in highland and lowland administrative health areas of Guatemala. The study was conducted from October 2001 through October 2002.

### *Study Sites*

We compared two areas in the highlands that serve mainly Mayan populations; Quiché's southern half was selected to parallel the high-altitude districts of Quetzaltenango. The other two areas, Jutiapa and Jalapa, serve Spanish-speaking Ladino populations in the lowland southeast. In each area, seven health centers and 13 health posts were included in the study, the selection criterion being maximum similarity between health areas within each region.

### *Measurement of Providers' Compliance*

We defined providers' compliance with the strategy as their use of the method cards and pamphlets according to service test reports. In the service test (León et al. 2001), services were sought by "mystery clients" trained to enact a specific client profile, to observe the provider's behavior, and to report their observations. The providers were unaware that they were under observation, and the mystery clients were unaware of the study's hypotheses.

### *Measurement of Quality of Care*

Two client profiles were developed for the service test. Profile A depicts a healthy woman aged 19–22 with a 10-month-old child who has never practiced contraception but who wants to control her fertility because her husband, an unskilled laborer, is returning from overseas. The profile requires her to reject all methods offered except depot medroxy-progesterone acetate (DMPA). The client has never had a Pap smear and has not had intercourse during the past year. Profile B depicted an older woman, 30–40 years old, with four children. She and her husband feel that they cannot afford to have a fifth child. The woman has experienced one miscarriage, but is otherwise healthy. She has practiced periodic abstinence, used condoms and pills, and her last three children were conceived as a result of contraceptive failures. Currently, the client would prefer a permanent method, specifically tubal ligation.

The two most prevalent contraceptive methods in Guatemala are tubal ligation and DMPA (MPHSW et al. 2003). Other methods offered by the MPHSW were vasectomy, the IUD, oral contraceptives, and condoms. Information was also available about lactational amenorrhea and periodic abstinence.

Each mystery client learned to enact one of the client profiles and to fill out a checklist as she left the premises. The checklist of items for profile A referred to 65 behaviors expected of providers in the management of a case such as client A. The profile B checklist included 64 items. The mystery client marked whether or not she observed the behavior at any time during the consultation. The dichotomous item scores (1 = observed, 0 = not observed) produced a total summary score and the following three subscores:

(1) Relations/needs included interpersonal relations (for example, “The provider treated me amiably”) and needs diagnosis (“The provider asked what methods I had used in the past”).

(2) Choice/contraindications involved the offer of relevant method options (“The provider told me that the IUD can protect me for up to 10 years”) and screening for the chosen method’s contraindications (“The provider asked if I had a lump or swelling in my breasts”).

(3) Method details pertained to information provided about the chosen method, including instructions for use (“The provider told me that I would have to go to a hospital for the tubal ligation”), side effects (“The provider said that I could have partial or total absence of menstruation”), alarm signs (“The provider said that I should see a provider immediately if I had very strong headaches”), and follow-up (“The provider asked if I had understood what he/she had told me”).

### *Measurement of Unintended Effects*

The mystery clients learned to time the length of the session by paying attention to the facility’s clock and registering the time at which the consultation began and finished. (Women in these areas of Guatemala generally do not wear watches.) The session included any clinical procedures performed for the client and excluded waiting time.

Information about the clinic’s client load—the number of clients (women, men, children) attended during the month of data collection—was obtained from each facility’s service statistics. The data were obtained by a study coordinator who requested the facility’s service statistics weeks after the pretest and, months later, after the post-test.

### *Intervention and Data-collection Process*

The job aids were printed in Spanish. The 22-step algorithm was disaggregated and presented with detailed verbal instructions. Eight method cards and one to discard the option of pregnancy were printed. The method cards presented on one side four essential attributes of the method and on the other its contraindications. The contents of the job aids were derived from the National Reproductive Health Care Guidelines (MPHSW 2000).

Providers were approached individually prior to data collection and invited to participate in the research. To take account of cultural differences between the study sites, *ad hoc* teams of mystery clients were recruited in the Mayan and Ladino regions. Two supervised two-person teams (one mystery client per client profile) operated in each region. Once the two mystery clients completed their visits to a clinic, they moved on to the next clinic. They visited whichever person provided family planning services at each facility. (Following-up with individual providers for the post-test would have been obtrusive.) Observational bias was controlled within each region by having one of the teams first visit ten facilities in one (for example, control) health area and then cross to the other (for example, experimental) area to visit the ten facilities that had not been covered by the parallel team. To avoid the possibility that providers might recognize the mystery clients, the teams that participated in the pretest were replaced by different teams in the post-test.

Quiché and Jutiapa were randomly assigned as the strategy’s intervention sites and Quetzaltenango and Jalapa were used as controls. Four health professionals were coached as trainers and produced agendas for training providers. The provider-training sessions emphasized role playing, lasted six hours, and targeted the health

center's medical director, nurse, and nurse auxiliaries as well as providers from all the dependent health posts. On average, eight participants attended each training session.

The training team made four more visits of variable length to each provider at each experimental facility to reinforce their knowledge and to train those who had been absent during the main training session. If the provider was consulting with a family planning client, the counseling session was observed and the provider was given feedback. If family planning clients were not available, as was often the case, the trainer asked the provider to role-play the use of the algorithm and supplied immediate feedback.

### Uncontrolled Intervention

A parallel intervention was performed at the time of our study by a local organization, Calidad en Salud, which trained MPHSW providers in contraceptive technology and counseling nationwide. The intervention encompassed both our experimental and control groups. Calidad en Salud followed the GATHER approach to counseling (see Rinehart et al. 1998), which is not inconsistent with the balanced counseling strategy.<sup>1</sup> The strategy adds precision to the GATHER postulates. Calidad en Salud produced the method pamphlets required by our conceptualization and distributed them nationwide. The Calidad en Salud training began before our pretest and its reinforcement continued throughout our project's life. Table 1 summarizes the inputs in control and experimental sites.

## Results

To assess providers' compliance with the strategy, we use percentages based on the total number of service test protocols filled out. The mystery clients' reports from the post-test showed that in 72 percent of the observations made at experimental clinics the providers used

**Table 1** Controlled and uncontrolled inputs in experimental and control sites, Guatemala, 2001–02

Inputs	Experimental sites (Quiché, Jutiapa)	Control sites (Quetzaltenango, Jalapa)
Training of providers in balanced counseling strategy. Delivery of algorithm and method cards to providers. Reinforcing visits.	Yes	No
Training of providers in contraceptive technology and GATHER counseling. Delivery of method pamphlets to facilities.	Yes	Yes

the method cards and the pamphlet during counseling. One reason for noncompliance was the unavailability of pamphlets at the facilities. In their consultations with mystery clients, the great majority of providers (85 percent) used the method cards distributed by project staff.

### Consistency of Quality-of-care Indicators

The summary scores of the service test for both profiles showed normal distributions. Their consistency was evaluated by means of product-moment correlations entailing each clinic's scores and subscores. Table 2 presents the intercorrelations among the quality-of-care indicators. The maximum possible number of protocols was 20 clinics x 4 health areas x 2 client types x 2 measurement occasions (N = 320), but the actual number of the correlations reached a maximum of 137 pairs, that is, 274 observations, because of the temporary or permanent closing of facilities at the time of the mystery clients' visits. The part-total same-profile correlations, ranging from 0.79 through 0.96, suggest a high level of internal-consistency reliability and justify exclusive use of the total score (full set of items).

The correlations involving different profiles were lower. This finding can be attributed in part to the different contents of the respective checklists. Another source of variability was the heterogeneity of observers:

**Table 2** Coefficients of Pearson correlations between quality-of-care indicators, balanced counseling strategy intervention, Guatemala, 2001–02

Service test profiles and components	Service test components			
	Relations/needs	Choice/contraindications	Method details	Total (full set of items)
Same profile <sup>a</sup>				
Relations/needs	1.00	0.71	0.68	0.79
Choice/contraindications	0.66	1.00	0.87	0.96
Method details	0.64	0.82	1.00	0.96
Total (full set of items)	0.78	0.97	0.92	1.00
Different profile <sup>b</sup>				
Relations/needs	0.58	0.59	0.56	0.63
Choice/contraindications	0.52	0.75	0.66	0.75
Method details	0.54	0.74	0.69	0.75
Total (full set of items)	0.58	0.77	0.70	0.78

**Note:** All the correlations are significant at  $p < 0.01$ , using a two-tailed test.

<sup>a</sup> The triangle of coefficients above the diagonal presents correlations between profile A indicators (injectable choice), whereas the triangle below the diagonal shows the correlations between profile B indicators (tubal ligation choice). The diagonal contains 1.00s because these coefficients correspond to the correlation between the indicator and itself. <sup>b</sup> This set of coefficients refers to the correlations between profile A and profile B indicators. Here, the diagonal contains coefficients pertaining to the parallel indicators (for example, choice/contraindications from profile A with choice/contraindications from profile B). The upper triangle contains the correlation between a component of profile A (for example, relations/needs) and another component of profile B (for example, choice/contraindications), whereas the lower triangle contains the correlation between the same components inverted (for example, relations/needs of profile B with choice/contraindications of profile A).

different mystery clients enacted profiles A and B of the service test. Nonetheless, the total-score correlation between the two profiles ( $r = 0.78$ ) was high enough to justify aggregating them and using a single quality-of-care indicator in the study as a basis for decisionmaking concerning the adoption of the new strategy.

### Effects on the Quality of Care

Figure 1 presents the quality-of-care results for the Mayan and Ladino regions. Reflecting the concentration of past international cooperation and Calidad en Salud's current emphasis on Mayan populations, the results were better in the highlands. Nonetheless, the trends are virtually identical. In both cases, in the experimental health area quality of care was improved to a substantially greater extent than it was in the control area. To assess statistically the effects of our controlled intervention while maximizing parsimony, the Mayan and Ladino regions were integrated. An integrated experimental group was formed by combining data from Quiché and Jutiapa, and Quetzaltenango and Jalapa data were combined to form an integrated control group. The quality-of-care means increased from 20.2 to 24.4 in the control

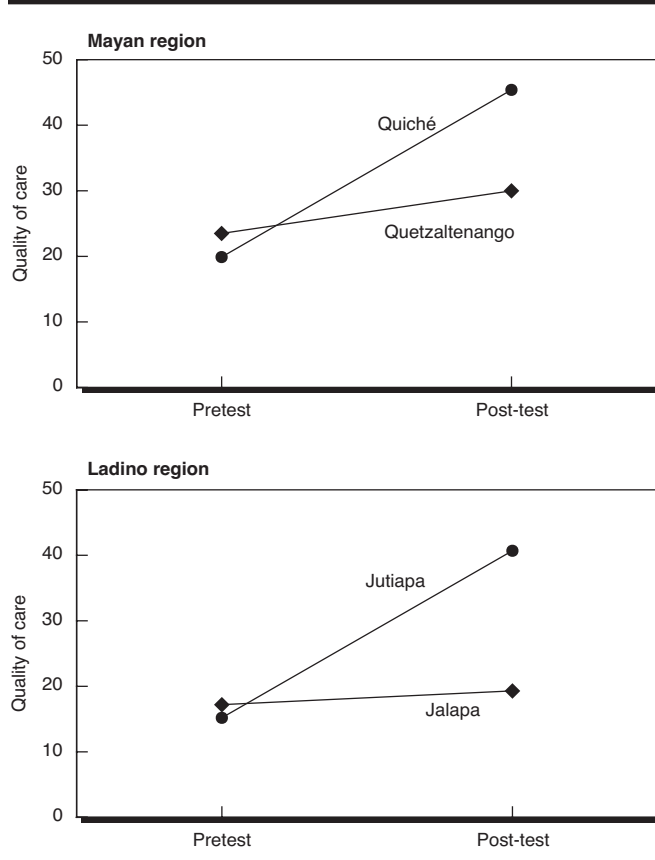
group and from 17.7 to 42.8 in the experimental group.

The simplest statistical analysis for a nonequivalent control-group quasi-experiment with pretest and post-test components is the comparison of gain scores (Reichardt 1979). Another reason for using gain scores is that both the experimental and control groups received interventions and were expected to show improvements. Hence, a post-test minus pretest difference in quality of care was calculated for each case, and the gain scores were subjected to a  $t$ -test for differences in means between independent samples. Table 3 presents the results. The statistically significant finding reappeared when the analysis targeted each of the client profiles of the service test and was replicated in each region.

To obtain a standardized measure of the amount of change that allowed a comparison between the results for the dependent variables, we reverted to the effect size (Lipsey 1990; Kline 2004), which expresses the difference between the experimental and control means in pooled standard-deviation units. In the calculation of the effect size, the difference between the experimental ( $Mean_{\Delta_e}$ ) and control ( $Mean_{\Delta_c}$ ) mean gains was divided by  $([s_{\Delta_e}^2 + s_{\Delta_c}^2] / 2)^{1/2}$ , where  $s_{\Delta}$  is the sample standard deviation for the gains. As shown in Table 3, use of the strategy improved the quality of care by 2.4 standard deviations while the unintended effects amount to less than one standard deviation.

To attain further insights, we calculated standardized regression coefficients ( $\beta$ ) from an ordinary linear regression using the quality-of-care gain score (post-test minus pretest) as the dependent variable. The predictors, jointly evaluated in the same equation, are the quality-of-care score obtained in the pretest (a continuous variable), the region (Mayan = 1, Ladino = 0), and the strategy factor (experimental group = 1, control group = 0). As shown in Table 4, the greatest coefficient pertains to the region, indicating greater gains in the Mayan than in the Ladino health areas. This difference is expressed in Figure 1 by the difference in steepness between the Mayan and Ladino control-group curves, which we interpret as a consequence of the priority accorded by the Calidad en Salud intervention to Mayan populations. The significant pretest factor reveals that greater gains were achieved within each regional group and within the experimental and control levels of the strategy factor by the providers who had poorer quality-of-care scores at baseline. The poorest achievers at baseline apparently took greater advantage of the interventions and gained the most between pretest and post-test. Finally, the regression model predicts greater gains for experimental than for control clinics when the effects of region and pretest are removed from the quality-of-care scores. An independent effect emerges indicating that the strategy im-

**Figure 1** Mean quality-of-care scores per health area before and after the intervention, Guatemala, 2001–02



**Table 3** Effects of the balanced counseling strategy intervention on mean gains for quality of care, session length, and client load, Guatemala, 2001–02

Dependent measure	Control group <sup>a</sup>			Experimental group <sup>b</sup>			t-value	Effect size
	Number of cases	Mean gain per clinic	Gain standard deviation	Number of cases	Mean gain per clinic	Gain standard deviation		
Quality of care	71	4.30	9.19	64	25.17	8.27	13.81*	2.39
Session length (minutes/seconds)	71	1/59	9/47	68	10/40	15/27	3.94*	0.67
Client load	37	4.54	18.76	34	3.02	25.64	-0.28	-0.07

\*Significant at  $p < 0.001$ , using a one-tailed test.

**Note:** Generally, two cases per clinic were used with the service test (injectable and tubal ligation profiles). In the case of client load, the number of cases is the number of clinics. The variability in sample size is due to the loss of clinics or lack of measurements. The difference between t-value and effect size is that the latter does not consider the degrees of freedom.

<sup>a</sup> Data for Quetzaltenango and Jalapa combined. <sup>b</sup> Data for Quiché and Jutiapa combined.

proved the quality of care regardless of the provider's performance level at the baseline and regardless of ethnic or regional differences.

### Unintended Effects

Similar operations were performed for the length of sessions (in minutes). Although session length was nonnormally distributed, the distribution of the gain scores is normal. Counseling session length increased from 12 minutes 12 seconds to 14 minutes 8 seconds in the control group and from 13 minutes 18 seconds to 24 minutes 16

seconds in the experimental group, which amounts to a nine-minute difference in mean gains between the groups. The difference between mean gains is significant, but the impact of the strategy on session length alone is about one-third of its impact on quality of care (see Table 3).

In the regression analysis, the greatest increases in session length are shown for the Mayans (within intervention groups and levels of pretest session length) and by the providers who had invested the least time in the consultations at the baseline (within regional and intervention groups; see Table 4). The standardized coefficient for the strategy is not statistically significant. This finding can be understood as an indication that the residuals lacked any further intervention effects. The strategy and Calidad en Salud interventions had their greatest impacts in increasing the dedication of the providers who had invested the least time in the consultations at the baseline, but these effects were controlled by the pretest factor in the regression.

The quality-of-care improvements attributable to the strategy were observed at every level of session length. Figure 2 compares the consultations conducted by experimental providers in the post-test with their consultations in the pretest averaged with those of the control providers in pretest and post-test. None of the former lasted less than six minutes, whereas one-fourth of the latter did so. This finding reveals that all the experimental providers who in the pretest showed brief session lengths increased the time they invested in the consultation in the post-test.

To obtain an average daily number of clients served at each clinic (C), we computed  $C = N/D$ , where N is the number of clients attended at the clinic during the month that the simulated clients visited that clinic and D is the number of working days at that clinic in that month. To obtain the mean for each group, we aggregated the Cs for a given group and divided that number by the number of clinics in that group. Daily client load increased from 21.5 in the pretest to 27.6 in the post-test in the control group and from 30.9 in the pretest to

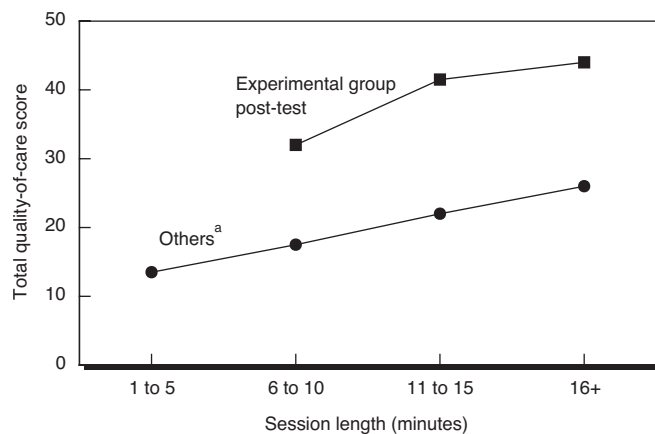
**Table 4** Standardized regression coefficients from the regression of gain scores on pretest (continuous variable), region, and strategy, by quality of care, session length, and client load, Guatemala, 2001–02

Dependent variable/predictor	$\beta$	Standard error	t-test
Quality of care <sup>a</sup>			
Pretest	-0.35	0.05	-4.64***
Region			
Mayan	0.51	1.22	7.00***
Ladino	0.00	0.00	0.00
Strategy			
Experimental	0.15	1.28	1.96*
Control	0.00	0.00	0.00
Session length <sup>b</sup>			
Pretest	-0.33	0.10	-3.63***
Region			
Mayan	0.33	2.01	3.80***
Ladino	0.00	0.00	0.00
Strategy			
Experimental	0.06	1.93	0.68
Control	0.00	0.00	0.00
Client load <sup>c</sup>			
Pretest	-0.70	0.06	-7.64***
Region			
Mayan	-0.09	4.04	-1.03
Ladino	0.00	0.00	0.00
Strategy			
Experimental	0.06	3.96	0.64
Control	0.00	0.00	0.00

\*Significant at  $p < 0.05$ ; \*\*\* $p < 0.001$ .

<sup>a</sup> Adjusted R square = 0.32. <sup>b</sup> Adjusted R square = 0.11. <sup>c</sup> Adjusted R square = 0.44.

**Figure 2** Quality-of-care score at various levels of session length: post-test of experimental group versus others, Guatemala, 2001–02



<sup>a</sup> Pretest of experimental group + pretest of control group + post-test of control group/3.

33.9 in the post-test in the experimental group. These results reveal important differences between the experimental and control groups at the baseline and thus confirm that the assumption of nonequivalence between the groups was justified and that the gain was the natural unit of analysis in this study. The gain in client load was greater for the control than for the experimental group, but the *t*-ratio computed for the gains was not statistically significant ( $p < 0.39$ , with a one-tailed test).

Only the standardized regression coefficient for the pretest factor was significant (see Table 4). The smaller the number of clients in the pretest, the greater was the gain from pretest to post-test. This gain can be attributed to the *Calidad en Salud* intervention, which sought to increase coverage in addition to making quality-of-care improvements.

## Discussion

Compared with 37 percent of trained providers who used the strategy's method cards and pamphlets at experimental clinics in the Peru study (León 2002), 72 percent did so in this study. Had some logistical problems with the pamphlets been avoided, the proportion could have been greater: 85 percent of the MPHSW providers used the method cards distributed by project staff.

The wide use of the balanced counseling strategy was reflected in the quality of care received by the mystery clients. The providers trained in the strategy enhanced their quality of care significantly above that of providers who had been trained only in contraceptive technology and the GATHER approach to counseling.

The regression analysis revealed that the strategy improved the quality of care offered by providers regardless of their level of performance at the baseline even though the providers with lower baseline scores showed greater improvement. This finding indicates that important service-delivery improvements can be achieved if certain manpower goals are met. We assume that foremost among the factors relevant to improvement were the sense of institutional legitimacy accorded to the new counseling model and the providers' feeling of self-efficacy resulting from their solid grounding in the strategy. Other factors that contributed to the enhancement of quality of care were the precise instructions accompanying the 22-step algorithm, the focused training, and the strong reinforcement of the training. Because these were handled as a package, we cannot ascertain whether the effect of one key determinant predominated.

In the Peru study, roughly similar results were obtained with the full set of service test items and a reduced set (30 percent) of essential items (León 2001). Yet, to take advantage of the higher reliability associated with a larger number of items, we used a large list in Guatemala, incurring a risk of recall problems for the mystery clients. The service test showed a high level of internal-consistency reliability, however. The high correlations observed among the different components, the total score, and the different profiles suggest that the amount of recall error was limited. As for the test's validity, Luck and Peabody (2002), working with medical services that did not involve family planning, demonstrated a high level of consistency between the reports of mystery clients enacting a standard-patient profile and what transpired in consultations as revealed by recorded audio tapes.

The strategy improved the quality of care by 2.4 standard deviations. Because the unit of analysis was the gain from pretest to post-test, this size of effect must be interpreted accordingly. It does not reflect the variability of the quality-of-care scores, but rather the variability of the changes in quality of care that occurred in the experimental and control groups.

That the behaviors of providers provoked by mystery clients will be sustained with real clients is a conclusion that merits discussion. Our confidence in the construct validity of the findings is based on the interpretability of the results according to the assumption that, subjectively, the providers believed they were caring for real clients. Two widely different client profiles were employed in the service test to encompass a broad range of client types, and the differences in quality of care observed between experimental and control groups held across the client profiles. Therefore, we can expect that the increase in quality of care will hold across a wide variety of real clients.

Three other sets of findings are consistent with the conclusion that, subjectively, the providers believed they were caring for real clients. First, quality of care increased from pretest to post-test in the control group, which can be explained by the parallel intervention performed by *Calidad en Salud* that trained providers in contraceptive technology and the GATHER approach to counseling nationwide at the time of our study. Second, the Mayan providers benefited the most from the interventions, a result consistent with the concentration of the international technical cooperation in family planning in Mayan areas in the past and with the current stronger emphasis of *Calidad en Salud*'s intervention on Mayan areas. Third, more generally, the providers who offered lower quality of care at the baseline gained the most from the interventions, which indicates that the minimal quality-of-care level was raised.

An alternative interpretation would consider that the providers were aware of the identity of the mystery clients and that, consequently, what was measured was their maximal rather than their typical performance. Providers improve their performance when they sense that they are being observed (León et al. 2003). No evidence exists to suggest that the providers in this study felt observed, however. Moreover, the session lengths shorter than six minutes suggest that the observed behavior was typical.

The *Calidad en Salud* intervention conducted simultaneously with our study poses an interpretive problem. The results reveal that the balanced counseling strategy improved the quality of care above what the GATHER approach could achieve in the control group, and no evidence contradicts the parsimonious interpretation that the effects of the two interventions were additive and that the difference between the experimental and control groups is entirely attributable to use of the balanced counseling strategy. One question that remains unanswered, however, is whether the strategy would achieve the same results in the absence of the *Calidad en Salud* intervention. The treatment for the experimental group included the strategy as well as the *Calidad en Salud* intervention. Because the latter is a type of intervention that has been used throughout the developing world in the past decade, we conclude that the findings of this study can be generalized to a diverse array of settings. Nevertheless, whether the strategy can produce similar results in the absence of *Calidad en Salud*-like interventions remains unclear.

Quality assurance is a legitimate program goal in accord with the assumption that improved quality of care results in positive outcomes for clients. The small number of family planning clients at the Guatemalan clinics studied prevented us from evaluating this hypothesis. Nevertheless, our quality-of-care findings are indirectly relevant. In the Peru study (León 2002), an effect size three

times as large for providers' behavior appeared to be necessary to produce a one-unit improvement in clients' knowledge of their chosen method. In light of the level of effect size for quality-of-care gains obtained in the present study, positive knowledge effects amounting to 0.8 standard deviations can be extrapolated to the clients of MPHSW services. Further research is needed to confirm this hypothesized impact on client outcomes.

The wide implementation of the balanced counseling strategy had an unintended effect: the providers in the experimental clinics increased the length of their counseling sessions by nine minutes. Because this increase represents only one-third of the total impact of the intervention on the quality of care, the results of the study suggest that only part of the quality-of-care improvements required increased session length. The regression analysis revealed that the interventions caused greater increases in session length among the providers who had invested the least time in the consultations at the baseline, yet quality-of-care improvements were observed for every session length recorded.

That the increased session length had no negative effects on daily client load can be easily explained. Of approximately 30 clients who visit a typical Guatemalan health center or health post per day, only one, at best, is a new family planning client. Because providers may not be fully occupied with service or administrative functions during their shifts, each day's work can accommodate the extra nine minutes used in consultation with a new family planning client.

Session length could become a problem, however, if, as the MPHSW intends, the load of family planning clients increases substantially in the future. In such a contingency, spending nine extra minutes per family planning client could compromise providers' ability to attend to other clients, and the MPHSW would need to increase clinics' staff; otherwise, providers might return to old practices to save time. This potential policy dilemma could be avoided, however, if providers learn to save time by gaining greater expertise in using the balanced counseling strategy. Another solution is suggested by the finding that improvements in quality of care occurred at every level of session length. The MPHSW could set a time standard—for example, an average of 15 minutes per family planning session—and still be able to obtain substantial improvements in quality of care without causing unmanageable increases in session length. Based on this study, the MPHSW expanded the use of the strategy nationwide in 2002–03. Another problem the MPHSW might have to consider is the logistics of distributing the cards.

The balanced counseling strategy was developed to address an issue peculiar to Peru, where an excess of ir-

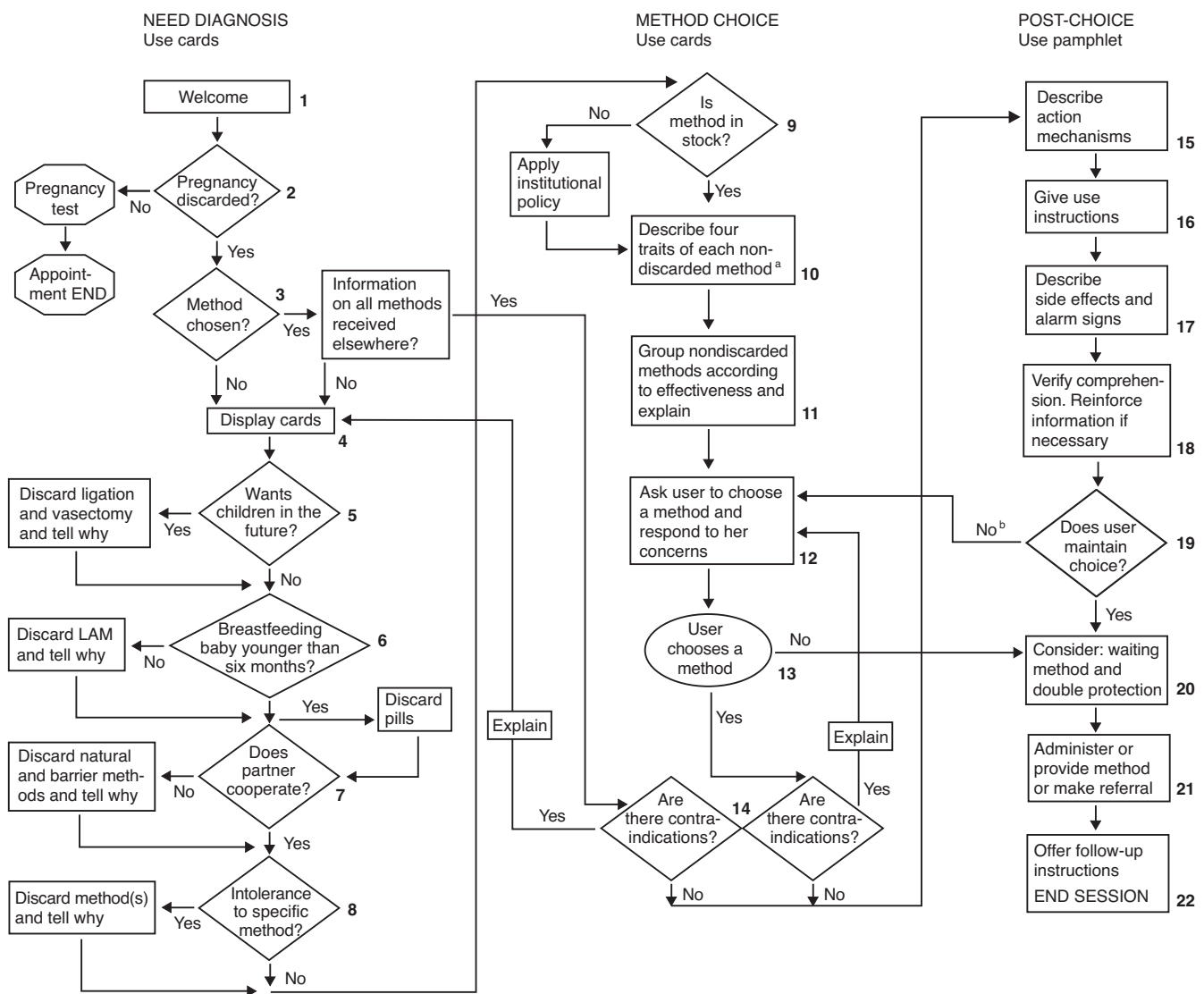
relevant information was provided to family planning clients. Will the strategy be useful in settings characterized by a limited exchange of information? In community-based family planning programs, the number of methods offered is small, and numerous studies in both community-based and clinical settings have shown that providers place limits on the information they offer to clients. Our study showed more positive impacts of the intervention in Guatemala than in Peru, although the number of methods offered by the Guatemala MPHSW was smaller (eight versus 11) and previous family planning training for providers was less intensive (in the La-

dino areas, it was nonexistent). The findings from this study suggest that the balanced counseling strategy may be universally useful if it is adequately implemented, and they point to the value of testing the strategy's effectiveness in less-developed settings.

Finally, a limitation of this study pertains to the selection of the two most prevalent methods, DMPA and tubal ligation, for the client profiles. Because we did not use client profiles other than these, it is impossible to establish whether the providers' instructions to a client who chose a prevalent method differed with respect to those delivered to a client who chose a less popular method.

## Appendix

**Figure A1** Job aids—assisted balanced counseling algorithm



LAM = Lactational amenorrhea method.

<sup>a</sup> Explain that IUD and injectables are specially recommended if long-term contraception is desired. <sup>b</sup> Display cards if the user has already chosen method.

## Note

- 1 The GATHER acronym stands for: Greet the client in a friendly manner, Ask questions and listen carefully, Tell the client tailored and personalized information, Help her make an informed choice, Explain to the client how to carry out her decision, and invite her to Return whenever she wishes.

## References

- Cook, Thomas D. and Donald T. Campbell (eds.). 1979. *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally.
- Kline, Rex B. 2004. *Beyond Significance Testing: Reforming Data Analysis Methods in Behavioral Research*. Washington, DC: American Psychological Association.
- León, Federico R. 1999. *Peru: Provider Compliance with Quality of Care Norms*. Final operations research project report of Frontiers in Reproductive Health Program. Lima: Population Council.
- . 2001. "The Job Aids–assisted Balanced Counseling Strategy Causes Positive Effects on Essential Provider's Service Behaviors." *Peru PCI-QoC Project Bulletin* No. 8. 15 June.
- . 2002. "Introducing the job aids–assisted balanced counseling strategy: A test in Peru MOH clinics." Paper presented at the 130th Annual Meeting of the American Public Health Association, Philadelphia, 9–13 November.
- León, Federico R., Rosa Monge, Adriana Zumarán, Ingeborg García, and Alex Ríos. 2001. "Length of counseling sessions and amount of relevant information exchanged: A study in Peruvian clinics." *International Family Planning Perspectives* 27(1): 28–33, 46.
- León, Federico R., Verónica Espinoza, Agustín Espinosa, and Betty Meza. 2003. "Expectancy of being observed and quality of care offered." Paper presented at the 131st Annual Meeting of the American Public Health Association, San Francisco, 16–19 November.
- Lipsey, Mark W. 1990. *Design Sensitivity: Statistical Power for Experimental Research*. Newbury Park, CA: Sage Publications.
- Locke, Edwin A. and Gary P. Latham. 1990. *A Theory of Goal Setting and Task Performance*. Englewood Cliffs, NJ: Prentice-Hall.
- Luck, James and John W. Peabody. 2002. "Using standardized patients to measure physicians' practice: Validation study using audio recorders." *British Medical Journal* 325: 679.
- Ministry of Public Health and Social Welfare (Guatemala) (MPHSW). 2000. *Guías Nacionales de Salud Reproductiva*. Guatemala City: Ministerio de Salud Pública y Asistencia Social de Guatemala.
- Ministry of Public Health and Social Welfare (MPHSW), Instituto Nacional de Estadística et al. 2003. *Encuesta Nacional de Salud Materno-infantil, Guatemala 2002*. Guatemala City: Editorial Oscar de León Palacios.
- Peru Ministry of Health (MOH). 1999. *Normas del Programa Nacional de Planificación Familiar*. Lima: Ministerio de Salud del Perú, Dirección General de Salud de las Personas.
- RamaRao, Saumya and Raji Mohanam. 2003. "The quality of family planning programs: Concepts, measurements, interventions, and effects." *Studies in Family Planning* 34(4): 227–248.
- Reichardt, Charles S. 1979. "The statistical analysis of data from non-equivalent group designs." In *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Eds. Thomas D. Cook and Donald T. Campbell. Chicago: Rand McNally. Pp. 147–205.
- Rinehart, Ward, Sharon Rudy, and Megan Drennan. 1998. "GATHER Guide to Counseling." *Population Reports, Series J, No. 48*. Baltimore: Johns Hopkins University School of Public Health, Population Information Program.
- Simon, Herbert A. 1955. "A behavioral model of rational choice." *Quarterly Journal of Economics* 69: 99–118.
- . 1978. "Rationality as process and as product of thought." *Journal of the American Economics Association* 68: 1–16.
- Vroom, Victor H. 1964. *Work and Motivation*. New York: Wiley.

## Acknowledgments

The authors are indebted to Edgar Reyes, Genard Méndez, Alba Ramírez, and Maritza de Andrino for coordinating and supporting the intervention in Quiché and Jutiapa; Irma Ramos for her assistance in the coaching of trainers; Rosa Monge in the training of simulated clients; Jorge Solórzano and Berna Salas in the coordination with Calidad en Salud; and Orly Lam, Clara Luz Barrios, and Julio Hernández in the training of providers. Ricardo Vernon made useful suggestions during the intervention. John Bratt conducted a cost study that is not reported here because of space limitations. Ed Scholl supported the project at USAID/Guatemala. The study was funded by the United States Agency for International Development.