CP191

Technology evaluation: five case studies from West Africa

Peter J. Matlon, International Crops Research Institute for the Semi-Arid Tropics, Ouagadougou, Upper Volta

On-farm tests of technology are usually distinguished in the literature as either on-farm trials or farmers' tests. In on-farm trials, the researcher manages the trial in an effort to control variation. Examples include multilocational testing of advanced varieties or tests of new and promising intercropping combinations. In farmers' tests, the farmers manage all (or most) test operations. Even management may be a test factor, with the researcher simply monitoring how the test is executed by the farmer.

Between these extremes, the researcher and the farmer are co-managers. How much of the testing should be managed by the researcher and how much by the farmer depends on what is already known about the technology to be tested, what one wishes to examine, what control is required on the levels of treatments, and how precise the data must be.

At ICRISAT (International Crops Research Institute for the Semi-Arid Tropics), I have distinguished between six levels of tests to reflect variation in the inputs, degrees of management and risks absorbed by the farmer, as well as the possible analyses and types of conclusions that can be drawn (Table 1). In levels 1 and 2, all management is provided by the researcher, and land and labour are rented from participating farmers. The value of such trials is to verify agronomic performance of technologies in a wider range of soils and rainfall conditions than are present on the research station and (in the case of level-2 tests) to get early feedback from farmers on the appropriateness of test factors. Level-3 tests, in which researchers introduce and control certain treatments but farmers manage all other operations on the fields and keep the yields, are designed to obtain precise information about response to treatments under farmers' conditions. This approach is appropriate if management (planting date and density, thinning, intensity of first weeding, etc.) is likely to affect treatment response and if it would be difficult to simulate farmers' management. It is preferable to tests that are totally managed by farmers (levels 4, 5, and 6) if exact precision is needed for treatment doses.

In levels 4 and 5, all test inputs are provided to farmers, and recommended practices are explained, but all farm operations, including treatment applications, are done by the farmers. The farmers choose the plots and are free to modify recommended practices within the designated plots. All modifications are recorded so that researchers can identify reasons for change and quantify their effects on performance. The objective of this approach is to duplicate as closely as possible the conditions faced by farmers

568

Table 1. Levels of farmer participation in on-farm tests of technology.

		Farmers' participation						
		Provision of inputs		Management				
Level	Description	Test factors	Nontest factors	Test factors	Nontest factors	Evaluation	Risk	Scale of observation, analysis
1	On-farm trial	None	Land, labour — fully reimbursed	None	None	None	None	Plot
2	On-farm trial with evaluation by farmer panel	None	Land, labour — fully reimbursed	None	None	Subjective commentary	None	Plot
3	Test of technology exogenously introduced into farming system	None	All — not reimbursed (guarantee possible)	None	All	Objective results, subjective commentary	Limited	Plot
4	Farmers' test	Control-treatment inputs only	All — not reimbursed (guarantee possible)	All	All	Objective results, subjective commentary	Limited	Plot
5	Farmers' test in context of baseline study	Control-treatment inputs only	All — not reimbursed (guarantee possible)	All	All	Objective results, subjective commentary	Limited	Whole-farm
6	Adoption and impact study as follow-up to farmers' tests	All — not reimbursed	All — not reimbursed	All	All	Objective results, subjective commentary	All	Whole-farm

who have just adopted a technology. Baseline surveys of all farming activities are an integral part of level-5 tests so that researchers can examine the effects of the technology at the household level, employing analytical techniques such as complete farm budgeting and optimization modeling.

Level-6 tests closely relate to adoption and impact. All inputs are purchased by farmers, although researchers may find it necessary to make the inputs more readily accessible than under normal conditions of poor transport, inadequate extension, etc. The aim in level-6 tests is to identify in what ways farmers actually incorporate the new technology into their farming systems — e.g., on what soil types, substituting for what enterprises, what level of management is provided to the technology, and what performance is achieved. Results from this stage provide the most realistic base from which to predict performance, adoption patterns, and consequences. Final conclusions, even regarding the agronomic performance of a new technology, will probably take several years — much longer than at other test levels — because the sample group is likely to be small initially and because it often takes years for farmers to switch from experimental use of new technology to full production.

The ICRISAT West Africa program of on-farm testing

Beginning in 1981, ICRISAT initiated a set of long-term studies in six villages of Upper Volta. The six villages represent three distinct agroclimatic zones, with two representative villages located in each zone. A stratified random sample of farmers was selected, with strata defined by the ownership or nonownership of animal-powered equipment for cultivation. The objective of the sampling procedure was to support comparative analyses of both cultivation systems. Similar studies were initiated in four villages in Niger in 1982. The studies involve an intensive monitoring of the production, marketing, and consumption by about 250 farm units, with 25–30 farmers participating in each village.

Following the first year's baseline study in Upper Volta and during the first year of studies in two of the Niger sites, the on-farm trials (researcher-managed) and farmers' tests (farmer-managed) began. Coordinated by the economics program, these tests involve ICRISAT scientists in agronomy, sorghum improvement, and millet improvement.

The long-term program of on-farm testing provides, first, for a limited number of researcher-managed trials (levels 1 and 2 in Table 1) in study villages believed to represent the zones in which the technology could be adopted. The objective of this phase is primarily to verify regional adaptation and to solicit comments from farmers in each village. If results of the on-farm trials warrant, the technology is advanced to farmer testing (levels 3–6 in Table 1) to confirm performance under farmers' conditions and fit within local production systems.

ICRISAT farmers' tests at levels 3 and 5 last at least 1 year. Level-6 testing begins as early as the second year and involves continual monitoring of how participants incorporate new technologies into their farming systems.

Baseline studies complement the farmers' tests and involve all the participant farmers: they provide data on all production activities — a base into which test results from single enterprises can be placed for whole-farm analyses. But also, by marginally disturbing local systems with new technical alternatives, one should be better able to understand objectives and constraints in the system and, consequently, the direction and rates of possible change.

An enumerator living in each village is responsible for following 25 farmers. Farmers are interviewed weekly, and the test plots are observed as needed. In addition, a technician living in each zone is responsible for conducting researcher-managed on-farm trials in two villages as well as assisting enumerators in taking agronomic observations on the farmers' tests.

The principal audience for the results of the on-farm tests is other scientists in ICRISAT technical programs. The tests are designed not only to examine technologies that are in a final stage of development but also to examine the concepts and objectives on which the technologies are based. Results are intended to help scientists appreciate the conditions that technologies must satisfy if they are to be widely adopted. Thus, the tests are not a final, preextension screening but an integral part of technology development.

Evaluation criteria

The questions that ICRISAT staff ask and the methods they use to answer them include:

- *What technical performance can be expected under farmers' conditions?* Yield germination, stand establishment, disease and pest prevalence, tillering, and lodging are some of the indicators of performance. For yield, both the means and the modes are identified as measures of central tendency, and the risks associated with adoption are forecast from the variance and frequency distributions of yields, compared across treatments. Particular emphasis is given to the probability of low yields.
- *What factors in the farmers' environment determine yield variability?* Yield-function analysis is the principal tool employed in attempts to identify the sources of variation in yield. Independent variables include both environmental factors (soil type, slope, rainfall, disease and pest prevalence) and management factors (field history, soil preparation, timing of seeding and weeding, manuring, and plant density). This analysis can lead to an identification of the particular conditions in which a new technology has technical superiority, can help specify needed changes in extension advice, and can aid in the identification of technical problems that require further research.
- *Does the technology require farmers to change the level or timing of their resource use, and, if so, do the changes conflict with their capacity or with their other production activities?* Because all farmers participating in the ICRISAT farmers' tests are also included in the baseline studies, the data on inputs and outputs are comprehensive for all farming activities and provide a picture of the entire production system — the context within which resource-use conflicts can be identified and quantified. At a preliminary stage, ICRISAT staff use activity budgets and, later, programing models, to analyze the data.
- *What returns can be expected from the new technology, and how do*

these compare with those from alternative activities? Inputs and outputs are costed so that the returns from each input can be calculated at both the farm level and the societal level. From the baseline data, one can identify constraints that are in effect at specific times on different types of farm units and compare returns accordingly.

- *Is the technology consistent with farmers' consumption goals?* In the case of improved varieties or hybrids, ease of processing, storage, taste, timing of harvest, and quality and quantity of by-products are important.
- *Will the technology be adopted and what are the likely impacts?* In other words, under what conditions (environmental, technical, and economic) will farmers find the new technology profitable, substituting for what other activities, with what level of management, and at what scale?

Case one: cereal–legume intercrop

Information derived from the baseline studies in Upper Volta had shown t cowpea intercropped with sorghum or millet is the most common crop mixture. Densities for the cowpea intercrop tend to be low, generally between 1000 and 8000 plants/ha, although results of on-station experiments in both Upper Volta and Mali have shown optimal densities to be much higher, about 15 000 plants/ha. Researchers also consider increased cowpea to be a means for maintaining soil quality through soil cover, organic-matter production, and nitrogen fixation

Baseline survey data had also identified sorghum and groundnut mixtures as common in areas of 850 mm or more annual rainfall. These mixtures were characterized by low sorghum densities and relatively high (near-pure stand) groundnut densities.

Against this background, a researcher-managed trial (level 2) was prepared. Its objectives were:

- To measure, in zones of 950- and 750-mm rainfall, the returns to land at low (3000 plants/ha) and high (15 000 plants/ha) densities of cowpea intercropped with sorghum sown at the density found in pure stands;
- To observe how sorghum type, fertilizer treatment, and insecticide use interact and affect intercrop returns;
- To explore the feasibility of increasing sorghum density in sorghum–groundnut mixtures and of introducing the combination in areas where rainfall is less than 800 mm; and
- To solicit farmers' critiques of the trials and their suggestions for alternative means of increasing legume density in cereal-based mixtures.

The trials, designed by ICRISAT agronomy staff and conducted in 1982, were exploratory demonstrations with single replications of each treatment combination. One demonstration was located in each of four villages, representing the 950-mm and 750-mm agroclimatic zones.

Farmers provided land and labour (for which they were reimbursed) and their comments on all aspects of the trial design. All operations were performed under the direction of a field technician.

Results were lost in both villages in the low-rainfall zone because of problems that plague on-farm experiments. In one village, animals damaged both the cowpeas and the groundnuts so heavily that the legume results were no longer valid. In the other village in the same zone, farmers were busy planting their own fields and were not available to be hired to plant the trial on a timely basis.

In the higher-rainfall zone, the results of the trials indicated that net returns to the land increased by an average of greater than 60% as cowpea density was increased (Table 2). Moreover, the response to density was consistently greater for the local variety than for the improved variety, whereas sorghum yields were higher with the latter. The dense canopy of the improved variety reduced the grain response to increased plant stand. Although the grain yield of cowpeas at high densities increased with an insecticide treatment, the value of the increase was insufficient to cover both the annual costs of the insecticide and the pump. That is, the losses caused by insects were less costly than were the available means of control. Finally, highest returns were obtained for the high-density sorghum–groundnut mixture.

Farmers visited the trials frequently to provide their comments. At the end of the season, all farmers participating in the village studies were assembled for a field day that included an extended walk through, and critique of, the trial. Their comments proved to be extremely valuable in interpreting the objective results of the trial and in deriving implications for subsequent research.

Farmers were generally unimpressed with the increasing aggregate production brought about by increased cowpea density. They pointed out that the risk of animal damage was considerably greater at high densities. They also pointed out that labour requirements for weeding would be substantially greater with a high population of the rampant local varieties of cowpea and that the use of animal traction for weeding and ridging would be impossible. Farmers also observed that the substantial reduction of yields for sorghum (in their view, the priority component in this cereal–legume mixture) was unacceptable. In short, they felt that the possibility of higher financial returns from cowpeas grown at high densities did not offset the disadvantages and that the traditional density better met their objectives and was more consistent with their available labour.

Commenting on the sorghum–groundnut mixture, farmers explained that they considered groundnut the priority crop in the system. They noted that competition for light at high densities of sorghum forced the groundnut plants to grow upward, with reduced rooting and nut formation. They also criticized the spatial arrangement of groundnuts as being too close to allow adequate nut filling. In conclusion, they recommended a planting pattern that would increase the proportion of groundnut in the mixture, give greater room for each groundnut plant, and substantially reduce shading from sorghum.

As a result of the input from farmers, together with the returns analysis, the accent in subsequent on-farm trials of intensified cereal–legume

Table 2. Costs an[illegible]turns (francs CFA/ha) from an on-farm trial of intensified cereal – legumes intercrops, Koho village, 1982.[a]

	Cowpea density						Groundnut	
	3000 plants/ha		15000 plants/ha					
			No insecticide		Sprayed with insecticide			
	Local variety	Improved variety[b]	Local variety	Improved variety[b]	Local variety	Improved variety[b]	Local variety	Improved variety[b]
Value of production[c]	21921 (32879)	24460 (57622)	37034 (67481)	32393 (44414)	45774 (56625)	34877 (46346)	53822 (60105)	48305 (73786)
Sorghum	8029 (20091)	11396 (49802)	5846 (18537)	8473 (26566)	5846 (18537)	8473 (26566)	4736 (14097)	7400 (27454)
Legume	13892 (12788)	13064 (7820)	31188 (48944)	23920 (17848)	39928 (38088)	26404 (19780)	49086 (46008)	40905 (46332)
Variable cost[d]	542 (11042)	542 (11042)	1057 (11557)	1057 (11557)	1657 (12157)	1657 (12157)	1342 (11842)	1342 (11842)
Sorghum seed	420	420	420	420	420	420	420	420
Legume seed	122	122	637	637	637	637	922	922
Insecticide	—	—	—	—	600	600	—	—
(NPK, 100 kg/ha 14 : 23 : 15)	— (6500)	— (6500)	— (6500)	— (6500)	— (6500)	— (6500)	— (6500)	— (6500)
(Urea, 50 kg/ha)	— (4000)	— (4000)	— (4000)	— (4000)	— (4000)	— (4000)	— (4000)	— (4000)
Capital costs[e]	—	—	—	—	800	800	—	—
Sprayer	—	—	—	—	800	800	—	—
Net margins	21379 (21837)	23918 (46580)	35977 (55924)	31336 (32857)	43317 (43668)	32420 (33389)	52480 (48263)	46963 (61944)

[a] Figures and items in parentheses indicate where the fertilized trial differed from the unfertilized trial. All other values and inputs were the same.
[b] SRN4841.
[c] Outputs were valued at mean farm-gate prices for a 3 month postharvest period: sorghum, 37 F CFA/kg; cowpea, 92 F CFA/kg; groundnut, 81 F CFA/kg.
[d] Seed was valued at mean farm-gate prices for May - July 1982: sorghum, 40 F CFA/kg; cowpea, 120 F CFA/kg; groundnut, 75 F CFA/kg. NPK fertilizer was valued at 65 F CFA/kg and urea at 80 F CFA/kg.
[e] Depreciation for spraying equipment: 5-year pump life, 3 ha/pump.

mixtures has been shifted to groundnut-based systems. Planting patterns were modified to reflect the objectives expressed by the farmers, and early maturing varieties of sorghum and millet were sown late in some treatments (an alternative not now available to farmers) in an attempt to increase sorghum densities without adverse effects on the groundnut.

Case two: measuring fertilizer response

Farmers' tests conducted in Upper Volta in 1982 had measured the profitability and risks associated with the recommended dose of NPK (14 : 23 : 15) cotton complex fertilizer when used with both local and improved cereal varieties. The analysis did not answer the question of whether the recommended dose was optimal by financial and economic criteria and whether the risks were the same at levels other than the recommended dose. To answer these questions required data from tests that would allow a comparison of yield responses at different fertilizer levels and the calculation of profit distributions. Moreover, the profitability of urea in combination with cotton complex fertilizer had not yet been tested in Upper Volta under farmers' conditions.

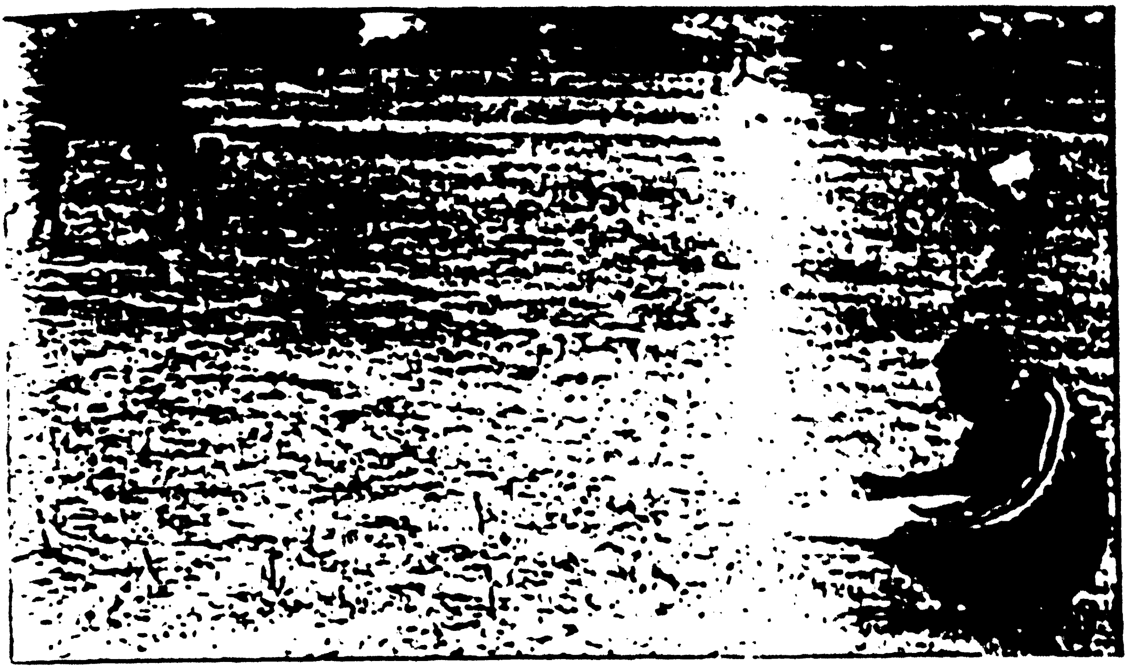
A joint researcher- and farmer-managed trial (level 3) was set up with the objectives to:

- Estimate response functions to cotton complex fertilizer in each of the three agroclimatic zones, and, based on these results, calculate levels that maximize financial and economic profitability in the short term;
- Calculate the probability distribution of gains and losses associated with a range of fertilizer doses applied to local and improved varieties in different regions;
- Measure the profitability of applying urea at a recommended dose and the probability of losses and gains, again by variety and region; and
- Identify and measure the effects of management factors (e.g., soil preparation, fertilizer use) and microenvironmental factors (e.g., soil type) on returns.

The trial was designed to combine researcher and farmer management because the amounts of fertilizer applied had to be precise, whereas, in previous farmers' tests, farmers had modified recommended fertilizer doses in up to 30% of all cases.

A level-3 fertilizer-response trial combined with a level-5 varietal test seemed to be the most workable. Field assistants would intervene to apply fertilizer on plots demarcated within farmers' tests of improved and local cereal varieties, and all other operations were to be performed by the farmers.

Six fertilizer doses were selected. Included was the recommended dose (100 kg/ha) of cotton complex fertilizer with and without urea. The number of treatments/farmer was limited to four so that errors in reporting would not be unacceptably large. All farmers received three treatments (0; 100 kg NPK/ha; 100 kg NPK/ha plus 50 kg urea), and the remaining three treatments were randomly distributed, with each farmer receiving one (50 kg NPK/ha; 200 kg NPK/ha; or 400 kg NPK/ha). Detailed data on operations



ICRISAT technician records labour data for the framework within which the effects of new technologies can be measured

were collected for each of the eight test plots. Yields were measured by field enumerators, harvesting each plot completely. Because the trial is being carried out in 1983, results are not yet available.

Case three: varietal tests

Between 1980 and 1983, the ICRISAT economics program in Upper Volta and Niger tested 14 of the most promising sorghum and millet varieties from each country's crop-improvement programs. The approaches used in the tests (level 5) have evolved and illustrate how a fairly uniform design can, with only minor modifications, address a relatively wide range of issues in technology evaluation.

The major objectives have been:

- To assess new varieties for agronomic performance, fit into local systems, and consumer acceptability;
- To evaluate the economics of agronomic practices and inputs in combination with local and improved varieties; and
- To measure yield losses caused by pests and diseases.

These various objectives can be satisfactorily met with a split-block design, which permits the researcher to examine both the main effects and interactions of varietal and agronomic treatments. Each farmer cultivates a single replication of the four-treatment block, with sites serving as replications for subsequent analysis. Plots employed in farmers' tests should be large enough to provide insight into performance under nontest conditions but not so large as to impose an unreasonable burden or risk on the farmer. In 1981 and 1982, for varietal tests on treatment plots of 250 m² farmer-used levels of labour and nonlabour input that did not significantly differ from their traditional fields. Smaller plots (100 and 150 m²) are being tried in 198[illegible] as a test of whether an increased number of treatments can be satisfactori[illegible] introduced on about the same total area. For farmers' tests of agronomic practices where labour inputs are changed or economies of scale are expected, 250 m² is a minimum. Larger, and perhaps various-sized, plots stratified across sites might be necessary.

Sites are selected by each farmer on soils suitable for the crop being tested. To facilitate farmer recall and staff observations, colour-coded stakes indicate treatment locations, and plot placement is not randomized. Data on labour use and nonlabour inputs are obtained in weekly interviews. Cropping histories for each plot are also obtained. The microenvironment (soil type, slope, etc.) is observed during staking, and the findings are recorded. Agronomic observations (seedling establishment, insect and disease damage, lodging, etc.) are noted at appropriate times in the season. The densities of plants and heads as well as yield are determined at the end of the season by field staff who harvest the entire crop.

Agronomic treatments represented farmers' current practices for the crop being tested (zero tillage, no fertilizer) and the package recommended by the extension service (preplanting plowing and 100 kg NPK, 14 : 23 : 15/ha).

Farmers generally have had little problem in following the recommended treatments for varietal tests in a systematic split-block design with colour-coded inputs and stakes for the plots. However, because the farmers perform all operations and are free to modify the recommended practices, field staff must visit the plots regularly with the farmer to verify the treatments. These visits are particularly crucial during operations early in the season when fields are planted, manure and fertilizers applied, etc. so that information elicited in interviews can be verified and, when necessary, corrected.

A sample of results drawn from several varietal tests demonstrates the types of analyses and conclusions that can be supported by such farmers' tests of varieties.

Agronomic performance and fit

Major criteria employed in evaluating the agronomic performance of new varieties are seedling establishment, mean yields, yield variability, and yield determinants. Tests of the improved white sorghum variety E 35-1 in 1980 and 1981 and the red sorghum Framida in 1982 provide useful examples of the first three criteria.

Results of farmers' tests in 1980 showed that, with low tillage, seedling emergence was significantly ($P < 0.05$) lower for E 35-1 than for local varieties and consequently that soil preparation by animal traction was essential for a full stand of E 35-1. However, the baseline survey had shown that plowing requires nearly 200 person-hours/ha by hand hoe and 60 person-hours by donkey traction. This labour requirement and the need to delay plowing until immediately after a rain would bring E 35-1 into conflict with the timely planting of local varieties.

Confirmation of these results for E 35-1 and for other elite sorghum varieties in subsequent farmers' tests led to the initiation of systematic laboratory screening of promising sorghum varieties for emergence. As for E 35-1, a crossing program was begun to incorporate improved emergence and seedling vigour.

Table 3 Mean yields (kg/ha) of improved and local sorghums by position along the toposequence at two levels of management in level-5 farmers' tests, Nakomtenga and Nabitenga, 1981.

	Low management				High management			
	E 35-1	38-3	CSH5	Local	E 35-1	38-3	CSH5	Local
Plateau								
Mean yield (kg/ha)	—	318	144	189	–	185	813	273
Observations	0	1	1	1	0	1	1	1
Upper slope								
Mean yield (kg/ha)	268	305	773	605	966	1018	1256	1102
Standard deviation	286	395	377	473	568	693	480	551
Observations	8	7	9	12	8	7	9	12
Mid slope								
Mean yield (kg/ha)	685	311	537	626	1405	915	1369	1197
Standard deviation	609	376	374	459	763	362	583	454
Observations	17	16	15	24	17	16	15	24
Lower slope								
Mean yield (kg/ha)	810	516	602	606	1389	1106	1202	1150
Standard deviation	645	655	313	525	1162	799	1033	588
Observations	4	6	4	7	4	6	4	7

Because variability between sites is typically wide, a comparison of mean yields from all sites rarely gives significant results. Alternative approaches that can be used in the absence of computer equipment include t-tests of mean differences with paired observations for each site and the poststratification of sites according to principal site and management characteristics. The advantage of poststratification is that one can examine differences in response to the stratifying factors and thus identify the conditions under which particular varieties are best adapted.

Poststratification analysis (Table 3) of mean yields for two improved sorghum varieties, one hybrid, and a local variety suggested that local varieties and, to a lesser degree, the hybrid CSH 5, were more widely adaptable than E 35-1 but that E 35-1 was best adapted to fields on the lower half of the slope under low-input management and to both mid- and lower slope fields under high management.

Combining poststratification analysis with data on labour use and factor returns (for the test varieties and for all other farm-level activities included in the baseline survey) can elucidate probable adoption patterns and fit within existing systems. For example, in 1980, an analysis of yields across field locations showed that E 35-1 achieved significantly ($P < 0.05$) greater yields only on fields where it received large amounts of organic refuse — that is, fields adjacent to family dwellings. As baseline data showed that these plots are predominantly sown with maize and red sorghum, budgets were calculated, and the returns to both land and labour for E 35-1 were compared with those for the alternative crops sown near the compound. The analysis revealed that, on highly manured soils, E 35-1 was significantly more profitable than local sorghums but not more profitable than maize. Moreover, because maize is harvested 1 month earlier than E 35-1, it serves a critical role in providing calories before the major cereal harvests. This source of food during the hunger period would be foregone if E 35-1 were substituted for maize. Also, technical budgets showed that soil preparation and planting of the shorter-cycle and later-planted E 35-1 conflicted with the

Table 4 Financial budgets (CFA/ha) for E 35-1 and local sorghum under seven management classes[a], level 5 farmers' tests, Nakomenga and Nabitenga, 1981[b]

	Zero tillage		Traction line tracing		Traction plowing		Hand plowing, chemical fertilizer		Traction plowing, chemical fertilizer		Hand plowing, chemical fertilizer, manure		Traction plowing, chemical fertilizer, manure	
	E 35-1	Local	E 35-1	Local	E 35-1	Local	E 35-1	Local	E 35-1	Local	E 35-1	Local	E 35-1	Local
Value of output	44806	39265	28441	[illegible]	70165	57990	52873	71128	58061	69472	143115	81401	96310	95533
Variable costs	1546	1431	1337	[illegible]	1337	1273	5793	5147	5327	5068	20111	18400	18960	18745
Gross margins	43260	37834	27105	[illegible]	68791	56717	47081	65981	52734	64404	123004	62994	77350	66788
Animals and equipment	1068	950	2152	2052	2437	2671	515	515	3220	3329	515	795	3247	3400
Net margin to household, labour, management[b]	42192 (43420)	36884 (30262)	24952 (26659)	26362 (18545)	66354 (52270)	54046 (39684)	46566 (29054)	65446 (33271)	49514 (39178)	61076 (32682)	122489 (67,092)	62200 (41021)	74103 (54981)	65358 (38760)
Total labour (CFA/h)	106 (125)	95 (77)	75 (80)	79 (77)	148 (105)	141 (882)	58 (18)	155 (106)	131 (109)	155 (98)	150 (21)	80 (48)	131 (106)	144 (159)
Production labour (CFA/h)[d]	149 (125)	121 (95)	116 (139)	115 (121)	241 (189)	227 (140)	64 (21)	213 (156)	173 (137)	207 (139)	178 (43)	95 (60)	183 (181)	203 (254)
Marginal rate of return to total costs over lowest-cost management class	—	—	-1970%	-1124%	2015%	1098%	118%	871%	123%	402%	445%	151%	163%	144%
Observations	11	15	10	16	5	7	3	9	9	13	4	8	14	17

[a] Management classes appear in ascending order by cost, with zero tillage being least expensive and traction plowing, chemical fertilizer, and manure being most expensive

[b] In parentheses is standard deviation

[c] Outputs and variable costs were valued at mean farm gate prices

[d] Total labour time, unweighted for age or sex, less labour used in harvest

first weeding of local sorghums. The conflict would be eliminated if E 35 were substituted for local varieties. Thus, the improved variety would probably be adopted primarily on the most fertile soils, but as a replacement for local sorghums rather than maize. Subsequent analysis of adoption patterns has supported the early projection.

Depending on the distribution of yields, over time and across sites, the mean may be an inadequate tool to evaluate yields and to project adoption patterns. Examination of yield distributions can provide valuable additional information on stability across soils and management conditions and on risks associated with adoption. In both 1981 and 1982, for example, the distribution of yields from farmers' tests of local varieties were more peaked and concentrated around the mean, whereas those for improved varieties, which were responsive to management, were substantially more positively skewed. With a positively skewed distribution, adoption patterns projected from the mean alone would likely be unrealistic because the probability of yields below the mean exceeds that for yields greater than the means

Agronomic practices

The early designs of ICRISAT farmers' tests of varieties provided for two discrete management levels, representing local and recommended practices. Because of modifications introduced by farmers (e.g., use or nonuse of manure, tillage equipment, etc.), however, the number of management "packages" were often substantially more. Given a sufficient number of observations, one can analyze these management packages to determine incremental changes in returns with the evolution to more complex and costly systems.

One such budget analysis (Table 4) showed no consistent or significant differences between E 35-1 and the local variety in returns to either land or labour and no trend in differences as one moved from low- to high-cost management. Although the low number of observations and the high variation in data make conclusions somewhat suspect, the local variety appears to be at least as responsive as E 35-1. For example, in several management classes, the local variety responded relatively more to chemical fertilizer than did E 35-1. Also, the rate of return to incremental costs over the base management class (zero tillage and no fertilizer) tended to fall with the adoption of higher cost systems. Nevertheless, the marginal return to total costs in the fully developed system (traction plowing, chemical fertilizer, and manure) remained attractive for both varieties at between 140% and 180%.

Another example of how data from tests of improved varieties can be used to evaluate the economics of agronomic treatments is drawn from farmers' tests conducted in 1982 when rainfall was below average. Data were analyzed to determine the average financial and economic returns to the recommended dose of NPK fertilizer as well as the risk of financial loss by zone, variety, and price conditions. The results (Table 5) showed that, for local sorghum varieties, average financial returns to fertilizer were highest (80%) when applied in the high-rainfall zone, declined systematically (40%) in the intermediate-rainfall zone, and were negative when applied to the dominant cereal, millet, in the lowest-rainfall belt. Returns for improved varieties were consistently higher than those for local varieties and were positive.

Table 5 Returns to 100 kg NPK (14 · 23 · 15) fertilizer/ha with and without subsidy by variety and region, level 5 farmers' tests, 1982

	Average return to cost of fertilizer over 6 months (%)*		Plots where return was less than break even (%)			Minimum cereal yield increments necessary to break even (kg/ha)			Grain prices[c] (CFA/kg)	Number of paired observations
	With subsidy	Without subsidy	With subsidy	Without subsidy	FAO 2:1[b] criterion	With subsidy	Without subsidy	FAO 2:1 criterion[b]		
Djibo										
Souna 3	19	-39	56	72	72	94	184	188	69	18
Local millet	-16	-57	61	72	72					18
SPV35	190	49	46	62	62	100	195	200	65	16
Local sorghum	71	-12	62	77	77					16
Yako										
SRN4841	44	-26	54	69	69	163	318	325	40	13
Local sorghum	42	-27	69	69	69					13
Boromo										
SRN4841	153	30	26	44	44	176	343	352	37	18
Local sorghum	77	-9	44	61	61					18

* Not annualized.

[b] Increment needed to produce a benefit : cost ratio greater than 2 at financial prices with subsidy included

[c] Cereal prices are the average, for 3 months postharvest in each region, fertilizer prices are 65 CFA/kg with subsidy and 127 CFA/kg without subsidy

The results also clearly demonstrated the high risks associated with fertilizer use in semi-arid conditions under farmers' management. Thus, even with mean financial returns of 77% and 42% in the high- and middle-rainfall zones, the percentages of fields where incremental yields did not cover subsidized fertilizer costs were 44 and 70 for the local varieties. Costing fertilizer at its unsubsidized price found average negative returns for all cases except improved sorghum varieties in the high-rainfall zone and under lowland conditions in the lowest-rainfall zone. An important question left unanswered was whether the recommended dose (100 kg/ha) of the available NPK fertilizer was the optimal dose. A farmers' test was subsequently designed to address this question.

Although tabular analyses of yields stratified by management and environmental factors can point toward likely causes of yield variation, yield-function analysis by computer can be a more powerful tool to measure the independent effects of a range of yield determinants. For example, regression analysis of an improved red sorghum variety, Framida, tested by farmers in two agroclimatic zones provided useful information concerning varietal response, fit, and the economics of various management factors (Table 6).

Table 6. Regression coefficients for yield determinants and varietal effects of the improved sorghum variety Framida, level-5 farmers' tests 1982 *

	SRN4841			
	Yako/Ziniare		Boromo	
Improved variety ×				
Alone	1.31	(0.01)	181	(1.05)
Plowing	235	(1.21)	349	(1.35)
Fertilizer	1.64	(0.93)	0.19	(0.09)
Plateau soils	-63	(-0.18)	-270	(-1.12)
Lower slope soils	-110	(-0.13)	107	(0.32)
Lowland soils	-111	(-0.47)		
Management factors				
Plowing — local variety	135	(0.7[illegible])	1[illegible]5	(1.01)
Chemical fertilizer — local variety	1.61	(1.25)	2.93	(2.02)
Plowing × fertilizer interaction				
— local variety	-0.31	(0.16)	-0.03	(-0.21)
Manure	0.04	(2.36)	—	—
Date of planting	5	(1.06)	121	(0.95)
Date of planting squared	-0.02	(-1.50)	-0.16	(-1.06)
Field location				
Village dummy 1	-90	(-0.66)	-76	(-0.61)
Village dummy 2	-151	(-1.39)		
Plateau soils	-132	(-0.46)	130	(0.73)
Lower slope soils	-79	(-0.42)	491	(2.01)
Lowland soils	91	(0.43)		
Field history				
Sorghum preceding crop	-64	(-0.66)	-169	(-1.08)
Legume preceding crop			-105	(-0.33)
Fertilizer applied preceding year	17	(0.24)	121	(0.76)
Constant	1039		-21587	
R^2	0.33		0.37	
F	2.98		3.21	
Degrees of freedom.	117		88	

* t-statistics are included in parentheses.

In brief, the analysis suggested that, under conditions of low management, Framida yields were essentially identical to those for local varieties in the low-rainfall zone but probably superior in the high-rainfall zone. Yield response to plowing for the improved variety was significantly greater than the locals. The results also showed that the improved variety was less well-adapted to shallow plateau soils than were local varieties but probably superior on mid-slope fields (the *reference* soil type) on the toposequence. Combining the technical coefficients on fertilizer response with input and price data, the analysis also suggested that, at recommended doses, the NPK fertilizer was financially profitable when applied to local varieties only in the high-rainfall zone. In contrast, application of fertilizer was profitable for the improved variety in both zones.

Pests and diseases

Although an accurate assessment of the potential gains to investment in research on crop protection requires detailed estimates of the yields that would be lost without protective measures, such estimates are rarely available under farmers' conditions. With adequate resources for numerous observations on disease and pest prevalence, farmers' test plots provide an extremely useful medium for such an assessment.

Methods to evaluate the economic cost of factors causing yield losses have been developed in the context of farmers' tests conducted in the ICRISAT Niger program. The procedure used has been to score at appropriate times for the presence of bird damage, *Raghuva*, downy mildew, wild millets *(Chibra)*, *Striga*, and stem borers. The scores are then included as independent variables in regression equations of yield functions. To arrive at the value of foregone output, one multiplies the estimated regression coefficients by the mean values for each factor responsible for losses and in turn by the postharvest price of millet (CFA/kg).

The results of such an analysis for the farmers' tests conducted in 1982 showed clearly that bird damage, *Striga*, and downy mildew were of no economic importance (Table 7). Stem borer had one large and significant ($P < 0.05$) loss value for the local variety but was otherwise insignificant. The outstanding causes of yield loss were *Raghuva* and *Chibra* millets. In all except one case, they resulted in statistically significant yield losses of more than 4900 CFA/ha, representing between 11% and 25% of the gross value of output. Combined, the two reduced output 27–37%.

Farmers' assessments

Farmers were initially overly positive when asked to evaluate production and consumption qualities of materials introduced by researchers. For example, in the 1980 tests of E 35-1, when farmers were asked to compare yields of the new variety with their local, only 70% responded correctly; that is, their responses agreed with the results of the yield-plot results. Moreover, of the farmers who responded incorrectly, 70% erred in favour of the introduced variety. In 1982 tests carried out with a separate sample of farmers, only 54% of farmers answered correctly. And, of those who answered incorrectly, 66% erred in favour of the test variety. Both ratios are significant at the 5% level. Similarly, when one of the varieties tested that year suffered widespread lodging, farmers in a group session were extremely reluctant to admit the deficiency.

Table 7. Values (CFA/ha) of yield-reducing variables, farmers' tests, 1982

	Equation number[c]		
	1	2	3
Birds			
Value[a]	—	704	—
% of revenue[b]	—	1.7	—
Raghuva			
Value	4419*	8451**	5712**
% of revenue	11.9	19.9	14.3
Downy mildew			
Value	2098*	—	768
% of revenue	5.7	—	1.9
Chibra millets			
Value	9260**	6391**	5214**
% of revenue	25.0	15.0	13.1
Striga			
Value	—	715	423
% of revenue	—	1.7	1.1
Stemborers			
Value	16002**	—	3493
% of revenue	43.1	—	8.8
Mean millet yield for equation	301.6	346.3	324.0

[a] Values are the regression coefficients multiplied by the mean value of the variable times the market price of millet, 123 CFA/kg.

[b] Revenue is the mean yield of millet for the equation times the millet price of 123 CFA/kg

[c] Significance values * = 0.10, ** = 0.05.

These experiences advise one to be cautious in giving a great deal of weight to farmers' assessments until they fully understand the experimental nature of the tests and until they feel at ease in criticizing technologies brought to them by the researcher. One is also well advised to combine subjective assessments with objective tests of the same elements whenever possible to identify the presence and direction of such biases.

Case four: farmers' tests of sorghum–cowpea intercrop

As a complement to on-farm trials of intensified cereal–legume intercropping systems conducted in 1982, farmers' tests of sorghum–cowpea systems were simultaneously conducted in identical village locations.

The tests (level 5) were carried out with two objectives:

- To measure the increased labour demands for planting and cultivating cowpea intercropped at high densities with sorghum; and
- To determine how returns to labour varied with changes in cowpea density.

A split-block design was used with two levels of cowpea density (3000 and 15 000 plants/ha) and two levels of fertilizer (0 and 100 kg 14 : 23 : 15 plus 50 kg urea/ha) as the test treatments. Each of the four possible treatment combinations occupied an area of 250 m^2 demarcated by colour-coded stakes. Sorghum density was to be constant at 60 000 plants/ha, and only local varieties of both sorghum and cowpea were sown.

Data were collected from farmers in weekly interviews and verified through frequent observations. Sorghum yields were measured through two systematically placed 10 m^2 plots in each treatment, and for cowpea through the complete harvest of each 250-m^2 plot.

In marked contrast to experiences with varietal tests, farmers generally did not respect recommendations concerning the major test factor, cowpea density. Densities varied widely, often irrespective of plot designation. And in a portion of cases, the sorghum and cowpea seeds were sown together in the same hill, as per local practice. Reasons for these departures from recommendations were not satisfactorily determined, although many farmers had difficulty understanding and remembering the guidelines. Moreover, many farmers did not appear to view changes in cowpea density as a discrete "new technology" needing to be tested and saw no point in planting cowpea and sorghum in separate hills — a practice that requires additional labour.

Not anticipating such a wide variability in cowpea densities nor a substantial loss in cowpea plants during the season (as conventionally occurs on farmers' fields), field staff observed the plant stands only once at the time of harvest. This error in design led to later analytical problems relating cowpea density to net aggregate returns.

Because of farmers' modifications in execution, the test results could not be analyzed on the basis of two discrete density levels. Rather, the variation in density required a poststratification of plots into three ranges of cowpea density (2000 – 4999; 5000 – 10 999; and 11 000 plants/ha) for labour-use analyses.

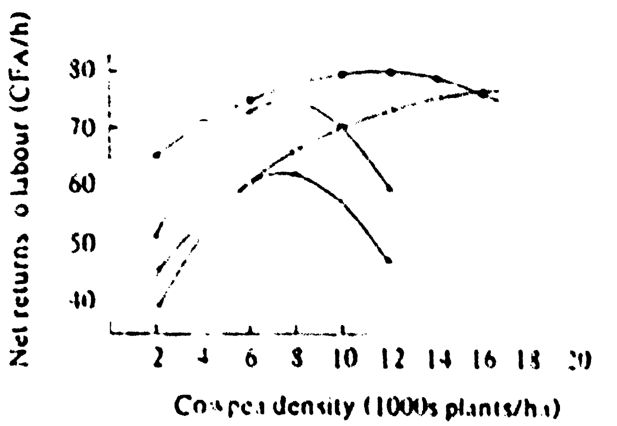
The labour data confirmed farmers' comments during the critique of the on-farm trials. The change from planting sorghum – cowpea together to seeding them on separate hills increased planting labour by at least 20%. As the cowpea density increased to between 5000 and 10 999 plants/ha, labour time for cowpea planting alone increased by an additional 50% over all sites. The additional care required to weed high-density cowpea also resulted in an increase (25 – 50%) in total labour use for first and second weedings in the different village sites. Finally, the data showed that the frequency of ridging also declined directly with higher cowpea densities, suggesting any advantages from ridging would be foregone in a high proportion of fields if high-density cowpea systems were to be introduced.

When data were pooled across sites, labour for the peak period of planting and of weeding increased by more than 40% overall for a shift from sole sorghum to an intercrop with cowpea at a moderately high density (5000 – 10 999 plants/ha). At an opportunity cost of roughly 35 CFA/h during June – July, this represents an additional labour cost of nearly 5000 CFA/ha.

Thanks to the wide variability in cowpea densities introduced by farmers, the independent functional relationship between cowpea density and factor returns could be estimated by regression analysis. A profit function was fit with factor returns as the dependent variable and a range of profit determinants (including cowpea density and cowpea density squared) as independent variables.

The partial relationship linking returns to labour and cowpea plant stand (Fig. 1) indicated important differences between agroclimatic zones with respect to the optimal cowpea density. Interestingly enough, the optimal ranges were relatively stable with or without fertilizer.

The principal conclusion drawn from this farmers' test is that, under farmers' conditions, optimal densities of spreading local cowpea varieties intercropped with sorghum are probably substantially lower than suggested from trials on the experimental station and not greatly different from farmers' current practices. A promising direction for additional research, suggested by the farmers' test results, is the possibility of higher density cowpea intercropped with sorghum using upright cowpea varieties sown in the same pocket with sorghum. This approach would eliminate the additional labour demands for planting and weeding that were present in the system tested.



	Sorghum density (1000s plants/ha)			
	Boromo (800 mm)		Yako/Ziniare (600 mm)	
Fertilized	70	—•—	52	·—·—
Unfertilized	56	—·—	46	——

Fig. 1. Changes in cowpea density in farmers' tests clearly affected the returns to labour in a sorghum–cowpea mixture in two rainfall zones, 1982.

Case five: follow-up — patterns and consequences of adopting a new variety

Because all ICRISAT sample farmers participate simultaneously in tests of technology and baseline studies, farmers are automatically followed up in an effort to determine to what extent they adopt elements of the test technologies. Because of possible biases in level-5 tests, this subsequent stage in the farmers' tests is believed to give the most accurate information on adoption potential and impact. As such, results drawn from follow-up studies (level 6) serve to verify provisional projections made on the basis of level-5 test results. Follow-up of farmers who had participated in 1980 tests of sorghum E 35-1 as a possible substitute for local varieties or for maize is a good example.

Activities on all cereal fields cultivated in 1981 by 44 participating households were followed through weekly interviews. Farmers were asked to estimate, from recall, yields and applications of inputs such as seed and manure. They used local units for quantities and these were converted to metric weights, later, from samples. All fields were measured by compass and chain.

The major problem in implementation derived from the high yield variance caused by differences in environmental and management factors. Lack of computer facilities in 1981 meant that some types of analyses were not performed. In particular, the independent effect of variety on yields and returns could not be determined by means of regression models.

Distributing seed for sorghum trials under farmers' conditions.

The results nevertheless indicated that the adoption and management of E 35-1 corresponded remarkably closely to projections from level-5 tests. As had been predicted, early adoption was more common among farmers who had animal traction and, thus, added capacity to prepare the soil and added access to manure.

All nonadopters had experienced significantly lower yields than adopters for both E 35-1 and the local sorghum check during the previous year's tests. This finding reflected a greater propensity for early adoption among efficient farmers. Moreover, the difference of E 35-1 yields in 1980 less yields of the local check was positively (but weakly) correlated (r = 0.26) with the area of E 35-1 sown in 1981. Although farmers were clearly influenced by the relative performance of each variety, other factors were more important so that plantings continued to follow an exploratory, experimental mode.

The farmers' evaluations of E 35-1 after the 1980 harvests were poor predictors of early adoption. Although the percentage of low scores given E 35-1 on a wide range of performance and consumption criteria was generally higher among nonadopters than adopters, in no case was the difference significant. This result has been confirmed in subsequent seasons in other locations: namely, that farmer evaluations obtained in interviews tend to be positively biased toward the test materials, and, as such, result in poor projections of subsequent behaviour.

Nevertheless, the 1981 data on cropping patterns and field management showed that farmers had correctly assessed the management requirements of E 35-1. Thus, they tended to concentrate fields for E 35-1 close to their dwellings for ease in management and manuring. As a result, the E 35-1

Cowpea experiments in Upper Volta.

fields received 4 times the amount of manure and 10 times the amount of plowing labour devoted to the average local white sorghum field. Also reflecting farmers' recognition of the responsiveness of E 35-1 to fertilizer, E 35-1 was sown more often on plots previously in fallow or sown to legumes than was the local.

In level-6 tests, where farmers provide all inputs and modify recommended practices to fit their resources, multivariate analysis is essential to reduce unexplained variance in nontest factors (such as soil quality, timing and intensity of operations, etc.) and to isolate the independent effects of response parameters. Although regression techniques are the most powerful tools for this purpose, lacking computer capacity, one can learn much from budget analyses that poststratify cases by environmental or management variables.

For example, poststratification of results in the 1981 follow-up studies provided a good means to evaluate the financial performance of E 35-1 compared with local varieties. The 63 sorghum fields cultivated by participant farmers were poststratified according to method of soil preparation, fertilizer application (with or without), and variety. Further poststratification according to level of fertilizer applied or field type was not possible because of insufficient observations.

Poststratified test data support several other types of analyses that provide useful insights into possible patterns and consequences of adoption. For example, further analysis of the poststratified data from the 1981 follow-up showed that the highest-cost management package together with the local variety should be the preferred treatment and that the adoption o higher cost management was generally associated with increasing returns to both land and labour. Thus light animal traction plus fertilizer may be appropriate for both land- and labour-scarce households. Moreover, for E 35-1, the rate of increase in returns to labour was in fact somewhat greater than for returns to land, suggesting that the technical packages compared were probably somewhat labour, rather than land, biased.

Concluding observations

Three of the major problems posed by on-farm tests of technology are high variance, bias, and insufficient field staff who are adequately trained and supervised. There are a number of approaches to reduce their impact.

High variance

The principal sources of variability in on-farm tests are environmental differences between and within sites and the differences in management by participants. Rather than masking intersite soil variability through uniform basal doses of fertilizer as in on-station trials, on-farm tests have as one of their objectives explaining performance variability as a function of environment. This can be done if one characterizes the microenvironment and incorporates such site characteristics in yield and returns analysis.

The method normally used to reduce the effects of within-site variance in researcher-managed trials on farms is increasing the treatment replications, whereas this approach is too complex for farmer-managed tests, the sites themselves often serving as replications. Thus, a more workable approach is to include large plots and to harvest treatment plots completely rather than to use yield samples to estimate production. As is the case for different sites, soils for individual treatment plots need to be characterized and included as performance determinants in subsequent analyses.

Although farmer modifications in recommended practices constitute an essential element in farmer-managed tests, they generally increase substantially the variability between sites. Consequently, the quality and the timing of all key operations on the farms need to be identified through interviews and frequent observations at the sites.

As farmer participation in tests increases, analytical methods based on traditional experimental designs become increasingly less appropriate and are replaced by methods developed for the analysis of data from cross-sectional surveys. Multivariate approaches that identify the direct effects as well as interactions of environment and management become essential. Depending on the availability of computing equipment, these approaches can vary from simple tests of mean differences with poststratification of cases to complex multiple-regression analysis. The number of observations (sites) to support these types of analysis must be large to preserve adequate degrees of freedom.

Bias

At least three types of bias, often present in on-farm tests, can seriously jeopardize the validity of the results: biased behaviour in the management of farmers' tests, biased reporting by farmers of operations performed, and biased subjective assessments of new technologies.

The first source of bias occurs when production objectives differ between farmers' test plots and farmers' traditional fields. If, for example, farmers believe that special status is to be gained through high yields on the test plots, additional inputs and management attention may be provided that would not be replicated if the technologies were adopted. If, in contrast, farmers consider the tests not as their own fields but rather as additional work imposed on them by "outsiders," the opposite bias would occur.

The misreporting of activities performed and biased subjective assessments derive from farmers' misconceptions of researchers' objectives and, consequently, from their desire to respond to questions in a way that they believe will please the researchers. Thus, despite being assured that modifications in recommended treatments are perfectly acceptable, farmers are often reluctant to report such changes.

Bias in farmers' subjective assessments of technologies usually stems from exposure to "development" interventions brought by outsiders. Most farmers initially fail to understand the experimental nature of on-farm tests and that they can actively critique technologies without offense to researchers and without jeopardizing their continued participation.

For each type of bias, the problem for the on-farm researcher is, first, to identify the presence, direction, and magnitude of biases, and, second, to reduce their effects. Identifying the biases requires close objective verification of all key on-farm test data. For example, to identify biases in behaviour requires systematically comparing test-plot management with management in other fields; to identify biases in the reporting of work performed requires frequent on-site verification; and to identify biases in farmers' subjective assessments requires the use of checks through which subjective assessments can be compared with objective measures of identical elements.

Over time, these biases tend to disappear as farmers understand more clearly the purposes of the on-farm tests and as they perceive these tests more as their own. Thus, researchers need to be patient as well as cautious in interpreting early results. Also, they should regularly explain the nature of their work and interact with farmers in a way that encourages open and frank dialogue.

Staffing and supervision

Most types of on-farm research pose substantially greater problems in staffing and supervision than are encountered in on-station research. Whereas researchers can daily direct and correct the work of staff at the research station, field staff assigned to villages must often work independently and be able to take appropriate decisions without consulting researchers. In addition to taking technical observations, village staff must be skilled in developing and maintaining both social and professional rapport with farmers. Finally, such staff must be willing to live for prolonged periods under village conditions.

For all of these reasons, field staff must be recruited carefully and trained well. Their responsibilities must be precisely defined and their workloads sufficiently flexible to allow for changing seasonal requirements and unexpected problems. At ICRISAT, for example, a ratio of about 25 farmers/field agent is nearly maximum if observations of farmers' tests an collection of baseline data are to be done weekly. And an incentive system that reflects differences in living and working conditions between field and station-based staff is necessary to maintain morale and motivation.

Perhaps most essential in maintaining accuracy and efficiency in a program of on-farm testing, however, is that the researchers themselves frequently visit and stay in the villages. There is no substitute for personal input in following the seasonal evolution of the tests, in verifying observations and data registation, and in discussing with farmers and field staff their problems and impressions. On-farm testing programs cannot be directed from a distance. Rather, the researchers' close, frequent, and personal contact is absolutely necessary to ensure accurate data and valid interpretation and to maintain the commitment of field staff and, most importantly, of the farmers.