Government of Guyana
Ministry of Agriculture
European Commission

Guyana Rice Project Management Unit

AGRIFOR Consult

On-Farm Research Manual

Bert Meertens – Research & Extension Management Specialist

FINAL DRAFT

March 2008

Support to the Competitiveness of the Rice Sector of the Caribbean
(REG/7641/000 EDF IX Support to the RPR 006)

Table of content

List of Annexes

List of Figures

List of Tables

List of abbreviations

ACP	African Caribbean and Pacific countries
AMRR	Acceptable Minimum Rate of Return
ANOVA	Analysis of Variance
BRRS	Burma Rice Research Station
CARIFORUM	Caribbean Forum of the ACP States
CIAT	International Center for Tropical Agriculture
CIMMYT	International Maize and Wheat Improvement Center
EU	European Union
FFS	Farmer Field School
GRDB	Guyana Rice Development Board
GRPMU	Guyana Rice Project Management Unit
IITA	International Institute for Tropical Agriculture
IRRI	International Rice Research Institute
LSD	Least Significant Difference
MRR	Marginal Rate of Return
NGO	Non-Governmental Organization
RCBD	Randomized Complete Block Design
REMS	Research and Extension Management Specialist

Preface

One of the activities of the Research and Extension Management Specialist (REMS) in the EU/CARIFORUM Rice Project (Support to the Competitiveness of the Rice Sector of the Caribbean) is to assist in reorienting rice research in Guyana towards more adaptive on-farm research involving extension staff and farmers.
Rice researchers in Guyana have already some experience with the conduct of research in farmers' rice fields. However, still too many research activities are carried out on-station under serious substandard conditions. One questions how valid the results of these on-station research activities will be for the farmers. More on-farm research activities are therefore required in Guyana to test the main selected research themes under farmers' conditions. These on-farm activities should involve small-scale, medium-scale and large-scale rice producers.

The objective of this manual is to supply rice researchers and extensionists in Guyana with some basic practical information on planning, implementing, analyzing and reporting on-farm research activities.[1] All kind of statistical books and manuals are available in bookstores and on the internet. Some of these books and manuals are specifically for on-farm research activities. This manual presents the most common characteristics (planning, design, statistical analysis, economic analysis) of on-farm research in a practical way. For more detailed information and less common types of on-farm research the rice researchers and extensionists can consult the supplied reference list.

Georgetown, 28 March, 2008
Dr. Bert Meertens
REMS Guyana/Suriname
EU/CARIFORUM Rice Project

[1] In 2007 the Guyana Rice Project Management Unit (GRPMU) supported a training programme with manual on statistical analysis and report writing conducted by Dr. Oudho Homenauth. This training programme concentrated on general statistical and report writing guidelines and was less specific and detailed for executing and analyzing on-farm research activities.

1 Introduction

Agricultural research has traditionally been undertaken on research stations where facilities for experimentation are usually available and accessibility to researchers is favorable. It was assumed long time that the best technology in research stations is also the best in farmers' fields. However, especially in the developing countries of the humid tropics there is a high variability between farmers' fields and response to improved crop management is less favorable than that in the research stations. In the case of such technology performance inconsistencies the technology selection process should be done on-farm in comparison to the farmers' existing practice under the farmers' growing conditions (Gomez and Gomez, 1984).

On-farm research is, however, not invented by researchers who wanted to have a confirmation that new, on-station developed technologies would perform equally well in farmers' fields. Since the advent of agriculture farmers have always been engaged in trials on their farms. They have experimented with new cultivation techniques and inputs, explored alternative and diversified markets, and initiated innovative pest and disease management practices (Colley and Myers, 2007).
Apart from being on-farm researchers farmers are also the decision makers and managers of agriculture enterprises. Farmers are the adopters, the adapters, and often the innovators of new farming techniques (Wuest et al.). It would therefore be unwise to not undertake on-farm research and to not involve the farmers in the research process as much as possible.

Next to the initial stimulus for on-farm research to broaden the range of validity of conclusions beyond the narrow confines of a research institute setting it has now been recognized that farmer participation is important and that successful programmes must incorporate farmers' abilities to experiment and innovate (SSC, 1998).
In fact, from the 1980s onwards agricultural research has been increasingly conducted on-farm in collaboration with the farmers. This farmer participatory on-farm research has shown that this type of research leads to more appropriate, site-specific technology, broader and faster adoption, and increased producer ability to adapt and develop farming technologies. Using accepted methods of on-farm testing, farmers can achieve experimental precision comparable to those of intensive university research trials (Veseth et al., 1999).

The move towards participatory on-farm research means that many researchers, such as breeders and agronomists, who have been trained in techniques of on-station research, are now under pressure to move on-farm. It is therefore perhaps not surprising that the design of such trials has, too often, resulted in miniaturized research institute experiments. Conditions for on-farm trials are typically less controlled than those for research institute fields, and this means that more thoughtful designs are needed (SSC, 1998).
Several criticisms have been raised concerning the promotion of on-farm tests for farmers. The first criticism is that on-farm tests are too site and manager specific to be able to generalize the results in order to make recommendations for other farmers, or to share it in peer reviewed journals. Although this may be true, the primary goal of on-farm research is to foster adoption, adaptation, and innovation by farmers, not to further the science of agriculture. A second criticism is that on-farm tests are likely to miss fine points of understanding. In large scale, low budget experiments the control of variables is poor, measurements are few, and conclusions are based on gross effects without understanding what may be the cause of observed effects. This criticism, like the first, is based on a direct comparison of farmer oriented on-farm tests to researcher oriented research. On-farm tests are

not a substitute for more basic, fact finding research. Tests implemented by farmers are bound to be focused on performance, and that is why on-farm testing is appropriate as a technology transfer tool as well as a tool for gaining information on farming practices. Nothing prohibits the use of on-farm tests for more intensive research, however, and there are more and more cases where researchers are making use of on-farm tests as research sites (Wuest et al.).

If conducted in the proper way on-farm research will give high quality results regarding the suitability of the investigated technological innovations under small-, medium-, and large-scale farmers' conditions. The use of appropriate farmer participatory on-farm research methodologies and proper involvement of extensionists have to be used to capture the high potential results from on-farm research. The formation of farmer-research groups would be good for improving farmers' research capacity.

The next chapter will elaborate on the specific characteristics of different types of on-farm research activities and the current status of on-farm research in Guyana. Chapter 3 deals with the planning and implementation of on-farm research. Some practical guidelines will be supplied on the choice of statistical designs and the collection of data next to some basic statistical principles as replication and randomization. Chapter 4 shows how to analyse the results of typical on-farm research. Chapter 5 presents simple economic analysis tools as partial budgets and marginal analysis in order to assess on-farm research activities economically. Chapter 6 is about farmer assessment and chapter 7 about reporting on-farm research.

2 On-Farm Research Background

2.1 On-farm research in general

Farmers have often evaluated a new practice by applying it to a small field and comparing the results with nearby fields, or by splitting a field and applying the new practice on one side and their normal practice on the other. This allows for a local comparison of how a practice "looks." It can be an important first step. The problem comes when you want more than a "look." It is simply not possible to make reliable comparisons of yields and other "quantitative" data without a scientific approach (Veseth et al., 1999).

On-farm research is in fact that “look” in combination with a scientific approach. An on-farm trial aims at testing a technology or a new idea in farmer’s fields, under farmers’ conditions and management, by using farmer’s own practice as control. It should help to develop innovations consistent with farmer’s circumstances, compatible with the actual farming system and corresponding to farmer’s goals and preferences. An on-farm-trial is not identical to a demonstration field, which aims at showing farmers a technology of which researchers and extension agents are sure that it works in the area (RNR-RC Jakar, 2001).

Some basic principles for on-farm research are:

- ✓ *Keep it real:* The experiments must address problems that are important to farmers. The farmers’ practice should be included as one of the treatments in the experiment. The non-experimental variables of an experiment should reflect farmers’ actual practice (CIMMYT, 1988).
- ✓ *Keep it simple:* Limit your experiment to a comparison of two (or maximum three) treatments. The research objective has to be a single researchable question. Many test plots are too complicated because they are looking at too many treatments. On-farm research must be practical for the farmer by using plots that are field-scale with standard farm implements (Sundermeier, 1997).
- ✓ *Go step by step:* Farmers usually do not adopt entire new systems of production; they go step-by-step adapting components of the technology. Therefore the on-farm-trial should not include too many new steps at once. The success of innovation is measured by scale of adoption by farmers (RNR-RC Jakar, 2001).
- ✓ *Remain objective:* The results may not turn out as you hoped or planned. Be prepared to accept and learn from negative results (RNR-RC Jakar, 2001).

During the 1980s CIMMYT identified five stages of an on-farm research program (Figure 1) which are still valid today. The first step is diagnosis. If recommendations are to be oriented to farmers, research should begin with an understanding of farmers’ conditions. This requires some diagnostic work in the field, including observations of farmers’ fields and interviews with farmers. The diagnosis is used to help identify major factors that limit farm productivity and to help specify possible improvements.
The information from the diagnosis is used in planning an experimental research program that includes experiments in farmers’ fields. The on-farm experiments should be planted on the fields of representative farmers. After the first year, the experimental results form an important part of the information used for planning research in subsequent crop cycles. Other

diagnostic work continues during the management of the experimental program as researchers continue to seek information about farmers' conditions and problems which will be useful in planning future experiments (CIMMYT, 1988).

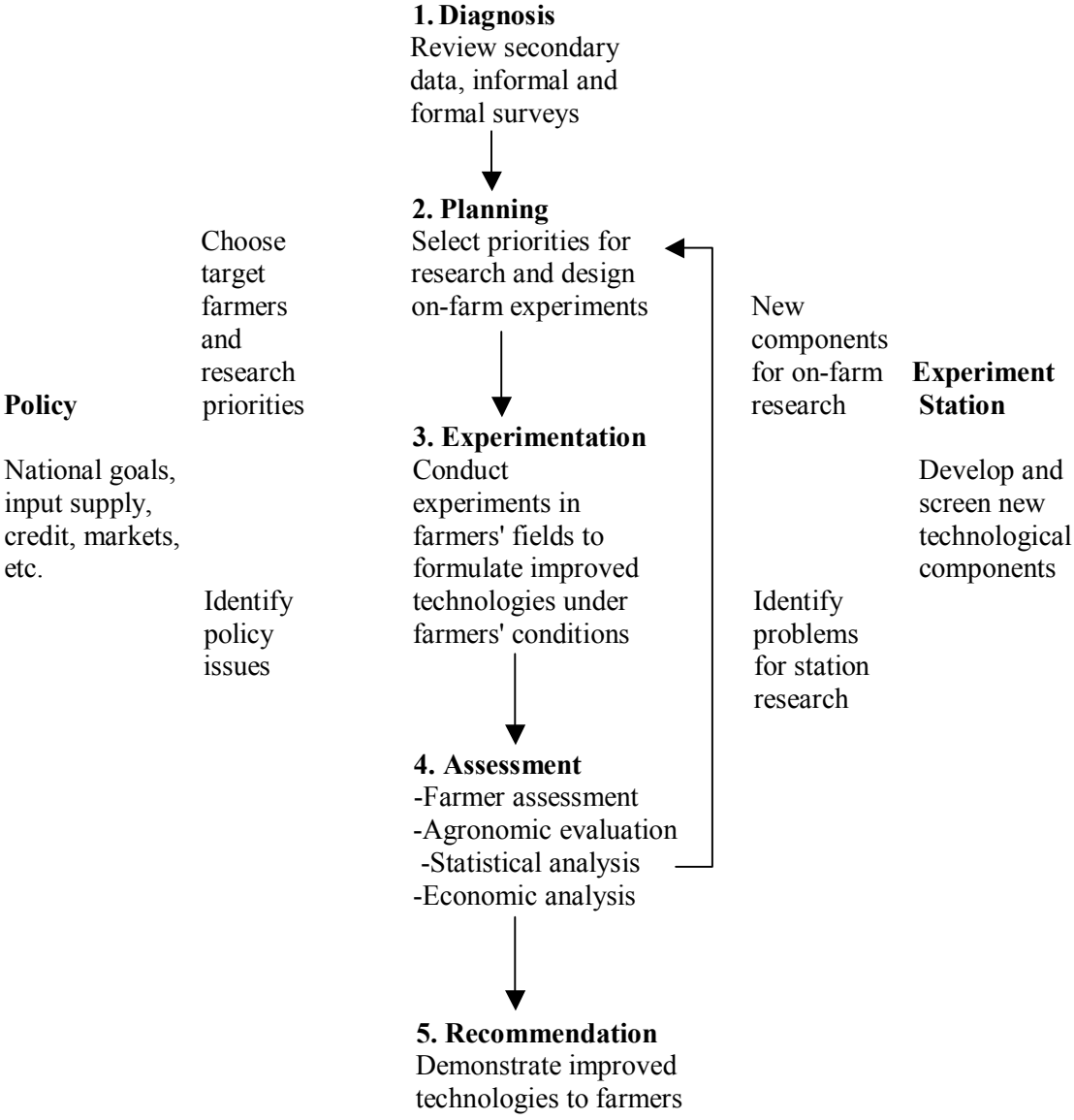


Figure 1 Stages of on-farm research (CIMMYT, 1988)

The results of the on-farm experiments must be assessed. There are several elements in such an assessment. First, researchers must discuss the results with farmers to get their opinions of the treatments they have seen in their fields. The farmers' assessment is very important. The experimental results must also be subjected to both an agronomic evaluation and if possible a statistical analysis. Finally, an economic analysis of the results is essential. The economic analysis helps researchers to look at the results from the farmers' viewpoint, to decide which treatments merit further investigation, and which recommendations can be made to farmers.

The results of an assessment of on-farm experiments can be used for several purposes. First, they may be used to help plan further research. Some experiments will have as their goal the

clarification of production problems: Is production limited by the availability of phosphorus? Will improved weed control give an important increase in yields? The answers to such questions provide researchers with information for further work. As Figure 1 shows, that information can be used to plan subsequent experiments. It also may help orient work on the experiment station.
Second, the results may be used to make recommendations to farmers. Some experiments will compare possible improvements to farmers' current practices: Which level of phosphorus should be applied? Which weed control method gives the best results? The answers to these questions provide information to guide farmers in their management decisions.

Finally, the results of on-farm experiments may sometimes be used to provide information to policymakers regarding current policy toward such matters as input supply or credit regulations. Experimental results can be used to help analyze policy implementation: Given a significant response to phosphorus, is the appropriate fertilizer available? Do local credit programs allow farmers to take advantage of new weed control methods? Although the emphasis in this manual will be on the economic analysis of on-farm experiments for guiding further research and making recommendations to farmers, at several points links between on-farm research and policy implementation will be mentioned (CIMMYT, 1988).

2.2 Types of on-farm research

On-farm research can differ according to the chosen objective, the selected location, the type of design used and the degree of farmer involvement. Table 1 presents three types of on-farm research according to different objectives (biophysical versus socioeconomic), which require different degrees of researcher control and farmer control.

Table 1 Three types of on-farm research

Objective	Trial Type	Trial design	Trial management
Biophysical feasibility	1	Researcher-led	Researcher-led
Profitability, farmer assessment of prototypes	2	Researcher-led	Farmer-led
Acceptability: Farmers' own innovations, assessments	3	Farmer-led	Farmer-led

Source: (Rudebjer, 2001)

In type 1 researcher-designed and managed trials the objective is to assess the performance of the new technology under various biophysical conditions. Research has shown promising results in on station trials. Now the concerned researcher wants to evaluate the new technology in multi location as the on station trial does not represent the wide range of conditions (e.g. soil fertility, weed flora, altitude, rainfall, farmers' conditions). Usually plots are of small size.
In type 2 researcher-designed and farmer-managed trials the objective is to get farmer feedback on specific prototypes and to conduct economic analysis. Type 1 trials have confirmed that the new technology will work in farmers' conditions, therefore it is planned to implement the trial on a wider scale with active involvement of the farmers. Researchers are interested in getting the information on biophysical, economical and farmers assessment of the technology. Plots are often larger than in type 1.
In type 3 farmer-designed and –managed trials, in which each farmer experiments on their own, the objective is to assess farmer innovation and acceptability. Farmers are aware of a given technology, they like what they see and would like to experiment it by themselves or

farmers are aware of a problem and would try some methods to solve them. Researchers want to know to which extent and how a technology is adapted by farmers. Constant monitoring and recording of farmers' comments is necessary (RNR-RC Jakar, 2001).
These trial types are not strictly defined, rather they are points along a continuum. However, you do not need to start with type 1 and proceed to type 3. Farmer participation is furthermore relevant to type 1 trials as well.

Box 1 presents a checklist for evaluating on-farm trials according to their degree of farmer involvement. The degree of farmer involvement does not only depend on the attitude of the researchers and extension workers but also on the research capacity of the farmers. On-farm research can build that capacity but appropriate structures at the community level are also important. For example groups of farmer experimenters who regularly visit each others farms and help in the review of each others work or support activities that can nurture farmer interest in experimentation e.g. cross visits, nurseries and study visits (Rudebjer, 2001).

Box 1 Checklist on how decisions are made in participatory on-farm trials

DESIGN

1. Who decides who the target group is?
2. Who decides what the most important problems are?
3. Who decides which options should be included in the trial as treatments? Who decides what the control(s), if any, should be?

TESTING

1. Who decides which farmers will collaborate in testing?
2. Who decides what design to use, that is, layout of treatments, plot size, no. replications, etc?
3. Who decides at what level to fix non-experimental variables?
4. Who manages the trial, that is, who provides the labour and other inputs?

EVALUATION

1. Do farmers get together and evaluate results during trial? After trial? Is this done only informally or do they actually write a report?
2. Does researcher share farm-specific analyses with each farmer? Does researcher share overall analyses with farmers as a group?
3. Who decides what should be recommended and to whom?

Source: (Rudebjer, 2001)

2.3 Stakeholders of on-farm trials

There are various stakeholders in an on-farm trial:

- ✓ The *farmers* who are the clients for the out-coming results.
- ✓ The *extension agents* who should help the farmers to overcome their problems and improve their economical situation. On-farm trials can give them valuable information in this respect.

- ✓ The *researchers* who need to apply promising on-station results under farmers' conditions before releasing the technology to the extension service.
- ✓ The *policy makers* that are interested in an efficient and participatory technology development, since other more top-down approaches have proven futile.

The different stakeholders have different interests, expectations and positions. Therefore the collaboration is not easy and smooth.

In the past **the farmers** have been treated with top-down approaches. Provision of free inputs was mostly the reason for the adoption of new technologies. The farmers accepted many trials without really having a genuine interest in the new technology. However farmers' participation in the on-farm trials needs to be created not by just free delivery of inputs, but by providing the relevant information that will suite their complex, diverse farming system. It is necessary that extension agents find out the real farmers needs and interests by creating a good dialogue with them.
Unfortunately, farmer communities are often still not informed that there is going to be a trial to test a new technology. The objectives of the trial are often not mentioned to the farmers, and the results are not discussed with the farmers and with the community; no action for next steps is taken. Furthermore it happens also frequently that no prior agreement has been made on compensation measures after crop failure. To improve this, one has to involve the farmers in planning, implementation, monitoring and evaluation of the on-farm research activities. A community meeting should be organized to discuss the objectives of the on-farm trials and select the participants. Incentives to attract the farmers participation should be avoided; too many inputs free of cost falsify the economical aspect of the activity. The results have to be discussed with the farming community and it has to be decided whether compensation has to be accorded in case of failure of the trial (RNR-RC Jakar, 2001).

The extension agents are in a difficult position for concrete decision. They get the order to implement on-farm trials by their superiors, often without the necessary information, training and participation in the planning process. As a result most of the extension agents are not informed on the upcoming trials and have no understanding of the concept of on-farm trials. On the other hand they can not give orders to the farmers on who's land the trial should take place. They must convince them first. They do not receive enough follow up by the research workers and the policy makers during the installation of the trial. At the end, the credit for the results stays with the researchers and/or the policy makers. To solve these problems one has to involve the extension workers right from the beginning in planning, implementation, monitoring and evaluation of the on-farm research activities (RNR-RC Jakar, 2001).

The researchers are interested to take the new technologies in the field for testing the proven results of on-station trials but they usually do not have direct contact with the farmers. They need to depend on the goodwill of the extension agents and the local policy makers. In order to get the agreement of the farmers and the collaboration of the extension agents, they are inclined to give whatever incentive or promise to them. Sometimes researchers do not follow up the trial and pretend to be handicapped to visit the field by transport problems. To make on-farm research a success researchers have to devote sufficient time in the fields and have to try to understand the farmers, their circumstances and interests (RNR-RC Jakar, 2001).

The (local) policy makers have their own strategies and development goals. Increase of production and productivity as well as import substitution are drives to promote new crops and technologies. Often the planned outcome of the programs is more important than the

participatory process with the farmers. Policy makers are often absorbed by non-research activities and can not dedicate enough time to the on-farm trials. At least one has to involve them at the beginning of the on-farm research process. If well-implemented, on-farm trials can meet the interest of all stakeholders. For that matter one has to define clearly the roles and responsibilities of each stakeholder (RNR-RC Jakar, 2001).

2.4 Status of on-farm rice research in Guyana

For the 2008 spring crop season the GRDB researchers at Burma Rice Research Station (BRRS) planned only few on-farm research activities. Apart from on-farm advanced yield trials of the breeding department there is an on-farm comparison of different phosphorus sources from the soil agronomy department. That means that almost all research activities are conducted on-station and take place at BRRS during the 2008 spring crop season. In previous seasons a similar concentration of research at BRRS took place. Clearly more on-farm research activities are required in Guyana to test the main selected research themes under farmers' conditions. These on-farm activities should involve small-scale, medium-scale and large-scale rice producers. On-farm research is normally more expensive than on-station research due to the higher transport costs involved. It also requires more time from the researchers because they have to travel regularly to sometimes distant fields. Consequently, on-farm research in Guyana has to be undertaken only for the priority research topics identified in the research steering committee meetings with proper farmer involvement. Ideally, forthcoming on-farm research themes should be identified during Farmer Field School sessions and previous on-farm activities with adequate farmer involvement.

The lack of on-farm research activities in the current and previous rice growing seasons has as consequence that rice cultivation recommendations to the farmers have to come mainly from on-station research activities. However, due to the unfavourable research field conditions at BRRS (fields not level; variable water stands) regularly on-station trials do not have statistically significant results and can not form the basis for recommendations to the farmers. This lack of valid research results is probably the reason that the so called six strategic points from Venezuela are currently introduced as a recommendation to the farmers without prior on-farm testing in Guyana.[2] The normal procedure would be on-station testing followed by on-farm testing and eventually on-farm demonstrations.

A field survey covering 580 rice producers from all major rice producing areas in Guyana was conducted by GRPMU during October 2007. This survey revealed that about 40% of the producers found that the rice research supplied by GRDB was not good, 43% found it acceptable and 17% found it good.[3]
The farmers advised GRDB to do more research on their fields in order to interact more with them and to pay more attention to their real needs. The small-scale farmers requested GRDB to include them more in their research activities. These remarks from the rice producers show clearly that GRDB can improve their rice research through increasing the number of farmer participatory on-farm trials involving small-scale, medium-scale and large-scale rice producers.

[2] The six strategic points in Venezuelan rice management refer to date of sowing, density of plants, treatment of seeds, weed control, fertilization (40 kg P_2O_5/ha; 60-90 kg K_2O/ha; 120-210 kg N/ha) and water management. This package has been demonstrated to farmers participating in Farmer Field Schools (FFS) from 2007 onwards.
[3] These are preliminary results; final results will be published soon by GRPMU.

3 Planning and Implementing On-Farm Research

3.1 Objectives of on-farm research

All experiments have certain things in common, so designing an on-farm experiment usually includes the following steps:

1. *Establish your goal and objectives*
2. *Determine what treatments you will use and what your check plot or control will be*
3. *Select a site*
4. *Determine how best to lay out your plots on the selected site*
5. *Determine what data you will collect and how you will collect it*
6. *Determine how the data will be evaluated*
7. *Determine how the data will be shared with others*

You must decide what question you want to have answered. This is the goal of the experiment. Goals are statements of the overall theme of your experiment. Objectives are statements of the problem you wish to evaluate in your project. Objectives are measurable and relate to your overall goals. The goal of the experiment will dictate what to include in the experiment to help you answer your question. The individual things that you wish to test in your experiment are called “treatments” and the physical areas to which the treatments are applied are called “plots”. Then you need to decide how the treatments should be physically arranged in the field. Technically, this is what is called the “experimental design”. Experiments answer your original question by allowing you to make unbiased comparisons among the treatments you selected. You will need some way to evaluate how well each treatment worked to make comparisons among treatments. The information you collect to help you make those comparisons (such as yield, insect counts, or disease severity) is called “data”. Finally, you need an objective way to evaluate the data. This is usually done through statistical analysis (Davis et al., 1999).

Trials should be designed to resolve specific research questions, and researchers need to be impartial to the perception that donors expect to see the words "on-farm" and "participatory" before they will consider supplying funds. Usually a careful assessment of the gaps in the current knowledge will show that a series of initiatives is needed. These may include a small survey, plus a number of trials, some on-station and others that are on-farm, possibly some researcher- and some farmer-managed (SSC, 1998).

As with any scientific investigation, it is crucial to specify the objectives of the study clearly. Time must be allowed for this planning phase and the objectives need to be re-assessed during the planning of the trial, to see whether they need to be revised. This is particularly challenging in on-farm trials, where researcher and farmer are now working together, often with extension staff and NGOs. It is important that the objectives are clearly identified from all perspectives (SSC, 1998).

The topics for possible on-farm research should be put forward by the farmers. Information from diagnostic activities will help in transforming the topics mentioned by the farmers into researchable objectives for on-farm research. This forms part of the planning process of on-farm research. Steps in this planning process are (Tripp and Woolley, 1989):

1. *Identify problems* → 2. *Rank problems* → 3. *Identify causes* →
4. *Analyze interrelations among problems and causes* → 5. *Identify solutions*
→ 6. *Evaluate solutions*

In step 6, the list of possible solutions will be narrowed by evaluating each solution according to a number of criteria. These criteria involve (Tripp and Woolley, 1989):

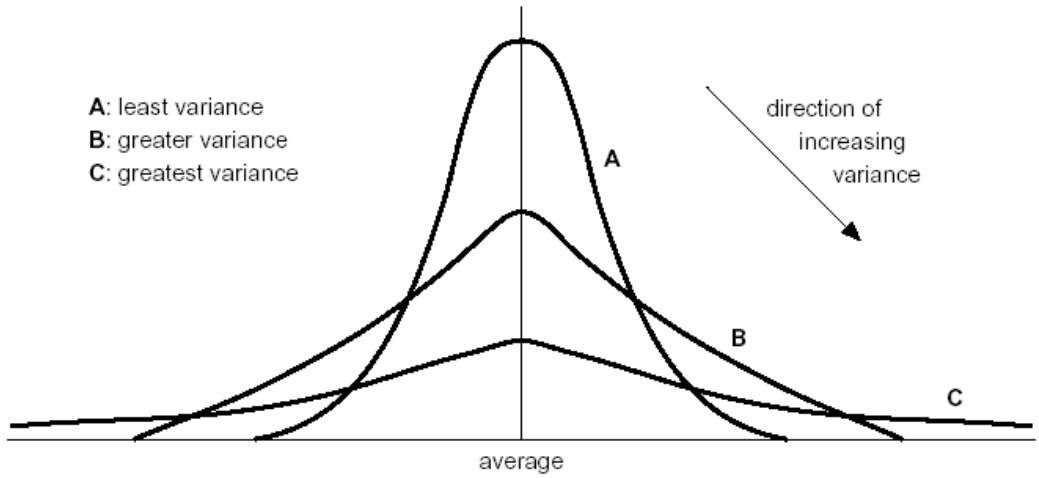
- ✓ *probability that the technology will function*
- ✓ *profitability*
- ✓ *compatibility with the farming system*
- ✓ *contribution to reducing risk*
- ✓ *need for institutional support (extension, provision inputs, credit)*
- ✓ *ease of testing by farmers and*
- ✓ *ease of carrying out the on-farm research activity*

Remember, the question is where it all starts. Narrow your idea or inquiry down to its simplest form — a researchable comparison. For example: ridge-till versus current tillage system, 100% nitrogen rate versus 75% nitrogen rate, or soil-health additive versus no soil-health additive objectives. No matter what your question is, keep it simple, especially at first. The more simple a research project is, the easier it is to conduct (Anderson, 1993).

3.2 Basic statistical principles

The importance of having a statistical analysis of the on-farm research activity decreases from researcher-led type of trials to farmer-led type of trials. However, even on-farm activities which were originally meant as demonstrations can be statistically analyzed in certain cases. This manual will show what type of designs are appropriate in such cases.

Much of the mathematics of statistics is based on what's called the ***normal distribution***.
A frequency distribution that is shaped like a bell (see Figure 2) is called a normal distribution. In many situations, this normal curve describes the frequency of events or characteristics that occur in populations.
A thing to notice about the bell curve is the variation. Variation is due to many factors (genetic, environmental, managerial, and so forth) acting upon individuals or individual plots within a particular population. Some of these factors are obvious and can be accounted for; others cannot. In statistical terms, the measure of this naturally occurring difference is called the ***variance***.
The variance is described by how wide the bell-shaped curve spreads and is a measure of how far from the average the population can go. The variance is a small number when the population is grouped tightly about the average and a large number when the population spreads widely (see Figure 2). When you compare two or more treatments, simple averages are not enough; you'll need to know if the differences are due to the treatments or to natural variation. That's why knowing the variance is so important. If the differences are due to the treatments, they are considered "real" or significant differences. If the differences are due to natural variation, they are not truly differences and are considered "not real" or not significant.



A, B, and C are three populations with the same average, but with different variances.

Figure 2 Normal distribution and variance (Anderson, 1993)

Variety may be the spice of life but research demands rigid standardization. In doing an experiment, you want to control all external sources of variation as best you can. This helps to ensure that observed differences are more likely to have been caused by treatments you applied. It is impossible to have complete control of a research activity, especially one being done outdoors on a working farm, but it is possible to minimize variation in two ways: to establish research plots on relatively uniform ground, and to treat all plots exactly the same except for the treatments you are testing (Sooby, 2001).

Replication, ***randomization***, and ***use of a control*** are essential in designing an experiment because they help to separate out treatment effects from natural levels of background variation. Without these three factors, any data you gather will be just about worthless. In research, ***error*** refers to anything besides your treatment effects that is measured. Setting up the experiment in a structured way helps to reduce the amount of error in your experiment; replication, randomization, and use of a control allow you to use statistics to actually separate out error from treatment effects in your measurements (Sooby, 2001).

3.2.1 Replication

In an experiment, replication or repetition means that individual treatments have been applied to more than one plot. Replication is necessary because all test plots are not identical, and that leads to variation in the data you collect; you will not get exactly the same results from two plots that received the same treatment. You can take steps to minimize the effect of variation if it has an identifiable cause, but there will always be some variation among plots that cannot be controlled. The purpose of replication is to allow you to make a more accurate estimate of how each treatment performed even though there is uncontrolled variation in the experiment. Replication of treatments increases your ability to detect differences in treatment means. Having more replications allows you to identify (statistically) smaller differences in treatment means than you could identify with fewer replications (Davis et al., 1999).

One of the simplest means of increasing precision is increasing the number of replications. Beyond a certain number of replications, however, the improvement in precision is too small to be worth the additional cost. When such a point is reached and the required precision is still

not attained, other means besides increasing the number of replications must be used. Four replications are commonly used in rice field experiments at IRRI. The number of replications you need may differ greatly from this figure (Gomez, 1972).
For field tests in plant pathology, nematology, weed science, soil fertility studies, and entomology, a minimum of four replications is suggested, but five or six replications are much better. If treatment means are close together or variation is relatively large among the plots that received the same treatment, then you may need more replications to detect differences among treatments (Davis et al., 1999).
Three replications are less precise in determining treatment differences than four replications, but may be adequate for some management practice comparisons. The danger of starting with three replications is that if you loose data from one plot, you no longer have an effective trial (Veseth et al., 1999).

There are different ways to replicate an experiment. One way is to have multiple plots at one location. Another way is to replicate the experiment on many farms. Another way is to replicate across time, performing the experiment in multiple years. Each type of replication adds to the confidence you can have in the results (Sooby, 2001).
It is usually preferable to have more farms and fewer repeats of the same treatment per farm, rather than fewer farms and more replication within a farm. Consequently, in on-farm experiments, it is frequently the case that there are many farmers but each farmer has only one replicate of each treatment. The problem with having no within-farm replication, is that the farmer-treatment interaction is then normally used as the random (or residual) variation. However, the treatment effects may really be different for the different farmers and understanding this interaction, e.g. which treatments are most effective for which types of farmers, may be an objective of the research. In such cases one would like to distinguish between the interaction and the residual, and having some within-farm replication is the obvious way to do this. If within-farm replication proves impossible, then it is still possible to carry out some investigation of the farmer-treatment interaction, provided there is information on the characteristics of the farms (SSC, 1998).
This will be shown through an example further on in this manual.

3.2.2 Randomization

Randomization is one of the basic principles of experimental design. It makes valid the estimate of experimental error which is essential for comparing treatments. It is a procedure for allocating treatments so that each experimental plot has the same chance of receiving any treatment. The process of randomization can be done with a table of random numbers or by drawing lots (Gomez, 1972).

Randomization in an experiment means that the treatments are assigned to plots with no discernable pattern to the assignments. The reason randomization is important is that the positioning of treatments within the block (other word for replicate) may affect their performance. One example of this is an experiment testing five corn hybrids (labeled 1 through 5) in which you plant the hybrids in the same order in each block: 1, 2, 3, 4, then 5 (see Figure 3). If hybrid 2 is naturally much taller than the others, it can slightly shade the hybrids planted next to it (hybrids 1 and 3) and unfairly make them look a little bit worse than they would look if they were not planted next to hybrid 2 (Davis et al., 1999).
Apart from the competition for light there is namely also the competition for nutrients (Colley and Myers, 2007).

Another example is a field in which soil fertility gets progressively lower as you cross the field from east to west so that productivity is reduced as you go from one side of the field to the other. If two corn hybrids are planted side by side but within a block, hybrid 1 is always planted on the west side of hybrid 2, then hybrid 1 is always planted in slightly less fertile soil and therefore has an unfair disadvantage (Davis et al., 1999).

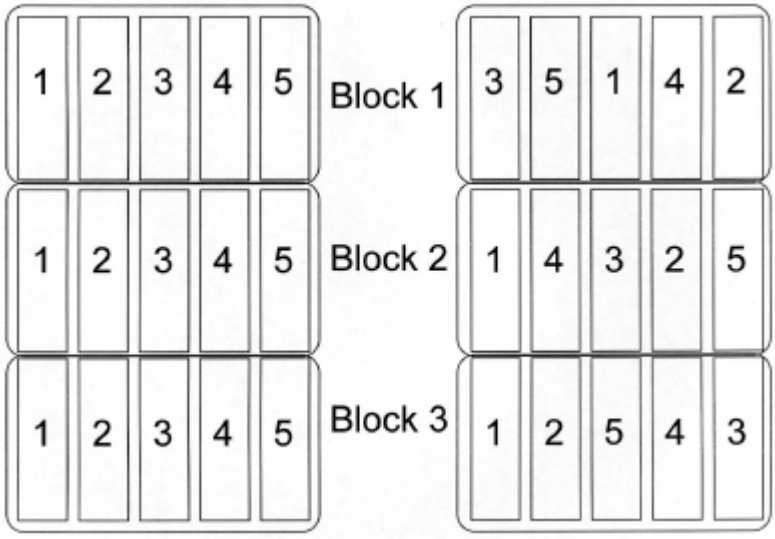


Figure 3 Example of randomization and non-randomization (Davis et al., 1999)
Test plots on the left are not randomized. Plots on the right are randomized. The numbers (1-5) represent the five treatments in this test.

In both of the preceding examples, randomization could have prevented the unintentional bias because the arrangement of the treatments would have been different within each block. Because you cannot anticipate all the influences that may introduce bias into a test, **all experiments should be randomized**. There are many ways to randomize treatments within a block, but the simplest is literally to pull the numbers out of a hat. Assign each treatment a number, write the numbers on individual pieces of paper, mix the slips of paper up, and then select the slips one at a time without looking at them first. The order in which the numbers are drawn is the order in which they will be arranged in a block. Repeat these steps for each block in the experiment (Davis et al., 1999).

If you wish to use the experiment as a demonstration plot (such as for a field day), it is common that one block not be truly randomized. This is done so that particular treatments can be seen side-by-side to facilitate comparisons and highlight differences for casual observers. Though it is better to randomize all blocks and not intentionally arrange treatments, arranging the treatments in one block is unlikely to affect the test's results as long as the other blocks are truly randomized. (Davis et al., 1999).

3.3 Choice of treatments

The objective, or purpose, of the study will determine the treatments included in an experiment. Writing down the test objectives is helpful because this forces you to define them precisely. A test may have more than one objective, although multiple objectives should be closely related.

The selection of treatments is usually logical if you can define the purpose of the study; all treatments necessary to address the test's objective should be included. For example, if the purpose were to determine which of five fungicides works the best, then the treatments would include all five of those fungicides. If the purpose were to determine if any of the five fungicides works better than the current farmers' choice, then the treatments would include the five fungicides plus the fungicide the farmer is currently using. Accurately stating the purpose of the test **before** the treatments are applied in the field is critical. After the treatments have begun, it will be too late to add other treatments to answer the question you really wanted to address.
The selection of treatments and the experimental design get more complicated as the question you are trying to answer gets more complex. It is common to want to test in the same experiment two (or more) things that influence crop production. For example, you may want to test chicken litter as a fertilizer and test five corn hybrids to maximize yield. To do that, the treatment list must include each hybrid without chicken litter and each hybrid with chicken litter; a total of 10 treatments (Davis et al., 1999).

There are no general rules prescribing the number of treatments needed. The minimum may be one, when a single new variety is distributed in an adoption study. However, normally there are at least two treatments being compared. Researcher-led on-farm trials usually have more treatments than farmer-led on-farm trials. When selecting treatments for a farmer-led on-farm trial one can advise farmers to keep them simple and few, no more than 3, including the check plot. As treatments increase in number, so do the number of plots, and the complexity of the on-farm research programme.

Treatments should vary enough to detect differences visually and/or by measuring yield. For example, observing a visual or yield response between 45 and 55 kg of nitrogen per ha may not be possible. The difference between 0 and 45 kg of nitrogen per ha, however, may be more obvious. Some comparisons of treatments may result in little or no difference. When this occurs, it is difficult to conclude which treatment is superior. A difference of 50 or 100 kg of rice per ha between treatments does not represent a true yield advantage (Havlin et al., 1990).

Always include an appropriate check or control plot. Having control plots where the treatment is not applied gives you a basis for comparison. If you are researching a new practice or variety, the control would be a plot that receives the normal practice or variety. For example, in variety trials the control is usually the traditional, established variety whose performance is known. If you are researching the effects of a particular input, the control would be a plot that isn't treated with that input. In designing an experiment, it is essential to include a control so that the effects of your treatment can be measured against something not receiving that treatment (Sooby, 2001).

The control treatments must be justified as treatments in their own right. In on-farm trials the control is often the farmer's normal practice. Since this is likely to be different for each farmer, it cannot be regarded in the usual sense of "control", i.e. as a baseline treatment for the experiment as a whole, against which other treatments are compared. The farmer's normal practice will be useful as a baseline for each farmer, but the researcher may also wish to have a common baseline in addition (SSC, 1998).

There are some things you should consider before you apply your treatments. Your experiment is designed to compare the effects that different treatments have on a particular

measure such as yield. Therefore, you want to be sure that the differences measured are due to the treatments, not some other factor. Apart from the intentional variation created by the treatments, the plots should be as identical as possible. Uniform plots will minimize natural variation and isolate any measured differences so they can be attributed directly to the treatments. Some practices that will keep your plots uniform are to drive over all the plots the same number of times; till all the plots in the same way; and control weeds in the same way in all plots. *What* you do is important; so is *how* you do it. As much as possible, standardize the techniques by which the field work is done: For example, if different rates of a treatment are being tested, set the equipment to apply the first rate and do all the reps of that rate, then change the setting and do all the reps of the second rate, and so on (Anderson, 1993).

3.4 Site selection

The development of recommendations for farmers must be as efficient as possible. The conditions under which farmers live and work are diverse in almost every respect imaginable. Farmers have different amounts and kinds of land, different levels of wealth, different attitudes toward risk, different access to labor, different marketing opportunities, and so on. Many of these differences can influence farmers' responses to recommendations. But it is impossible to make a separate recommendation for each farmer.
As a practical matter, researchers must compromise by identifying groups of farmers who have similar circumstances and for whom it is likely that the same recommendation will be suitable. Such a group of farmers is called a **recommendation domain**. Recommendation domains may be defined by agro-climatic and/or by socioeconomic circumstances. Recommendation domains are identified, defined, and redefined throughout the process of on-farm research. They may be tentatively described during the first diagnosis. Experimentation adds precision to the definition of domains. The final definition may not be developed until the recommendation is ready to be passed to farmers (CIMMYT, 1988).

Apart from the recommendation domain, for which results are intended, the selection of farms must be closely related to the objectives of the research. The large variation that generally exists between farms means they must be selected with care to ensure that conclusions will apply to the appropriate group of farmers. An initial survey is valuable in identifying how farms may be grouped, for example according to their socio-economic characteristics and environmental conditions. Decisions have then to be made whether research results will be relevant to all groupings or only to a subset of such farming groups. A representative (usually a random) sample of farms is then selected from the relevant group(s) of farms. Enough farms have to be used to have a reasonable estimate of between-farm variability. Stratified sampling may be recommended, to ensure that a wide range of farms is included in the sample. The sample of farmers must be large enough for a valid analysis when split into different groups, for example by soil type, tenants and owners, access to credit or not. Where resources seriously limit the number of farms in the study, the objectives of the study may have to be re-examined. For example, for a new topic, the first year may become a pilot study, from which ideas and objectives are refined for the following year's research. When selecting farms any restriction of the sampling scheme so that only 'good' farmers are included will restrict the recommendation domain in the same way. One justification for using 'good' farmers is that they set an example for their neighbours. This argument is at best weakly relevant because we are concerned with research not demonstration (SSC, 1998).

The most common problem with on-farm trials is lack of recognition that field variation can mask or conceal treatment differences. Take special care to plan and organize the field plot

layout to assure that all treatments have an equal opportunity to perform. Choose a field site with the greatest possible uniformity. Regardless of the size of the plot, a uniform field location is critical. When choosing a site consider previous crop history (fertilizer rates, herbicides, tillage, etc.), drainage, soil texture, soil depth, topography, pest infestations, and bordering influences such as trees, runoff from neighboring fields, lack of fencing from animals, and other factors. Avoid placing trials in runoff areas, near fence lines or in field corners. These areas are often subject to multiple or irregular applications of fertilizer and herbicides.
Consider site access when selecting a plot location. Is the site easily accessible for mid-season treatment applications and data collection? If early or differential harvest is likely (such as with an alternate crop), can you get at the site with harvest equipment without destroying other crops? Will you hold a tour of your site? If so, is there ready access for visitors and their vehicles (Miller et al.)?

3.5 Choice of statistical design

Experimental design is a way to arrange treatments in a field so that error and bias are reduced and the data may be accurately analyzed using statistics. Design and analysis fit together to make a meaningful whole. If an experiment has a poor design, you can't have confidence in what the data are telling you.
Standard experimental formats or designs are usually used in on-farm research. The criteria used to select which design fits which experiment depends on the number of treatments under investigation. If you want to compare two levels of a treatment, you can use a design called a **paired-comparison**. An easy statistical analysis, the t-test, can be performed on the data to detect any significant differences. If you want to add more levels of treatments, you can use a **randomized complete block design**. A **split plot design** allows you to see how different treatments interact. These designs can provide you with more information than the paired-comparison but also require more sophisticated statistical analyses and more space in the field (Sooby, 2001).

3.5.1 The paired-comparison

The **paired-comparison** trial can meet many of the needs of on-farm research. If treatment strips of one or two equipment-widths are run completely across the field, no extra stopping, turning or adjustments are required. Because pairs of these narrow strips lie close to one another, field variability between treatment strips is minimized. The pair of treatments is replicated at least six times in the field to overcome chance field differences. The order of the two practices in each pair is chosen at random, avoiding a source of unintentional bias. Each data pair yields one difference. These differences can be analyzed using a simple, single-sample technique to generate the *LSD* (least significant difference) with which to evaluate a two-treatment trial (Exner and Thompson, 2005).

Since there will be at least two treatments in a paired-comparison trial, position of the treatments should be a consideration. All treatments should have equal opportunities. If the field is uniform, simply arrange treatments to accommodate easy observation of each. If the field slopes in one direction, arrange treatment plots with the slope (Figure 4).
If the field has two different soil types or conditions, arrange the plots at right angles to these conditions (Figure 5). Harvesting the whole area is appropriate only if the different soil conditions exist to the same extent across treatments and replications. If non-uniform areas

exist, treatment observations and yield measurements should be random from the most uniform areas of similar soil condition within each plot (Figure 6).

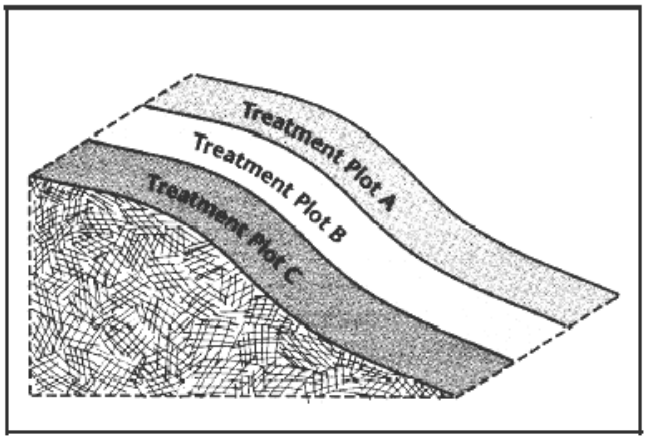


Figure 4 Example of treatments or replications planted with the field slope (Havlin et al., 1990)

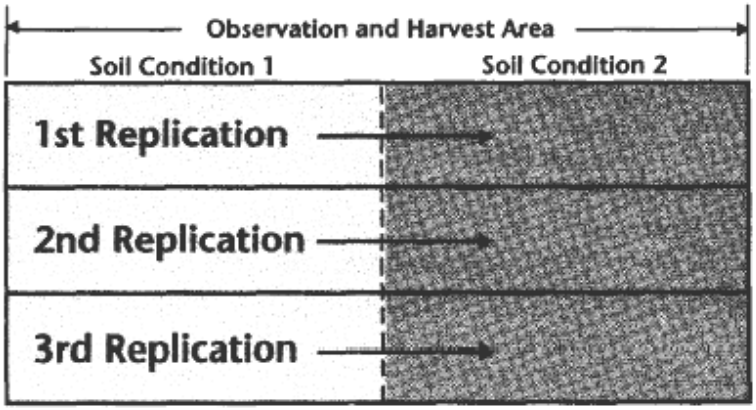


Figure 5 Example of treatments planted perpendicular to direction of two soil conditions *Observations and harvest samples can be taken from both soil conditions provided the two occur across all treatments or replications sufficient to allow observations.* **(Havlin et al., 1990)**

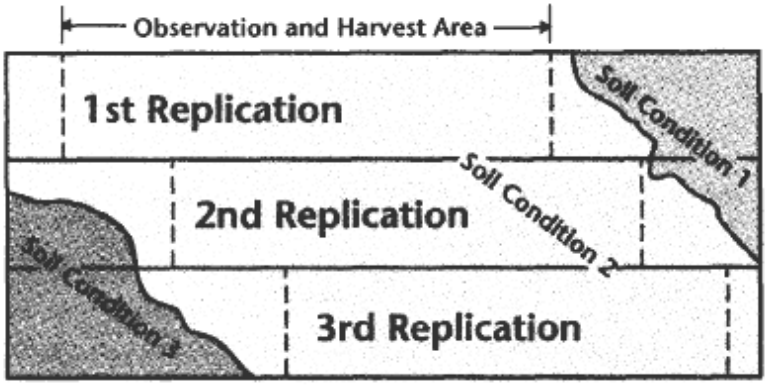


Figure 6 Example of treatments planted across three soil conditions *Observations and harvest samples should be taken from similar areas in the most uniform soil condition. Harvest area is shown between the dotted lines in each replication.* **(Havlin et al., 1990)**

Provide a sketch of the field showing soil variability, tree lines, terraces, topography, plot location within the field, and intended and actual harvest locations (Havlin et al., 1990).

The paired-comparison is an excellent way to assess the effects of separate components on a crop. Growing corn with and without starter fertilizer or mulch (see Figure 7), comparing two varieties, cover cropping compared with fallowing--in a homogeneous field, any pair of treatments can be effectively compared using this design. If there is a large amount of variation in the terrain, some kind of blocked design is required to remove or "block out" the effect of this variation on the measurements (Sooby, 2001).

Mulch	No Mulch	No Mulch	Mulch	Mulch	No Mulch	No Mulch	Mulch	Mulch	No Mulch	No Mulch	Mulch

Figure 7 Paired block comparison measuring the effect of mulch compared with no mulch using 6 replications (Sooby, 2001)

In the so-called single replicate on-farm test, four or more farmers establish a single replicate, that is, one complete set of treatments. A combined paired-comparison analysis over all farmers will then be possible when at least four farmers are involved.
The single replicate on-farm test is useful for developing recommendations about a variety or production practice for a broad production or climate area. The single replicate on-farm tests are at least as powerful as the university's variety evaluation trials at detecting treatment differences. These tests are very popular with farmers and are an important technology transfer tool for variety evaluation in eastern Washington, USA (Wuest et al.).

3.5.2 The randomized complete block design

The completely randomized design is the simplest experimental design. In this design, treatments are replicated but not blocked, which means that the treatments are assigned to plots in a completely random manner (as in the left side of Figure 8). This design is appropriate if the entire test area is homogeneous (uniform in every way that can influence the results). Unfortunately, it is rare that you can ever be confident of a test site's uniformity, so a completely randomized design is rarely used in field tests. The completely randomized design is used more commonly in greenhouse tests, though blocking is often useful even in the more controlled environment of a greenhouse (Davis et al., 1999).

The randomized complete block design (RCBD) is the most commonly used design in agricultural field research. In this design, treatments are both replicated and blocked, which means that plots are arranged into blocks and then treatments are assigned to plots within a block in a random manner (as in the right side of Figure 8). This design is most effective if you can identify the patterns of non-uniformity in a field such as changing soil types, drainage patterns, fertility gradients, direction of insect migration into a field, etc. If you cannot identify the potential sources of variation, you should still use this design for field research but make your blocks as square as possible. This usually will keep plots within a block as uniform as possible even if you cannot predict the variation among plots.

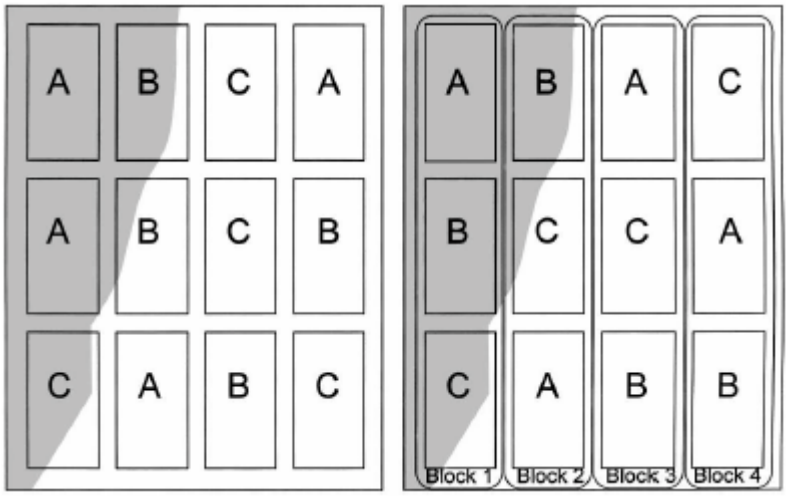


Figure 8 Completely randomized design with and without blocks (Davis et al., 1999)
The shaded area represents an area of the field that is different from the un-shaded area. Treatments (A, B, and C) are replicated but not blocked in the field on the left. On the right, treatments are replicated and blocked; each block contains one plot of each treatment.

The process of blocking follows a logical sequence. First, you determine that there is something (weeds, drainage, sun/shadow, water, soil type, etc.) that is not uniform throughout the experimental area (field, greenhouse, etc.) that may influence whatever you are measuring (yield, plant height, etc.). Then you can arrange your treatments into blocks so that the area within each block is as uniform as possible. Though the area within a block should be relatively uniform, there may be large differences among the blocks, but that is what makes blocking effective. **Your goal is to maximize the differences among blocks while minimizing the differences within a block** (Davis et al., 1999).

Blocking refers to physically grouping treatments together in an experiment to minimize unexplained variation in the data you collect (referred to as experimental error). This allows the statistical analysis to identify treatment differences that would otherwise be obscured by too much unexplained variation in the experiment. The goal in blocking is to allow you to measure the variation among blocks and then remove that variation from the statistical comparison of treatment means (Davis et al., 1999).
In this design, blocks are synonymous with replications because a complete set of treatments is replicated in each block. The statistical test *analysis of variance* is used to analyze the data from an RCBD. Analysis of variance can be calculated by hand but because of the large number of arithmetical steps it is usually done using a computer program. Statistics software is widely available (Sooby, 2001).

The shape of the blocks is not important as long as the plots within a block are as uniform as possible. Ideally, the only differences among plots within a block should be due to the treatments. Blocks in field experiments are usually square or rectangular, but they may be any shape. Blocks in the same experiment do not have to be the same shape; the shape of individual blocks will be determined by variation in the field that you are trying to minimize. If you are not sure what shape your blocks should be, square or nearly square blocks are usually a safe choice. Blocks may be arranged through the field in many ways. If the field is wide enough, an easy way to arrange blocks is to place them side-by-side all the way down

the field. But blocks do not have to be contiguous and may be scattered through the field in any way that is convenient for you (Davis et al., 1999).

Layout of plots within each farm will primarily be guided by perceived or known variation within the farming area. The farmers' knowledge about the variation in their fields should be used to determine the location of the plots and any blocking scheme, and to avoid using particular patches of the field where necessary (SSC, 1998).

Conduct all management operations and data collection "on a per-block basis" to control any variation that may occur in the management and operation processes as well as in the data collection. In other words, whenever a source of variation exists, attempt to have the major portion of the variation separated by blocks.
For instance, when an operation (for example, application of treatments, measurement of data) can not be completed for the whole experiment in 1 day, at least complete the work for all plots in a block. In this way, the difference, if any, from day to day can be controlled by "blocking." The same technique can be used to handle variation among operators. Whenever an operation i**s** known to be affected by the operator (for example, application of fertilizer or insecticide, measurement of plant height), each operator should be assigned to different blocks (Gomez, 1972).

A **factorial arrangement of treatments** is not an experimental design, though you will often hear it referred to as a factorial design or a factorial experiment. A factorial arrangement of treatments means that the experiment is testing two or more factors at the same time, and that the experiment includes all combinations of all factors. The term "factor" is used to describe a group of treatments that have something in common. Fungicides, sources of nitrogen, or rice varieties could be considered factors in an experiment. Factors may be defined broadly or narrowly in different experiments. All herbicides may be grouped as a factor in one experiment, but pre-plant and post-plant herbicides may be treated as separate factors in another experiment. A single-factor experiment tests one factor at a time; a two-factor experiment tests two factors at once.
It is sometimes useful to test two or more factors at once. For example, a two-factor experiment would allow you to compare the yields of five rice varieties at three planting dates. This accomplishes three things at once:

1. It allows you to compare the rice varieties to each other.
2. It allows you to evaluate the effect of planting date.
3. It allows you to determine if varying the planting date changes the relative performance of the varieties (e.g. one variety may only perform well if planted early).

The first two could be done in separate single-factor experiments, but the third can only be achieved by having both factors in a single experiment. This becomes especially important if one factor can have a significant influence on the effect of the other factor. When one factor influences the effect of the other factor, there is said to be a significant interaction between the two factors. It can be very important to know if there is an interaction between factors, because if there is an interaction, you can make predictions or recommendations based on the results of single-factor experiments only when all other factors are at the same levels they were at in the experiment. If you change some factor not included in the experiment, the results from your single-factor experiment may no longer be valid.
With a factorial arrangement of treatments, all values (or levels) of each factor must be paired with all levels of the other factors. If you have two herbicides and five rice varieties, then your treatment list must include each variety with each herbicide for a total of 10 treatments. This

would be referred to as a "two by five factorial" to denote how many factors were present in the experiment and how many levels of each factor were used. The number of treatments increases quickly when you add more levels for a factor, so choose your levels carefully or the experiment can get too large to manage.

A factorial arrangement of treatments can be a very powerful tool, but because the number of treatments can get very large it is best used when some reason exists to believe that the factors may influence each other and have a significant interaction. **If there is no suspicion that the factors may influence each other, it is frequently easier and more thorough to test the factors in separate experiments.** A factorial arrangement of treatments can be used with a completely randomized experimental design or a randomized complete block design. The top half of Figure 9 shows a factorial arrangement of treatments in a randomized complete block design.

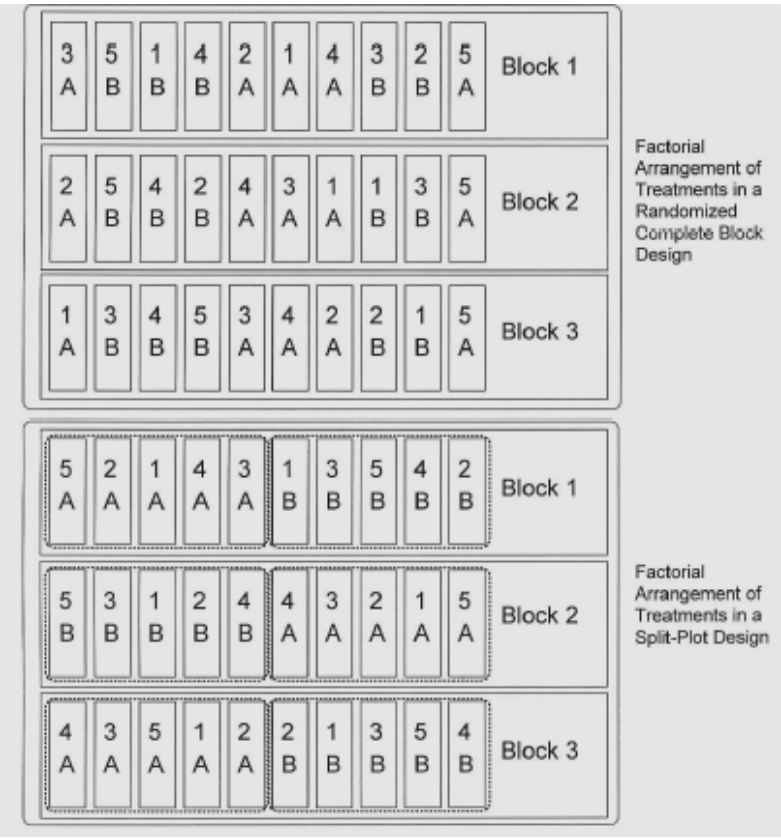


Figure 9 A 2x5 factorial arrangement of treatments in a randomized complete block design and in a split-plot design (Davis et al., 1999)

A and B represent two levels of one factor, and the numbers (1-5) represent five levels of a second factor. The combinations (e.g., 4A, 5B, etc.) denote individual treatment combinations. Either experimental design could be used, but the randomized complete block design is preferred unless the split-plot design is required by some limitation on randomization.

3.5.3 The split-plot design

A split-plot experimental design is a special design that is sometimes used with factorial arrangements of treatments. This design usually is used when an experiment has at least two factors and some constraint prevents you from randomizing the treatments into a randomized

complete block design. Such a constraint may be based on equipment limitations, on biological considerations or when certain treatments require a larger plot size than others. For example, the equipment you have may make it difficult to put out a soil fumigant in randomized complete blocks, but you may be able to put out the fumigant so that all treatments within a block that get the fumigant will be clustered together rather than scattered throughout the block. You can use a split-plot experimental design to work around this limitation as long as you are able to randomize the other factors. There are other situations when this design is appropriate, but a constraint on randomization is the most likely to occur. In split-plot designs, the terms "whole plots" and "subplots" refer to the plots into which the factors are randomized. As the names imply, whole plots are subdivided into subplots. In Figure 9, a whole plot would be the areas designated with A or B, and the subplots, the subdivisions within the whole plots, are designated 1, 2, 3, 4, or 5. In this example, A and B could represent two varieties (two levels of one factor) and the numbers could represent different fungicides (five levels of a second factor). Each whole plot serves as a block for the subplot treatments. To assign treatments in a split-plot design, start by identifying where each block will be. Then randomize the whole plot treatments within each block. The whole plot treatments will be the treatments that you are unable to randomize into a randomized complete block design. The subplot treatments can then be randomized within each whole plot treatment (Davis et al., 1999).

Statistically speaking, you sacrifice precise information on the main treatment for more precise measurements of the sub-treatments simply because the sub-treatments are replicated more than the main treatments. Though fairly easy to set up in the field, analyzing the data of a split-plot experiment can be somewhat complex. Because of the greater number of treatments, adequately replicating a split-plot experiment can take much more space in the field. Work with someone knowledgeable in statistics to set up a split-plot experiment (Sooby, 2001).

3.6 Plot size

Farmers have an abundance of ideas. Sometimes the ideas (and the new practices and products that may accompany them) can be implemented quickly and easily. At other times, a major investment is necessary. But no matter how simple or complex an idea, a certain amount of risk is involved. To minimize the risk, it's smart to test these ideas on a small scale first (Anderson, 1993).
Therefore experiments are usually done on portions of the farm, seldom on the entire farm. The mathematical techniques of statistics are used to calculate the odds that what you are measuring on one part of the farm will hold true for the whole farm (Sooby, 2001).

Field experiments are seldom carried to a conclusion if they aren't designed to be relatively easy to maintain. In the planning stages, decide what size your plots will be and where they will fit best. On-farm research typically uses plots that are field length and one or two tractor passes wide. This makes it easier to apply treatments along the entire strip without having to start or stop in the middle of the field. Flags or fence posts are useful to mark where one treatment ends and the next one begins at planting, when applying treatments, and at harvest. Such markers can easily be knocked over or ripped out with machinery, so be careful and immediately replace any that are moved (Sooby, 2001).

The size of a plot can greatly affect the magnitude of experimental error in a field trial. Too small plots may give unreliable results; unnecessarily large plots waste time and resources. In

general, experimental error decreases as plot size increases, but the reduction is not proportional. Plot size not only affects variability but may also bring about bias in the experimental results. Plots should be wide enough to permit the removal of border rows when necessary. Whatever size and shape of plots you choose, make sure that an area not smaller than 5 m^2, free from all types of competition and border effects, is available for harvesting and determining plot yield.[4] The cultural practices related to the experiment can dictate the size and shape of plots for ease of operations. Fertilizer trials require larger plots than variety yield tests. Irrigation studies may require even larger plots. In insecticide or herbicide trials where the chemicals are to be sprayed, the width of the plot may be governed by the range of coverage of the sprayer used. When soil heterogeneity is patchy a large plot should be used. In experiments where border effects might be appreciable, square plots are desirable because they have minimum perimeter for a given plot size. In varietal yield tests where varietal competition is expected, plots with at least six rows should be used to allow exclusion of one row on each side of the plot, thus leaving four center rows for harvest (Gomez, 1972).

To get an accurate measurement of the effect of pest and weed management treatments, the plot must be large enough to account for uneven initial distribution of the pest (pathogen, insect, etc.) and weed. Some areas may start with the pest/weed present, but the pest/weed may occur in other areas only after it has spread from its initial location. This is very important for pests that spread very slowly (such as most soil-borne organisms). Some diseases and pests are highly mobile and spread very rapidly (such as many insects). In an insect management trial, measuring the effect of a treatment can be very difficult if your plots are too small because the insects that you see in the plot may have simply spread from the plot next to it. To minimize this problem, you can increase your plot size and then collect data from the middle section of the plot. For example, you might have an eight-row plot but only collect data from the middle four rows. The rows from which you do not collect data are often referred to as “buffer rows” because they buffer the effect of the neighboring plots. If you do not use buffer rows when they are needed, you may fail to detect differences among treatments and incorrectly conclude that many treatments were ineffective. Buffer rows are frequently used when there is uncertainty whether treatments can influence nearby rows. A similar concept involves the use of border rows along the edges of your test area. A significant “border effect” commonly exists at the edge of a field where the plants may grow differently than plants not at the edge. Although you may be able to minimize this problem with blocking, it is often better to eliminate the problem by not using the rows at the edge of a field in your experiment (Davis et al., 1999).

It is often assumed that the plot size should be larger for on-farm than for on-station trials. This is the result of past on-farm trials usually having been at the validation stage of the research. There is no general justification, on statistical grounds, for preferring large plots in on-farm trials. The most efficient use of a given area, or of a given number of trees, is normally achieved with more small plots, so more replications, rather than fewer larger plots, unless you are dealing with some agro-forestry systems, or areas with nutrient or water movement. Normally, there is a balance between the preference of farmer and researcher for larger plots on the basis of realism, or ease of treatment application, and the statistical benefit of improved precision from more, smaller plots.
The cases for realism when seeking farmer opinion regarding treatments, or when comparing treatments with regard to labour requirements are examples of compelling reasons for using large plots. However, the case for large plots should be made in relation to the objectives of

[4] Such small plot sizes can occur in researcher-managed on-farm trials. Treatment plots in farmer-managed on-farm trials are, however, often between 400 and 2000 m^2 (Miller et al.).

the experiment; merely stating that on-farm experiments require large plots is not enough (SSC, 1998).

3.7 Collection of data

You can collect an almost infinite amount of data in any experiment, but not all of it will be useful. Proper planning will ensure that you collect the right data to address your test's objective. The "right" data to collect can usually be determined by examining the stated purposes of the experiment. It is useful to ask yourself, "How can this data be used?" If you have trouble answering that question, then collecting that data may be a waste of time. **It is much more common for people to collect too little data than to collect too much data.** Deciding what data to collect is only part of the process. You also have to decide when to collect that data and if you need to collect the same type of data on more than one occasion. So, how much data is enough? The answer is "enough data to fully address the test's objective." If you understand the biology of the organisms involved and how your data addresses the test objective, then you should be able to tell if you are collecting enough data. You should take photographs of any differences among treatments that are easily visible. To most farmers, a picture is more convincing than a graph or data table (Davis et al., 1999).

In on-farm trials, there is often still too much on-station type of measurements, because the implicit assumption is that the methods of analysis will be the same as on-station. Whilst these data may still be of interest, we suggest that more attention be given to the collection of measurement of concomitant variables (waterlogging, rainfall, soil type, dates of sowing and weeding) and farmers' opinions. The main reason for devoting time to the concomitant information is that we still need to try to understand the causes of as much of the variability as possible. In on-station trials the plots may be smaller and are likely to be more homogeneous. In on-farm trials there may be more variation within a farm than on-station and in addition there is variation between farms. In on-station trials there is a consistent management structure, whereas here there can be large differences in management practice between the farms. As a general guide, what is not controlled should be measured, both at the plot level and at the farm level, if it is of direct interest or if it might explain some of the variation in the data. The direct and concomitant measurements to be taken are normally decided at the planning stage. **Often too much data is collected that is never analyzed.** Our encouragement to measure potential concomitant variables is not intended as support for the measurement of all possible data, just in case they may be useful (SSC, 1998).

The most important data for an experiment is usually the yield, but that doesn't mean you shouldn't inspect the plots during the course of the season. Look the plots over from time to time and note things that you may want to remember later. For example, weed or insect infestation may be especially heavy in one or more of the plots. Such information can help you explain your final results.
Although data collection sounds complicated, it's really just observation, something that every farmer does naturally every day. Recording your observations is vital to research. If you always carry a pencil and small notebook, you can easily jot down observations and ideas and record the date as you write. Be sure to note from which plots the observations were taken. Additional useful observations are weather, weed types and pressure, insect and other pest damage, soil condition, dates of operations, **things that went wrong**, diseases, chemicals applied, if more than one person worked on the plots, and who worked on which plots.[5]

[5] Specially for irrigated rice the water levels in the different treatment plots have to be monitored regularly.

If more than one person collects the data, everyone involved should use the same technique. Uniformity of technique reduces the human factor as a source of variation. For pre-planned measurements such as plant heights and harvest, it may be easier to use a worksheet (Anderson, 1993).

It is useful to draw out a plot map or plan to help visualize the project and to keep track of which treatment has been applied where. Make sure that any changes you make in the field are reflected on your map. Make at least one copy of the plot map and keep it somewhere safe so that you don't lose all your work if you lose your working copy of the map.
Compile a field history for the past 5 years that documents crops grown, tillage operations, inputs applied, and yields in the experimental fields. Past management can strongly influence present performance and provide valuable clues to why things turned out the way they did (Sooby, 2001).
Annex 1 gives an example of general (not rice-specific), baseline data needed to document and interpret a valid, unbiased test. This information is easily entered on record sheets (Miller et al.).

It is critically important to collect unbiased data. The only way to ensure this is to collect data without knowing what the treatment was in that plot. That would be difficult to do if the treatment were written on a stake in front of each plot. It is beneficial to use some type of code on the plot stakes so that you have to decode the stake number to determine what the treatment was. You can make up any code you like just so long as the person collecting the data cannot tell from the plot stake what the treatment was. For example, you can number the plots sequentially (1, 2, 3, etc.) and have a sheet of paper listing what treatment was applied to plot 1, plot 2, etc. When you collect the data, you write down your observation for plot 1 and later look at your list to see what treatment was in that plot.
If you know what treatment was in a plot, or which plots were the untreated controls, your evaluations (disease severity ratings, insect damage ratings, etc.) may inadvertently be influenced. Your subconscious may slightly increase the ratings for untreated plots and decrease it for the plots with treatments that you think should work well. You will probably not even be aware that it is happening, but these subtle influences can change the data enough to affect your ultimate conclusions from the test. **If you do not collect unbiased data, you cannot be certain that your conclusions are correct** (Davis et al., 1999).

A participatory on-farm trial is not really participatory, unless a record is kept of the important contributions made and opinions stated by the farmers. These may be the actual farmers who use the land, or others who view the fields (SSC, 1998).
An important factor in the evaluation of a new technology for farmers is the productivity of the labour involved. Therefore you have to be precise in the collection of data on labour use. Rudebjer (2001) advises for that matter to use a purposive sample (not random); not to use enumerators; to use a sample size of approximately 20; to interview in the field, just after the activity has taken place; to break down your questions; to use local measures (e.g., for area); to triangulate (use different ways to get the same info); and to report variability.

In research making measurements would take a prohibitively long time to, for example, separately weigh all the millet heads in a plot, so you sample a certain number of them to represent the entire population of millet heads. Sample size should be determined using statistical methods if your experiment is for publication in a scientific journal. Otherwise, you can be guided by plot size and time constraints. You need to collect samples from each region in the plot in order to have representation from the entire area, but you don't want to spend

days at the scale, either. You trade a degree of accuracy for pragmatic concerns, a trade-off that is made constantly in research (Sooby, 2001).

The greatest difficulty in sampling in broadcast rice plots lies in the identification of sampling units. In transplanted rice, the sampling unit used is based on hills which is not applicable to broadcast rice. In broadcast rice, the sampling unit must be identified in terms of area. But demarcating sampling areas for measuring various rice characters is an additional problem. The non-uniform plant density in broadcast rice plots also causes higher variability than in transplanted rice. Thus, for most characters the sample size in broadcast rice should be larger than in transplanted rice (Gomez, 1972).

Yield estimates are needed to make production and economic comparisons between treatments. To be valid, yield measurements must be taken from comparable areas in each treatment plot. For that matter you must measure the size of the harvest area.
To **measure the grain yield** you will have to harvest as large an area as possible after discarding any border areas on all four sides of a plot. Thresh, clean, dry, and weigh all grains harvested from each plot separately. Immediately after the grain from a plot is weighed, determine its moisture content. Adjust grain weight to 14 % moisture using the formula:

Adjusted grain weight = A x W

where A is the adjustment coefficient and W the weight of the harvested grains. The coefficient A can be computed:

$$A = \frac{100 - M}{86}$$

where M is the moisture content (percent) of the grains.

When a plot is damaged, do not take samples from plants damaged by pests and diseases for measuring any agronomic character (unless your objective is to evaluate damage). When incidence in each plot does not involve a large number of plants, say, when not more than 20 percent of the plants in any plot is damaged, exclude damaged plants from the harvest of each plot. Later, convert the grain weight of each plot based on the number of plants in the normal plots. When only a few plots are heavily damaged, do not collect samples from these plots. They should be treated as "missing data" when the data are analyzed. When damage is moderate to heavy, quite non-uniform from plot to plot, and measurable, harvest all plants in each plot in the usual manner. Collect data on pest and disease incidence for every plot. Use covariance analysis with the data on the incidences as covariates (Gomez, 1972).

4 Analyzing the Results of On-Farm Research

Data collected in the trial must first be converted to commonly used units prior to analysis and summary (Miller et al.).
Much can be learned from simply looking over the data without using statistical methods. Statistics are an additional tool to assess the data which can be used to assess the effects of natural variability between plots and the likelihood that the differences measured are due to the treatments (Colley and Myers, 2007).
Statistical analysis may not be necessary if treatment differences are very large and consistent; treatment means may then be sufficient (Davis et al., 1999).

However, often it can be difficult to tell just by looking at the data whether any differences are due to random variation or to treatment effects. Statistical analysis re-orders the data that were randomized in the field and performs mathematical computations to determine the probabilities that the differences were caused by normal variation or by the treatments. The results of the data analysis give you the basis for making conclusions on the effects of the treatments.
Analyses can be performed using different confidence levels. Professional researchers typically choose a 95% confidence level, which means that there is a 95% chance that the measured differences are due to the treatments rather than to random variation or error. There is a 5% chance that the analysis will lead you to conclude there were differences where there were none or vice versa. In scientific literature, this chance of wrongly interpreting the data is indicated as $p < 0.05$ (probability is less than 5% that the analysis is picking up on non-existent differences or not measuring real differences). In reality, you will never be 100% sure that you have proved or disproved your hypothesis. Statistics are based on tendencies and likelihoods, never on certainties (Sooby, 2001).
The probability level can feasibly range from 0.001 to 0.2 (0.1 - 20%), depending on the economic or environmental implication of choices between management options being tested, interactions with other managements practices, the grower's experience and other factors (Veseth et al., 1999).

A common statistical tool for on-farm tests is **LSD**, which is calculated through standard analysis of variance (ANOVA) statistical analysis software or can be calculated by hand. The LSD is a calculation based on the variability of treatment results within the trial and is used to help separate the effects of natural field variability from the treatment effects. If the difference between treatment means in a trial are equal to or larger than the LSD, the difference is statistically significant and believed to be due to the treatment effect and not natural field variability. On the other hand, if the difference between treatment means is smaller than the LSD, the differences are more likely due to natural field variability.
The LSD is only used if the ANOVA analysis determines that treatment differences are significant (Veseth et al., 1999).

Proper statistical analysis can be done if your experiment was designed according to the principles outlined in this manual, but proper analysis can be complicated greatly if these principles were not followed. Specialized statistical software is available, but most spreadsheet software can calculate simple statistics. Excel is a helpful program for inputting and analyzing data collected (Davis et al., 1999).

Data analysis largely depends on how the project was designed and conducted (Miller et al.).

The following presents the types of data analysis of the main designs used in this manual.

4.1 The t-test

As stated before you can use an easy statistical analysis, **the t-test**, on the data from a paired-comparison to detect any significant differences.
The paired-comparison trial compares two levels of a treatment. Each pair of treatments is replicated at random at least six times in one field to overcome chance field differences. Each data pair yields one difference. These differences can be analyzed to generate the **LSD** with which to evaluate a two-treatment trial (Exner and Thompson, 2005).

Anderson (1993) supplies an analysis example in six steps of a paired-comparison trial with six replications in which the two treatments are pre-plant and split-applied nitrogen application at the same rate.

Step 1. Analysis
Transfer the data from your data collection sheet to the table beneath.

	Treatments		Difference (C)	Deviation (D)	Deviation squared (D^2)
Blocks (r)	A (preplant)	B (split)	C = (A - B)	D = C - C_{avg}	D^2 = D x D
I	141	150	-9	-2.2	4.84
II	147	156	-9	-2.2	4.84
III	149	155	-6	0.8	0.64
IV	151	157	-6	0.8	0.64
V	149	150	-1	5.8	33.64
VI	142	-1	-10	-3.2	10.24
Totals	879	920	-41		D^2_{tot} = 54.84
Averages	A_{avg} = 146.5	B_{avg} = 153.3	C_{avg} = -6.8		

- *Value Treatment B-Block VI is 152 instead of the indicated -1 in the table above.*

Step 2. Calculate the variance.
variance = D^2_{tot} / (r-1) D^2_{tot} = 54.84 (r-1) = (6-1) = 5
54.84 / 5 = 10.97

Step 3. Calculate the variance of the means.
variance of the means = variance/r
10.97 / 6 = 1.83

Step 4. Calculate the standard error.
standard error = square root of the variance of the means
√1.83 = 1.35

Step 5. Calculate the least significant difference (LSD). Take the answer from step 4 and multiply it by the appropriate t-value.
standard error x t-value = LSD
1.35 x 2.57 = 3.47
(An alpha level of α=0.05 and n = (r-1) = 5; therefore a t- value of 2.57 ; see Annex 2)

Step 6. Look for a significant difference.
Take the answer from step 5, the LSD, and compare it to the box from the table labeled C_{avg}.
LSD = 3.47 C_{avg} = -6.8
C_{avg} is a negative value, so you ignore the negative sign. The C_{avg} (6.8) is greater than the LSD (3.47); therefore, the observed difference is significant (at an alpha level of 0.05). The yield increase observed under the split-application treatment can be considered real.

A "Student's" *t* table can be found in the back of most statistics textbooks. If the experimental question is: "Is practice 'A' *different* from practice 'B' (less *or* more)?" then you want a "two-tailed" *t* table that divides the chance for error (α) between the lower and the upper tails of the bell curve (see Figure 2). If your question is specifically "Is practice 'A' *greater than* practice 'B,' then you would use a one-tailed *t*-table.
Most *t* tables are two-tailed and don't even bother to say so (see for example the more elaborate *t* table presented in Annex 3). Notice how *t* diminishes as the number of replications (*n*) increases. Other things being equal, a smaller *t* means a smaller LSD, and a smaller LSD means the trial is more sensitive to the treatment differences you are looking for (Exner and Thompson, 2005).

A blank analysis worksheet for a paired-comparison trial involving the six steps used by Anderson (1993) is supplied in Annex 4.

4.2 Analysis of variance

Statistical analysis of the randomized complete block design and the split-plot design is computed using an analysis of variance. An ANOVA may be calculated for each trait individually or for the overall score assigned to each plot. The ANOVA provides an estimate of the experimental error or variation among the replications. It also provides a way to test the significance of the differences measured between treatments – or the likelihood that the differences measured are truly due to difference between treatments rather than differences in the environmental conditions (Colley and Myers, 2007).

A free, easy-to-use statistics computer program called AGSTATS02 (for IBM compatible computers) is available on the internet (see references for the internet hyperlinks) for analyzing the results of simple on-farm trials (Veseth et al., 1999).
After logging in with a password, one can use this program for analyzing randomized complete block designs with various numbers of treatments and replications/blocks at four different levels of significance. After inserting the research plot yield (or other trait) data, the program supplies the details of an ANOVA as in the example involving 6 treatments and 3 replications at the 5% level of significance presented in Table 2.

The source in Table 2 refers to the source of variation and the observed F-value is calculated by dividing the mean square of treatments or blocks/replications by the error mean square. The P-value gives the percentage chance that such an observed F-value would occur due to

normal random variation (error). If the P-value for the observed F-value is lower than your chosen level of significance (5% in our case) then you can state that the concerned source of variation (treatments, blocks/replications etc.) has led to statistically significant differences in yield or other trait. In our example this is the case for treatments but not for blocks.
The required F-value matches exactly the P-value corresponding with your chosen level of significance. The presented correction factor is used in the calculation of the sum squares.
The presented standard deviation (or in other words standard error) of the treatment means is an indication of the degree of precision in the on-farm trial.
The value of the coefficient of variation is an indication of the magnitude of experimental error likely to be obtained in the experiment. For on-station rice experiments at the IRRI experimental farm, the average coefficient of variation based on grain yield is about 8 percent for varietal yield tests and 10 percent for other agronomic trials (Gomez, 1972). In on-farm trials the level of the coefficient of variation is likely to be higher than 10 percent (as in the example) and sometimes even much higher than that (going to 30 to 40 percent).

Table 2 Example of ANOVA results computed by AGSTATS02

Source	Degree of Freedom	Sum Square	Mean Square	Observed F	P value	
Total	17	18.30	1.08			
Treatments	5	16.16	3.23	15.37	0.02 %	Significant
Blocks	2	0.04	0.02	0.09	91.72 %	Not signif.
Error	10	2.10	0.21			
Required F	3.33	*RCBD Example* *Level of significance 5%* *Treatments 6* *Replications/Blocks 3*				
Correction Factor	239.15					
Standard Deviation	0.46					
Coefficient of Variation	12.58 %					
LSD 5%	0.83					

That treatment yields are statistically significant different at the 5% level of confidence is not sufficient information for the evaluation of the on-farm experiment. One wants to know which treatments are significantly different from each other. For that matter AGSTATS02 calculates the treatment means and puts a, b, c or d behind them. Treatments with different letters are significantly different from each other. In this example this is the case when the difference between the treatment means is larger than the LSD of 0.83. Note that the LSD is the result of standard deviation[6] x t-value; in this example $0.46 \times t_{10;\,0.05} = 0.46 \times 1.812 = 0.83$.
The t-value is based on the 10 error degrees of freedom and a 2-tailed test of difference (either greater or less), which gives an α-level of 0.10 (2 x 5% = 10%).

Results from factorial arrangement of treatments and split-plot designs can not be statistically analyzed by AGSTATS02. In these cases one has to use more detailed statistical software packages such as the free GENSTAT-Discovery and IRRISTAT/CROPSTAT packages, or any of those numerous available commercial statistical software packages.

[6] The standard deviation = √mean square error = √0.21 in this example

4.3 Combined statistical analysis

On-farm experiments with sufficient within-site replication and detailed measurements of yield response can have **separate within-site analyses initially, then a combined analysis**. This is usually only the case for researcher-designed and managed trials. Others will use the data within a single analysis. However, there are two main differences between on-station and on-farm trials that have a bearing on the analysis. One is that with on-farm trials we expect a farm by treatment interaction, and one of the objectives of the trial is often to explore this interaction. The other difference is that there is now variation at different levels - there is variation between farms because of characteristics such as different agro-climatic conditions, management practices, as well as variation between plots within farms. As always, any analysis should try to explain as much of the variation as possible (SSC, 1998).

A combined statistical analysis of on-farm trials is often used to evaluate relationships between biophysical responses and environmental, management and social variables. The results of such combined analyses can also help to understand farmer assessments and may be turned into decision trees for farmers or maps of recommendation domains (SSC, 1998).
Due to the variation between farms it is, however, not always allowed to perform a combined analysis of variance over all farms. One has to check first if certain farms belong to the same "group of error variance". Or in other words; to check if certain farms are enough similar to each other. More details on how to perform this check can be found in Mutsaers et al. (1997) and Neeley et al. (1991).
If not all farms can be included in the analysis one has to perform an analysis on subsets of groups of similar farms.

Gomez and Gomez (1984) provide an example of a step-by-step procedure for examining interaction effects between treatments and farms for a combined analysis of variance over four farms with five weed-control treatments and three replications per farm. Table 3 presents the results of the combined analysis of variance for this example.

Table 3 Example of a Combined Analysis of Variance over four farms with five weed-control treatments and three replications per farm

Source of Variation	Degree of Freedom	Sum of Squares	Mean Square	Computed *F*
Farm (*F*)	3	60.8298	20.2766	25.05**
Replications within farm	8	6.4760	0.8095	
Treatments (*T*)	4	73.7773	18.4443	35.07**
F x *T*	12	69.8161	5.8180	11.06**
Pooled Error	32	16.8306	0.5260	
Total	59	227.7298		

Coefficient of variation = 21.0 % ** = significant at the 1 % level
Source: Gomez and Gomez (1984)

Table 3 shows that in this example of a combined analysis of variance there is a highly significant interaction effect between farm and treatment. This means that the performance of the different weed-control treatments varied among the selected farms. Gomez and Gomez (1984) show step by step how one can examine further the nature of this interaction.

As stated earlier in this manual it is frequently the case in on-farm experiments that there are many farmers involved but each farmer has only one replicate of each treatment in his field. A combined analysis over farms can not be carried out to provide information about the interaction between treatment and farm. Quite often within-farm replication appears to be impossible or was simply forgotten while designing the on-farm experiment. Especially in farmer-managed trials the statistical analysis is far less important than the farmer assessment of the on-farm experiment. In such situations, **adaptability analysis** (Hildebrand and Russell, 1996) is an alternative way to see if treatment effects vary significantly across farms. In this method, all treatment yields in a given field are compared with the environmental index, which is the average yield of all treatments in that field. The environmental index is the product of the entire set of biophysical and socio-economic factors, and is assumed to reflect the overall growing conditions and the quality of management in a particular field (Mutsaers et al., 1997). In fact, adaptability analysis examines whether treatment effects differ significantly between fields with overall good conditions (high environmental index) and fields with overall poor conditions (low environmental index).

Figure 10 presents an example of an adaptability analysis for a rice-fertilizer on-farm trial with five non-replicated treatments on 15 farms. Figure 10 shows the regression lines of the five rice-fertilizer treatments on the environmental index (e).

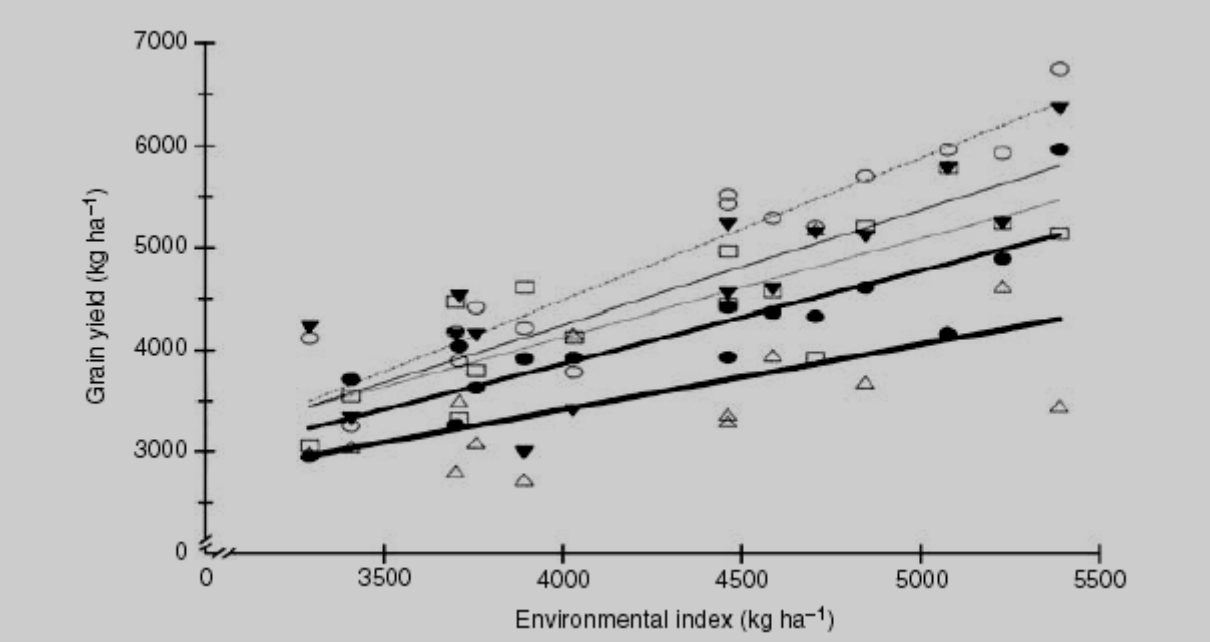


Figure 10 Regression of N-level trial treatments, 0N (△), 30N (●), 60N (□), 60N+17.5P (▼), 120N (○) on environmental index (e) over 15 rice fields (Meertens et al., 2003)

Fitted lines: △ $Y = 791\ (1000) + 0.660\ (0.231)\ e\ (kg\ ha^{-1})$, $r^2 = 0.39$;

● $Y = 186\ (672) + 0.927\ (0.155)\ e\ (kg\ ha^{-1})$, $r^2 = 0.73$;

□ $Y = 205\ (799) + 0.987\ (0.185)\ e\ (kg\ ha^{-1})$, $r^2 = 0.69$;

▼ $Y = -323\ (983) + 1.150\ (0.227)\ e\ (kg\ ha^{-1})$, $r^2 = 0.66$;

○ $Y = -1186\ (638) + 1.427\ (0.148)\ e\ (kg\ ha^{-1})$, $r^2 = 0.88$.

In this example, the environmental index varied according to differences in management (such as land preparation, type of planting, and weeding), soils, varieties, water depth and tillering stage at the time of urea application, rainfall, and position of the field in relation to the slope. From just looking at Figure 10 one gets the impression that especially applications of high nutrient doses ($N_{60}P_{17.5}$ and N_{120}) perform relatively better in fields with otherwise optimum conditions. However, in the combined analysis of variance on rice grain yields, the differences in regression of the treatments on the environmental index were not significant at the 10 % level of probability. This points to no clear interaction between treatments and fields. More details on performing an adaptability analysis are to be found in Mutsaers et al. (1997) and Hildebrand and Russell (1996).

Extension demonstration plots (such as the Field School Trials used in the FFS approach of Guyana) can, if implemented and monitored well, be used for a combined statistical analysis in case there are two treatments involved; the demonstrated treatment next to the control (often farmers' practice). The simple *t*-test can be used for that matter; each demonstration farm is then treated as a paired comparison. Adaptability analysis can further indicate if there is an interaction between the demonstrated treatment and type of farm.

Combined analysis of variance and adaptability analysis can be performed by most available commercial statistical software packages and by the free GENSTAT-Discovery and IRRISTAT/CROPSTAT packages.

5 Economic Evaluation

If the statistical analysis has shown that the average difference of the treatments is greater than the calculated LSD then we have some confidence that the difference is real. But is it the best choice economically? Is the advantage given to one treatment worth the cost difference between the treatments? An on-farm trial should be analyzed economically to determine where cost differences occurred such as seed, fertilizer, pesticides, tillage, and management time between the treatments. The additional income benefit can then be weighed against the possible increased costs for that treatment. Non-tangible benefits such as improved soil quality and environmental improvement also are to be considered (Sundermeier, 1997).

5.1 Partial budget analysis

One of the tools in economics used to compare the economic benefits of technologies is **partial budget analysis**. A budget is a farm management method that is intended to assist researchers, extension agents, and farmers in the decision-making process. It is a tool that aims at quantifying and comparing the effects of a proposed technology on crop production to those of other alternative technologies. Results from partial budget analysis assist agricultural scientists in identifying weaknesses (high cost and/or low income) of the technology being developed. Partial budget analysis shows also the level of profitability and helps to decide whether to adopt a new technology or not (IITA, 2003).

A partial budget shows the effect of change(s) in farm operations. For example, farmers know that fertilizer application will likely increase rice yield, and thus the gross income. The use of fertilizer also results in additional costs. To decide whether to use fertilizer for rice production or not requires a partial budget analysis. The term "partial budget" is a reminder that not all production costs are included in the budget. Only those costs that are affected by the alternative treatments being considered, or in other words the costs that vary are included. **Costs that vary** or **total variable input costs** are the costs (per hectare) of purchased inputs, labor, and machinery that vary between experimental treatments. It is very important to take into consideration all of the changes in labor implied by the different treatments in an experiment (CIMMYT, 1988).

To perform a partial budget one has to know some further economic descriptions in addition to the costs that vary. These economic descriptions are the following:

Opportunity cost. Not all costs in a partial budget necessarily represent the exchange of cash. In the case of labor, for instance, farmers may do the work themselves, rather than hire others to do it. The opportunity cost can be defined as the value of any resource in its best alternative use. Thus if farmers could be earning money as laborers, rather than working on their own farms, the opportunity cost of their weeding is the net wage they would have earned had they chosen not to stay on the farm and weed (CIMMYT, 1988).

Farm gate price (of an input). In some cases, farms are far from markets. While getting inputs to farms, additional expenses (such as transportation cost) must be added to the market price of the input. Farm gate price is the price of an input at the farm gate.

Farm gate cost (of an input). The farm gate cost of an input is the product of its farm gate price and the quantity of the input required for a given area.

With on-farm research you can obtain realistic input-output data for economic analysis.

Adjusted yield. The adjusted yield is the experimental average yield over a certain number of farms scaled down by a given proportion to approximate the yield that farmers can obtain on their farms. The scaling down is necessary to prevent overestimation of the returns that farmers are likely to obtain from a treatment. Experimental yields are higher than farmers' yields because of higher management level (recommended number of stands, timely planting, timely weeding, timely application of fertilizer, timely application and recommended doses of chemicals, etc.), smaller plot size, precision in harvesting date, and better harvesting methods. The difference between yields from experimental fields and those from farmers' fields in similar cropping conditions will be the basis for the scaling down. Scaling down may not be necessary in an on-farm experiment where the experimental design is very close to the farmers' practice (IITA, 2003).
As a general rule, total adjustments between 5 and 30% are appropriate. The yield adjustment should not be looked upon as a factor to be applied mechanically. Each type of experiment, each year, should be reviewed before deciding on an appropriate adjustment. If this is done, researchers will be able to make decisions about new technologies with a realistic appreciation of farmers' conditions (CIMMYT, 1988).

Farm gate price of the output. Distances between farms and markets account for the difference between the farm gate price and the market price of the output. The farm gate price of the output is the value (price) farmers receive or can receive for their harvested crops. By this definition, it is the price farmers receive at the end of the production process.

Gross field benefits. The gross field benefits is the product of the farm gate price of the output and the adjusted yield.

Net benefits. The net benefits is the difference between the gross field benefits and the total variable input costs.

An example of a partial budget is presented in Table 4.

Table 4 Example of a partial budget

	Hand weeding	Herbicide
Average yield (kg/ha)	2.000	2.400
Adjusted yield (kg/ha)	1.800	2.060
Gross field benefits ($/ha)	3.600	4.320
Cost of herbicide ($/ha)	0	500
Cost of labor to apply herbicide ($/ha)	0	100
Cost of labor for hand weeding ($/ha)	400	0
Total costs that vary ($/ha)	400	600
Net benefits ($/ha)	3.200	3.720

Source: CIMMYT (1988)

Dominance analysis. The process of eliminating dominated treatments from further analysis is called dominance analysis. A dominated treatment has the same or lower net benefit than other treatments of a lower total variable input cost. In a dominated treatment, a higher total

variable input cost is incurred to earn the same or less net benefit when compared with other treatments. Such a treatment should be eliminated at this stage from further analysis. An example of a dominance analysis is presented in Table 5, in which treatment 3 appears to be dominated by treatment 1 and will be eliminated from further analysis.

Table 5 Dominance analysis, weed control by seeding rate experiment

Treatment	Weed control	Seeding rate (kg/ha)	Total costs that vary ($/ha)	Net benefits ($/ha)
1	None	120	2,400	10,360
3	None	160	3,200	10,136 D
2	Herbicide	120	3,875	11,765
4	Herbicide	160	4,675	11,965

Source: CIMMYT (1988)

5.2 Marginal analysis

Although the calculation of net benefits accounts for the costs that vary, it is necessary to compare the extra (or marginal) costs with the extra (or marginal) net benefits. This comparison is important to farmers because they are interested in seeing the increase in costs required to obtain a given increase in net benefits. Higher net benefits may not be attractive if they require very much higher costs (CIMMYT, 1988).

Marginal rate of return. The marginal rate of return (MRR) is a ratio of the change in net benefits to change in total variable input costs between treatments.

The process of calculating the marginal rates of return of alternative treatments, proceeding in steps from the least costly treatment to the most costly, and deciding if they are acceptable to farmers, is called **marginal analysis**. Marginal analysis for a particular experiment should be done on the combined results from at least several locations over one or more years (CIMMYT, 1988). Table 6 presents an example of a marginal analysis.

Table 6 Marginal analysis, weed control by seeding rate experiment

Treatment	Costs that vary ($/ha)	Marginal costs ($/ha)	Net benefits ($/ha)	Marginal net benefits ($/ha)	Marginal rate of return
1	2,400		10,360		
		1,475		1,405	95%
2	3,875		11,765		
		800		200	25%
4	4,675		11,965		

Source: CIMMYT (1988)

Note that the marginal rates of return appear in between the two treatments. It makes no sense to speak of the marginal rate of return of a particular treatment; rather, the marginal rate of return is a characteristic of the change from one treatment to another (CIMMYT, 1988).

Net benefit curve. In a net benefit curve each of the treatments is plotted according to its net benefits and total costs that vary. The slope (gradient) between two consecutive treatments is the difference between the two consecutive net benefits divided by the difference between

the total variable input costs of the two consecutive treatments. The MRR is estimated from the slope of the curve. The steeper the slope, the higher the MRR. Figure 11 presents the net benefit curve for the marginal analysis example used in Table 6.

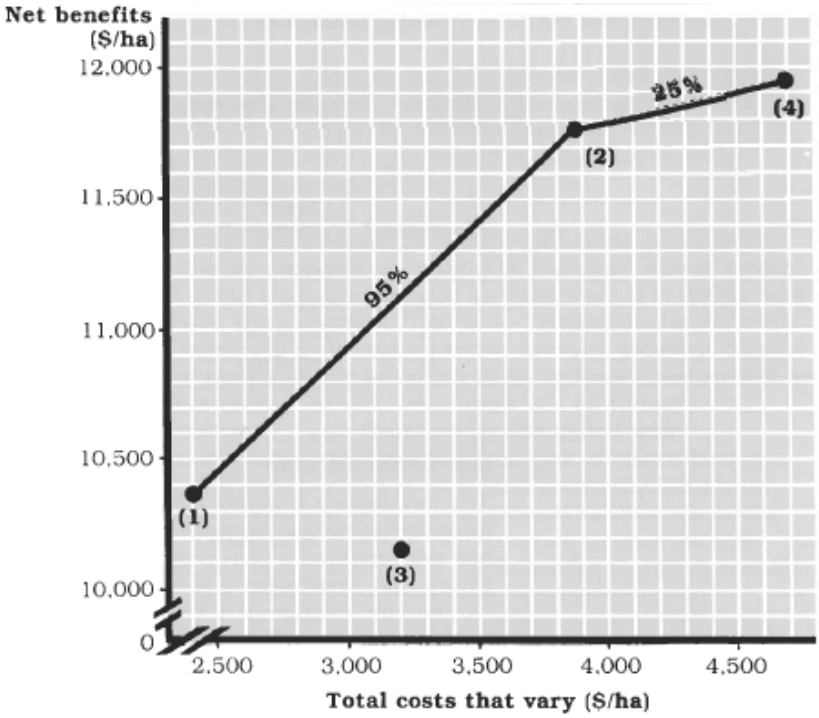


Figure 11 Net benefit curve, weed control by seeding rate experiment (CIMMYT, 1988)

A marginal rate of return of 95% means that for every G$ 1.00 invested in herbicide and its application, farmers can expect to recover the G$ 1.00 and obtain an additional G$ 0.95. Is a MRR of 95% high enough or is 25% already okay for farmers?

Acceptable minimum rate of return. The minimum return which farmers expect to earn from an enterprise or technology is the acceptable minimum rate of return (AMRR). Returns below this minimum makes the enterprise or technology a failure. AMRR is the sum of return to management and cost of capital.

Return to management. The return to management is the benefit which the rice farmer expects for managing a rice farm. In a new technology, it is the benefit which the farmer expects to receive for the time and effort spent in learning and using the new technology.

Cost of capital. This is the benefit forgone for tying up the working capital in one enterprise rather than in another enterprise. In the case of money, it is interest, or rent if the resources are land and equipment. The opportunity cost concept is used if owned inputs are used in production. In areas where the formal credit system is weak and the informal credit system is very active, the opportunity cost of capital should use the high interest rates that are usually applied by moneylenders.

Working capital. The working capital is the value of resources (purchased or owned) used in production with the expectation of returns in future. In rice production, the working capital is the value of land, labor, and capital.

In estimating an acceptable minimum rate of return, something must be added to the cost of capital to repay the farmers for the time and effort spent in learning to manage a new technology.
If the technology simply represents an adjustment in current farmer practice (such as a different fertilizer rate for farmers that are already using fertilizer), then a minimum rate of return as low as 50% may be acceptable. Unless capital is very easily available and learning costs are very low, it is unlikely that a rate of return below 50% will be accepted. It is important to note that this range represents an estimate for crop cycles of four to five months. If the crop cycle is longer, the minimum rate of return will be correspondingly higher. In areas where the inflation rate is very high, this range should be adjusted upward by the rate of inflation over the period of the crop cycle as well.
The minimum rate of return acceptable to farmers will rarely be below 50%, even for technologies that represent only simple adjustments in farmer practice, **and is often in the neighborhood of 100%**, especially when the proposed practice is new to farmers.
If a change in technologies offers a rate of return above 100% (which is equivalent to a "2 to 1" return, of which farmers often speak), it would seem safe to recommend it in most cases (CIMMYT, 1988).

In many areas, farmers do not have access to institutional credit. They must either use their own capital, or take advantage of the informal capital market, such as village moneylenders. The interest rates charged in this informal sector provide a way of beginning to estimate a minimum rate of return. If it turns out that local moneylenders charge 10% per month, for instance, then a cost of capital for five months would be 50%. To estimate the minimum rate of return in this case, an additional amount would have to be added to represent what farmers expect will repay their effort in learning about and using the new technology. This extra amount may be approximated by doubling the cost of capital (unless the technology represents a very simple adjustment in practices). Thus in this example, the minimum rate of return would be estimated to be 100%. It should be emphasized that this is simply a way of deriving a rough estimate of the level of returns that farmers will require.
If farmers do have access to institutional credit, the cost of capital can be estimated by using the rate of interest charged over the agricultural cycle. That is, the rate of interest should cover the period from when the farmers receive credit (cash or inputs) to when they sell their harvest and repay the loan. In addition, it is necessary to include all charges connected with the loan. There are often service charges, insurance fees, or even farmers' personal expenses for things like transport to town to arrange the loan, that must be included in the estimate of the cost of capital. Once the cost of capital on the formal market has been calculated, an estimate of the minimum rate of return can be obtained by doubling this rate. This will provide a rough idea of the rate of return that farmers will find acceptable if they are to take a loan to invest in a new technology (CIMMYT, 1988).

It is important to bear in mind that the calculation of the marginal rate of return is based on yield estimates derived from agronomic experiments and on estimates of various costs, often opportunity costs. Furthermore, the marginal rate of return is compared to a minimum rate of return which is only an approximation of the investment goals of the farmers. Discretion and good judgment must always play an important part in interpreting these rates and in making recommendations. If the marginal rate of return is comfortably above the minimum, the chances are good that the change will be accepted. If it is close to the minimum rate of return then caution must be exercised. In no case can one apply a mechanical rule to recommend a change that is a few percentage points above the minimum rate, or reject it if it is a few points below. Making farmer recommendations requires a thorough knowledge of the research area

and the problems that farmers face, a dedication to good agronomic research, and the ability to learn from previous experience. Marginal analysis is a powerful tool in this process, but it must be seen as only a part of the research strategy (CIMMYT, 1988).

Scientists are aware that certain factors usually beyond the control of the farmers could negatively affect (decrease) the net benefit and the marginal rate of return of shifting from the present practice to a proposed practice. Examples of these factors are technology failure and adverse weather condition (drought, flood, insect infestation, etc.) which reduce yield. Furthermore a change in market situation, price policies, or inflation can increase variable input price(s), decrease output price, or both.
Researchers cannot accurately predict the occurrence of these factors and their effects on net farm benefits before committing resources to the proposed technology. The factors bring about risk or uncertainty associated with the proposed technology. Farmers, most especially in developing countries, cannot afford to take risks because of lack of resources. Therefore researchers and farmers want to know the range of crop yields or prices for which the proposed practice may be recommended.
Sensitivity analysis is used to test a proposed technology for ability to withstand yield or price changes. Sensitivity analysis uses different prices or yields to determine what would happen to the net benefits and the choice of the recommended technology if it were to occur in different price or yield conditions to those expected (IITA, 2003).

Through using partial budget and marginal analysis agricultural researchers should be able to demonstrate, with little or no assistance from agricultural economists, the economic advantage of a proposed production technology over the existing methods (IITA, 2003).

6 Farmer Assessment

Even after a positive economic evaluation, showing that the marginal rate of return is well above the acceptable minimum rate of return expected by the farmers, there is still no guarantee that the technology will be adopted by the farmers.

Before changing from one production method to another, the farmer considers many factors, such as agro-ecological requirements, availability of required additional production resources (labor, credit, skill, farmland, equipment etc.), additional costs, and additional income resulting from the change. The farmer also considers the compatibility of the new technology with socio-cultural circumstances, goals, and the whole farming system (IITA, 2003).
To the extent possible, screening treatments for compatibility with the farming system should take place before experiments are conducted.

As farmers attempt to evaluate the net benefits of different treatments, they usually take risk into account. Farmers attempt to protect themselves against risks of loss in benefits and often avoid choices that subject them to these risks, even though such choices may on average yield higher benefits than less risky choices do. That farmers may prefer stable returns to the highest possible returns is referred to as **risk aversion**.
Another factor in farmers' decision making, related to risk aversion, is the fact that farmers tend to change their practices in a gradual, stepwise manner. They compare their practices with alternatives, and seek ways of cautiously testing new technologies. It is thus more likely that farmers will adopt individual elements, or small combinations of elements, rather than a complete technological package. This is not to say that farmers will not eventually come to use all the elements of a package of practices, but simply that in offering recommendations to farmers it is best to think of a strategy that allows them to make changes one step at a time (CIMMYT, 1988).

As researchers conduct on-farm experiments, they must constantly solicit farmers' opinions and reactions. Alternatives that seem to be promising both agronomically and economically may have other drawbacks that only farmers can identify. This information can be obtained from the farmers when researchers and extension agents interact sufficiently and appropriately with the farmers. That means that researchers and extension agents have to visit regularly the on-farm research plots and have to organize from time to time meetings with the participating farmers. At least a meeting at the end of the on-farm research process has to be held in which the farmers can give their feedback on the on-farm research results. The presentation of the on-farm results has to be understandable for the farmers. One should use the farmer vocabulary for plants, animals, insects, diseases and their measurement units. In the case of farmer-designed and -managed trials it is usually the farmers themselves who present the on-farm research results.

It should be clear that a farmer assessment of on-farm experiments is very crucial and can not be left out. The decision to adopt a certain technology is after all taken exclusively by the farmer. **It is the farmers who have the last word** (CIMMYT, 1988).

7 Reporting On-Farm Research

Proper conducted agronomic, statistical and economic analyses of on-farm research activities in combination with an adequate farmer assessment will enable stakeholders to discuss the obtained results in various ways and to draw solid conclusions.
The results have then to be reported. Reporting must be to all interested parties, with the collaborating farmers having high priority (SSC, 1998).

In the previous section on farmer assessment it was pointed out that reporting to farmers can best be done in general meetings, in which the results are presented in a way that farmers can understand. Reporting to other interested parties such as policy makers, fellow researchers, senior extension agents etc. can best be done through documenting the on-farm research activity.

Documentation of the on-farm research activity has several advantages. First of all it is needed to inform others, who were not involved in the on-farm research activity. Secondly it allows others to duplicate your experiment to verify it, and finally it gives you a record to look back on when you're trying to figure out what went wrong—or right (Sooby, 2001).

The on-farm trial report should ideally follow the lay-out of a scientific research article and include the following sections and subsections:

Introduction

- Background; problem formulation (consultation stakeholders)
- Previous research conducted related to identified problem
- Rationale/Justification (prioritization and targeting) and objectives of the research
- Expected impact of research (eg cost-benefit analysis)

Materials and Methods

- Experiment area (location and description)
- Treatments (number, level, general description)
- Sites description and experiment design
- Non-experimental variables (type of data collected and observations made)
- Statistical analysis

Results

- Agronomic
- Statistical

Discussion

- Discussion of agronomic data
- Economic analysis and farmers' assessment

Conclusions

Just as the data may vary within a replicated treatment, the results may vary among experiments if the whole experiment is repeated in another season or year. This can happen because of different weather conditions, different disease or insect pressure, or many other factors beyond your control. This does not mean that the results of a single experiment are not valid, but it does make it dangerous to draw conclusions from a single experiment. The one set of results you have may indicate treatment differences, but if you repeated the on-farm trial several times you might not see those treatment differences again. If the on-farm trial is repeated (and that means you cannot change any of the treatments) and you get similar results, then you can be much more confident that your conclusions are correct (Davis et al., 1999).

It is critical that final conclusions about a new practice be made only after being evaluated over several years and/or at several locations (Miller et al.).

References

Anderson, D., (**1993**), **On-Farm Research Guidebook**, University of Illinois Urbana-Champaign, Department of Agricultural Economics. www.ag.uiuc.edu/~vista/abstracts/aGUIDEBK.html or aslan@uiuc.edu

CIMMYT, (**1988**), **From agronomic data to farmer recommendation: An economic training manual**, CIMMYT Economics Program, Mexico, DF, http://www.cimmyt.org/english/docs/manual/agronomic/manual/contents.htm

Colley, M. R. and J. R. Myers, (**2007**), **On-farm Variety Trials: A Guide for Organic Vegetable, Herb, and Flower Producers**, Organic Seed Alliance, www.seedalliance.org

Davis, R. F., G. H. Harris, P. M. Roberts, and G. E. MacDonald, (**1999**), **Designing Research and Demonstration Tests for Farmers' Fields**, **Bulletin 1177**, University of Georgia, College of Agricultural and Environmental Sciences, Cooperative Extension Service. http://pubs.caes.uga.edu/caespubs/pubcd/B1177-w.htm

Exner, R. and D. Thompson, (**2005**), **The Paired-Comparison: A Good Design for Farmer-Managed Trials**, Practical Farmers of Iowa, Revised version, dnexner@iastate.edu or www.pfi.iastate.edu/ofr/OFR_worksheet.htm

Gomez, K.A., **(1972)**, **Techniques for field experiments with rice.** International Rice Research Institute, Los Baños, Philippines.

Gomez, K.A. and A.A. Gomez, (**1984**), **Statistical Procedures for Agricultural Research**, 2nd Edition, An International Rice Research Institute Book, John Wiley & Sons, New York.

Havlin, J. L., J. P. Shroyer and D. L. Devlin, (**1990**), **Establishing On-Farm Demonstration and Research Plots**, Kansas State University, Cooperative.Extension Service, MF-966, www.oznet.ksu.edu/library/crpsl2/

Hildebrand, P. E. and J. T. Russell, (**1996**), **Adaptability Analysis. A Method for the Design, Analysis and Interpretation of On-Farm Research-Extension**. Iowa State University Press, Ames.

IITA, (**2003**), **Partial budget analysis for on-farm research**, IRG 65, Ibadan, Nigeria, www.iita.org/cms/details/trn_mat/irg65/i65.pdf

Janke, R., D. Thompson, C. Cramer and K. McNamara, (1991), **A Farmer's Guide to On-Farm Research**, Rodale Institute Research Center.

Miller B., E. Adams, P. Peterson, and R. Karow, **On-Farm Testing: A Grower's Guide (EB1706)**, Washington State University Cooperative Extension. bulletin@wsu.edu or http://pnwsteep.wsu.edu/onfarmtesting/oftman.htm

Meertens, H.C.C., G. J. Kajiru, L. J. Ndege, H. J. Enserink and J. Brouwer, (**2003**), **Evaluation of On-Farm Soil Fertility Research in the Rainfed Lowland Rice Fields of Sukumaland, Tanzania**, *Experimental Agriculture*, volume 39, pp. 65–79. hccmeertens@hetnet.nl

Mutsaers, H. J.W., G. K. Weber, P. Walker and N. M. Fisher, (**1997**), **A Field Guide for On-Farm Experimentation**, IITA/ACP-EU (CTA)/ISNAR, Ibadan.

Neeley, D., J. Matata and F. Kisyombe, (**1991**), **Statistical Techniques for the Analysis of On-Farm Crop Trials. A Manual**, CIMMYT, Harare/Nairobi.

RNR-RC Jakar, (**2001**), **On-farm trial - Guidelines for extension agents,** Renewable Natural Resources Research Centre Jakar, Extension Sector, Bhutan.

Rudebjer, P. G., (**2001**), **Training Course on Participatory On-Farm Experimentation and Integrated Approaches to Land Management**, Bandar Lampung and Kutabumi, Indonesia, 17-23 November 2000, Proceedings, International Centre for Research in Agroforestry, Southeast Asian Regional Research Programme.

Sooby, J., (**2001**), **On-Farm Research Guide**, Organic Farming Research Foundation.

SSC, (**1998**), **On-Farm Trials - Some Biometric Guidelines**, Statistical Good Practice Guidelines, Statistical Services Centre, University of Reading, http://www.rdg.ac.uk/SSC/publications/guides/topoft.html , http://www.rdg.ac.uk/SSC/publications/publications.html

Sundermeier, A., (**1997**), **Guidelines For On-farm Research**, Ohio State University Factsheet Extension, ANR-001-97, http://ohioline.osu.edu/anr-fact/0001.html

Tripp, R. and J. Woolley, (**1989**), **The Planning Stage of On-Farm Research: Identifying Factors for Experimentation**, CIMMYT/CIAT, Mexico, D. F. and Cali.

Veseth, R., S. Wuest, R. Karow, S. Guy and D. Wysocki, (**1999**), **On-Farm Testing – A Scientific Approach to Grower Evaluation of New Technologies**, Pacific Northwest Conservation Tillage Handbook Series No. 9, Chapter 10, Economics and Application of New Technology, http://pnwsteep.wsu.edu/tillagehandbook/chapter10/100999.htm

Wuest, S., B. Miller, S. Guy, R. Karow, R. Veseth, and D. Wysocki**, Use of On-Farm Research by Farmers for Technology Development and Transfer**, Washington State University, University of Idaho, Oregon State University, http://www.ianr.unl.edu/ianr/csas/VOL3-3.htm

Free Statistical Software

AGSTATS02. An easy-to-use, Windows-based software for statistical analysis of simple on-farm field experiments, http://pnwsteep.wsu.edu/onfarmtesting/ or http://pnwsteep.wsu.edu/agstatsweb/index.html

GENSTAT-Discovery. http://www.genstat.co.uk/downloads/discovery/

IRRISTAT/CROPSTAT. Specific for rice field experiments. http://www.irri.org/science/software/irristat.asp

Annex 1 Baseline data needed to document and interpret on-farm trials

On-Farm Trial Description. *Clearly state the goals, objectives, treatments and experimental design of the trial.*

Field History. *Record differences in soil type and other obvious variations within the test site and the previous cropping history. Include crop rotation, tillage practices, previous crop and variety, fertilizer and pesticides applied. Make a diagram showing the layout of the field trial.*

Soil Test and Fertility Program. *Sample soil from the intended harvest areas using university guidelines. Send samples to a reputable laboratory for analysis. Make fertilizer applications based on soil test results. Record the quantity and form of fertilizer used.*

Soil Moisture at Seeding. *If your soil samples are taken near the time of seeding, record the depth to moisture and depth of moisture. Have your soil samples evaluated for available soil moisture.*

Planting Conditions. *Record the crop, variety, seeding rate (pounds per acre and seeds per pound), planting date, soil temperature, type of planter, seeding depth, row spacing, residue levels and any other conditions that might influence the stand establishment and crop production.*

Field Operations and Observations. *Record all field operations in diary format. Take notes on the methods of your field operations, such as the type of equipment, depth of tillage operations and materials applied to either the whole field or to just one treatment.*

Weather. *General observation of growing season weather conditions is all that is required. If practical, place a rain gauge at the test site. After each storm, record rainfall in the rainfall record sheet and empty the rain gauge. A little oil in the rain gauge will prevent the water from evaporating before you can get out to the field to measure it.*

Insects, Weeds and Disease. *Make notes on the presence and density of insects and diseases, date of infestation, and extent or severity of damage. Record similar observations for persistent weeds. Note differences between treatments, if any, due to pests. If pesticide treatments are being compared, take more detailed data to evaluate crop injury and level of control of different pest species.*

Crop Growth and Development. *During the growing season make and record observations of plant growth and development. Record the date each treatment reaches a critical growth stage. It is just as important to record that you see no differences among treatments at a certain growth stage as it is to record obvious differences. For example, critical stages in cereal crop development include emergence, tillering, stem elongation, booting, and heading. Record crop stage at the time of treatment applications, such as spraying or top dressing. When abnormal conditions occur, such as drought, note the differences in plant growth or response among treatments.*

Source: *Miller et al.*

Annex 2 STUDENT'S *t*-DISTRIBUTION CRITICAL POINTS

"2-TAILED" TEST OF DIFFERENCE (EITHER GREATER OR LESS)

Number of Pairs	Degrees of Freedom	α =.20	α =.10	α =.05	α = .01
2	1	3.078	6.314	12.706	63.657
3	2	1.886	2.920	4.303	9.925
4	3	1.638	2.353	3.182	5.841
5	4	1.533	2.132	2.776	4.604
6	5	1.476	2.015	2.571	4.032
7	6	1.440	1.943	2.447	3.707
8	7	1.415	1.895	2.365	3.499
9	8	1.397	1.860	2.306	3.355
10	9	1.383	1.833	2.262	3.250
11	10	1.372	1.812	2.228	3.169
12	11	1.363	1.796	2.201	3.106
13	12	1.356	1.782	2.179	3.055
14	13	1.350	1.771	2.160	3.012
15	14	1.345	1.761	2.145	2.977

When α = .20, the confidence level = 80%. (Chance of wrong conclusion = 20%.)
When α = .10, the confidence level = 90%. (Chance of wrong conclusion = 10%.)
When α = .05, the confidence level = 95%. (Chance of wrong conclusion = 5%.)
When α = .01, the confidence level = 99%. (Chance of wrong conclusion = 1%.)

Source: Exner and Thompson (2005)

Annex 3 Distribution of *t* Probability

n	.9	.8	.7	.6	.5	.4	.3	.2	.1	.05	.02	.01	.001
1	.158	.325	.510	.727	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	.142	.289	.445	.617	.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.598
3	.137	.277	.424	.584	.765	.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	.134	.271	.414	.569	.741	.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	.132	.267	.408	.559	.727	.920	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	.131	.265	.404	.553	.718	.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	.130	.263	.402	.549	.711	.896	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	.130	.262	.399	.546	.706	.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	.129	.261	.398	.543	.703	.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	.129	.260	.397	.542	.700	.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	.129	.260	.396	.540	.697	.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	.128	.259	.395	.539	.695	.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	.128	.259	.394	.538	.694	.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	.128	.258	.393	.537	.692	.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	.128	.258	.393	.536	.691	.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	.128	.258	.392	.535	.690	.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	.128	.257	.392	.534	.689	.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	.127	.257	.392	.534	.688	.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	.127	.257	.391	.533	.688	.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	.127	.257	.391	.533	.687	.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	.127	.257	.391	.532	.686	.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	.127	.256	.390	.532	.686	.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	.127	.256	.390	.532	.685	.858	1.060	1.319	1.714	2.069	2.500	2.807	3.767
24	.127	.256	.390	.531	.685	.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	.127	.256	.390	.531	.684	.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	.127	.256	.390	.531	.684	.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	.127	.256	.389	.531	.684	.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	.127	.256	.389	.530	.683	.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	.127	.256	.389	.530	.683	.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	.127	.256	.389	.530	.683	.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646
40	.126	.255	.388	.529	.681	.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
60	.126	.254	.387	.527	.679	.848	1.046	1.296	1.671	2.000	2.390	2.660	3.460
120	.126	.254	.386	.526	.677	.845	1.041	1.289	1.658	1.980	2.358	2.617	3.373
∞	.126	.253	.385	.524	.674	.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291

Annex 4 Analysis Worksheet for a Paired-Comparison Trial

	Treatments		Difference (C)	Deviation (D)	Deviation squared (D^2)
Blocks	A	B	C = (A - B)	$D = C - C_{avg}$	$D^2 = D \times D$
I					
II					
III					
IV					
V					
VI					
Totals					D^2_{tot} =
Averages	A_{avg} =	B_{avg} =	C_{avg} =		

Step 1. Analysis
Transfer the data from your data collection sheet to the above table.

Step 2. Calculate the variance.
variance = D^2_{tot} /(r-1)

Step 3. Calculate the variance of the means.
variance of the means = variance/r

Step 4. Calculate the standard error.
standard error = square root of the variance of the means

Step 5. Calculate the least significant difference (LSD). Take the answer from step 4 and multiply it by the appropriate t-value.
standard error x t-value = LSD

Step 6. Look for a significant difference. Take the answer from step 5, the LSD, and compare it to the box from the table labeled C_{avg}.
LSD = C_{avg} =

Conclusions:
If the C_{avg} is a negative value, you ignore the negative sign. If the C_{avg} value is less than the LSD, then the two treatments are not significantly different.

Source: Anderson (1993)