

## CHAPTER 6

### METHODS

#### Timing and location

The study consisted of two related parts: a mail survey and a series of personal interviews conducted before and after the survey. This chapter deals first with the interviews before turning to the mail survey.

The study was conducted in Indonesia between 9 March and 8 November 1991. During this time I was based at the Research Institute for Animal Production of the Central Research Institute for Animal Science, Bogor, West Java. I also made two extended field trips outside the Bogor/Jakarta area to pretest questionnaires and collect data: one around West Java, and the other to North and South Sumatra.

#### Interviews

I conducted personal interviews with two types of individuals: extension subject-matter specialists (SMSs) (members of the population from whom the mail survey sample was drawn), and other officials concerned with research-extension links. All interviews were conducted in the Indonesian language.

#### Interviews with extension specialists and local officials

I interviewed a total of 66 local agricultural officials (37 SMSs and 29 others), using a semi-structured, open-ended questionnaire. The non-SMS officials were all directly involved in extension work; most were heads and extension section leaders of *Dinas* offices.

These interviews consisted of two types: personal visits, and group interviews. I visited 42 livestock and food crops personnel employed in a total of 22 provincial and district offices and Agricultural Information Centers (AICs) in three provinces: West Java, North Sumatra, and South Sumatra. I also conducted a series of structured group interviews with 24 livestock officials, SMSs and Agricultural Staff Training Center lecturers attending a livestock training course in Ciawi, West Java. These officials came from eight of Indonesia's 27 provinces (West and South Kalimantan, North and South Sulawesi, Aceh, Bali, South Sumatra and Jambi).

Before each interview, respondents were requested to complete a written questionnaire -- either the pretest (if the interview took place before the survey instrument had been finalized) or the final form of the questionnaire.

The interview questions focused on the respondents' extension activities, information seeking behavior, and problems they faced in their work.

### **Other interviews**

I interviewed a total of 35 other persons connected with research-extension linkages, both as part of the research planning process and after the mail survey had been sent out. These included Ministry of Agriculture officials, Agency for Agricultural Research and Development (AARD) communication personnel, faculty at Bogor Agricultural University and Pajajaran University (Bandung), and expatriates working in agricultural research and administration. I interviewed several of these individuals more than once. Several interviews were with groups rather than individuals.

I also talked with field extension agents and farmers where possible, though I made no attempt to do so on a systematic basis.

Most of these interviews were conducted on an informal basis. Where possible, I made written notes during the interview; otherwise I made tape-recorded or written notes immediately afterwards.

### **Analysis**

I have not made a formal analysis of the interview material for this study. However, the interviews were valuable for two reasons:

- They provide the necessary background for the quantitative part of the study. This was especially valuable in designing the survey questionnaire and in interpreting its results.
- They provide a rich store of examples from which to illustrate the survey findings.

### **Mail survey**

The mail survey used two four-page questionnaires, both in the Indonesian language. The first, the Information Sources questionnaire, explored respondents' information sources and their extension activities (Appendices 1 and 2). I prepared four different versions of this instrument, each focusing on a different information source. The second, or Publications questionnaire, gathered data on problems the respondents faced, their use of AARD publications, and their information needs (Appendices 3 and 4). The two questionnaires were distributed to several different samples of extension specialists, as explained later in this chapter.

For convenience, I have numbered references to questions from the Information Sources questionnaire as IS1, IS2, etc., and those from the Publications questionnaire as P1, P2, etc. Translations of the questionnaires are given in Appendices 2 and 4.

### **Questionnaire design and pretesting**

The instruments were designed after considerable discussion with communication professionals in AARD research institutes, Ministry of Agriculture extension agencies and training institutes, the Ministry's public relations division, Bogor Agricultural University, and representatives of the private agricultural press.

These discussions served to identify problems in research-extension linkages and methods of evaluating them. The pretest questionnaire combined items of theoretical interest with questions of practical value for AARD and extension personnel. The instrument contained a series of identical closed-ended questions about four different source types: AIC publications, AARD publications, Directorate-General recommendations, and other SMSs. Other questions included both closed- and open-ended items. The instrument was edited for language and content by native Indonesian speakers.

Forty SMSs and Ministry of Agriculture local officials involved in extension and training completed the pretest questionnaire during visits to six provincial and district agricultural offices in West Java; the AICs at Ciawi and Lembang, West Java; and the livestock training course mentioned above. Most of the 40 officials also were interviewed individually or in groups about their extension and information seeking activities.

I analyzed the pretest using the Statistical Package for the Social Sciences (SPSS) PC+ version 4 (Norusis/SPSS 1990). This analysis included simple frequency tabulation to ensure adequate variation in responses, and reliability and factor analysis to test the usefulness of information flow and independent variable measures. In order to discover any reliability problems specific to any one of the four sources (AIC publications, etc.), I analyzed each separately. I dropped and added several items and made changes in question wording as a result of this reliability analysis.

Because another question in the pretest indicated that the agricultural press was an important information source, I replaced the items dealing with Directorate-General recommendations with equivalents relating to the agricultural press.

Several respondents complained that the eight-page pretest questionnaire was too long. I therefore dropped some items and split the questionnaire into two four-page instruments: one focusing on information sources, and the other on usage of publications.

I reduced the instrument length by creating four separate versions of the information sources questionnaire -- one focusing on each of the four sources. This meant that it was not possible to compare the four different source types within subjects, as had been planned. Instead, comparisons must be made between groups of respondents responding to each of the four versions. This requires the assumption that the four groups are similar in all respects except for the questionnaire they were sent. I achieved this through random assignment of respondents to groups receiving each version and tested this assumption by comparing variables among the four groups. The four-fold versions also necessitated increasing the sample size to ensure a large enough  $n$  for significance in statistical tests within each version.

At the suggestion of AARD communication officials, several items were added to the publications questionnaire to provide data on specialists' information needs, their use of AARD publications, and their knowledge of AARD-developed innovations.

### **Information Sources questionnaire**

The *Questionnaire about Extension Specialists' Information Sources* (Appendices 1 and 2) aimed to discover why extension specialists use certain information sources.

Four versions of this instrument were distributed to random subsets of the respondent sample. The four versions were identical except for Questions IS7 and 8, on the third page of the questionnaire. I first discuss the questions common to all versions before describing questions IS7 and 8.

#### **Common questions**

Questions IS1 and 2 measured the respondent's personal characteristics, such as institutional affiliation, specialization, education, experience, gender, provenance (urban or rural), and outside work. Question IS2.9 attempted to measure the respondent's wealth by asking if he or she owned various items.

Questions IS3, 4, and 9 asked about the respondent's extension activities: the allocation of time to various tasks, the frequency of seeking information, providing information to clients, and ties with research.

Questions IS5 and 6 asked about actual and desired information flows from 24 sources. These sources included the mass media, scientific and extension publications, categories of individuals such as farmers and extension colleagues, training, and the respondent's own experience. The sources listed were those mentioned by extensionists and other experts during the planning phase of the study; the list was substantially revised in the light of the pretest results. Question IS5 measured how much information respondents obtained from each source. Question IS6 asked which they would like to use in an ideal situation.

Questions IS9 to 11 measured the number of times the respondent had engaged in various extension and information-seeking activities (such as visits to farmers and seeking answers to clients' questions) in a specified period. Several items in this group measured specific linkages between research and extension, such as joint meetings, research seminars, training, and reading of publications. The time period specified ranged from three months to three years, depending on the expected frequency of the activity and respondent's ability to recall it.

Question IS12 focused on the respondent's

attitudes toward centrally developed recommendations. This question was intended to distinguish specialists who viewed themselves as conduits of information generated by higher authorities from those who thought they had more autonomy in generating and adapting technology recommendations to suit local needs.

### **Questions different in each version**

I was unable to discover any comparable studies that tried to measure the Information Flow and FP SCORES variables, so I had no standard set of questions that would give reasonably "clean" indices. I therefore based questions IS7 and 8 on Havelock and Lingwood's (1973) coding sheet for responses to open-ended questions in personal interviews. In addition, space limitations meant it was possible to include only a small number of questions to measure each concept. I attempted to overcome possible reliability problems by testing questions in the pretest and selecting those that had the best construct validity and reliability (as revealed by factor analysis)

The wording of these questions differed among the four versions of the questionnaire. In each version, these questions explored the level of information flow from and the respondent's relationships with one of four sources:

- Publications from the Agricultural Information Centers (Version 1).
- Agricultural newspapers and magazines (Version 2).
- AARD publications (Version 3).
- Information from other extension specialists (Version 4).

Each version posed the same questions about one of the sources, with near-identical wording except for the name of the source type.

Questions IS5 and IS7.1 to 7.4 measured the amount of information flow from the

**Table 6.1** Questions in the Information Sources questionnaire used to measure Information Flow.

Variable	Q #	Question wording (translated) <sup>a</sup>
Information Flow	5.	For extension purposes, how much information do you obtain from [source]? <sup>b</sup>
	7.1	Frequency you read [source]. <sup>c</sup>
	7.2	Frequency you obtain information useful for extension from [source]. <sup>c</sup>
	7.3	Frequency you use information originating from [source] in your extension activities. <sup>c</sup>
	7.4	Usefulness of [source] as a source of extension information for you. <sup>d</sup>

<sup>a</sup> In each of the four versions of the instrument, "[source]" was replaced by one of four sources: AIC publications, agricultural press, AARD publications, and other extension specialists. Question wording differed minimally among the four versions.

<sup>b</sup> Information from AIC publications was measured by question IS5.4, the agricultural press by question IS5.1, AARD publications by question IS5.7, and other extension specialists by question IS5.13. Scale: 1 = very little, 2 = little, 3 = somewhat little, 4 = medium, 5 = somewhat much, 6 = much, 7 = very much.

<sup>c</sup> 1 = never, 3 = once a year, 5 = once a month, 7 = once a week, 9 = every day.

<sup>d</sup> 1 = not useful, 3 = somewhat useful, 5 = rather useful, 7 = very useful.

source, while questions IS7.5 and IS8.1 to 8.15 measured the FP SCORES variables (Familiarity, Proximity, etc.). The order of items in questions IS8.1 to 8.15 was randomized to conceal the nature of the variables of interest, and several items were worded negatively to reduce the risk of monotonous responses. The relationships between the items in questions IS7.1 to 8.15 and the concepts they were intended to measure are shown in Table 6.1 and Table 6.2. I discuss the validity and reliability of these measures later in this chapter.

**Table 6.2** Items in the Information Sources questionnaire and the FP SCORES variables they were intended to measure.

Variable	Q #	Item wording (translated) <sup>a</sup>
Familiarity	8.11	You know the type of information [source] has.
	8.14	You know [source] well.
Proximity	7.5	Nearest place where [source] is usually located. <sup>b</sup>
	8.2	[Source] is difficult to find.
Structure	8.3	Obtaining information from [source] is an important part of your job.
Capacity	8.4	[Source] is the most credible source.
	8.10	[Source] has more complete information than do other sources.
Openness	8.5	Compared with other sources, [source] is easy to use for extension purposes.
	8.9	[Source] provides information in a ready-to-use form.
	8.15	Compared with other sources of information about agricultural technology, [source] is easy to understand.
Reward	8.3	Much information from [source] is not relevant to problems faced by farmers in your area.
	8.8	[Source] often discusses topics that are closely related to conditions in your area.
	8.12	[Source] often provides information that is not new to you.
Energy	8.7	You devote a lot of effort to obtain information from [source].
Synergy	8.1	Information from [source] often does not agree with information from other sources.
	8.6	[Source] usually has information that is currently needed (timely).

<sup>a</sup> See footnote *a* to Table 6.1. Unless otherwise indicated, responses were measured on a scale from 1 to 7 (1 = disagree, 3 = somewhat disagree, 5 = somewhat agree, 7 = agree).

<sup>b</sup> 1 = in own office, 2 = in other room in same building, 3 = in other building, 4 = in same town, 5 = < 50 km away, 6 < 150 km away, 7 = >150 km away.

### **Publications questionnaire**

The *Questionnaire about Publications Used by Extension Specialists* (Appendix 2) evaluated extensionists' problems, their use of AARD publications, and their information needs. A single version of this instrument was produced.

Questions P1 and 2 measured the respondent's personal characteristics: institutional affiliation, specialization, number of years experience, and gender. These questions were identical to items in the Information Sources questionnaire.

Questions P3 and 4 sought to identify major problems extensionists faced in their work as a whole (question P3) and in obtaining information (question P4). The wording of question P3 was based on Sigman and Swanson's (1984) evaluation of problems facing extension systems worldwide. Changes from their questions included (1) simplified wording and translation into Indonesian, (2) division of their category of "linkage" into two: "obtaining information" and "feedback," (3) the addition of the category "reward," and (4) the use of a 7-point response scale instead of one containing 3 points.

Items in question P4 sought to measure the relative importance of various problems mentioned by research and extension personnel during planning and by respondents during pretest interviews.

Questions P5 and 6 compared the frequency respondents read and the usefulness of seven publication types. These included Indonesia's agricultural newspaper (*Sinar Tani*), three publication series produced by the Agricultural Information Centers (*Liptan*, *Buletin Informasi Pertanian*, and booklets), and three AARD publications types (*Warta Litbang*, scientific journals, and books). The pretest had included the farmers' magazine *Setia Kawan*, but so few of the pretest respondents had heard of this magazine that I excluded it from the full survey.

Questions P7 and 8 sought to measure the actual information flow from AARD to respondents. Question P7 listed six items of food cropping technology recently developed by AARD. Respondents specializing in food crops were asked if they had obtained information about each item, and if so, from which source. Question P8 posed the same questions about six livestock technologies to livestock specialist respondents.

Question P9 aimed to identify topics about which specialists most needed information.

### **Sampling**

#### **Sample size**

A sample size is a compromise between the desirable and the possible. The desired size depends on the amount of variation within the population and degree of accuracy and certainty required. The possible size depends on the resources available for the sample collection.

In order to calculate the desired total sample size, it was necessary to determine the

size of subgroups that would be compared (McCall 1982:202). For the Information Sources questionnaire, the design called for a maximum of four subgroups (corresponding to the four versions of the questionnaire) to be compared within three institution types (province, district, and AIC). For the Publications questionnaire, a maximum of three subgroups (province, district, and AIC) were to be compared.

I determined the desired sample sizes for each questionnaire were determined using a procedure given by Steel and Torrie (1980):

$$r \geq 2(Z_{\alpha/2} + Z_{\beta})^2 (\mathbf{s}/\mathbf{d})^2,$$

where

$r$  = The number of replications (= respondents per subgroup).

$Z$  = The value of the standard normal variable.

$\mathbf{s}$  = The population standard deviation. The pretest indicated that standard deviations for Information Flow and the FP SCORES variables were in the range 0.95 to 1.77 on a scale from 1 to 7. I selected 1.77 as the value to use for  $\mathbf{s}$ .

$\mathbf{a}$  = The desired confidence level (set to 0.05). This is divided by two for a two-tailed test.  $Z_{\alpha/2}$  thus equals 1.96.

$\beta$  = The assurance with which it is desired to detect the difference  $\mathbf{d}$ . ( $1 - \beta$  = the power of the test).  $\beta$  was set to 0.2;  $Z_{\beta}$  thus equals 0.85.

$\mathbf{d}$  = The minimum size of the difference to be detected. I set this arbitrarily to equal  $\mathbf{s}$ . In other words, the smallest difference that could be detected with confidence would be equal to the standard deviation of the variable being tested.

These calculations yielded an  $r$  of 15.79 (rounded up to 16) for the smallest subgroup.

**Information Sources questionnaire** Based on the lists of SMSs collected (see below and Table 6.3), about 61% of SMSs were based at district offices, 28% at province offices, and the remaining 11% at AICs. The small numbers at AICs meant it was necessary to stratify the sample, over-selecting AIC specialists in order to obtain enough AIC respondents to respond to each version of the Information Sources questionnaire.

It was neither possible nor necessary to sample province- and district-level specialists separately. It was not possible because of the address lists I had: for numerous SMSs I had an address "in care of" the provincial *Kanwil* or *Bimas* office only, without any indication of whether the specialist worked at the province or district level. It was not necessary because the numbers of province-level specialists (Table 6.3) were sufficient to allow pooling of province and district personnel and draw a common sample, being

reasonably confident of obtaining the minimum of 16 respondents in each subgroup.

Multiplying by the minimum sample size of 16 by the proportion of district to province specialists (956/437) yielded 35, the number of district specialists desired in each subsample. The desired sample size for each questionnaire version was thus  $16 + 35 = 51$ . Multiplying by four gave 202 as a total desired sample size for the Information Sources questionnaire. Multiplying this figure by 1.5 to compensate for an expected response rate of 67% yielded a total sample of 303 province and district specialists.

For the AIC SMSs, I also aimed at a minimum subgroup size of 16. Multiplying this figure by three (only three of the four questionnaire versions were sent to AIC specialists) and by 1.5 to allow for non-responses yielded a total of 72 AIC SMSs to be sampled.

Publications questionnaire Increasing the sample size above the minimum enables more confident estimation of smaller differences between means. The resources available allowed me to send the Publications questionnaire to 150 province and district specialists and 67 AIC specialists. After allowing for expected non-responses, this would yield a minimum sample size of 31 province SMSs, allowing differences of 0.7  $s$  to be detected with the confidence levels given above.

### **Developing a list of specialists**

Respondents for the mailed survey were selected from a list of all Indonesian extension subject-matter specialists. Because no one organization at the national level is responsible for coordinating the specialists, no complete, up-to-date national list of them existed. It was therefore necessary to develop one.

I obtained lists of specialists in each province from various branches of the Ministry of Agriculture: *Bimas*, the Secretariat-General and AAET in Jakarta, and the Agricultural Training Institute (*Balai Penataran dan Latihan Pertanian*) in Ciawi, West Java. These lists varied in their date and the types of information they contained: some included names and specializations but no addresses, while others showed names and addresses but no specializations. The oldest lists dated from 1988, the newest from 1991.

By comparing these lists, I developed an aggregate list for each of Indonesia's 27 provinces, containing the specialists' names, work addresses and areas of specialization. These lists were mailed to the corresponding provincial offices of the Ministry of Agriculture (*Kanwil*, the body responsible for coordinating the SMSs' activities in that province), with a request to update and return it. Updates were received from 10 of the 27 provinces. Revisions for West Java were obtained by interviewing the specialists' coordinator in Bandung. A repeat request to the remaining 16 provinces failed to generate any response. A total of 1560 names and addresses were collected in this way.

Because of the difficulty in compiling this master list and rapid turnover among specialists, doubt must be cast on the validity of the list. An indication of its accuracy can be obtained from the number of corrections received from the 11 provinces providing these: Of 575 SMSs in the original lists sent to the *Kanwil* offices in these provinces, 118 (20%)

**Table 6.3** Numbers of specialists sent questionnaires.

Questionnaire/version	Province <sup>a</sup>	District <sup>a</sup>	AIC	Total
<b>Information sources questionnaire</b>				
AIC publications	22	53	0 <sup>b</sup>	75
Agricultural press	24	53	28	105
AARD publications	25	53	28	106
Other specialists	31	45	27	103
Total	102	204	83	389
<b>Publications questionnaire</b>	44	108	67	219
<b>Grand total</b>	146	312	150	608
Estimated population	437	956	167	1560

<sup>a</sup> Province and district specialists were combined during the drawing of the samples.

<sup>b</sup> None sent because AIC specialists author AIC publications.

had changed address within the province, 29 (5%) were no longer extension specialists, and an additional 119 SMSs (20%) had been appointed.

We might expect similar proportions to apply to the 985 specialists in the remaining 16 provinces which had not provided updated information. Taking the 11 provinces' corrections into account, this would mean that overall, some 184 of 1560 specialists (12%) in the master list had moved, 45 (3%) were no longer extension specialists, and another 185 newly appointed personnel (12%) are not represented.

The total of 1700 specialists these calculations yield is an underestimate in light of President Soeharto's statement in August 1991 that the Ministry of Agriculture then employed 2247 specialists. This may mean that recently hired SMSs are under-represented in the sample. However, this causes little concern because recent hires are least likely to be familiar with local conditions or available information sources or to have developed firm information seeking habits.

### Sample selection

The pretests and interviews suggested that specialists at province and district offices and at AICs obtained information from rather different sources. However, AIC specialists were much less numerous than those at the other two institutions. This meant it was necessary to oversample AIC specialists to obtain a sufficiently large  $n$  for statistical tests comparing them with personnel at the other institutions.

I therefore sampled the two groups separately, selecting nearly 90% of all AIC specialists to complete either one or the other of the instruments, but only one-third of all provincial and district specialists (Table 7.2).

I drew the samples at random from the master list of 1560 names. Specialists who had been interviewed or participated in the pretests were excluded before the samples were drawn. All samples were mutually exclusive. The sampling scheme is summarized in Table 6.3.

AIC specialists Of the 1560 specialists in the master list, 167 worked at AICs, and 1393 at provincial and district level agricultural offices. I randomly selected 67 of those working at AICs as respondents for the publications questionnaire.

Of the remaining 100 AIC specialists, 83 received the Information Sources questionnaire (the other 17 had been interviewed or participated in the pretest). I randomly divided these 83 persons into three groups: 28 received version 2 of the questionnaire (asking about the agricultural press), 28 received version 3 (AARD publications), and 27 received version 4 (other specialists). I did not send version 1 (AIC publications) specialists at the AICs because these specialists are responsible for producing these publications.

Province and district specialists I deleted names of those specialists known to specialize in topics other than food crops and livestock (e.g., in fisheries or estate crops) from the list of 1393 specialists working at provincial and district level offices. Some province lists classified specialists by discipline (e.g., economics, soil and water conservation) rather than by commodity. This made it difficult to determine whether the specialist was engaged in food crops or livestock activities. In such cases, included only those disciplines clearly related to food crops (e.g., agronomy) or to livestock. While this meant that some provinces (e.g., East Java) are under-represented in the sample, I felt this was better than risking the bias of including large numbers of respondents from outside the food crops and livestock areas.

From the remaining exactly 1000 names, I randomly selected 152 to receive the publications questionnaire. I chose another 306 specialists at random from this same list to receive the information sources questionnaire. Of these, a random 75 were sent version 1 (focusing on AIC publications), 77 received version 2 (agricultural press), 78 got version 3 (AARD publications), while 76 were mailed version 4 (other specialists).

### **Distribution logistics**

The questionnaires were mailed on 17-24 September 1991. Respondents at AICs received theirs via the directors of their institutes. In addition to the questionnaire, each mailing envelope contained:

- A copy of a letter from the Directorate of Food Crops Extension (or Livestock Extension, depending on the recipient's area of specialization) indicating the importance of

the research

- A letter from the director of the Central Research Institute for Animal Science, requesting respondents to reply to the questionnaire.
- Brief instructions on completing and returning the questionnaire.
- A stamped, addressed return envelope bearing the letterhead of the Central Research Institute for Animal Science.

Each questionnaire bore a unique identification number. The anonymity of all responses was guaranteed in the instruction sheet.

A small number of specialists completed questionnaires during personal interviews in North Sumatra and West Java. These have been added into the totals shown in Table 6.3.

### **Data cleaning and additional variables**

I entered coded data on computer disk using the dBase III+ database program (Ashton-Tate 1986). I imported the data into SPSS PC+ version 4 (Norusis/SPSS 1990) and cleaned them using the SPSS Data Entry program. All analysis used SPSS PC+ and Lotus 1-2-3 (Lotus Development Corporation 1991).

### **Missing values**

Respondents occasionally omitted to answer a question or gave more than one answer. I coded such instances as missing and excluded them from analysis involving that variable. A response of "Don't know" (for instance in question IS8) was coded as missing.

One exception was necessary because of an error in the question wording in Indonesian: a response of "0" ("don't know") to question IS8.11 ("You know the type of information [source] has") was coded as "Disagree": someone who did not know the type of information could correctly circle either "don't know" or "disagree." It was necessary to recode 25 responses (9%) in this way. (This item was later dropped from the analysis anyway because of reliability problems.)

### **Outliers**

Questions IS3 and 9 to 11 invited respondents to indicate the number of times they had performed a particular activity within a given period. A small number of specialists wrote unusually or unrealistically large numbers in response. I identified such cases using the SPSS EXAMINE procedure (Norusis/SPSS 1990:C -60) and treated them as missing for the variable involved.

### Data transformations

I reversed the coding of negatively worded items in the Information Sources questionnaire to ensure that high scores corresponded to high expected information flows. For uniformity, I converted responses to questions IS7.1 to 7.3 from a scale from 1 to 9 to one from 1 to 7 using a simple linear transformation:

$$\text{new score} = ([\text{old score} - 1] \times 6/8) + 1.$$

### Levels of measurement

Several of the variables used in this study were measured on ordinal scales: questions IS5, 7, 8 and 12, and P3 to P6). Throughout, I assumed that these were measured at the interval rather than the ordinal level. This assumption is justified for four reasons:

- The scales used approximate an interval scale, and their use as such is common in the social sciences.
- It enables the construction of indices composed of multiple variables.
- It allows the use of powerful parametric techniques such as multiple regression, correlation and reliability analysis
- Such techniques as multiple regression, etc., are relatively robust with regard to minor violations of some assumptions.

### New variables

I calculated or entered the following additional variables:

#### Information Sources questionnaire

- Specialization, taken from question IS2.2.
- Wealth index, calculated as the total number of items the respondent reported owning in question 2.9.
- Province, taken from the respondent's address.
- Inner Island location, reflecting the respondent's location in the densely populated and relatively well developed "Inner Islands" of Java, Bali and Madura, as opposed to the sparsely populated Outer Islands.
- Province category, reflecting beliefs commonly held in Indonesia about the level of agricultural development in each province: 2 (high) for Java and Bali, 1 for North and South Sumatra, Lampung, North and South Sulawesi; 0 (low) for all other provinces.

- Total information obtained, the mean of scores from all 24 information sources listed in question 5.
- Information flow and FP SCORES indices (see below).

I was interested in the effect distance has on the SMSs' use of sources. For each respondent, I therefore measured map distances from the respondent's office location to the nearest known source of each type. For simplicity, I used direct straight-line distances (i.e., not road or true air or sea distances). The maximum distance measured was 650 kilometers (from Merauke in Irian Jaya to Jayapura, the provincial capital).

The distance to the nearest known location for each source depended on the source type the respondent was questioned on:

- For AIC publications: the number of kilometers to nearest Agricultural Information Center (located in or near the provincial capital).
- For the agricultural press: All respondents are assumed to receive *Sinar Tani*, so the distance was set to 0 km (but see below).
- For AARD publications: the number of kilometers to the nearest AARD food crops (or livestock, depending on the respondent's specialty) research institute or sub-institute. The distance to the provincial AIC was used for respondents in other specialties, or if the AIC was closer than the AARD site. This was because the AICs are supposed to receive all AARD publications.
- For other SMSs as sources: the number of kilometers to the nearest other SMS in the respondent's commodity group (e.g., food crops). The location and specialty of other SMSs was taken from the master list of 1560 addresses and plotted on the map to allow distances to be measured. Respondents with colleagues in the same institution were assigned a distance of 0 km; those with colleagues in a different institution in the same town were assigned a score of 5 km.

I took the natural logarithm of the number of kilometers to approximate the non-linear change in friction of distance at greater distances. Distances of zero (natural log undefined) were allocated a score of 1 km. The logarithm scores were linearly transformed and reversed to form a 7-point scale, so high scores corresponded to high proximity.

#### Publications questionnaire

- Specialization, taken from question P2.2.
- Number of items on which respondents had information: the sum of answers to the first part of questions P7 and 8, asking whether the respondent had obtained information about six recent food crops (question P7) or livestock (question P8) technologies. Similar sums were calculated for each of the sources named in second parts of these questions.

- Information needs indices for each of the commodity groupings (rows) and disciplines (columns) in question P9. For instance, the index for rice information need was the number of boxes checked in the first row in this question; that for cultivation was number checked in the first column under the "food crops" rubric. I also calculated total information needs for food crops and for livestock information: the food crops index included the 24 food crops boxes, plus those for "estate and industrial crops," "machinery and equipment," and "irrigation." That for livestock included the 24 livestock boxes, plus "livestock fodder plants," "milk handling," and "livestock post-harvest."

### **Initial analysis**

#### **Significance level**

A level of  $p = 0.05$  was selected as the criterion for significance.

#### **Frequencies and point estimates**

I calculated frequencies and appropriate descriptive statistics (e.g., mean, standard deviation, mode, median) for all variables. I ranked comparable variables (such as information flows from various sources, questions IS5.1 to 5.24) by their mean value for ease of analysis and presentation.

The survey used a stratified sample containing two subgroups of specialists (those working at AICs and those at provincial and district offices), and different proportions of the two groups were sampled. Because of this, it is necessary to weight responses according to their proportions of the total population when making point estimates of the whole population of extension specialists (e.g., when estimating the proportion of all specialists who are women). The presentation of results distinguishes between the various samples as necessary. I weighted means when calculating point estimates for the population of specialists as a whole. Table 7.2 gives the weights used for this.

#### **Differences among groups**

I used two main statistical techniques to test for differences among groups of respondents. These were the chi-square for count data and one-way analysis of variance for ordinal (interval) data. If the analysis of variance identified significant variation, I used Student-Newman-Keul's and Scheffé's tests or contrasts using  $t$ -tests to test for differences among pairs of groups (Norusis/SPSS 1990:C-45, Steel and Torrie 1980:177-187).

I used one-way analysis of variance to test for differences among respondents grouped by situation (institutional affiliation, institutional level, province, Inner Island location, province development category) and demographics (sex, education, rural background, specialization, engagement in outside work).

I used regressions to gauge the effect of continuous variables such as wealth (index developed from question IS2.9) on specialists' responses.

## **Influences on information sources**

### **Uniformity of subsamples**

The analysis of factors affecting information flows rests on the assumption that the subsamples of respondents to the four versions of the questionnaire did not differ in important ways from each other. I tested this using chi-square (for categorical data) and oneway analysis of variance (for "interval" data) (see the earlier discussion of the assumption of interval-level data). I tested all major variables (51 in all) other than those in the Information Flow and FP SCORES indices (see below) for differences among the four subsamples.

### **Subsets of respondents**

The survey design required that analysis take into account two categorical variables:

- The respondents' institution (provincial vs district vs AIC). This is necessary because AIC specialists were oversampled and because pretests and interviews showed that these three groups have rather different characteristics.
- The source type each respondent was questioned on (AIC publications vs agricultural press vs AARD publications vs other specialists). This must be taken into account because a missing cell (AIC respondents were not questioned about AIC publications) in the data may otherwise bias results.

It is thus possible to compare subsets of responses: across source types or across institutions. While they are necessary for the reasons outlined above, comparisons among subsets of respondents are also valuable:

- They can reveal differences among subsets. For instance, respondents at AICs may consistently score higher on a variable than their colleagues at district offices. Discovering this is a first step to finding out why.
- They provide internal replications for tests of hypotheses. If a relationship holds over several subsets of respondents, it can be considered robust. If it fails to hold, doubt must be thrown on its wider applicability. Most studies fail to include such internal replications.

### Scale construction

Two problems were encountered in constructing multi-item indices for FP SCORES variables. First, items in some of the *a priori* scales (Table 6.1 and Table 6.2) were relatively poorly correlated with other items in the same scale. This raised questions about the validity of the indices, and meant that scales constructed solely on the basis of the *a priori* groupings would have low reliability. The second problem was that some items in an *a priori* scale were relatively highly correlated with items in other scales. This raised the danger of multicollinearity among independent variables, which would cause problems in interpreting regression results. These problems are discussed further in Chapter 9.

I used five criteria to deal with these problems and ensure a consistent approach to index construction. I included an item in an index if:

- It contributed to the face validity of the index.
- It was significantly correlated with all other items in the index.
- It had an item-total correlation greater than  $r = 0.3$ .
- Cronbach's  $\alpha$  of the index would not be raised by deleting the variable.
- All items in an index loaded onto the same factor in factor analysis.

Details on these criteria are given below.

- Face validity was based on the groupings in Table 6.1 and Table 6.2. I did not consider an item for inclusion in any other scales than the one for which it is listed in these tables. If a choice had to be made between deleting one variable or another, I retained the item with the higher face correspondence with the concept in question.
- Inter-item correlations I calculated simple Pearson correlation coefficients for pairs of items in each *a priori* scale. An item was retained in the index only if it was significantly (at  $p = 0.05$ ) correlated with all other items in the scale.
- Item-total correlation SPSS's RELIABILITY procedure (Norusis/SPSS 1990:C-77) was used to calculate item-total correlations. The item-total correlation measures the correlation between each item in the scale with the sum of all variables in the scale, excluding the item in question. I considered for possible deletion variables with such correlations of less than 0.3.
- Cronbach's  $\alpha$  RELIABILITY also calculates Cronbach's  $\alpha$ , a measure of the overall mutual correlation among scale items. I removed items from an index if doing so would raise the value of  $\alpha$ .

- **Factor analysis** I performed factor analysis on the responses to Questions IS7.5 and 8.1 to 8.15 (the hypothesized influences on information flows). The analysis used principal component extraction and the varimax rotation with pairwise deletion of missing variables. I considered for deletion any item from an *a priori* scale that loaded most heavily on a factor different from others in the scale.

The five criteria were occasionally in conflict. For instance, two variables in the same *a priori* index might load onto the same factor but have a mutual correlation lower than the criterion level. Or an item loaded onto a different factor from others in its *a priori* scale, but deleting it would significantly decrease the Cronbach's  $\alpha$  of the scale as a whole. Judgements in such instances were based primarily on the face validity and item-total correlation of the scale. They are described below in Chapter 9.

I constructed the Information Flow and FP SCORES indices by summing responses to items remaining after evaluating these criteria and dividing by the number of items in the index. The result was a series of indices with a potential range from 1 (low) to 7 (high). Because of the small number of missing scores (a maximum of 13 cases, or 3.4%), I made no attempt to estimate any missing values.

### **Differences among source types and institutions**

I used one-way analysis of variance to test for differences in Information Flow and each of the FP SCORES variables among respondents at the three institutions (province and district offices and AICs) and among the four source types (AIC publications, agricultural press, AARD publications, and other specialists).

I also tested for differences between source types within each institution, and for differences between institutions within each source type. This was necessary because one source-type-by-institution combination was not surveyed: AIC specialists' use of AIC publications. This was because the AIC specialists themselves author these publications. This omission might bias any comparison that includes either AIC specialists or publications.

### **Influences on information flow**

I ran several regression models with the information flow index as the dependent variable, using SPSS procedure REGRESSION (Norusis/SPSS 1990:C-53).

**Individual influences of FP SCORES variables** If an SMS knows a source well, will he or she get a lot of information from it? If we assume that everything else is constant, we can answer this question by testing the relationship between Familiarity and Information Flow.

I did this by running a simple regression using Familiarity as the predictor and Information Flow as the dependent variable. I did this for all respondents.

I wished to know if specialists in the three institutions differed. For instance, is Familiarity with a source important for AIC specialists but not for their district counterparts? If specialists at the three institutions respond in similar ways, then the three groups can be

combined. If not, they must be treated separately. I tested this by running separate simple regressions for each institution for Familiarity as a predictor of Information Flow.

Similarly for the four source types: Familiarity may be important for AIC publications but not for AARD publications, for instance. I tested this by running a similar set of simple regressions for each of the four sources.

In order to gauge the overall effect of Familiarity on Information Flow after controlling for the source type, I ran a multiple regression using Familiarity and three dummy variables representing the source types as independent variables.

To avoid the potential bias due to the missing source-type-by-institution combination (AIC specialists and AIC publications), I also ran a series of simple regressions using Familiarity to predict Information Flow for each of the 11 remaining combinations of sources and institutions. I used the  $z$ -test (Walker and Lev 1953:255) to discover differences among regression coefficients within each source type or institution.

I repeated the above analyses using each of the other seven FP SCORES indices as a predictor of Information Flow.

Combined effects of FP SCORES variables How do the FP SCORES variables affect Information Flow when all are considered together? To answer this question, I ran several multiple regression models. I first investigated the correlations among the FP SCORES variables to check for possible multicollinearity problems. I then ran a multiple regression model using the FP SCORES indices as predictors of Information Flow. First, FP SCORES variables were entered by stepwise forward regression with a minimum significance level of  $p = 0.05$  for inclusion in the equation. Finally, all remaining FP SCORES variables were entered into the model.

The same model was run a second time, controlling for the source type and institution. Dummy variables for these variables were entered into the equation before the stepwise inclusion of the FP SCORES indices. Entry of remaining FP SCORES variables was then forced.

Comparison among sources and institutions To discover any relationships specific to individual sources or institutions, I also ran multiple regressions for seven subsets of cases: for each of the four sources and the three institution types surveyed. For each subset, I ran these regressions with and without controlling for source or institution.

Violations of assumptions I inspected scatterplots and histograms of residuals for evidence of violations of regression assumptions such as non-normality of residuals and heteroscedasticity. I used variance inflation factors and other collinearity diagnostics to detect multicollinearity in the model.