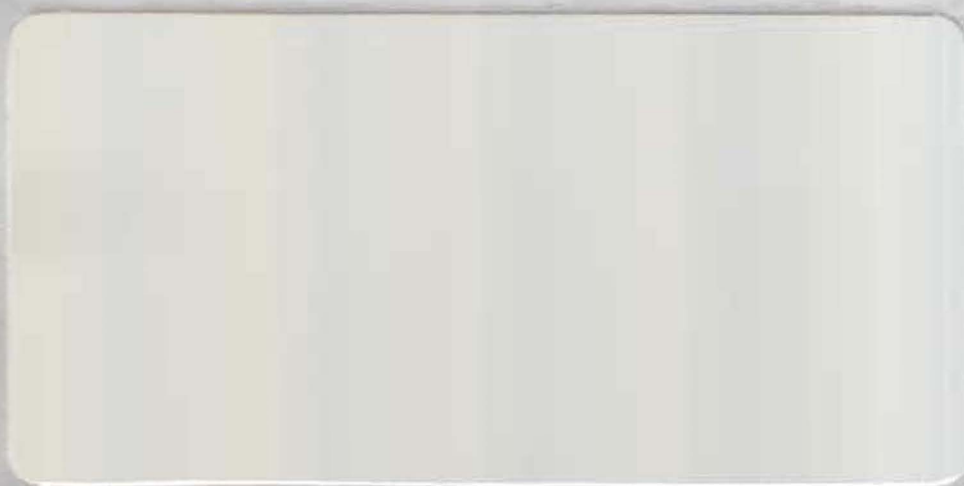Departemen Pendidikan dan Kebudayaan

The British Council

Contract No: CNTR97 2735A

Curriculum Development Centre
Curriculum Capacity Project
Department for International Development

Pusat Pengembangan Kurikulum
Jl. Gunung Sahari Raya No.4, Jakarta 10002
Tel: (021) 350 9022, Fax: (021) 345 3440, e-mail: emsweet@indo.net.id

CURRICULUM CAPACITY PROJECT

**CREATING METHODOLOGIES TO
INVESTIGATE THE CONTENT, PRODUCTION
AND ROLE OF QUARTERLY TESTS
IN INDONESIA SCHOOLS**

Richard S. Sandman

August 1999

Contents

1. Introduction

Quarterly tests are typically given in Indonesian classrooms three times a year to assess the learning progress of the children. Locally, these quarterly tests are called cawu tests, after the Indonesian word "caturwulan" meaning academic quarter. These cawu tests are meant to reflect the syllabi and curricula currently in use in the schools, namely Curriculum 94.

However, the present syllabi and curricula are now in the process of being revised to reflect the current interest in basic competencies (kemampuan dasar). When this work is completed, it will also be necessary to alter the cawu tests to better represent this new curricular emphasis. To determine the kinds of changes needed for these tests, and the best way to implement these changes, it seems useful to investigate the current status of cawu tests in Indonesia's education system. In particular, the following questions are of interest:

1. How well do the current cawu tests match the current curriculum and textbooks?
2. What kinds of questions are used in the current cawu tests?
3. How do the types of questions used influence what goes on in the classroom?
4. By what process are the current cawu tests produced?
5. What use is currently made of the cawu test results?

The thought is that by having answers to these questions, education authorities will be in a better position to decide how to go about altering the cawu tests so that they will be maximally effective in representing and furthering the new curricular ideas, and in encouraging an improved teaching-learning process within the classroom. Improved procedures for producing cawu tests, and for using cawu test results, may also result from possessing accurate current information on these activities.

The purpose of the work presented here was to develop methodologies for obtaining answers to these questions, and to try out these methodologies. Below, the new methodologies are described, along with the procedures employed in developing them and trying them out. Some initial findings from applying the new methodologies are also given. Finally, some recommendations regarding follow-through activities are offered.

The work described here was carried out by eight staff members of the Curriculum Development Centre (Pusat Kurikulum) of the Office for Research and Development (Balitbang) of the Indonesian Ministry of Education and Culture. I served as a consultant and adviser to the group. The activities were carried out over the three-week period from 26 July to 13 August, 1999.

2. Initial Discussions

Initially the group, with my guidance, held a discussion concerning the purposes of educational assessment, and the uses to which educational assessment data could be put. The distinctions between formative and summative uses of data were described. We then related the results of this discussion to the possible uses of cawu tests within the Indonesian education system. The importance of discovering which of these uses were actually being implemented was confirmed by the group, showing the need to answer question 5 above.

We then considered the form of questions that could be included on a cawu test: multiple-choice, essay, short answer, and matching. We then talked about the levels of thinking that different questions require, and reviewed the six levels of cognitive thinking included in Bloom's Taxonomy. It was admitted that the kinds of questions included on the cawu tests might possibly influence the teaching program in the classroom (question 3, above), and thus that it was important to discover the types of questions included on current cawu tests (question 2, above). An exercise followed where each member of the group classified the questions of one cawu test according to Bloom's Taxonomy.

It was also considered important to determine how well cawu tests match the current curriculum and textbooks – to see how far off the mark these tests are in covering required material (question 1, above). Where changes in cawu tests are in order, effective ways of introducing these changes need to be devised. To do this successfully, information on the current process for producing cawu tests would seem to be required (question 4, above).

Thus, although the potential questions for investigation were drawn up in advance, the group eventually came to agree on their importance for enabling improvements to be made in the cawu test system.

A discussion then ensued on the best way to find out answers to these questions. It was decided to employ two methodologies, one to find answers to the first two questions above, and the other to answer the last three questions. The first methodology was to be a review process applied directly to the cawu tests themselves, determining the types of questions used, and the match between the included questions and the relevant curriculum and textbook. The second methodology was to be an interview schedule, administered to selected education personal to find out how cawu tests are produced, how they affect the teaching-learning process, and how the results from them are used. Details regarding the development and tryout of these two methodologies, along with some initial findings, are given in the next two sections.

3. Methodology 1: Reviewing the Content of Cawu Tests

Development and Description of the Methodology

This methodology was devised to check on the match between a specific cawu test and the curriculum and textbook material that the test is supposed to be measuring. Also included is an enumeration of the types of questions appearing on the test. After these information needs were discussed, one member of the Pusat Kurikulum (Puskur) team I was working with prepared a draft form for reviewing a cawu test. This form was discussed, revised, and expanded as a group effort by the Puskur team, and a meaningful way of filling in the form was determined. I provided some suggestions along the way.

The form arrived at for implementing this methodology is given in Appendix A. The principles behind its use are really quite simple. In the left-hand column (column 1) is written the number of the cawu test item being looked at (1, 2, 3, ...), while column 2 gives the basic material covered by the item. Column 3 gives the cawu when this material is meant to be covered. This is necessary because a test given at the end of cawu 2, for example, is supposed to cover material from both cawu 2 and cawu 1. Column 4 gives the format of the question – multiple-choice, essay, short answer, or matching – and column 5 gives the relevant cognitive level from Bloom's Taxonomy.

In Indonesia, the mandated curriculum for each subject is given in outline form in a document called the GBPP (Garis Besar Program Pengajaran or teaching program outline). The specific objectives of instruction are given in the GBPP, grouped by academic term (cawu). In most subjects these objectives are numbered. Column 6, then, is to be filled in with the number of the specific objective in the GBPP that the item currently being looked at is testing. Where numbers are not assigned to the objectives in the GPBB, it is suggested that some numbering system be superimposed on the document to facilitate use of this methodology. Of course, if no relevant objective is found in the GBPP, column 6 is left blank for that particular item in the cawu test. In this case, it is not possible to determine an appropriate cawu for the material tested in the item, so column 3 would also be left blank.

Columns 7-10 evaluate the match between the cawu test items and the textbook being used in the class. If the material of the test item is found in the textbook, then the relevant chapter and page are noted in column 7. In this case a comparison is then made between the difficulty of the cawu test item, and the text and exercise items found in the textbook that relate to the same material. Depending on whether the cawu test item is easy, medium, or difficult in relationship to the textbook items, column 8, 9, or 10 is checked.

After all items in the test have been analysed in this fashion, totals are prepared at the bottom of the columns. Thus, at the bottom of column 4, we have the number of items in each item format, while the bottom of column 5 gives the number of items in each Bloom classification. The bottom of column 6 gives the number of items that fit, and do not fit, the GBPP, calculated simply by counting the items for which relevant objectives were found, and not found. Similarly, the bottom of column 7 gives the number of test items for which relevant material was found in the textbook, and not found. Finally, the totals of easy, medium, and difficult items at the bottom of columns 8, 9, and 10 show how the difficulty of the cawu test items compares with the difficulty of the items in the textbook. It is suggested

that the totals at the bottom of the columns also be prepared separately for the items of each cawu (column 3), so as to provide information on individual cawu coverage.

Thus the row of totals at the bottom of the table gives a summary of the types of items included in the cawu test, and the degree to which the cawu test items are represented in the GBPP and the textbook. From this summary we still don't know the extent to which the cawu material included in the GBPP and the textbook is represented in the cawu test. To obtain this information, we use the simple summary tables on the second page of the form. For the GBPP, for each cawu separately, we count the number of specific objectives that are tested by the cawu test items, and the number that are not. The former number is easily obtained by counting the number of specific objectives that are filled in on the first page of the form under the relevant cawu. The second number is the number of remaining objectives listed in the GBPP under that cawu – those not covered by the cawu test. Similarly, for each cawu separately, we count the number of topics in the textbook that are represented in the cawu test, and the number that are not.

Thus from the summary row on the cawu test on page 1, and the summary information on the GBPP and textbook on page 2, we can get a good idea of the overall degree to which the cawu test matches the curriculum and textbook.

Tryout of the Methodology

Due to limitations of time, and the lack of availability of team members, this methodology was tried out only minimally during the three-week period of my consultancy. The methodology was applied to only two cawu tests, one in Indonesian language for the second quarter of class 3 of primary school, and one in Social Studies for the second quarter of class 5 of primary school. The two team members involved reported that they could carry out the cawu test review without much difficulty. There were, however, a few problems that were mentioned or noted:

1. There was a tendency to carry out the methodology first without using the form, and then to fill in the form afterward.
2. One of the reviewers used an additional category "composition" in specifying the item format in column 4. There is some question as to how this type of item differs from an essay item.
3. There was some confusion as to what should be entered into column 6 of the form. The reviewer of the Indonesian language test was entering "sesuai" and "tidak sesuai" (suitable and not suitable) instead of the number of the relevant specific objective from the GBPP. This is understandable, as the GBPP for Indonesian language has no numbering system for its objectives. Some such numbering system would have to be devised to fully utilise this methodology.
4. There was uncertainty over whether the "easy" and "difficult" labels for columns 8 and 10 should be interpreted as "too easy" and "too difficult," given the textbook material, or simply as within the easy and difficult ranges of the textbook material.
5. There were difficulty ratings given to some cawu test items for which relevant material could not be found in the textbook. This shows some confusion over the appropriate use of the methodology
6. In using the summary for the GBPP on page 2 of the form, the reviewer of the Indonesian language test found many objectives in the GBPP that were not

represented on the test. However, several of these objectives were for speaking and listening skills, not normally included on current cawu tests. There is a question as to whether inclusion of these objectives in the tally biases the review results in an undesirable way.

7. In using the summary for the textbook on page 2 of the form, one reviewer interpreted the term "topic" to mean chapters of the textbook. Information on the fit of the textbook to the cawu test was thus very general. It is possible that the use of smaller units of textbook material as topics might produce more useful information.

It seems likely that other such problems will reveal themselves upon further tryout of the methodology. Thus this further tryout needs to be carried out, the problems revealed discussed, and the methodology itself refined accordingly. Further trials need to include cawu tests from other subjects and from other levels of the education system, including the junior secondary school. This tryout and discussion process should continue until a standard agree-upon procedure for applying the methodology is arrived at. Then, anyone new who is planning to use the methodology to review cawu tests will need to be trained in this procedure.

The overall conclusion from the tryout, however, seems to be that, after some refinement, this methodology could be meaningfully used to review the content of cawu tests.

The summary results from the two cawu test reviews mentioned above are given in Table 1. These are provided only to serve as an example of the kind of information obtainable from using this methodology. From the information on the Indonesian language test, we may conclude that the items are about evenly split between multiple-choice and short answer, that the Knowledge category of Bloom's Taxonomy predominates, that about three-quarters of the cawu test items fit the GBPP and the textbook, and that of the items fitting the textbook, most are of medium difficulty. We also see, however, that a large number of specific objectives in the GPBB, and topics in the textbook, are not covered in the cawu test.

Similarly, from the information on the Social Studies test, we see that multiple-choice items predominate, all items are in the Knowledge category, most items fit the GBPP and the textbook, and that the items fitting the textbook average somewhat on the easy side. We see, also, that only one of the test items pertains to cawu 1 material. Perhaps for this reason, the GBPP objectives and textbook topics from cawu 1 are poorly covered by the test, while the objectives and topics from cawu 2 show much better representation.

Again, as is quite obvious, these interpretations are given only for illustrative purposes, and cannot be applied to anything beyond the two cawu tests reviewed.

Table 1

Summary Results from the Tryout of Methodology 1

A. Cawu Test in Indonesian Language for Primary School, Class 3, Cawu 2

	Item Format	Bloom's Taxonomy	Fit to GBPP	Fit to Textbook Present	Easy	Medium	Hard
Total Items=41	MC=20 SA=20 Comp=1	Knowl=34 Compr=4 Synth=3	Yes=33 No=8	Yes=30 No=11	0	28	2
Cawu 1 Items=4	MC=1 SA=3	Knowl=2 Synth=2	Yes=4 No=0	Yes=2 No=2	0	2	0
Cawu 2 Items=29	MC=17 SA=12	Knowl=25 Compr=4	Yes=29 No=0	Yes=27 No=2	0	25	2
No Cawu Items=8	MC=2 SA=5 Comp=1	Knowl=7 Synth=1	Yes=0 No=8	Yes=1 No=7	0	1	0

	Summary for GBPP	Summary for Textbook
Cawu 1	Specific Objectives = 15 Present on Test = 4 Not Present on Test = 11	Topics =7 Present on Test = 3 Not Present on Test = 4
Cawu 2	Specific Objectives = 18 Present on Test = 7 Not Present on Test = 11	Topics = 7 Present on Test = 2 Not Present on Test = 5

Key: MC = multiple choice
   SA = short answer
   Comp. = composition or essay (possibly with a stimulus)

B. Cawu Test in Social Studies for Primary School, Class 5, Cawu 2

	Item Format	Bloom's Taxonomy	Fit to GBPP	Fit to Textbook Present	Easy	Medium	Hard
Total Items=50	MC=35 SA=10 Essay=5	Knowl=50	Yes=43 No=7	Yes=43 No=7	20	23	0
Cawu 1 Items=1	MC=1	Knowl=1	Yes=1 No=0	Yes=1 No=0	0	1	0
Cawu 2 Items=42	MC=30 SA=8 Essay=4	Knowl=42	Yes=42 No=0	Yes=39 No=3	20	19	0
No Cawu Items=7	MC=4 SA=2 Essay=1	Knowl=7	Yes=0 No=7	Yes=3 No=4	0	3	0

	Summary for GBPP	Summary for Textbook
Cawu 1	Specific Objectives = 16 Present on Test = 1 Not Present on Test = 15	Topics =13 Present on Test = 1 Not Present on Test = 12
Cawu 2	Specific Objectives = 19 Present on Test = 13 Not Present on Test = 6	Topics = 12 Present on Test = 11 Not Present on Test = 1

4. Methodology 2: Determining How Cawu Tests Are Produced and Used

Development of the Methodology

This methodology is an interview schedule that is to be administered to education personnel to find out how cawu tests are developed and used, and how they affect the teaching-learning process. The interview schedule was constructed by the Puskur team through a step-by-step process, with my guidance. The first step was to decide on the types of information we wanted to obtain from the interview, and the likely respondents. A provisional list of ten categories of relevant information was developed, while the respondents were seen to be primarily teachers, head teachers, and school supervisors who were somehow involved in the cawu test implementation process.

The team was then divided into three groups, with each group having responsibility for writing possible interview questions for particular assigned categories of information. The large group then came together again, reformulated and ordered the categories of information, and discussed each interview question that had been prepared, adding additional questions along the way. The result was an initial draft of the interview schedule. A further review resulted in some additional changes, and it was then felt that the interview schedule was ready to be tried out.

Tryout and Refinement of the Methodology

The interview schedule was first tried out in Tangerang District (Kabupaten), in West Java, on 4 August 1999. A total of 12 education personnel were interviewed by the Puskur team: three teachers from a local primary school, representing classes 1, 3, and 5; the head teacher of that primary school; five teachers from a local junior secondary school, representing five core subject areas (mathematics, science, social studies, Indonesian language, and English language); the head teacher of that junior secondary school; a primary school supervisor from the district education office; and the head of the local sub-district education office (Ka kancam). Chosen as respondents were only those who were involved in the cawu test development or management process. The hope was to also interview a junior secondary school supervisor, but we found that these were working out of the provincial education office (Kanwil), rather than out of the district education office (Kandep). While in the field, a sample of cawu tests used in the district was also collected.

I observed several of the interviews in progress, and, in general, they appeared to run smoothly. An initial problem occurred when many of the team members paraphrased questions in their own words, instead of using the form of the questions that had been agreed upon. I pointed this out, and henceforth the interviewers used the appropriate wording for the set questions. They could, however, use their own words for any needed follow-up questions. The interviews averaged about one hour in length.

The following day we met back in Jakarta to discuss the Tangerang interviews. It was felt that some of the questions were not entirely clear to the respondents, so the wording of these was changed. A few new questions were added to obtain some additional information that was considered important. In addition, some notes to the interviewer were added to the schedule, to cover situations where, based on previous responses, certain questions didn't need to be asked.

The revised interview schedule was tried out in Bekasi District, in West Java, the following day, 6 August. Here again, 12 education personnel were interviewed: three teachers from different primary schools, representing grades 2, 4, and 6; one primary school head teacher; five teachers from a local junior secondary school, representing the five core subjects; the head teacher of that junior secondary school; a primary school supervisor from the district education office, and the head of primary education (Kasi Dikdas) at the district education office. Again, junior secondary school supervisors were not available. As before, a selection of cawu tests was obtained from the field.

These interviews went even more smoothly than those in Tangerang, due to the improvements made in the interview schedule, and the increased experience with that schedule of the participating team members. Again, the interviews averaged about an hour in length.

Upon returning to Jakarta that same day, we went through the interview schedule once again, and made some additional changes and clarifications based on the Bekasi experience. The resulting version can be considered the final one from this exercise. It is attached as Appendix B. The following list gives the translation of the main categories of information listed within the interview schedule:

I.   Respondent Identity
II.   Respondent Involvement and Role
III.   Cawu Test Development Process
       A. Classes, Subjects, and Level of Management
       B. Sources and Materials
       C. Item Development: Persons Involved and Their Roles
       D. Item Development: Time Involved and Place Carried Out
       E. Item Tryout and Revision
       F. Obstacles in the Test Development Process
IV.   Item Difficulty
V.   Item Suitability
VI.   Influence of Items on Teaching-Learning Process
VII.   Use of Cawu Test Results
VIII. Training in Test Development
IX.   Other Obstacles

Some Initial Findings

Members of the Pusat Kurikulum team tallied the responses from the interviews at Tangerang and Bekasi. This was done separately for the primary school respondents and the junior secondary school respondents. The tallies from the two districts were then looked at together, and two summary reports were prepared, one for each level of schooling. It turned out that the summary report for primary school was in somewhat more detail than the report for junior secondary school.

The basic content of these reports is given in Table 2. The information in the table is organised in accordance with the categories found within the interview schedule, and listed above. It must be remembered that this information comes from only two districts out of a total of more than 300 districts in Indonesia. Indeed, even within the two districts themselves, only a few schools and a handful of education personnel, unscientifically

selected, were utilised. Thus the information in Table 1 can be considered only as illustrative of the type of information obtainable through applying this methodology. It should not be considered as indicative of the status of cawu testing in the country. The purpose of this activity was to develop and try out the methodologies. The methodologies have to be applied in a more general fashion in order to obtain information that can be acted upon.

It is worth noting at this point that the participating schools and respondents for the interviews were selected by the district offices themselves, and not by the investigators. This may possibly explain to some degree the seemingly conforming type of responses that were sometimes heard. In addition, the decision to interview only personnel involved in the cawu test process may also have affected the results, as these people may have a stake in defending the system. Thus, when the methodology is generally applied, it would probably be better to use a more scientific selection procedure, controlled by the investigators, and including as participants some personnel not closely connected with the cawu test process.

Table 2

Some Initial Findings from the Tryout of Methodology 2

Category	Primary School	Junior Secondary School
I. Respondent Identity	1 head of primary education in district office 1 head of sub-district office 2 supervisors, 2 head teachers, 6 teachers In general, teaching experience of more than 10 years	2 head teachers, 10 teachers Most with either first degree or diploma, and with teaching experience of more than 15 years
II. Respondent Involvement and Role	Head of primary education teaches how to write good items; supervisor checks on quality; teachers write items Criteria for teachers as item writers: understand curriculum, master material, can write items, good Indonesian Head teacher or KKG (Teachers Working Group) recommends teacher to supervisor; head of primary education decides	Criteria for item writers: senior teacher with experience, active in MGMP (Subject Teachers Working Group), master subject material, received training often
III. Cawu Test Development Process A. Classes, Subjects, and Level of Management	All subjects and classes (1-6); for class 1 begins with cawu 2 Managed by head of primary education, helped by supervisors; organisation relatively fixed	All subjects and classes except class 3, cawu 3 (EBTANAS given instead) Carried out by school for cawu 1; carried out by KKKS (Head Teachers Working Group) for cawu 2 and 3

Category	Primary School	Junior Secondary School
B. Sources and Materials	GBPP, textbook, supporting text, guide to writing questions, guide to good Indonesian Material and learning objectives to be tested, and types of items, decided as a group; then kisi-kisi constructed Majority of items multiple-choice – influenced by EBTANAS	GBPP, textbook, supporting text, kisi-kisi, book of test items (item bank) for subject Sources very useful as a model for writing items Material/learning objectives and item types decided by provincial education office (Kanwil) because considered as training for EBTANAS Kisi-kisi made by item writing team or by MGMP
C. Item Development: Persons Involved and Their Roles	Work in group of 3 for each subject; group responsible for all classes (1-6) Test submitted to team of head teachers and supervisors (7 people) for review and revision Test writers and reviewers relatively fixed for each cawu	For cawu tests managed by KKKS, 3 roles: Item writer – choose essential material, make kisi-kisi, write items Co-ordinator of item writers – co-ordinate collection of items, review items, deliver to KKKS, watch over security of items Cawu test manager (from KKKS) – prepare accommodation, collect funds Item writing for each subject and each class is done by 2 people; each person prepares about 40 items Co-ordinator then reviews items and returns them to item writer for revision
D. Item Development: Time Involved and Place Carried Out	Usual time for preparing test: 3-7 days Initial meeting to decide on work and divide it up – work done at home In general, they want special place for preparing tests – safe, calm, can concentrate	Usual time for writing items: 2-7 days Done at a school fixed by the cawu test manager In general, item writers want special place to prepare tests where item security is attended to

Category	Primary School	Junior Secondary School
E. Item Tryout and Revision	Cawu test implemented directly after being reviewed and revised – not trial-tested	No tryout because of loss of security for the items
F. Obstacles in the Test Development Process	Cawu test restricted to two pages – limits creativity of teachers Errors in printing and errors in pictures	Too little time available for writing items
IV. Item Difficulty	Item writers try to construct easy items 25%, medium 50%, difficult 25%; test like this too easy for city area but hard for rural parts – Indonesian not used as much as in city Item difficulty fits what has been taught Test results are satisfactory	Items are considered easy for the students Item difficulty suitable for the ability of the students and the material that was taught
V. Item Suitability	Test fits GBPP, but not always textbook since different textbooks have different emphases and ways of presenting Test constructed to be suitable to the stage of development of the children	In general, items fit the GBPP, the textbook, and the level of development of the students
VI. Influence of Items on Teaching-Learning Process	Item types used in cawu tests influence teaching activities In general, teaching neglects process and emphasises product Teachers will drill students on item types used in cawu tests	Most respondents said the item types very much influence the teaching activities

Category	Primary School	Junior Secondary School
VII. Use of Cawu Test Results	Report to parents, measure success of teaching, find out how well curriculum targets are achieved, measure achievement of children, find out which schools need improvement Head teachers, supervisors, teachers, and pupils use cawu test results Test results kept at school, sub-district office, and district office Suggestions: Try out test before using, and use results to make test better; help schools with unsatisfactory test results to improve	Most respondents said: make a report, find out the progress of the students All respondents said that teachers use the test results Test results kept in grade book and in the school's register of grades
VIII. Training in Test Development	No special training for writing tests In KKG sometimes have discussion about writing questions, but not deep enough	In general, training in developing tests given each cawu at MGMP in the form of general training Key teachers give the training
IX. Other Obstacles	School quality varies, so hard to decide on difficulty of questions to include Teachers' skill at writing items not satisfactory – many copy items from other books In rural schools contributions are a problem – school can't do cawu tests well because too few teachers, and teacher quality is low	None reported

5. Constraints

The main constraint in carrying out this consultancy was the frequent unavailability of members of the Pusat Kurikulum team. Other work activities continually intruded upon their commitment to this one. Some members were assigned work in other cities. Others were pulled away for what was considered to be higher priority work. In particular, a previously-unannounced general Puskur activity taking place during the third and final week of the consultancy severely limited the amount of work that could be done on our task during that week. For this reason Methodology 1, for reviewing cawu tests, could not be fully tried out, and subsequent refinements made, during the period of the consultancy.

6. Recommendations

Below are some recommendations for follow-through activities to the work that has been presented here:

1. Methodology 1 needs to be further tried out and refined. Initially some discussion should take place concerning the problems mentioned in Section 3, above, and provisional decisions made about how these can be resolved. Then the methodology should be tried out with a variety of cawu tests from different locations, different levels of primary school and junior secondary school, different subject areas, and different cawu (1, 2, and 3). Puskur already has in its possession sufficient cawu tests to carry out this tryout.

Further problems and issues in implementing the methodology are likely to arise from these tryouts. These need to be discussed, decisions made, and the methodology refined accordingly. The resulting product should be an agreed-upon methodology for cawu test review that can be applied unambiguously.

2. To obtain the desired information about cawu testing – information that can be acted upon – the two methodologies need to be applied to a representative sample of provinces, districts within provinces, and, perhaps, sub-districts within districts. Thus such a sample of administrative units needs to be selected. As time and resources for carrying out this investigation will undoubtedly be limited, the sample selected should be no larger than is necessary to provide useful generalisable information.

3. Puskur staff that are going to be involved in this general application of the methodologies will need to receive some training on how the methodologies are implemented. An explanation of the cawu test review process, plus a practice review of a couple of cawu tests, would probably be sufficient training for Methodology 1. A discussion of the interview process, plus a careful reading of the interview schedule, would probably suffice as training for Methodology 2.

4. When a selected education office is visited, whether at the provincial, district, or sub-district level, both methodologies can be applied. The interviews can be conducted during the visit, while a selection of cawu tests can be collected for review at a later time. As the applicable textbooks are used in the review process, information about these textbooks should be obtained during the visit. For tests that are actually reviewed, a copy of the textbook must somehow be procured.

5. During each visit, a representative group of education personnel should be interviewed. Where possible, the investigators should choose the respondents, rather than the education office. Good schools and poor schools should be represented, city schools and rural schools, primary schools and junior secondary schools. Teachers interviewed should be from different grade levels and from different subject areas. Some personnel interviewed should be closely connected with the cawu testing process, while others should not. While it is not possible, nor advisable, to cover all varieties of personnel in a given visit, the idea is to get a good sampling of these varieties, so that the information obtained provides a reasonably accurate picture of cawu testing in the area.

6. The majority of interview respondents will be teachers, head teachers, and school supervisors. It should be remembered that, in the case of our fieldwork, the junior secondary supervisors were working out of the provincial education office, rather than the district office. Also of importance is the head of primary education at the district office. Mentioned as a possible source of information, although not interviewed in our case, is the head of educational planning (Kasi PRP) at the district office. Finally, as the various working groups of teachers and head teachers were often mentioned in our interviews (KKG, MGMP, KKKS), it would be well to include representatives of these groups among the respondents.

7. For each administrative unit visited, the cawu tests chosen for review under Methodology 1 should be a representative sample of those in use in the area. Tests included should be from a variety of locations, from different levels of primary and junior secondary school, from different subject areas, and from different cawu. Here again, as time does not allow the review of all possible types of cawu tests from a given administrative area, the idea is to get a good sampling of these types, so that summary information about cawu tests in the area will be reasonably accurate.

8. As both methodologies are applied in each administrative area visited, it is possible to investigate the relationship between the way cawu tests are produced and the characteristics of the cawu tests that result. It is suggested that some such study be carried out on the information obtained. By so doing, it might be possible to discover some key variables in the production process that affect the quality of the product. If so, this information can be useful in planning for change in cawu testing.

9. The information collected from the various administrative units has to be integrated, so that a coherent summary of cawu testing in the country can be prepared. Such a summary will include an accounting of the variety of ways cawu tests are produced and used, as well as the varied features of the resulting products. Any relationships found between process and product should also be described.

10. The next step is to use the information obtained from applying the two methodologies in planning for change in cawu testing. The summary of cawu testing mentioned above needs to be disseminated to the appropriate education officials. These need to come together to decide on the kinds of changes needed in cawu tests, given the revised curriculum being prepared, and the types of teaching-learning activities desired in the classroom. Then strategies for implementing these changes need to be devised, employing as a basic resource the information on current processes for producing and using these tests.

7. Conclusion

Toward the end of the three-week period of my consultancy, the Puskur group I was working with presented, in an open seminar, an account of the activities they carried out, and the results of their work. An opportunity for questions followed. One of the principal questions asked was "What is all this for? What is the use of having these methodologies?" Well, the answer, of course, is that there is no use at all if the methodologies are not applied, or, even if applied, if the information collected is not used in making the cawu testing system more effective.

The hope, then, is that these methodologies will be applied, in accordance with the recommendations given above, and that cawu tests will then be improved to better accord with a revised curriculum, and to stimulate desirable kinds of teacher-learner interaction in the classroom. In that case, the work presented here will be seen as a useful first step in creating some meaningful forward progress in the Indonesian education system.

Appendix A

Form for Reviewing Cawu Tests
(Methodology 1)

METODOLOGI REVIEW TES CAWU

No. Soal	Materi soal	Cawu	Bentuk soal	Taxonomi Bloom	Sesuaian dengan PB/ Pembelajaran di GBPP	Kesesuaian dengan buku pelajaran			
						Bab, hal	mudah	sedang	sulit
1	2	3	4	5	6	7	8	9	10
1									
			PG = Uraian = Jwb singkat = Menjodohkan=	1 = 2 = 3 = 4 = 5 = 6 =	Sesuai : Tidak sesuai :	Sesuai: Tidak sesuai:			
					Jumlah:	Jumlah:	Jumlah:	Jumlah:	

Keterangan:
1. Kolom 1 diisi dengan nomor soal
2. Kolom 2 diisi dengan materi soal
3. Kolom 3 diisi dengan nomor cawu yang berkaitan dengan soal (1, 2, atau 3)
4. Kolom 4 diisi dengan bentuk soal (uraian, pilihan ganda, jawaban singkat, menjodohkan)
5. Kolom 5 diisi sesuai dengan tingkap taxonomi Bloom (1-6) (1= pengetahuan, 2 = pemahaman, 3 = penerapan, 4 = analisis, 5 = sintesis, 6 = evaluasi)
6. Kolom 6 diisi dengan PB/pembelajaran yang ada di GBPP
7. Kolom 7 diisi dengan nomor bab dan halaman dalam buku pelajaran yang dipakai (buku paket atau buku penunjang lain)
8. Kolom 8, 9, atau 10 diisi dengan cek menurut tingkat kesulitan soal, kalau dibandingkan dengan soal yang termasuk di dalam buku pelajaran

1

Rekap untuk GBPP

PB/Pembelajaran
Ada   Tidak Ada

Rekap untuk Buku Pelajaran

Topik
Ada   Tidak Ada

2

Appendix B

Interview on Cawu Testing
(Methodology 2)

WAWANCARA TENTANG TES CAWU

I. IDENTITAS RESPONDEN

A. Nama   : ....

B. Jabatan   : ....

C. Pendidikan terakhir : ....

D. Pengalaman : 1. Mengajar : .... Th

                2. Kepsek   : .... Th

                3. Pengawas : .... Th

Butir Pertanyaan	Responden		
	KS	Guru	Pgws
II. KETERLIBATAN DAN PERAN RESPONDEN 1. Kapan saja Saudara terlibat dalam penyusunan tes cawu? 2. Apa peran Saudara dalam penyusunan tes cawu tersebut? 3. Apa tugas yang Saudara kerjakan dalam peran tersebut? 4. Apakah kriteria seseorang untuk peran tersebut? 5. Bagaimana Saudara dipilih untuk peran tersebut?	v	v	v

III. PROSES PENYUSUNAN TES CAWU
A. Kelas, Mata Pelajaran, dan Tingkat Pengelolaan
  1. Pada kelas berapa tes cawu diberikan?
  2. Mata pelajaran apa yang diteskan?
  3. Siapa yang mengelola tes cawu? (Sekolah, Kancam, Kandep, Kanwil atau yang lainnya)
  4. Apakah setiap tes cawu pengelolanya sama?
  5. Apa jabatan orang yang memimpin pengelolaan tes cawu?
	v	v	v

B. Sumber dan Bahan
  1. Sebutkan sumber yang Saudara gunakan dalam menyusun soal tes cawu? (GBPP, Buku Paket, Kisi-kisi, Buku lainnya)
  2. Apa manfaat Saudara menggunakan sumber-sumber tersebut?
  3. Bagaimana cara menentukan materi/pokok bahasan dalam penyusunan soal tes cawu? Jelaskan!
  4. Bagaimana cara menentukan bentuk soal tes cawu? Jelaskan! (Pilihan ganda, uraian, hafalan, penalaran tingkat rendah, penalaran tingkat tinggi)
  5. Apakah bentuk soal yang dipakai dalam tes cawu dipengaruhi oleh bentuk soal yang dipakai dalam EBTANAS? Jelaskan!
  6. Apakah proses penyusunan soal tes cawu tersebut menggunakan kisi-kisi?
  7. (Jika ya) Siapa yang menyusun kisi-kisi?
  8. Dari mana kisi-kisi tersebut diperoleh?
  9. Adakah persiapan tambahan yang dilakukan sebelum menyusun soal tes cawu? Jelaskan!
	v	v	v

	KS	Guru	Pgws

C. Siapa Yang Terlibat dan Peranannya dalam kegiatan Penyusunan Soal
   1. Siapa saja yang terlibat dalam penyusunan soal tes cawu? (guru, guru inti, Kepsek, instruktur, pengawas, pemandu bidang studi)
   (**Catatan**: Jika responden penyusun soal, pertanyaan no. 2 & 3 tidak usah diajukan)
   2. Bagaimana seseorang menjadi penyusun soal tes cawu?
   3. Siapa yang menentukan seseorang sebagai penyusun soal tes cawu?
   4. Apakah penyusun soal tes cawu bekerja sendiri atau secara berkelompok (Tim)?
   5. Berapa orang yang menyusun soal tes cawu untuk setiap mata pelajaran dan untuk setiap kelas?
   6. Berapa kira-kira butir soal tes cawu yang disusun oleh setiap orang?
   7. Apakah soal tes cawu yang telah disusun ditelaah dan diperbaiki?
   8. Siapa yang menelaah dan memperbaiki soal tes cawu yang telah disusun? Bagaimana caranya?
   9. Apakah para penyusun atau penelaah berganti-ganti atau tetap untuk setiap cawu? Mengapa, jelaskan!
	v	v	v

D. Lamanya dan Tempat Penyusunan Soal Tes Cawu
   1. Pada umumnya berapa lama penyusunan soal tes cawu untuk masing-masing mata pelajaran?
   2. Di mana penyusunan soal tes cawu dilakukan? Mengapa?
   3. Apakah sebaiknya penyusunan soal tes cawu dilakukan di tempat khusus? (di rumah, hotel, sekolah, dll). Mengapa, jelaskan!
   4. Apakah waktu (lamanya) dan tempat penyusunan soal tes cawu berpengaruh terhadap kualitas tes cawu? Mengapa, jelaskan!
	v	v	v

E. Uji coba dan Penyempurnaan
   1. Apa yang dilakukan setelah soal-soal tes cawu ditelaah dan diperbaiki?
   2. Apakah soal tersebut diujicobakan? Mengapa?
   3. (Jika ya) Bagaimana cara mengujicobakan soal tersebut? Jelaskan! (kapan, di mana, siapa respondennya)
   4. Apakah ujicoba soal dilakukan pada saat tes cawu dilaksanakan?
   (**Catatan**: Pertanyaan no. 5 dan 6 diajukan jika pertanyaan no. 2 dijawab ya)
   5. Bagaimana cara menyusun tes ujicoba?
   6. Bagaimana memproses hasil tes uji coba menjadi tes cawu?
	v	v	v

	KS	Guru	Pgws
F. Kendala Selama Proses Penyusunan Tes Cawu 1. Kendala apa saja yang dialami selama proses penyusunan tes cawu? 2. Bagaimana cara mengatasinya?	v	v	v
IV. TINGKAT KESULITAN SOAL 1. Pada umumnya, bagaimana tingkat kesulitan soal tes cawu? (mudah, sulit) 2. Pada umumnya apakah tingkat kesulitan soal tes cawu sesuai dengan kemampuan siswa? Jelaskan! 3. Pada umumnya apakah tingkat kesulitan soal tes cawu sesuai dengan bahan yang diajarkan? Jelaskan! (**Catatan**: Jika jawaban no. 1, 2 dan 3 tidak ada masalah, maka pertanyaan no. 4 & 5 tidak perlu diajukan) 4. Mengapa terjadi hal seperti itu? 5. Bagaimana cara mengatasi hal tersebut? 6. Pada umumnya bagaimana hasil tes cawu di tempat Saudara?	v	v	v
V. KESESUAIAN SOAL 1. Secara umum, apakah soal tes cawu sesuai dengan GBPP? (Jika tidak) Jelaskan! 2. Apakah soal tes cawu sesuai dengan materi Buku Teks yang Saudara gunakan? (Buku Paket, Buku Pelengkap) (Jika tidak) Jelaskan! 3. Secara umum apakah soal tes cawu sesuai dengan tingkat perkembangan umur siswa? Jelaskan! (hafalan, penalaran tingkat rendah, penalaran tingkat tinggi)	v	v	v
VI. PENGARUH BENTUK SOAL TERHADAP KBM 1. Apakah bentuk soal tes cawu berpengaruh terhadap pola mengajar guru? Jelaskan! (Uraian, pilihan ganda → KBM) (Hafalan, penalaran → KBM)	v	v	v
VII. PENGGUNAAN HASIL TES CAWU 1. Hasil tes cawu digunakan untuk apa saja? Jelaskan! (Contoh: pelaporan, membuat peringkat, ketercapaian tujuan pembelajaran, memperbaiki KBM, kenaikan kelas, penilaian terhadap guru) 2. Siapa yang memanfaatkan hasil tes cawu? (Guru, Kepala Sekolah, Kancam, Pengawas, Kandep, Kanwil, Walikelas, orangtua siswa) 3. Di mana hasil tes cawu didokumentasikan? 4. Apa saran Saudara agar hasil tes cawu lebih bermanfaat?	v	v	v

	KS	Guru	Pgws

VIII. PENATARAN PENYUSUNAN TES
    1. Apakah ada pelatihan tentang penyusunan tes?
       Berapa kali dalam setahun?
    2. Siapa yang mengelola pelatihan tes itu?
    3. Siapa yang menyajikan materi dalam pelatihan itu?
    4. Berapa kali Saudara pernah mengikuti pelatihan tes?
    5. Bagaimana bentuk dan materi pelatihan yang pernah diikuti? (pelatihan umum atau khusus tes, ceramah, diskusi, simulasi, praktek dan lain-lain)
    6. Apakah pelatihan tersebut membantu Saudara dalam menyusun tes cawu? Jelaskan!
    7. Apa saran Saudara agar pelatihan penyusunan tes lebih bermanfaat?
	v	v	v

IX. KENDALA LAIN
    1. Sebutkan kendala lain yang Saudara hadapi berkenaan dengan tes cawu?
    2. Apakah saran Saudara untuk mengatasi kendala tersebut?
    3. Adakah sesuatu yang ingin Saudara kemukakan berkaitan dengan tes cawu? Sebutkan!
	v	v	v

Metodologi Review Tes Cawu

No. Soal	Materi soal	Cawu	Bentuk soal	Taxonomi Bloom	Sesuai dengan PB/ Pembelajaran di GBPP	Kesesuaian dengan buku pelajaran		
						Bab, hal	mudah	sedang	sulit
1	2	3	4	5	6	7	8	9	10