Using Research in Farm Publications

How does a farm paper happen to get started on research in the field of readership? Probably because a farm paper editor is likely to think in terms of experiments. Experiment station data on corn yields and hog feeding are the editor's daily diet. Why not apply the same methods to readers?

The only surprising thing about readership surveys in farm papers is that they came so late. It has been said, "Without readership surveys a farm paper editor is like a farmer who throws feed through a hole in the fence to hogs he never sees. He doesn't know whether they eat the feed or reject it. He doesn't know whether the hogs are gaining or losing."

Henry A. Wallace, from 1904 until he left Wallaces Farmer for Washington in 1933, was continually running tests on different strains of corn. Why not use similar methods on readership?

"We must get at it," said Wallace. But the actual work came after his time. What Wallace had done was
to make the staff alert to the experimental approach to any problem.

"Ted" Gallup, working in Des Moines on readership surveys in the 'twenties, started many people thinking about experiments of this kind. The Gallup and Roper opinion surveys in the election of 1936 helped to emphasize these possibilities.

It was 1938 before *Wallaces Farmer* started the *Wallaces Farmer* Poll and began to report on farm attitudes on elections and — more important to the paper — about farm reading habits.

Some editors insisted that readership surveys were not needed and that letters to the editor would give a picture of farm response to copy. We checked this several times. For instance, we asked farm people through the poll about their views on social security for farmers. A big majority approved. At the same time, we checked the letters on the subject. The letters only gave a 50-50 break to social security.

Ballots printed in the paper and sent in by readers also proved to be misleading. *Prairie Farmer* ran an experiment along this line and checked mailed-in-ballots against a personal interview survey. The two failed to match.

What *Wallaces Farmer* did, therefore, was to set up a polling system using a sample of around 400 interviews (200 men and 200 women) to check readership. The sample was distributed over the state according to the economic regions as defined by the U.S. Department of Agriculture. Interviews were made by farm women trained by the *Wallaces Farmer* staff.

What do we mean by readership? If the respondent (any adult on a farm into whose mailbox the publica-
tion is delivered) said that he remembered reading the issue in question and could identify one or more items as having been read, he was classified as a reader.

After a respondent was identified as a reader, the key question by the interviewer on each page of the issue was, "Did you HAPPEN to see or read anything on this page?" The word HAPPEN is stressed to support the interviewer's opening statement that there is nothing especially virtuous about readership. We do not want the non-reader of an article to feel guilty about being a non-reader.

The respondent's answer on any page is likely to fall into one or more of the classes below. In the first place, he will have a page score. Either he read or saw nothing on the page, or he did notice something. If he did notice something, he is given a score for "Any This Page."

Some possible reports on advertisements and articles are listed below:

<table>
<thead>
<tr>
<th>Advertisement</th>
<th>Article</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Nothing</td>
<td>1. Nothing</td>
</tr>
<tr>
<td>2. &quot;Any This Ad&quot; — Respondent has seen or read one or more features of this particular advertisement</td>
<td>2. &quot;Any This Article&quot; — Respondent has looked at head, picture, or read something in the article</td>
</tr>
<tr>
<td>3. &quot;Seen&quot; — Respondent has looked at a picture or a head</td>
<td>3. &quot;Seen&quot; — Respondent has looked at picture or head</td>
</tr>
<tr>
<td>4. &quot;Read Some&quot; — Respondent has read less than half of a particular piece of copy</td>
<td>4. &quot;Read Some&quot; — Respondent has read less than half of article</td>
</tr>
<tr>
<td>5. &quot;Read Most&quot; — Respondent has read half or more of a particular piece of copy</td>
<td>5. &quot;Read Most&quot; — Respondent has read half or more of article</td>
</tr>
</tbody>
</table>

An advertisement, therefore, might have one score for "Any This Ad;" another for "Seen" on the head; another "Seen" on the illustration; another score for
"Read Some" on a block of sales copy; another for "Read Most" on the same block of sales copy. Scores for men and women are always reported separately.

In the tables that follow, the figures given are always percentages of the sample used. When the men's sample is 200 cases, a score of 50 per cent, of course, means that 100 men responded in the way indicated. In split runs the A sample and the B sample each includes 100 men and 100 women. A score of 50 per cent means that 50 cases responded in the way indicated.

On opinion polls, the sample is larger and not so constant. On breakdowns of readership surveys, the sub-samples are smaller; and they vary. In each case, however, when there are exceptions to the rule noted in the paragraph above, the sample size is given.

While readership surveys (except in a few early surveys) always use a constant sample of 200 men and 200 women, opinion polls vary in size. The opinion sample ranges from 400 to 700 interviews in each state.

Opinion polls, of course, deal with a sample of all the farm men and women in each state. Readership surveys deal with a sample of the subscribers in each state.

One of the hazards of setting up your own survey machinery is that your interviewers may show a bias in favor of the paper that hires them. We tried to offset this in training sessions to point out necessity for keeping absolutely impartial approaches and comments.

We also checked our survey results against surveys made by independent operators. The Continuing Study of Farm Papers, conducted by the Advertising Research Foundation, ran a survey in the September 20, 1947 issue of Wallaces Farmer. (1)

Our survey crews checked the same issue independ-
ently. Was the *Wallaces Farmer* Poll getting higher readership scores than the Continuing Study? At our request, Professor Roscoe Giffin of Iowa State University went over the results and found that with men, in six cases our scores were higher than the Continuing Study. In 49 cases our scores were lower. In one case, they were exactly the same. With women, in eight cases our scores were higher than the Continuing Study and in 54 cases, lower. In one case, the score was the same.

Starch makes regular checks on readership of *Wallaces Farmer* and *Wisconsin Agriculturist*. Starch scores, as a rule, tend to run a little higher than ours, especially on ads. There is some difficulty in comparison because surveys are made in different months. In 1959, however, we had a readership survey of *Wisconsin Agriculturist* in October, and Starch had one of the same magazine in November.

Of 11 editorial departments, the Read Most scores compared as follows:

<table>
<thead>
<tr>
<th>OUR SURVEY</th>
<th>STARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
</tr>
<tr>
<td>Read Most</td>
<td>34%</td>
</tr>
<tr>
<td>Average of</td>
<td></td>
</tr>
<tr>
<td>11 departments</td>
<td>34%</td>
</tr>
</tbody>
</table>

This seems a reasonably close fit. On the whole, it does not seem that the readership scores in our surveys are biased because our interviewers do the work. It may help impartiality that our interviewers are trained in pre-election polls and know that errors from bias will show up.

A series of surveys by the Statistical Laboratory of Iowa State University at Ames gave us further information. These surveys, in 1947, 1951 and 1955, told where
farmers go to get information on different subjects. (2) They gave us a better picture of our subscribers, what they were like, and what kind of subjects interested them.

* * *

It should be kept in mind that throughout this book, the readership scores are given with readers of the issue as a base. Non-readers are also measured but, of course, in a different way.

A reader, as noted above, is any adult in the subscriber sample, who recalls one or more items in the issue. A non-reader is any adult in the sample of subscriber homes who reported that he or she had read nothing in the current issue.

While readership studies are built around readers of the current issue, the non-reader is also important. Professor Bryant Kearl, head of the Department of Agricultural Journalism at the University of Wisconsin, said, "A description of non-readers could be one of the most useful parts of a readership survey." (3) Chapter 9 goes into this.

* * *

What did we learn from the early readership surveys? Perhaps the first thing was to avoid jumps. We found that when an article started on page 10 and jumped to page 50, many readers were lost.

This seems obvious enough now. Since 1938, most publications in our field have abolished the jump. It is still used in magazines which use extra-long articles or stories. In these cases, however, the opening page or spread, featuring a big illustration and not much type, is really only an expanded plug.
In our early experiments, we found that we lost about 30 per cent of our readers when we had a sizable jump. If the jump were dressed up with a cut and a strong head, the loss was cut down.

In March, 1944, we ran another test. An article starting on page one scored 61.1 for men there; the runover on page 21 scored 49.7. Somewhat later, we tried starting an article on the right-hand page and continuing it on the following left-hand page. This lost readers, too.

A detailed report on this point came from the University of Iowa in 1958. Six articles in one publication started on the right-hand page and were continued on the following left-hand page. These runovers lost, on the average, 49 per cent with men and 25 per cent with women.

The answer, so far as we were concerned, seemed plain enough back in 1940. Complete every article on the page on which it starts. If the article is longer, let it run from a left-hand page to a right-hand one — but no farther. This policy, of course, meant more editing and more rewriting. But most farm publications, including ours, don’t do as much desk work on copy as they should anyway.

* * *

Readership scores were helpful in showing us which kind of copy was likely to be read and whether an article appealed to young readers, to old readers, to men, to women, to big farmers, to small farmers, to owners, to tenants, and so on.

What these surveys lacked was a definite comparison between different layouts, uses of color, styles of
cuts, placements of copy and other points. We could say that Ad A, in January, scored 40 per cent Noted with men, and that Ad B, in March, on the same kind of product, scored 30 per cent Noted with men. But did that mean Ad A was any better? A snow storm in March, a thaw in January, might have changed reader response to the two issues.

To give a fair test to editorial copy or to ads, it seemed necessary to expose this copy to readers at exactly the same time. This meant using what we called a "split run."

The term "split run" has since changed its meaning for many. Now it often indicates that Magazine A will run special copy in, for example, Illinois and Indiana. The same pages will carry special and different copy for Ohio and Pennsylvania.

Our "split run" goes back instead to the practice of newspapers with mail-order ads. With newspapers, every other one that came off the press would have different copy. Ad A would go to half the readers; Ad B to the other half. The results were measured by coupon return.

This worked well for mail-order ads. It was no help to other kinds of advertisers, to whom coupon returns were not important.

What we did was to set up two samples in each state. Think of Iowa, with its 99 counties, as a checkerboard. We sent A copies to the red counties, and B copies to the black counties.

Actually, we never used this big a sample. Ordinarily, we interviewed in about 20 A counties and 20 B counties. The interviews in the A counties were distri-
buted proportionately among the five economic regions of the state. The same was done in the B counties. We ended up with two samples, closely matched — 100 interviews with men and 100 interviews with women in the A counties and the same number in the B counties.

* * *

We tried our first split in 1946 to test readability levels. When Rudolph Flesch published his doctoral dissertation at Columbia on the subject back in 1944, we got one of the copies and began to wonder about its applicability to our problems. With his *Art of Plain Talk* (Harpers, 1946) we settled down to test his theories.

The Flesch hypothesis was that copy with short sentences, short words (few affixes) and “personal” words would attract and hold more readers than copy with longer sentences, longer words and fewer “personal” words.

In the March 1, 1946 issue of *Wallaces Farmer*, we ran three splits based on the Flesch formula. The main thing we learned from this was that we were shooting too high. We moved — in Flesch’s words — from a seventh grade level to a sixth grade level. Our readers didn’t notice the difference.

Only when we moved to a much simpler level did the new copy take hold. The Flesch index of 1.5 seemed to increase readership. What did this mean? The copy would average around 12 words per sentence, 20 affixes for 100 words and 10 personal references per 100 words.

In November, 1946, we tried again. This time we split three men’s articles and one woman’s article.

Remember that nothing was changed except the
style. The head, the illustration and the theme had not been altered. The copy was edited only to get different levels of readability in terms of the Flesch index.

In the four splits, the low Flesch copy ranged from 1.11 to 1.76. The high Flesch counts ranged from 2.48 to 4.27.

Of the eight comparisons (using men and women for each split), we couldn't use three. On two pieces of copy designed for men, the women's score was too low to provide any answers. On one piece of copy the men's scores were too low for us to use.

Of the five split comparisons, where the number of readers was fairly high, one piece of men's copy showed a loss of 9.4 per cent for the low Flesch score. With the other four, increases ranged from 7.3 to 66 per cent for the low Flesch score over the high score.

We have since tried similar splits in Wisconsin Agriculturist and in later issues of Wallaces Farmer. Experiments at the University of Iowa made another check. The results seemed clear enough. Other things being equal, simple language scored high.

Does this mean we try to write for morons? Not at all. An experiment at Iowa State University at Ames threw some light on this delusion. A split, using Flesch scores, was tried out on faculty members and students. Presumably the faculty members were the intellectuals. Yet the simpler Flesch copy did better with the faculty than with the students.

Why? The copy was in a field of more interest to the students than to the faculty. When readers are excited about a subject, they'll read difficult copy, printed in small type. When their feelings are neutral, they'll
respond better to readable copy. If you can pick subjects of overwhelming interest, you can write badly and get away with it. Nobody is that good a picker. It is better to assume that some of your readers may be indifferent.

As members of our staff promptly pointed out, you can write very bad copy that has a good Flesch score. "I see a cat. Do you see a cat?" scores well on the Flesch index.

To avoid disasters like this, our rule was to write the copy as well as we could. Remember what the teacher said in English 1—short, easily understood words, action verbs and not too many adjectives, specific and colorful descriptions, questions and names.

After writing—and often after publishing—we went over the copy with the Flesch index. The staff noted the scores and sometimes remembered them next time.

The extent of improvement in style may be measured in this way. Before we started testing, we took our usual copy (around 3.5 Flesch) and edited to bring it down to 1.5. After some months of education, we found that our usual copy was around 1.5 Flesch. To get a split, we had to edit to bring one version up to 3.5.

One source of confusion in using the Flesch index is the fact that the author changed his measuring device. In the earlier Flesch scoring system the low score (note 1.5 above) was the best. In The Art of Readable Writing Flesch uses a measuring stick called "Reading Ease." (6) Here the high score is the best. The Reading Ease score is based on syllables per 100 and on words per sentence.
A meritorious Reading Ease score would be 80 (much like the old Flesch index of 1.5). This would mean 12 words per sentence and 134 syllables per 100 words. A “difficult” Reading Ease score might be around 40 with 20 words per sentence and 173 syllables per 100 words.

Just to see if readability (in Flesch terms) was still important, we ran a split in *Wallaces Farmer* for January 16, 1960 on the article entitled “USDA Studies Soil Acidity” (Figures 1.2 and 1.3). Following is the Reading Ease Index and the resulting scores for men:

<table>
<thead>
<tr>
<th></th>
<th>A (Difficult)</th>
<th>B (Easy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words per sentence</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td>Syllables per 100 words</td>
<td>170</td>
<td>141</td>
</tr>
<tr>
<td>Reading Ease Index</td>
<td>56.0</td>
<td>74.6</td>
</tr>
<tr>
<td>Read Some on copy</td>
<td>32.0%</td>
<td>47.0%</td>
</tr>
<tr>
<td>Read Most</td>
<td>30.0%</td>
<td>43.0%</td>
</tr>
</tbody>
</table>

It may help to show what happened here if we quote the lead from each version of the copy: (A was a USDA release)

**Version A — 56 Reading Ease**

The strong trend to heavier nitrogen fertilization, coupled with reduced use of lime over the past decade, is making our nation's soil more acid. That's why USDA has expanded its research into the problem of acidity.

**Version B — 74 Reading Ease**

Are you planning to put a lot of nitrogen fertilizer on your fields this spring? If you are, make sure also that you've spread enough lime on these same fields.

How do you determine “Reading Ease,” according to Flesch? His *Art of Readable Writing* has a time-saving chart. The Flesch formula — if you don’t use the chart — goes like this:

Multiply the average sentence length by 1.015;
Multiply the number of syllables per 100 words by 0.846; Subtract the sum of the two items above from 206.835. What is left is Reading Ease.

It is a good deal easier to remember that you can get a Reading Ease score of 75 — which is pretty good — in the following ways:

15 words per sentence and 138 syllables per 100 words
or
12 words per sentence and 142 syllables per 100 words
or
18 words per sentence and 134 syllables per 100 words.

On the other hand, if your Reading Ease score falls below 50, you are probably losing readers. A score of 50 can be obtained in the following ways:

18 words per sentence and 164 syllables per 100 words
or
20 words per sentence and 162 syllables per 100 words
or
25 words per sentence and 156 syllables per 100 words.

It is still a good idea to get a copy of the book and use the Flesch chart. If you are hitting somewhere between 60 and 80 on the Reading Ease Index, you are doing fairly
well. But, note that we have had very high scoring copy — Read Most scores of 60 per cent or better — that had a Flesch rating of 65. And we have had low scoring copy — Read Most of 30 per cent or less — that had a Flesch rating of 80.

All this means is that subject matter is always more important than style. A hog article, in Iowa, will always outscore a sheep article. But a sheep article with a Flesch index of 80 will usually outscore another sheep article with a Flesch index of 55 or less.

There are, of course, other methods, and excellent ones, of scoring readability. We have stuck to Flesch because we happened to start with it and because it has stood up under split-run testing.

This was our start in split-run testing. I have reported it in some detail because it illustrates the methods used in many later experiments. The split-run device has thrown light on many problems in advertising and editorial customs.

Each split-run reader-interest survey for some years has had five to seven splits with ads and the same number or more with editorial matter. Some of these proved that the differences we expected did not exist. Some showed a sharp reader response to a change of layout or copy.

Succeeding chapters go into detail on some of the things that advertisers and editors learned.
Who Reads Editorials?

Do subscribers read the editorial page? What kind of subscribers?

Above is the first editorial page (Wallaces runs a spread of two) in a fall issue of Wallaces Farmer. Men read more editorials than women. As you might expect, the men's first choice was the editorial on hog prices, with a Read Most of 54.5 per cent.

Young men (21-34 years) read about as much as older men (50 and up). Farm men with gross incomes of $10,000 a year read more than men with smaller incomes.

Wallaces Farmer, November 19, 1960
USDA studies soil acidity

*Research is intended to show how acidity can cut yields*

The strong trend to heavier nitrogen fertilization, coupled with reduced use of lime over the past decade, is making our nation's soils more acid. That's why USDA has expanded its research into the problem of acidity.

While these experiments are not being conducted on Iowa soils, the results may prove helpful to Iowa agronomists. Scientists have taken only the first few steps in fully understanding soil-plant interactions. When they unlock enough of this information, fertilizers and fertilizer use can be adapted more closely to plant needs. And plants can be bred to make better use of their environment.

More than 50 field experiments on 23 agriculturally important soil types are now being conducted in cooperation with state experiment stations to fill in the gaps of our understanding of how acidity decreases yield. These same soils are also studied under uniform environmental conditions in greenhouse and laboratory to learn why they behave as they do.

Although this research is only in its second year, it has already given some interesting leads. For example, alfalfa was grown in the greenhouse on two silt loams that showed the same lime requirement when tested by accepted soil tests.

When limed, one soil gave a 400 percent increase in yield but the other gave only a 25 percent increase. Understanding the reason for such differences in behavior among soils would be one of the keys to accurate prediction of lime requirements.

Toxic concentrations of manganese and aluminum are probably the most common cause of poor plant growth in acid soils. There is no doubt in the scientists' minds that acidity is serious, but there are still questions as to how it causes known reactions.

Is aluminum or manganese toxicity the more important problem? Under what conditions and for which crops and soils are these toxicities of particular significance?

What are the interactions of nutrients such as phosphorus with toxic levels of aluminum and manganese?

How and to what extent does organic matter alleviate the toxicity of these elements? What are the critical levels of these elements in important crops?

Soil acidity is a serious soil management problem in the cornbelt, says ARS soil scientist R. W. Pearson.

From 1947 to 1957, the use of nitrogen fertilizer over the nation jumped from less than a million tons to 2.3 million tons a year. But the use of limestone dropped over a third, from 28 million to 18 million tons in 1954, and had only increased back to 22 million by 1958. Some 80 million tons are estimated to be needed annually to maintain optimum soil reaction.

No, They Aren't the Same!

The two articles here look alike, but they score differently. They have the same head, the same theme and the same structure. But they differ in the number of long words and long sentences.

In a series of splits, of which this is the most recent, copy with short sentences and short words has been shown to help readership.
USDA studies
soil acidity

ARE YOU planning to put a
lot of nitrogen fertilizer on
your fields this spring? If you
are, make sure also that you’ve
spread enough lime on these
same fields.
Sour soils take a lot of the
vigor out of fertilizers—especial-
ly nitrogen. Unless a field is
limed, you won’t get anywhere
near full value out of your fer-
tilizer.
Cornell soils are getting more
sour. At the same time that fer-
tilizer use has gone up, use of
time has gone down. In Iowa,
for instance, there was a 24 per-
cent drop in the use of lime-
stones from 1957 to 1958.
Nationally, the same thing is
happening. From 1947 to 1957,
the use of limestone dropped
ever a third. At the same time,
the use of nitrogen fertilizer
more than doubled.

To keep the nation’s soil
sweet, farmers need to
spread around 80 million
tons of limestone a year. In
1955, only 22 million tons
were spread.

Most farmers know that sour
soils make it harder to get a
stand of legumes. Not all realize
that the value of their fertilizer
may be cut heavily when applied
to sour soils.

To find out just what happens
to crops in sour soils, a number
of experiments are being car-
ded on. Over the nation, 50
different field experiments are
being carried on with 23 soil
types.

Some of the answers so far
just raise additional questions.
In one greenhouse experiment,
two silt loams were tested. Ap-
parently the two had the same
limestone requirements. But see what
happened:

When limed, one soil gave
a 400 percent increase in
yield. The other gave only
a 25 percent. Nobody yet is
sure why.

Two of the villains in sour
soils are probably manganese
and aluminum. Sour soils usu-
ally have too much of these
minerals. Crop growth slows down.

But does manganese do most
of the damage? Or is aluminium
a major culprit?

What happens when you add
phosphorus to a sour soil over-
loaded with manganese and
alumina?

Does it help sour soil of this
kind to plow under green ma-
nure to add a lot of barnyard
manure? Or does such action
make things worse?

These are some of the ques-
tions scientists are asking them-
selves. No final answers have
been determined yet.

In the meantime, it does seem
certain that sour soils don’t get
the big benefits out of nitrogen
fertilizer. When you plan your
program for 1960 crops, you may
want to use lime as well as
nitrogen.

The research is only in
its second year, many interesting
loans have already been pro-
duced. In the future, this re-
search should answer lots of
questions on soil acidity.

Wallaces Farmer, January 16, 1960

Here is the difference in the two pieces of copy:

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words per sentence</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td>Syllables per 100 words</td>
<td>170</td>
<td>141</td>
</tr>
</tbody>
</table>

Editing A copy to the B standard raised the Read Most score for men from 30 in A to 43 in B. Read Some scores showed a similar gain.
Recipes Pull Women In

Farm women continue to read food copy, especially if there is a local angle. This Home Department lead page addressed to "Young Cooks" actually scored almost as well with older women as it did with younger ones.

21-34 years 50 and up

Read Most . . . . . . 84.8% 76.3%

The photo (upper right) of the Wisconsin farm girl, Karen Ulness of Manitowoc County, drew the attention of 90 per cent of the women readers of the issue.

Wisconsin Agriculturist, April 15, 1961